

On Human-Computer Interaction in Complex Artefact Ecologies

Clemens Nylandsted Klokmose

PhD Dissertation



Department of Computer Science
Aarhus University
Denmark

On Human-Computer Interaction in Complex Artefact Ecologies

A Dissertation
Presented to the Faculty of Science
of Aarhus University
in Partial Fulfilment of the Requirements for the
PhD Degree

by
Clemens Nylandsted Klokmose
August, 2009

Abstract

The personal computer has lost its monopoly as our sole access-point to the digital world. Today we surround ourselves with a myriad of different interactive artefacts with vast computational capabilities. Modern music players, gaming consoles, entertainment systems, smart phones, netbooks, and e-readers offer internet access and functionality that a few years ago only were accessible on chunky beige desktop computers. Seen through the optics of an HCI researcher, this development is magnificent – but it also challenges the foundation of our field.

Our analytical vocabulary for describing human-computer interaction and the frameworks and tools we use to build interactive software rest on a conceptual foundation that takes personal computing for granted, and are therefore challenged by the plurality of heterogeneous interactive artefacts that are used interchangeably or in conjunction. In this PhD dissertation I present a *theoretical* and *technological* perspective of the challenges of human-computer interaction in these new landscapes of multiple, heterogeneous interactive artefacts, what I refer to as *Human-Computer Interaction in Complex Artefact Ecologies*.

I propose *The Human-Artefact Model*, which is an activity theoretical model for structuring design oriented analysis of interactive artefacts that are, or are to be, part of complex artefact ecologies. By reinterpretation of the activity theoretical foundation, I present a framework that helps better address mediators in plural. I show how the human-artefact model helps structure the understanding of the action possibilities of and in relation to the artefact ecology which surrounds it. Essential to the model is that it provides four interconnected levels of analysis and addresses the possibilities and problems at these four levels of activity.

From the technological perspective I propose an alternative interaction paradigm inspired by activity theory and based on Beaudouin-Lafon's work on *Instrumental Interaction* and I propose an architectural model for realising the paradigm. The interaction paradigm *Ubiquitous Instrumental Interaction* pushes forth an idea of interaction as being mediated by dynamic configurations of instruments applicable across different domain objects and interactive artefacts, hence creating an interaction environment that transcends the isolated interactive artefact. I present *VIGO* (Views, Instruments, Governors, and Objects), which is an architectural model embodying the principles of the new paradigm. I show that it offers new technical and interactional potential to HCI.

Acknowledgements

First of all I thank my supervisor Susanne Bødker for her guidance and support, and for creating and being the anchor point in our a great research community. I thank Michel Beaudouin-Lafon and Wendy Mackay for an inspiring 5 months in Paris. Furthermore I thank all the people in the lab in Paris for creating a both great personal and professional experience.

I thank Pär-Ola Zander and Olav Bertelsen for always being up of for a theoretical discussion. I thank Kasper Langer for encouraging me, being the realist critic of my ideas, and teaching me all sorts of programming tricks. I also thank Rasmus Berlin for programming on the initial prototypes.

I owe a great thank the co-authors of the included papers. And when I refer to what “I” have done in this dissertation, I of course also refer to my co-authors.

I would like to credit those who have given comments on earlier drafts on this dissertation. In alphabetic order: Susanne Bødker, Brian Bunch Christensen, Peter Dalsgaard, Tony Gjerlufsen, and Niels Mathiasen. Jonathan Bunde-Pedersen is thanked for tips on writing a dissertation, and for help with L^AT_EX. Hanne Kjærgaard Larsen is thanked for improving the English of the dissertation.

Finally thanks to Maria and my parents Kirsten and Arne for their endless love and support.

This PhD have been funded by The Danish Technical Research Council (Danish: Statens Teknisk-Videnskabelige Forskningsråd) under the project *Ubiquitous User Interface Design – a theory drive approach to ubiquitous user interfaces and their design* (26-04-0137).

*Clemens Nylandsted Klokmose,
Aarhus, August, 2009.*

Preface

This dissertation is divided into two parts: The first part provides a coherent overview of the work of this PhD project. The second part is a collection of five published papers and one submitted journal paper.

The overview part is divided into five chapters. Chapter 1 is an introductory chapter that motivates the research, presents research objectives, provides an overview of the central contributions, and discusses research approach and methodology. Chapter 2 presents the conceptual and theoretical foundation of my work and discusses the role of theory in HCI. Chapter 3 and 4 covers the work presented in the appended papers and discuss related work. The presentation of the work is broken up in two themes: *Complex artefact ecologies from without* (Chapter 3) covers work on analysis and conceptualisation of the type of HCI addressed in the dissertation. *Complex artefact ecologies from within* (Chapter 4) presents work on rethinking the way we build interactive systems to accommodate the new challenges facing HCI. Chapter 4 is slightly more content heavy than chapter 3 since it is based on a space-limited conference paper (Paper V) while the work of chapter 3 has been discussed in detail in the journal paper I. Chapter 5 concludes the overview part and discusses future work.

Included Papers

- [1] Susanne Bødker and Clemens Nylandsted Klokmoose. The human-artifact model – an activity theoretical approach to artifact ecologies. Submitted to Human-Computer Interaction.
- [2] Niels Olof Bouvin, Christina Brodersen, Susanne Bødker, Allan Hansen, and Clemens Nylandsted Klokmoose. A comparative study of map use. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 592–597, New York, NY, USA, 2006. ACM.
- [3] Christina Brodersen, Susanne Bødker, and Clemens Nylandsted Klokmoose. Ubiquitous Substitution. In *Proceedings of INTERACT 2007*, pages 179–192. Springer, 2007.
- [4] Christina Brodersen, Susanne Bødker, and Clemens Nylandsted Klokmoose. Quality of learning in ubiquitous interaction. In *ECCE 2007: Proceedings of the European Conference on Cognitive Ergonomics*, pages 121–128. ACM New York, NY, USA, 2007.
- [5] Clemens Nylandsted Klokmoose and Michel Beaudouin-Lafon. Vigo: instrumental interaction in multi-surface environments. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 869–878, New York, NY, USA, 2009. ACM.
- [6] Clemens Nylandsted Klokmoose and Par-Ola Zander. Rethinking laboratory notebooks with ubiquitous instrumental interaction. In *IRIS32: Proceedings of the 32nd Information Systems Research Seminar in Scandinavia*, August 2009.

Contents

Abstract	v
Acknowledgements	vii
Preface	ix
Included Papers	xi
I Overview	1
1 Introduction	3
1.1 Motivation	3
1.2 A word on complex artefact ecologies	5
1.3 Research Objectives	5
1.4 Overview of Papers and Contributions	6
1.5 Research context	8
1.6 Research Approach and Methodology	10
2 Conceptual Foundation	15
2.1 Why care about theory?	15
2.2 Activity Theory	17
2.3 Artefacts	21
2.4 The orienting basis of the user	25
2.5 Other postcognitivist frameworks	26
2.6 Instrumental Interaction	28
3 Complex Artefact Ecologies from Without	31

3.1	Chapter overview	32
3.2	Empirical point of departure	32
3.3	Three themes	33
3.4	The Human-Artefact Model	37
3.5	Related Work	45
3.6	Summary	50
4	Complex Artefact Ecologies from Within	51
4.1	Chapter overview	51
4.2	Elaborating the problem	52
4.3	Ubiquitous Instrumental Interaction	56
4.4	VIGO	59
4.5	Some implementation considerations in VIGO-based systems	72
4.6	Elaborated vision: Laboratory notebooks	74
4.7	Related work	80
4.8	Summary	86
5	Conclusion	89
5.1	Research objectives revisited	89
5.2	Future Work	91
II	Papers	95
I	The Human-Artefact Model – an Activity Theoretical Approach to Artifact Ecologies	97
1	Introduction	98
2	Artifacts	103
3	Design Case	120
4	Structuring a design-oriented analysis through The Human-Artefact Model	124
5	Discussion and related work	143
6	Conclusion	147
II	A Comparative Study of Map Use	149
1	The study rationale and focus	149

2	The three interaction technologies	150
3	The Experimental Setup	150
4	The Analysis	151
5	Examples	152
6	Conclusion and lessons for redesign	156
III	Ubiquitous Substitution	157
1	Introduction	157
2	Human Activity and Functional Organs	159
3	Mediation and Functional Organs	161
4	Analyzing Functional Organs	162
5	Ubiquitous substitution	169
6	Discussion	171
7	Acknowledgements	172
IV	Quality of Learning in Ubiquitous Interaction	173
1	Introduction	173
2	Theoretical Background	174
3	Unfolding the learnable interface	177
4	Designing the learnable artefact	182
V	VIGO: Instrumental Interaction in Multi-Surface Environments	189
1	Introduction	189
2	Ubiquitous Instrumental Interaction	190
3	Related work	192
4	The <i>VIGO</i> architecture	193
5	Using VIGO: Realizing Pick-and-Drop	199
6	Discussion and qualitative evaluation	206
7	Conclusion and future work	208
8	Acknowledgements	209
VI	Rethinking Laboratory Notebooks with Ubiquitous Instrumental Interaction	211
1	Introduction	212
2	Method and Participants	213

3	Thinking tool: Ubiquitous Instrumental Interaction	214
4	The laboratory notebooks	216
5	Rethinking the analogue laboratory notebook	218
6	What is a laboratory notebook?	223
7	Conclusion	228
8	Acknowledgement	229
	Bibliography	231

Part I

Overview

Chapter 1

Introduction

This dissertation concerns human-computer interaction that spans multiple and heterogeneous interactive artefacts and is not directed towards a personal computer alone. I refer to this phenomenon as *human-computer interaction in complex artefact ecologies*. I approach this phenomenon from two perspectives; a theoretical and a technological. The theoretical perspective is represented by the chapter *Complex artefact ecologies from without*, where the central contributions are *The Human-Artefact Model* and a revisiting of activity theory given the challenges of complex artefact ecologies. The central contributions from the technological perspective represented by the chapter *Complex artefact ecologies from within*, are the proposal of a new interaction paradigm, *Ubiquitous Instrumental Interaction*, and the architectural model *VIGO* (Views, Instruments, Governors, and Objects) for realising it.

The human-artefact model is an activity theoretically informed model for framing design oriented analysis of artefacts in complex artefact ecologies. The human-artefact model acknowledges that humans are embodied in ecologies of artefacts and encourages the user of the model to consider the dialectical relationship between humans and artefacts. *VIGO* is an architectural model based on *instrumental interaction* [18] for developing user interfaces spanning multiple interactive artefacts. *VIGO* is a reflection on the software architectural requirements for realising the *ubiquitous instrumental interaction* paradigm. Ubiquitous instrumental interaction is an interaction paradigm that embraces the activity theoretical emphasis on mediation and development and breaks with, what is today, the all pervasive and completely taken-for-granted application concept.

1.1 Motivation

The desktop computer has lost its monopoly as our only means of access to the digital world. Today we surround ourselves with a myriad of different

interactive artefacts with vast computational capabilities. Modern music players, gaming consoles, entertainment systems, smart phones, netbooks, and e-readers offer internet access and functionality that a few years ago only were accessible on chunky beige desktop computers. Seen through the optics of an HCI researcher, this development is magnificent – but it also challenges the foundation of our field.

Our analytical vocabulary for describing human-computer interaction and the frameworks and tools we use to build interactive software rest on a conceptual foundation that assumes personal computing in the form of one artefact – one application – one user. The move from the first to the second wave of HCI was based on an insistence on the human actor as a centre of attention [12]. Second wave theories (surveyed e.g. in [56]) provided a rich array of concepts and methods for designing and analysing interaction from the outset of the use situation. Bødker [37] argues that the third wave of HCI has brought forth a much welcomed focus on the cultural, aesthetic, and emotional aspects of interactive artefacts, however, hereby also drifting away from a commitment to users towards a more exploratory take-it-or-leave-it or lets-see-what-happens approach (e.g. through cultural probes [77]). We still need means for systematically structuring design and analysis of individual artefacts – while acknowledging that the artefacts are not alone. The aim of this dissertation is to contribute to a conceptual foundation for HCI that embraces the diversity and heterogeneity of our new artefact landscape. This conceptual framework must embrace that users navigate complex ecologies of artefacts, where the use of one artefact cannot be fully understood in isolation.

From the perspective of technology we now have almost all of the interactive artefacts proposed by Mark Weiser in his famous ubiquitous computing vision [170]. But we are still far from a situation where we can easily move our work between our interactive artefacts or exploit co-location of artefacts to distribute a user interface without significant overhead. Accessing documents and files created on a desktop computer on a smart-phone in a meaningful manner is not trivial. Nor is it possible to easily exploit co-location of artefacts for instance to use your laptop keyboard to write an SMS on your smart-phone [135] without having tailor-made software on both artefacts. This problem might stem from the fact that even though we have all these new artefacts, we still architect the interactive software running on them as were they personal computers – each a small isolated and self-contained entity. This leads to a situation where the handling and operation can differ radically between artefacts, creating the basis for potential breakdowns in use when artefacts are substituted for each other.

The motivation for this dissertation is twofold: Firstly to contribute to our analytical understanding of human-computer interaction in complex artefact ecologies. Secondly to question the way we architect interactive software today and propose an alternative approach that acknowledges that we have moved beyond the desktop computer, where interactive artefacts should not be understood as isolated and self-contained entities.

1.2 A word on complex artefact ecologies

Ask not what's inside your head, but what your head's inside of.

— William M. Mace

The word *ecology* is borrowed from Gibson's book *The Ecological Approach to Visual Perception* [79]. Gibson is famous for rejecting the mind-matter dualism and arguing that humans (or animals) should be understood in system with the environment. Mace [113] eloquently captures Gibson's approach to psychology in the above quote. In the Gibsonian sense, our (visual) perception is shaped by our physical ecology and cannot be understood in isolation. The ecology of a subject is the part of the physical world that it interacts with to realise its life. Following Gibson's definition of ecologies, Jung et al. [94] defines a person's *ecology of artefacts* as the artefacts applied to realise her activities. Jung et al. show in their study of people's ecologies of interactive artefacts that their use influence each other. A desktop computer is used differently if a user also has a laptop. Hence this provides fuel to the argument that artefacts cannot be understood in isolation.

However, Until recently this has not been a particularly prominent issue in HCI. With the introduction of the personal computer in the beginning of the eighties a radically new kind of artefact was brought into our artefact ecology. This artefact increased its mediation of our daily activities up to a point where many occupations, especially knowledge workers, could be almost completely mediated by this single artefact. Hence, in the optics of HCI, the artefact ecology of users has been simple. In order to understand HCI we only had to understand one artefact, the PC. During the last decade non-PC interactive artefacts have become more powerful and more common-place. The ecology of interactive artefacts has gone from being simple (a single personal computer) to being complex (music players, gaming consoles, PCs, entertainment systems, interactive walls etc.). When I refer to complex artefact ecologies it is implied I generally refer to *interactive* artefacts. Interactive artefacts have digital computing capabilities, take input from the user and provide feedback in some modality perceptible by the user (e.g. sound or vision). In this dissertation I will refer to interactive artefacts as both the physical artefact (e.g. a cell-phone) or the artefact and the application running on it (e.g. a map-application running on a cell-phone).

1.3 Research Objectives

The overall research agenda of this dissertation work has been to broaden our understanding of human-computer interaction beyond the desktop computer. I have approached this through two focus points, one theoretical, and one technological. These focus points can be formulated as two research objectives:

- **Theoretical Objective:** To enrich our analytical vocabulary of artefacts

that are, or are to be, part of complex artefact ecologies, and to provide means for structuring design-oriented analysis of such artefacts.

- **Technological Objective:** To reconsider the current predominant interaction paradigm and propose an alternative paradigm that embraces the heterogeneity of complex artefact ecologies. Furthermore to show that it is realisable in principle and that it provides new opportunities for HCI.

The theoretical objective is approached from an outset in activity theoretical HCI. The HCI community in Scandinavia and especially in Aarhus has a well-established tradition of activity theoretical HCI, with Bødker's *Through the interface* [38] as one of its seminal works. Dourish [66] points out that the move beyond the desktop computer calls for a revisiting of our theoretical understanding – one that acknowledges the user as embodied in the world and technology. The activity theoretical approach to HCI was also conceived in a context of desktop computing, therefore it also needs to be revisited. Where Dourish provides a philosophical understanding of the human role in the new technological landscape through Heidegger's (and other philosophers') phenomenology that leads to a number of design recommendations (discussed in Section 3.5), activity theory can provide applicable analytical concepts. Activity theory emphasises a dialectical understanding of the relationship between human beings and artefacts. Thus it is fundamental to the activity theoretical approach that artefacts cannot be understood in isolation.

I approach the technological objective by rethinking the way we interact with digital objects through the lens of *instrumental interaction*. Beaudouin-Lafon [18] proposes instrumental interaction as an alternative interaction model to direct manipulation [153]. Beaudouin-Lafon advocates viewing interaction on computers as being mediated rather than direct. Interaction is seen as mediated by an instrument (e.g. a scrollbar) onto a domain object (e.g. a document). Instrumental interaction is inspired by the activity theoretical emphasis on mediation. While instrumental interaction was originally conceived as a descriptive model, I explore applying the principles as an interaction paradigm for HCI in complex artefact ecologies, and as a principle for conceptualising a software architecture realising the paradigm.

1.4 Overview of Papers and Contributions

1.4.1 Theoretical

The central contributions of the theoretical part of the dissertation are presented in Paper I. The contribution is an activity theoretical foundation for HCI in complex artefact ecologies crystallised into the *human-artefact model*. The human-artefact model is a model for structuring design oriented analysis of artefacts that are to be integrated in complex artefact ecologies. What characterises the human-artefact model is its emphasis on the dialectical rela-

tionship between human beings and artefacts, and encouragement of the user (analyst, designer, or research) of the model to think along these lines.

The contributions of the papers [III](#) and [IV](#) are largely subsumed by [Paper I](#). However, I have included them as they provide insight into the iterative process leading to the human-artefact model.

- [Paper I](#) presents the major research contributions of this part of the dissertation. The paper collects and refines the theoretical developments of the two conference papers, [Paper III](#) and [Paper IV](#), and provides a activity theoretical foundation for HCI in complex artefact ecologies. This foundation is crystallised into the Human-Artefact Model. In the paper we illustrate the use of the human-artefact model by applying it on a fictional design case based on a study of map use on three different artefacts (presented in [Paper II](#)).
- [Paper II](#) is an extended abstract/work-in-progress paper that presents an empirical study on the handling of three kinds of maps, each on a physical artefact: a paper map, a tablet-PC based map, and a cellular phone based one. The study was executed before I began my dissertation work. However, I transcribed and analysed video-data from the study. In the paper we provide preliminary analyses, which are elaborated in paper [I](#), [III](#), and [Paper IV](#).
- In [Paper III](#) we present an activity-theoretically based framework for analysing what we refer to as *ubiquitous substitution*. In the terminology of this dissertation we address substitution of artefacts in complex artefact ecologies for mediating similar activities¹. With the framework we develop a four-layered analysis of artefacts where we highlight tensions between expected and intended uses and the conflicts encountered when attempting substitution between them. The empirical foundation is based on the study presented in [Paper II](#).
- In [Paper IV](#) we continue our development of an activity theoretical framework for addressing interaction in complex artefact ecologies by discussing *quality of learning*. We apply Gal'perin's [\[75\]](#) theory on learning on examples from the map study and address the design of learnable artefacts through a discussion of a previous design case on IT support at a wastewater plant (see e.g. [\[31\]](#)).

1.4.2 Technological

The central technical contribution of this part of the dissertation is the VIGO architectural model. VIGO is presented as an alternative to model-view-controller

¹During this research project we refined our vocabulary slightly – when the papers refer to *ubiquitous interaction* this should be understood as interchangeable with *HCI in complex artefact ecologies*.

(MVC) [143], it embeds the principles of the proposed interaction paradigm *ubiquitous instrumental interaction* and is the seed for a technical realisation of it. In [Paper V](#) we illustrate how VIGO enables us to easily implement ubicomp techniques that would be difficult to implement on an application-based interaction architecture.

A second contribution is presented in [Paper VI](#). Here we show that applying the vision of ubiquitous instrumental interaction and the concepts of the VIGO architecture enables a group of physicists to envision a laboratory notebook system that breaks with the limitations of traditional applications.

- [Paper V](#) presents the VIGO (Views, Instruments, Governors, and Objects) architectural model for realising multi-surface interfaces. VIGO is based on the vision of an alternative interaction paradigm to the application paradigm based on instrumental interaction [18]. In the paper we show how a classical ubicomp interaction technique, pick-and-drop [144], is easily implemented on a VIGO based architecture and augmented to support moving what resembles live applications between interactive artefacts.
- In [Paper VI](#) we present a pilot study where the ideas of [Paper V](#) are applied as thinking tools for re-envisioning laboratory notebooks. The study involves a group of five experimental physicists who we engage in future workshops and interviews. The paper discusses the role of laboratory notebooks conceptualised as a syntonic seed [174], why it challenges the application paradigm and how a VIGO based system could potentially overcome some traditional limitations.

1.5 Research context

In this section I give an overview of the research context my dissertation work is positioned in. For a detailed overview of related work I refer to the two chapters 3 and 4.

The theoretical part of this dissertation positions itself in a (relatively) long tradition of applying philosophical and psychological concepts to HCI. Books such as Card, Moran, and Newell's *Psychology of Human-Computer Interaction* [55], Winograd and Flores' *Understanding Computers and Cognition* [173] and Bødker's *Through the interface* [38] are examples of work doing exactly that. More recently Dourish has revisited phenomenology (as originally introduced to HCI in [173]) in his book *Where the action is* [66] in the context of ubiquitous, tangible, and social computing. The work of this dissertation builds explicitly on Bødker's bridging between HCI and activity theory. Activity theory is still a player in the field of post-cognitivist theories in HCI and contemporary work on activity theoretical HCI is represented e.g. by [174, 99, 49]. Another related strand of research I touch upon is the discussions of Gibson's affordance concept [79] in HCI. Norman introduced the concept to HCI in his book

The Psychology of Everyday Things and since it has been the subject of a heated debate [78, 4, 129, 82]. Recently Bærentsen and Trettvik [8] bridged activity theory and the affordance concept. This work provides foundation for many of my discussions.

The danger of a theoretical approach is dogmatism and what Dalsgaard refers to as translocutionarity [63] – when one’s language shapes what one says. The theoretical contributions of this dissertation should be seen as a complement to the recent discussions of experience, emotion, and culture as central dimensions of contemporary HCI [115, 46, 45, 128]. I acknowledge that these dimensions of technology play an increasingly large role, but I believe it is important to also maintain a focus on the more “dry” HCI virtues of actually theorising on how people use technology.

The work on reaching the technological objective is positioned in an intersection between HCI, user interface software technology, and software architecture. Hereby it intersects with multiple strands of research. Research on multi-device user interfaces is not a new thing [162, 123, 132, 139], however to my knowledge, no one has addressed whether the current application-centric interaction paradigm (elaborated in chapter 4) and the most common model for building applications (model-view-controller [143]) are the best and only candidates in this context. While the model-view-controller has been the dominant architectural model, alternatives have been developed [143, 62, 140, 83]. However, most of these are variations of the same conceptual separation of concerns (I discuss these and other work on multi-device user interfaces in more detail in Section 4.7). The visions of an alternative interaction paradigm positions itself in a history of the Dynabook [103] and the conceptual work laying the ground for the Xerox STAR [155]. Weiser’s ubiquitous computing vision [170] also needs mentioning yet he did not directly address how the vision would impact actual interaction with design of graphical user interfaces². Recently, work such as activity-based computing [15] proposes migrating activities as a paradigm for interaction with multiple heterogeneous interactive artefacts – as a layer on top of our current application-based paradigm. Kaptelinin and Czzerwinski’s recently edited book *Beyond the desktop metaphor* [98] collects a range of papers discussing post-desktop computing. However, by and large, the papers address problems with the desktop metaphor, whereas I look at problems in the application-centric paradigm as such. One paper from the book stands out, however, namely Karger and colleagues’ work on Haystack [101]. Karger criticises the application paradigm from the perspective of personal information management.

²Weiser touched upon the issue as one of the challenges in UbiComp research in [171]

1.6 Research Approach and Methodology

Human-computer interaction is a relatively young field and a unified research methodology for HCI does not exist. In this section I will provide an overview of the research approach and methodology that I have followed.

The HCI research community in Aarhus has a long tradition of participatory design [35] dating back to for instance the UTOPIA project [43]. Participatory design typically takes its starting point in a specific design case. In that respect my research project has been somewhat untraditional, given that it has mainly had a theoretical and conceptual focus. While it has touched upon specific application areas, geographical maps (Papers I–IV) and laboratory notebooks (Paper VI), the project has not been design oriented in the outset. The map study served as a means for iteratively developing and evaluating our analytical vocabulary while the workshops on laboratory notebooks served as an exploration of the potentials of a proposed interaction paradigm and architectural model.

My research approach shares with the participatory design approach a focus on gaining knowledge through creating something new and changing praxis. Engeström [71] refers to such research approach as *expansive developmental research*. Engeström argues that “*expansive research is not dealing with activities “in general” but with real activities realised by identifiable persons in identifiable locations*”. However, the field of HCI could do with an expansive solution given the challenges of complex artefact ecologies. In this dissertation I describe my work in terms of Engeström’s research programme even though my work does not take outset in concrete activities with identifiable persons and locations. Engeström describes the expansive process through his cycle of expansive developmental research. A central step in this cycle is the formation of instruments for change through a theoretical and phenomenological understanding of the problem area. In this dissertation I present two instruments for change: The human-artefact model and the VIGO architectural model.

Figure 1.1 illustrates the cycle of expansive developmental research. In short the iterative approach is comprised by five iterative steps. The *first* step deals with gaining a phenomenological insight into the nature of the area of concern. The *second* step consists of rigourously analysing the area of concern to identify potential problems that needs to be rectified. The *third* step, Engeström argues, is the most dramatic one in the expansive methodology. It deals with introducing new instruments to remedy the problems observed. The key aspect of this step is finding a *springboard* for change. A springboard is an original idea, which can put a new light on the problem area and foster change. A springboard could be a new way of looking at the area of concern. To give an example: If one dealt with e-governance a springboard could for instance be to view online government services as a web-shop rather than an information database. In the *fourth* step the instrument for change is applied in praxis to observe its impact. Finally the *fifth* step evaluates and documents the results.



Figure 1.1 Engeström's cycle of expansive developmental research

Figure 1.2 shows an adaption of my research to Engeström's cycle. The cycle should of course be understood as iterative and the different activities overlapped (which I have illustrated with some additional arrows).

The initial activity Figure 1.2A consisted of getting an overview over the area of concern. I approached this through classical literature surveys combined with formal and informal interviews with local ubiquitous and pervasive computing researchers. Furthermore I was provided with video-material from a recent experimental study of the use of geographical maps through three different artefacts (presented in Paper II). Hereby I obtained both general observations from the area of concerns and specific and concrete observations of real interaction.

As for Figure 1.2B I assumed the role as analyst and theoretician. Together with Susanne Bødker and Christina Brodersen I had the pleasure of transcribing, analysing, and theoretically explaining the video data from the map study. This led to an extensive revisiting of activity theory and related theory, building upon, among others, Bødker's previous work [38] on conceptualising HCI through activity theory. However, this time in a context of complex artefact ecologies. This work is discussed in chapter 3. Papers II–IV represent the trajectory of this work leading to the journal paper I.

As an HCI researcher I am interested in understanding our relationship to, and how we understand, artefacts. Psychology has a great deal to say about this however typically not in an HCI context. My role has been to aggregate scientifically sound theory, mainly from the field of psychology, and interpret

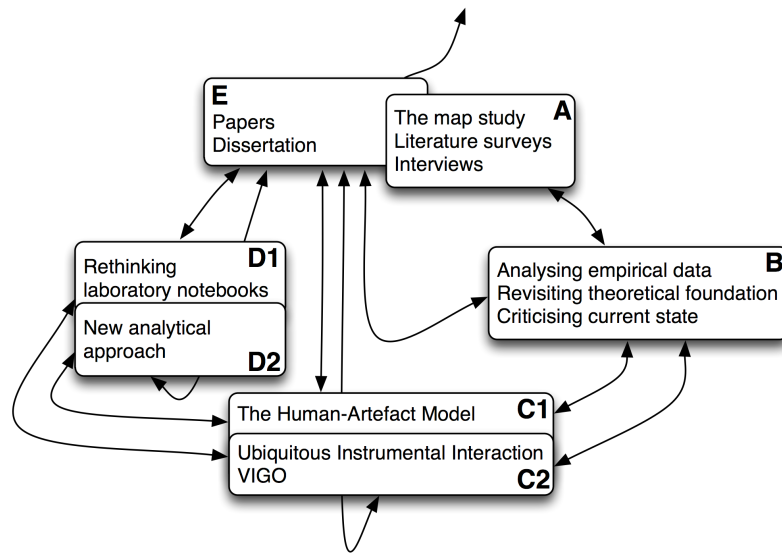


Figure 1.2 Adaptation of Engeström's cycle to my approach

it into an HCI context. What Rogers refers to as *knowledge transfer* [147]. The theoretical development was performed through iterative analysis of our empirical data, where I – from the ground up – enriched our vocabulary to explain the observed phenomena.

The work on aggregating theory led to a critique of the current application based interaction paradigm given the new reality of complex artefact ecologies – especially in the light of Gibson's [79] affordance concept as interpreted into an activity theoretical context by Trettvik and Bærentsen [8]. This critique was first documented in the short-paper [106] and later led to the development of ubiquitous instrumental interaction and VIGO presented in Paper V.

In the third activity Figure 1.2C I developed two instruments for change – C1 the Human-Artefact Model (presented in chapter 3 and Paper I) and C2 the vision of ubiquitous instrumental interaction and the VIGO architectural model (presented in chapter 4 and Paper V). The springboard for developing the Human-Artefact Model was a juxtaposition of multiple strands of theoretical work in both HCI, ecological psychology and activity theory. The springboard for developing ubiquitous instrumental interaction was the idea of applying instrumental interaction as an interaction paradigm to break the observed problems with the application paradigm.

The central method for developing the VIGO architectural model was prototyping. Throughout the process prototypes served, as Floyd puts it, as learning vehicles [74]. Prototypes were used to test, explore, and demonstrate ideas and technologies. The prototypes were architectural prototypes [14]. Hence they have been used to explore and investigate architectural qualities, principles and requirements rather than to explore a particular design. Prototyping in this research work has been used much in line with Lim et al.'s [112] dis-

cussion of prototyping as manifestations of an idea that filters the qualities in which the designers are interested, without distorting the whole.

The “final” prototype system presented in [Paper V](#) were also used as a vehicle for communication and evaluation. It was used for:

1. Demonstrating that our proposed interaction paradigm was realisable in principle.
2. Demonstrating new potentials to a technical audience (e.g. that a generalised version of the pick-and-drop interaction technique [144] can easily be implemented on top of a VIGO architecture).
3. Communicating potentials of technology.

The work presented in [chapter 4](#) addresses human-computer interaction on the level of interaction paradigms and architectural models. Hence there is a long way from demonstrating the realisability and technical potential to creating a prototype system that can be used to evaluate its impact on a concrete application area with actual users. E.g. methods like technology-probes [87] would be difficult to apply, in that it would take a massive effort to bring a VIGO based system to parallel our current interactive systems in regards to usability and functionality. It would require a substantial effort in interaction design, which is beyond the scope of this dissertation work. The prototypes therefore never met users in use – and they were not intended to either.

However, I still wanted and needed to provide some evidence that my work could have impact on a real world application area and not just a hypothetical scenario. I therefore conducted a study on applying the visions, concepts, and video-prototypes of ubiquitous instrumental interaction and VIGO as instruments for a participatory design workshop on re-envisioning electronic laboratory notebooks, hence addressing the forth element of Engeström’s cycle ([Figure 1.2D1](#)).

I gathered a group of 5 experimental physicists, who agreed to participate in individual interviews and two workshops on rethinking laboratory notebooks. The interviews were combined with guided tours in their laboratories and were conducted in order to understand their use of laboratory notebooks. The method applied for the workshops was a variant of Jungk and Müllert’s [95] future workshop. A method originally conceived for democratic brainstorming on social and community issues, however a method that has shown itself useful from an HCI and participatory design perspective [104]. The first workshop was held as prescribed by Jungk and Müllert but before the second workshop, the physicists were given a one hour introduction to the vision of ubiquitous instrumental interaction and the concepts of VIGO. Following the introduction, the second workshop was a reiteration of the fantasy phase of the future workshop. The laboratory notebook study can also be understood in terms of Engeström’s cycle where ubiquitous instrumental interaction and the VIGO model is injected into the process as springboards [Figure 1.3](#). The results can be found in [Paper VI](#).

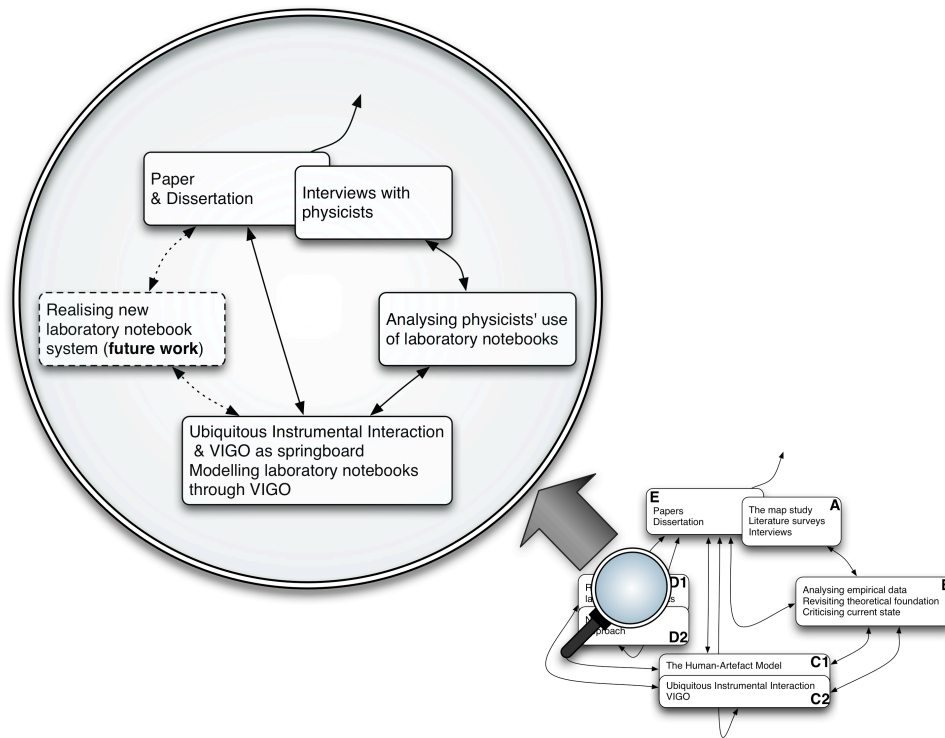


Figure 1.3 Laboratory notebook study expressed through cycle of expansive developmental research

To complete the description of the process: The human-artefact model (Figure 1.2C1) is an instrument introduced to provide a new analytical focus for HCI in complex artefact ecologies. Paper I propose how the model can frame design oriented analysis (Figure 1.2D2), hence changing the activity of interaction design for complex artefact ecologies. However, its true value will not be uncovered until it is applied as an instrument in a real design project.

Chapter 2

Conceptual Foundation

This chapter provides an overview of the conceptual foundation of the work presented in this dissertation. I will introduce and discuss the theoretical trends in activity theory that has been the foundation of most of my research the last couple of years. This chapter will not be an in-depth introduction to activity theory, for that I refer to e.g. Nardi and Kaptelinin's [99] introduction to activity theory and HCI. Instead I will highlight the concepts and understandings from activity theory and related approaches that have pervaded my work. Most of the material of this chapter has been presented in some form in the appended papers, however, in this chapter I will give a coherent overview and provide references to that of my work that has used or been inspired by the insights presented below. Furthermore I introduce *instrumental interaction* [18], which is an interaction model and a set of principles inspired by activity theory that has laid the ground for the work presented in chapter 4.

However, before delving into the flesh of activity theory I discuss why theory such as activity theory has a place in HCI at all.

2.1 Why care about theory?

In the early days of HCI theory was introduced as a means of assessing primarily performance of use in the interaction between humans and computer interfaces. Cognitive science was applied as a theoretical framework landmarked by Card, Moran and Newell's *The Psychology of Human-Computer Interaction* [55]. Methods for evaluating the potential execution time of given user interface tasks such as the key-stroke level analysis [54] were prevalent.

The early theoretical frameworks in HCI were based on a cognitive science understanding of the human mind and the computer as being equal in terms of them both being information processors (Figure 2.1). This has been criticised by many, first and foremost by Terry Winograd and Fernando Flores [173] and

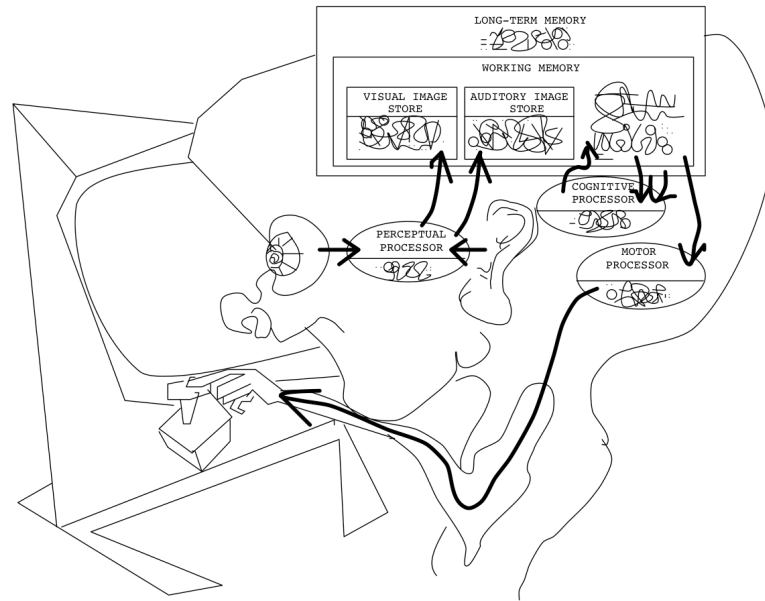


Figure 2.1 The model-human-processor [55] (illustrated by Mette Robak and Olav W. Bertelsen and reprinted with their permission)

by Lucy Suchman [158]. Both seminal works, in their respective way, in the renouncement of cognitive science as ample theoretical foundation for human-computer interaction. With a background in ethnomethodology Suchman introduces the concept of *situated action*, providing a useful insistence on a focus on the particular of specific unfolding interaction. This has been an eye-opener to many by its insistence on understanding how human beings cope despite messy environments and ill-suited technical solutions. Ethnomethodology renounces theory in general and argues that understanding should be built bottom up, and not abstracted to general terms doing away with the context that shape action. Winograd and Flores discuss human-computer interaction from a phenomenological perspective borrowing concepts from Heidegger. They especially insist on discussing the use of computers as tools and emphasise the notion of breakdowns as a focus of analysis. The notion of breakdowns is central in the activity theoretical approach to HCI as well (see e.g. [36]). Winograd and Flores are the first to present a more philosophical theory development for HCI in the mainstream of HCI research.

Nardi and Kaptelinin [99] counter-argues Suchman stating that a theoretical framework is important to HCI to avoid continuously going back to specific, detailed accounts of particular cases. What is needed, however, is a theoretical framework that is not reductionistic and encompasses the richness of the situated action. Such needs are even more predominant when insisting on studying use across interactive artefacts and interactive artefacts across use situations. Hence, we need a theoretical framework that enables us to look at human interaction with computerised artefacts beyond the isolated use situa-

tion and help conceptualise the dynamics in use of multiple artefacts.

It is important to note that theory should not be used as a static reference for explaining what we see when we look at the world. Theory should also be used to frame what we look for and while accepting that it cannot explain all we see. Theory is a plastic tool evolving through use.

The design of artefacts, whether digital or physical, embodies an understanding of use and human practice in the artefact. Activity conceptualises this as *externalisation* which I discuss in section 2.3.1. The foundations for our current user interfaces of our desktop computers were laid in a time where cognitive science was predominant as theoretical foundation – or perhaps no theory at all was the most common grounding (and perhaps still is). While it is, to quote Yvonne Rogers, “... foolish to assume or hope that theories do design” [147, pp. 116], theories such as cognitive science and activity theory embody an understanding of how we interact with and perceive the world – such an understanding can serve as explicit design principles. Figure 2.2 exemplifies such an approach quite elegantly. Here the philosophy of Nietzsche has been crystallised into a vending machine: If you want a candy bar from the machine you have to overcome its intimidating appearance and your submissive behaviour by shouting at it as loudly as possible. If you do not shout loud enough the machine will just clapper a bit.

While concepts of intimidation, submission and the *übermensch* is far from what we generally understand as central concepts in a usable interactive system, the vending machine illustrates that applying theory as a principle fosters creative thinking. A theory as activity theory has proven itself to be useful by numerous examples as a framework for analysing and conceptualising human-computer interaction. It would therefore be reasonable to apply it to interaction design in a manner parallel to the Nietzschean vending machine.

2.2 Activity Theory

The conceptual framework underlying this PhD project is activity theory as it has been interpreted into an HCI context by e.g. Bødker [38], Kaptelinin and Nardi [99], Bødker and Bertelsen [30] and Zander [174]. Activity theory is a psychological tradition that bridges the gap between materialism and idealism by letting activities compromised by both ideal and material tools and objects be the focus of investigation. The foundations of activity theory was laid by Lev Vygotsky through his work on cultural-historical psychology where he described human activity as being mediated by tools, whether material e.g. a hammer or psychological e.g. language [167]. The mediating tools were to Vygotsky the central component in understanding human activity. However, the founding father of activity theory as a psychological framework was Vygotsky’s student Alexei Leontiev. He conceptualised activity systems and developed the analytical vocabulary to encompass activities, actions and operations [111].



Figure 2.2 The Nietzschean vending machine. Developed at Technical University Eindhoven, Industrial Design by Bas Groenendaal, Rik Wesselink and Pei-Yin Chao. Picture printed with copyright holders permission.

An activity theoretical understanding of interaction made its entrance into the HCI community, especially in Scandinavia, as an opposition to the more engineering inspired adaption of cognitive science, which were predominant in the American HCI research in the seventies and eighties. It was found that efficiency was not the only crucial aspect of a good interface and that it was necessary to look beyond the then current theory drawn from an engineer's approach to cognitive science to understand the complexity of human-computer interaction [38].

Activity theory acknowledges that artefacts and materials carry a certain social practice and that material conditions are shared between groups of people while also being reflected in the consciousness of the individual [38]. Bertelsen and Bødker [30] summarise what activity theory focuses on and offers to HCI:

- *Analysis and design for a particular work practice concerning qualifications, work environments, division of work, etc.*
- *Analysis and design with focus on actual use and the complexity of multi-user activity. In particular the notion of the artifact as medi-*

ator of human activity is essential.

- *Focus on the development of expertise and of use in general.*
- *Active user participation in design and focus of use in general.* [30, pp. 3]

In my research activity theory has played a double role. It has acted explicitly as a framework for design oriented analysis as shown in chapter 3. It was also also more implicitly used in the design and rethinking of a new interaction paradigm and architectural model suited for HCI in complex artefact ecologies as presented in chapter 4. The aspect of activity theory that I have activated is its capability of conceptualising what constitutes an artefacts action possibilities as well as how these artefacts are perceived by human beings and the artefacts' interplay with user experiences, routines, etc. A central academic contribution that I have used in my work is Bærentsen and Trettvik's [8] bridging of activity theory and Gibson's affordance theory from the ecological psychology. My interest in activity theory has mainly been on the tool-oriented side of the theoretical framework, hence understanding how we perceive artefacts as tools, perceive the world through tools, and how tools are shaped together with practice.

2.2.1 Basic concepts

Activity theory has its foundation in the dialectic materialist psychology and distances itself from the isolated human being as a proper unit of analysis [30]. Instead activity theory suggests that human activity should be analysed as culturally and technically mediated. Hence the unit of analysis should include both technical artefacts and cultural organisation, both of which define the human being *and* are constantly developed by the activity of the human being. In the eyes of [167] human activity can be analysed from three fundamental characteristics:

1. The activity is directed towards a material or an ideal object.
2. The activity is mediated by artefacts.
3. The activity is socially constituted within a culture.

So basically activity theory is moving away from understanding a human being's interaction with the world as a relation between two isolated instances, the subject and the object, to interaction mediated by an artefact. The artefact is socially constituted and being either technical, a hammer for instance, or psychological, e.g. languages or concepts.

Human activity can be analysed according to three layers: Activity, action, and operation [111].

Activity is directed towards satisfying a need mediated through a material or ideal object. Hence analytically the activity layer addresses the question of *why* something takes place. In the realisation of an activity the motive or need driving the activity is not necessarily conscious to the subject. Activity is not a unit of analysis [23] given that it is partly subconscious to the subject. Bedny et al. [23] argue that to analytically recreate a holistic activity one must look at actions.

Human activity is realised through conscious actions. These actions are governed by the conscious goals of the subject, and accordingly the analytical layer focuses of *what* takes place. Goals are different from the motive or needs: They belong to the level where we immediately meet human activity in an analysis because they are conscious. This means that we as analysts, directly ask what is going on. Goals reflect the objective results of action.

Actions are realised through series of operations each triggered by the conditions and structure of the action. They are performed without conscious thinking but are oriented in the world by, what Kaptelinin [96] calls an unconscious orienting basis. According to [8], operations may be cultural-historically developed, naturally evolved, or ecologically determined. Accordingly, they may realise internalised cultural-historical patterns of conduct or inborn species-specific patterns of behaviour. This means that they result from appropriated use of tools, educated manners towards other human beings or movements in the physical world according to concrete physical conditions. Actions are dynamic structures where only the goal can be assumed to be conscious. Operations are never fixed but adapted dynamically to the conditions of the environment. Actions are thought of as a recursive structure consisting of both conscious and non-conscious sub-actions and operations. Table 2.1 which is adopted from [7, 8, 30] sums up the analytical layers of human activity.

Analytic layers	Mental representation	Realises	Level of description	Analytical question
Activity	Motive (need)-not necessarily conscious, but may become conscious	Personality	The social and personal meaning of activity, its relation to motives and needs	Why?
Action	Goal-conscious	Activities (systems of actions organised to achieve goals)	Possible goals, critical goals, particularly relevant sub-goals.	What?
Operation	Condition of actions (structure of activity)-normally not conscious, only limited possibilities of consciousness	Actions (chains of operations organised by goals and concrete conditions)	The concrete way of executing an action in accordance with the specific conditions surrounding the goal.	How?

Table 2.1 Human activity and its analytical layers

Activity theory is characterised by an insistence on understanding activity as composed by flexible hierarchies rather than understanding them as static decompositions. Actions may consist of mixtures of other actions and operations and operations may easily become actions themselves and perhaps even take over the activity. The violin-player might discover a deficiency in her otherwise operationalised bow movement, which will now define her practice the next many weeks¹. These involuntary transitions between operations, actions and activity, are typically called a *breakdown* [38]. Furthermore activity theory insists that activity is realised on all levels at the same time and not in a sequence (such as e.g. Norman's 7 stage model [126]) [30].

The activity theoretical tripartition lends itself well for framing analysis of human action. Since cognition and action are tightly connected and understood both as processes and as structured systems of action activity theory is suitable for systemic as well as functional analyses of human action (see e.g. Bedny's work [23]). The tripartition pervades all of our analytical work as presented in chapter 3.

2.3 Artefacts

One of the strengths of activity theory when applied to HCI is its emphasis of mediating artefacts. As HCI researchers mediating artefacts are what we either design or investigate. Activity theory was originally introduced in HCI to point out that the relationship between the human being and the computer is not a simple subject-object or subject-subject relationship: Instead of studying the relationship between the user and the computer as something that the user works *on*, or communicates *with*, Bødker [38] pointed out how we may more usefully see the computer as something that the user acts *through* on other objects or with other subjects: Designed artefacts and their role as mediators have been at the core of this way of thinking.

2.3.1 The dialectics in artefacts

Fundamentally artefacts are crystallisations of activity [111]. Human activity is crystallised into artefacts in two ways: Firstly, they are externalisations of operations with earlier artefacts, and secondly, they are representations of modes of acting in the given activity.

Bærentsen [7] illustrates the dialectical relationship between activity and tools in an artefact historical analysis of the development of hand-weapons from the time of bow and arrow to current automatic rifles. Bærentsen shows how human operations are gradually incorporated and refined into artefacts and how this development shapes both the goals and needs of human users.

¹Substitute violin-player with professional Counter Strike player, bow-movement with mouse-movement, and the example apply to HCI

The introduction of hand-muskets made re-loading far more complicated than placing an arrow on a string; nevertheless it was outweighed by the firepower and status in muskets. The operational level of the use of the weapons changes radically. The routine for reloading moves from putting an arrow on a bow-string over stuffing a musket, to inserting a clip into an automatic rifle. While the development of weapons has contributed to changing society the central goal of the development of the weapon has not changed, namely to be able to shoot down an enemy from a distance. Weapons are examples of externalisations that clearly represent the violent activity they are typically designed for.

In a dialectical understanding technology is both a product of society and a source of change in society. The World Wide Web is a good example of a technology introduced to remedy a practical problem of information distribution that turned out to change the world fundamentally.

However, the individual is also changed in appropriating an artefact as a tool for realising activity. Kaptelinin [97] argues that when an artefact is appropriated by the user in use to a degree where the user can think through the artefact it becomes a *functional organ*. The concept functional organ captures the functionally integrated, goal-oriented configurations of internalised operations *and* external mediation. The functional organ is i.e. a dialectical concept encompassing both aspects of the user and artefact. The fork and knife become functional organs in the activity of cutting and eating food. Hence the knife and fork augment the human capacity in regards to eating and handling. Mediators that become functional organs become part of human beings, moving the boundaries of the human being “outward”, similarly to how Polanyi [138] describes the blind man seeing with his stick. In [Paper I](#) we argue that the concept of functional organs provides a richer ideal for the design of interactive artefacts than e.g. the ideal of transparency.

Returning to externalisation; Karpatschhof [102] distinguishes between passive and active externalisations of human operation. Passive externalisations are supplements of our human morphology that are completely separated from our bodies. Their function is not to perform the operations that implement activity but to facilitate that a person can realise the operations of the activity. Applying an axe to cut down a tree is an example of passive externalisation. Another example of passive externalisation is the written language. The active externalisation, on the other hand, not only facilitates certain operations, it substitutes otherwise human operations. Comparing the throwing spear to the longbow: The spear is a passive externalisation that facilitates piercing a target from a distance while the longbow actively externalises the throwing operation (making it one of the earliest machines in history). While computer applications by nature are active externalisation [38, pp. 35] this distinction gives an interesting frame for discussing externalisations in computer applications and has provided inspiration from the separation between instruments and governors as presented in chapter 4.

Bødker [38] introduce a base for discussing the assumptions externalised

into mediating artefacts, user interfaces in particular, through three levels of aspects.

1. The **physical aspects** embody the assumptions of the user's physical morphology and motor functions. The physical aspects are the conditions for the physical handling of the artefact. Breakdowns [38, pp. 22] at this level typically result in that what otherwise were operations on a motor level becomes conscious actions, and object of reflection.
2. The **handling aspects** embody the assumptions of how the given artefact should be handled. The western forks and knives embody an assumption of, and are designed for, the western traditions of how to eat properly. The handling aspects are the conditions for transparency of the object. Breakdowns at this level can be the result of in-proper training of the user or false assumptions by the designer of the training of the user.
3. The **subject object-directed aspects** are the assumptions of the actions the artefact is designed to help the user realise. Breakdowns in use at this level are typically products of a mismatch between what the user actually wants to do and what is possible to do with the the given artefact.

In chapter 3 I present a further elaboration of these aspects incorporated into the Human-Artefact Model inspired by Bærentsen and Trettvik's take on Gibson's affordance concept, which is introduced in the next section.

2.3.2 Affordances

Since its mainstream introduction into the HCI community in Donald Norman's book *The Psychology of Everyday Things* [126] the concept of affordances has been the focus of some academic turmoil (e.g. [4, 78, 8, 129]). Affordances were introduced into HCI as a means of discussing the qualities of interface elements in regards to the immediate perceptibility, or lack of same, of their action possibilities. However, how affordances theoretically should be understood in the first place is a topic of debate.

The concept of affordances was conceived by psychologist James J. Gibson who was one of the leading characters in the ecological school of psychology. In the book *The Ecological Approach to Visual Perception* [79] Gibson establishes a theory of visual perception that rejects the duality of mind and matter present in both behaviourism and cognitive science.

The affordance theory is interesting from an HCI perspective, as the notion of affordances can help us understand the user's perception of tools – whether physical or digital. Understanding what artefacts, tools and instruments affords becomes especially interesting when one goes beyond the assumptions of a desktop computer

Gibson defines affordances in the following manner: *"The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good*

or ill.” An affordance is a relation between the organism and the environment. The affordance points both to the observer and the environment – therefore neither subject nor object². The affordances of an object to a subject are the perceived action possibilities by the subject of the object. The set of affordances of a given object to a given subjects is there and it is defined given the physical attributes of the subject and object respectively.

Where Gibson focuses on the more low-level perceptual aspects of human activity, Bærentsen and Trettvik bring the notion of affordances into an activity theoretical understanding of human activity. They claim that in order to understand what is actually afforded by an object you have to understand what influences the user’s perception on all three levels of activity – activity, action and operations. Bærentsen and Trettvik distinguish between *need related*, *instrumental* and *operational* affordances, where the operational affordances are subdivided into *conscious operational* and *adaptive operational* affordances. The need related affordances between an object and a human refer to *why* you would interact with this object, the instrumental affordances to *what* you perceive, you can do with the object to achieve your goals given the socio-cultural conditions. The operational relates to *how* you handle the object: The conscious operational affordances are the cultural-historically taught ways of using the object and the adaptive operational affordances relating to how we are actually able to interact with an object given our physique. To exemplify the consider guitar playing: To play on a guitar we are taught to hold it properly, and where to place our hands – the conscious operational affordances. But the relation between the thickness of the strings and our fingers will let us pluck the strings anywhere on the fret board we like – the adaptive operational affordances.

Bærentsen and Trettvik observe that while Gibson mainly talk about low-level affordances, he also argues that a mailbox has affordances to individuals in a letter-sending culture. Bærentsen and Trettvik argue that the perception of the affordances of a mailbox is no different than the perception of the affordances of a flat surfaces. The affordances of both objects have to be learned through what they call *tuning*, something that happens on all levels of activity. While all the affordances are present in the relation between the subject and the object, perceiving the relevant affordances to the subject’s given activity, is what Trettvik and Bærentsen call tuning. This tuning is governed by past experience and learning. Perceiving the affordance of putting letters, and only letters, in the mailbox requires knowledge of the postal system. A child, however, tends to put whatever fits into a mailbox – explaining why they are typically placed put above a child’s reach.

The concept of tuning into affordances has ties to the concept of functional organs, and appropriating artefacts as mediators of activity. The famous phrase of disputed origin, “*If all you have is a hammer, everything looks like a nail*”, catches the connection quite elegantly. The tuning of affordances are mediated through the instruments we hold. A simple tool like a hammer quite easily becomes appropriated as a functional organ so we can think and perceive affordances

²This stands in stark contrast to the dualistic view of mind and matter of earlier psychology.

as they were part of our own body. However, not all instruments are as simple as a hammer, and even instruments that appear simple, such as a pencil or a violin, can take years to master.

Kaptelinin [97] argues that when an artefact becomes a functional organ for a user, the user can think through the artefact even without possessing it. Hence learning to use artefacts will change the way we perceive the world around us.

The concept of affordances and especially the activity theoretical interpretation of the concept was the seed for the critique of contemporary graphical user interfaces laying the base for the software architectural developments presented in chapter 4. Furthermore it is fundamental to the development of the human-artefact model, albeit not used directly in the model.

2.4 The orienting basis of the user

Gal'perin [75] presents the orienting basis as central to supporting the application of actions and operations across different artefacts and other material conditions, in particular objects of work. The orienting basis guides the actions, it is shaped through experiences and would in Bærentsen and Trettvik's optics be what governs our tuning in to perceiving relevant affordances in our surroundings.

There are three ways of constituting this orientating basis. These reflect both how it is created and how it may potentially be applied: *Domain independent general actions; reusable knowledge within a specific domain* – and if everything else fails – *trial and error*. If an action has been developed through trial-and-error there is less chance that it can be reused to develop other quality actions since fundamental understanding may be missing. Accordingly we may address the issue of quality of actions and repertoires of actions (and hence operations) in terms of generality and reusability. According to Gal'perin [75] the quality of a specific action may be assessed through the following:

- Generality, i.e. the extent to which recognition of an object's properties relevant to completing an action is part of the given action.
- Support for abbreviation, i.e. the extent to which the action will help to consciously skip operations in the performance. Operations can be abbreviated either consciously or unconsciously. Conscious abbreviation is natural and non-harmful, e.g. skipping intermediate results in the process of solving a problem. Harmful abbreviation is the opposite unconscious skipping of intermediate results or actions because of lacking knowledge or experience with the consequences.
- Mastering, i.e. independent reproduction of an operation with new material, achieved within the action.

Paper IV discusses Gal’perin’s theory of learning from the outset of designing learnable artefacts. Furthermore the understanding of orienting bases plays a central role in our development of the human-artefact model Paper I. Also Gal’perin’s understanding of the quality of learning provided fuel for the motivation of devising an interaction paradigm where actions could be reused explicitly across artefacts as discussed in chapter 4.

2.5 Other postcognitivist frameworks

Championing one theoretical approach over another, without recourse to Popper’s scientific paradigm to back up one’s claims, often ends up being a matter of personal preference, stemming from one’s own background and values as to what constitutes good design practice and research. That is not to say that one cannot highlight the strengths and problems of a particular approach and how others have used it.

— Yvonne Rogers [147, pp. 117]

Activity theory is not the only postcognitivist theoretical framework that has gained ground in the HCI community. When using a theoretical framework of any kind, a valid question to raise is why one have chosen that framework in particular. Given my schooling both as a student of Susanne Bødker, and as a member of the HCI research community in Aarhus, I have been given a natural bias towards activity theory. Activity theory was the first theoretical framework I was properly introduced to.

A couple of years before activity theory began to get attention in the HCI community as a postcognitivist theory, Winograd and Flores [173] introduced the philosophical tradition of phenomenology as a counterweight to information processing understanding of human action from cognitive science. Phenomenology shares a lot of understanding with activity theory and emphasises that human action and perception should be understood as embodied in our daily life. Heidegger’s concepts of *zuhanden* and *vorhanden* gave inspiration to the concept of breakdowns, which have been central in the activity theoretical approach to HCI [38]. The phenomenological tradition has been followed by contemporary writers such as Paul Dourish [66]. While phenomenology emphasises the importance of being in the world, and acknowledges how perception and action is guided by tools or equipment in Heidegger’s terminology, activity theory is more radical and argues that human action should be understood and perceived with mediation as the first class entity. From the perspective of HCI this is interesting because this basically places the field in psychology. If tools are focal points of psychology so are computers.

Within recent years two good surveys of post-cognitivist theories have been published. Yvonne Rogers present one in the paper “*New Theoretical Approaches for HCI*” [147] and Kaptelinin and Nardi have a chapter on the topic in their recent book [99, pp. 195-236]. I refer to these two sources for an in-depth discussion of the topic. In Kaptelinin and Nardi’s survey three central approaches

stand out: Distributed cognition [149, 84], actor-network theory (ANT) [109, 53] and phenomenology [66, 173]. Rogers supplements with a theoretical framework called *external cognition*, which she herself has worked on [150], the ecological approach as inspired by Gibson [79] and the (atheoretical) situated action approach [158] from Suchman's book.

In the aesthetic branch of HCI. American philosophers such as Pierce, James, Dewey and Mead whose work constitute what is referred to as *pragmatism* have been used to conceptualise *experience* with technology (e.g. [115] and recently [63]). Pragmatism can therefore be seen as one of the theoretical frameworks for what Bødker calls the third wave of HCI. The focus for the researchers applying pragmatism to HCI is, however, different than mine. I focus on understandability and use rather than aesthetic experiences (although pragmatism is not only about aesthetics [63]). However, these two stances are not mutually exclusive and pragmatism and activity theory share many insights [72]. A bridging between these two approaches to HCI could be valuable to the conceptual foundation for third wave HCI.

Common for the post-cognitivist approaches (including activity theory) is that they are all critical of the mind body dualism, they all claim that individuals are not defined by the boundaries of their skin [99]. Distributed-cognition defines cognition as distributed between people and tools and in actor-network theory people and tools are placed in the same network where a clear boundary around the individual cannot be drawn.

Kaptelinin and Nardi argue that one of the central differences between on the one side phenomenology and activity theory and on the other side actor-network theory and distributed cognition is the relationship between technology and people.

Activity theory strongly argues that only humans have intentions, and technology are used to realise these intentions where actor-network theory argues that humans and technology both exude intentional behaviour. This stance can be attributed to the functional or provocative benefit of analysing technology as having agency, rather than actually believing that technology can be intentional subjects. In distributed cognition – Kaptelinin and Nardi argue – a doctor will distribute cognition onto, among other technologies, a thermometer to process the diagnosis of a patient. However, activity theory would argue that it might be the case that the function of the thermometer has been internalised by the doctor, perhaps she can feel the temperature with her hand, see it in the patient's skin or eyes, but she will still use the thermometer because it represents her activity to the patient.

In discussing the difference between theoretical foundations, parallels can be drawn to discussions of different programming languages in computer science. However with one big difference, in computer science Turing has made it possible prove that two languages can express the same. This is not the case with psychological theories. But there is a reason that there are different program-

ming languages even though they all basically do the same, namely that each language emphasises a certain way of understanding and doing things. The different postcognitivist approaches to HCI share a lot of insights, however, each emphasises different aspects. Activity theory emphasises that the relationship between humans and artefacts is asymmetric and dialectically evolving. This fundamental premise pervades all of this dissertation work.

2.6 Instrumental Interaction

The final concept that I introduce in this chapter is not directly activity theoretical – but inspired by activity theory. *Instrumental Interaction* is an interaction model presented by Michel Beaudouin-Lafon [18] in 2000, extending and generalising the principles of direct manipulation. The objective for Beaudouin-Lafon was to provide an interaction model that could encompass, at that time, novel interaction techniques such as e.g. Marking Menus [108], Toolglasses [34] or Zoomable interfaces [22]. These techniques were typically collected under the term post-WIMP user interfaces³.

Beaudouin-Lafon defines an interaction model as:

... a set of principles, rules and properties that guide the design of an interface. It describes how to combine interaction techniques in a meaningful and consistent way and defines the “look and feel” of the interaction from the user’s perspective. Properties of the interaction model can be used to evaluate specific interaction designs.

With instrumental interaction Beaudouin-Lafon emphasises that interaction, even in direct manipulation interfaces, is rarely *direct*. Scrolling a document is typically mediated by a scrollbar and not directly on the object. The direct manipulation interaction model lacks a vocabulary for discussing the mediated aspects of user interfaces – especially non-widget like interaction.

The central tenet of instrumental interaction is to view interaction in graphical user interfaces as mediated rather than direct. This tenet is inspired by the activity theoretical approach to HCI, which was prevalent in the HCI research community in Aarhus where Beaudouin-Lafon developed his ideas of instrumental interaction during a two-year sabbatical.

Instrumental interaction is inspired by our use of physical tools. The central concepts of instrumental interaction are *domain objects* and *interaction instruments*.

A domain object is defined as:

³A post-WIMP interface, according to [165], is “one containing at least one interaction technique not dependent on classical 2D widgets such as menus and icons. Ultimately it will involve all sense in parallel, natural language communication and multiple users.” Originally post-WIMP was thought in a context of desktop computing, however, the definition extends fairly well to e.g. tangible and ubiquitous computing (see e.g. [92]).

... domain objects form the basis of the interaction as well as its purpose: Users operate on domain objects by editing their attributes. They also manipulate them as whole, e.g. to create, move or delete them. [18, pp. 3]

Interaction instruments are described as:

... a mediator or two-way transducer between the user and domain objects. The user acts on the instrument, which transforms the user's actions into commands affecting relevant target domain objects. [18, pp. 3]

Instruments are composed by a physical part, an input device, and a logical part which transforms input to commands on domain objects. Instruments translate the response from objects to feedback to the user (Figure 2.3). Instruments are said to be activated when the user connects the physical part with the logical part. This can either happen spatially, e.g. moving a cursor to a scrollbar and pressing to activated, or temporal (or moded) by pressing a button to activate e.g. a selection instrument.

According to Beaudouin-Lafon, instrumental interaction introduces the potential of a second degree of reification in contrast to direct manipulation. Reification is the process of turning concepts into objects [18, pp. 3]. Primary domain objects such as text documents may have a style sheet. This style sheet can be reified as a domain object in its own right. Instrumental interaction reifies commands as instruments and makes it possible in principle for instruments to become objects themselves and be manipulatable by other instruments.

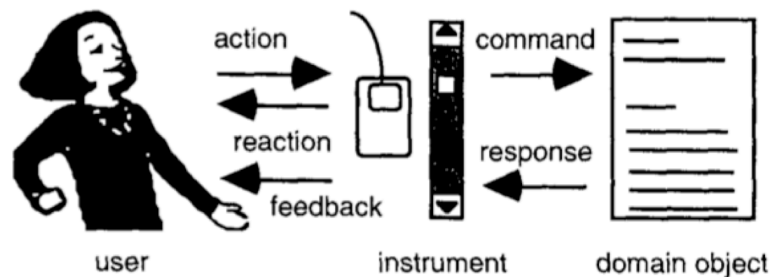


Figure 2.3 Beaudouin-Lafon's illustration of an interaction instrument mediating the interaction between a user and a domain object (reprinted with authors permission)

Beaudouin-Lafon furthermore used the instrumental interaction model together with a set of dimensions for evaluating interaction instruments to criticise the at that time dominant WIMP-based interaction techniques.

I visited Beaudouin-Lafon at Université Paris-Sud a day in late summer 2006. Among other things, discussed the potentials of reviving instrumental interaction as a general interaction paradigm for multi-device interaction. This meeting was the seed for the work presented in this chapter and for an ongoing collaboration.

In [chapter 4](#) I discuss the potentials in an interaction paradigm based on the principles of instrumental interaction. In section [4.4](#) I will present an architectural model for realising such interaction paradigm. In [Paper I](#) we discuss how Beaudouin-Lafon dimensions for evaluating instruments can be applied in conjunction with the human-artefact model.

Chapter 3

Complex Artefact Ecologies from Without

This chapter concerns analysing artefacts and human-computer interaction in complex artefact ecologies – hence complex artefact ecologies from *without*. In the papers relating to this chapter we have worked towards enriching our analytical vocabulary for analysing interaction with interactive artefacts that may differ from a personal computer and may be used in conjunction with many other heterogeneous artefacts.

Oulasvirta [134] describes how Ubicomp is facing a big divide where on the one hand practical applications and interfaces are pluggable and interchangeable only to a very limited extent. On the other hand, much of the research focusing on the visions for ubiquitous computing remains too simplistic and abstract. In Paper I we argue that with the third wave of HCI [37] follows a second divide. This is a divide between a HCI discourse focussing more and more on elements such as culture, emotion, and experience, and then a situation of increasingly complex artefact ecologies where we only have methods for understanding the aspects as the handling of artefacts one isolated artefact at a time. While the new trends in HCI give a much welcomed broadening of our scope, and a place where it is possible to move beyond the classical purpose-driven understanding of HCI, there is little focus on providing concepts and methods to help understand what users actually do in the messy ecologies of multiple and heterogeneous artefacts.

Almost a decade ago Dourish [66] called for a revisiting of our theoretical foundation given the challenges of ubiquitous computing, while other voices called for more applicable and comprehensible interpretations of post-cognitivist theories [147]. In the work of this chapter I have tried to accommodate both of these voices.

The overarching goal of the concepts, models, and methods of this chapter is to provide analysts, designers, and researchers a means of coping with the analytical complexity arising from moving beyond the desktop computer. The tangible contribution of the work in this chapter is the human-artefact model

– a simple model for structuring design-oriented analysis of artefacts, while drawing on a range of theoretical insights. Furthermore the papers provide a thorough interpretation of relevant activity theoretical concepts to the new challenges of HCI in complex artefact ecologies.

3.1 Chapter overview

This chapter is based on the four papers [I](#), [II](#), [III](#), and [IV](#). The journal [Paper I](#) provides a cohesive introduction to the human-artefact model going from motivation to a theoretical foundation, model to application. Hence, in order to avoid repeating the paper, this chapter gives an overview of three central themes pervading the work that have led to the model. Furthermore I introduce the model and give an example of back-of-the-envelope analysis through the model. Finally related work is covered.

In [Section 3.2](#) I present the study that has been the empirical point of departure for the theoretical work in this chapter. [Section 3.3](#) presents the three themes. [Section 3.4](#) gives an introduction to the human-artefact model and exemplifies its use with a simple example. [Section 3.5](#) surveys the related work, and [Section 3.6](#) summarises the chapter.

3.2 Empirical point of departure

The examples used throughout the three research papers comprising this chapter are largely based on a study (first document in [Paper II](#)) of geographical maps used on three different technologies: A paper map in a telephone book, a web based digital map on a tablet-PC, and a digital map on a GPS enabled Nokia smart phone. The motivation for the study was originally to create a more controlled setup for studying the handling of maps on different technologies. In an earlier project [\[47\]](#) kids were sent out to do project work in urban areas equipped with annotatable maps on smart-phones and tablet-PC's. However, it became clear that even with extensive video recordings it was very difficult to capture the detailed handling of the artefacts while the kids were moving about.

Hence the motivation of the study was to examine handling of maps in a controlled environment, however it showed to be a rich resource for general observations of interaction with heterogeneous non-PC artefacts. It is this facet of the study that has fed my theoretical developments.

The study was performed with six groups of 2-3 members, all adults. Each group was given eight assignments, e.g., locating north or identifying a local school and the distance and direction to it, one half to be done with one technology, the other half with another. The participants stood at a large panorama window in our department with a good view over the neighbourhood. All assignments relate to objects (buildings, landmarks, schools, etc.) that are (more

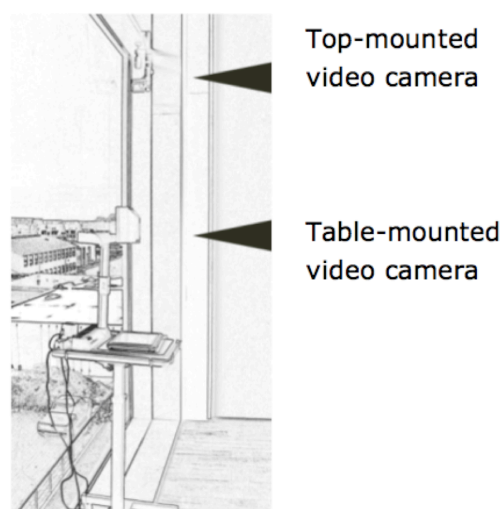


Figure 3.1 The Experimental Setup

or less) visible through the window. The experiments were recorded using the setup shown in [Figure 3.1](#) plus a tripod-mounted digital camera. The top-mounted camera provided a general view of the participants' posture and gestures and the table mounted camera provided a close-up of their interactions with the map technology at hand. The sessions were recorded, transcribed and analysed in activity theoretical terms.

From the perspective of my research, the map-study provided me with an opportunity to study actual use of three different artefacts that interchangeably were used to mediate the same activity. Especially it provided material for studying the breakdowns that happen on different levels of activity when users appropriate an artefact new to them.

3.3 Three themes

In the following I highlight three themes that have pervaded my work in the papers constituting this chapter. Firstly how to conceptualise the ideal of an artefact in use. Secondly a revisiting of the concept of affordances, and thirdly a focus on learning across artefacts.

3.3.1 The ideal of artefacts

The concept of *functional organs* (introduced in [Section 2.3.1](#)) has played a central role in framing our work on conceptualising human-computer interaction in complex artefact ecologies. The functional organ concept provides a rich ideal for the use of artefacts (which we elaborate in [Paper I](#)). In [Paper III](#) the purpose is to study the extent to which artefacts can be appropriated by users as functional organs. The concept of functional organs address by and large the

same qualities as for instance talking about transparency in interfaces, or when, in Heideggeren terms, talking about *vor* or *zu-handen* (see e.g. [173]). However, as with the concept of affordances, functional organs encompass both the artefact and the user. Talking about an artefact becoming a functional organ as a quality assumes that one talks about the user as well – something that is not implied by the concept of transparency. Hence the focus of the work presented in this chapter has largely been on understanding how artefacts can (or cannot) be appropriated to a degree where the user is capable of thinking through them and reasoning on use without experiencing severe breakdowns between the capabilities of the artefact and their expectations towards it.

Since functional organs encompass both artefact and user we need to explain the relationship between the artefact and the user's capabilities and experiences. This is basically what is the main theme of [Paper I](#), and the human artefact-model is a simple model for structuring an analysis that exposes this relationship.

3.3.2 Affordances revisited

In our work on analysing the empirical material from the map study we revisited Gibson's affordances in the light of Bærentsen and Trettvik's delineation of the concept (see [Section 2.3.2](#)). Instead of using affordances directly as an analytical concept (which is not without problems [127]), our analytical aim has been to look at what might constitute potential affordances. In Gibson's terminology affordances is a concept encompassing both subject and object. Hence in principle as an analyst one can only conjecture about what affordances a subject (user) might perceive. Our aim was to equip analysts with means of breaking down each side of the affordance-relation, to be able to make qualified conjectures about perceived affordances.

In [Paper III](#) we analyse breakdowns between the intended action possibilities in the artefacts and the action possibilities expected by the user. In the paper we derive a four-layered analytical approach based on a re-iteration of Bødker's aspects (see [Section 2.3.1](#)) through Bærentsen and Trettvik's [8] activity theoretical delineation of Gibson's affordance concept ([Table 3.1](#)). Hence instead of talking about affordances of an artefact we on the one hand talk about aspects of an artefact, which are the assumptions of use embedded in the artefact and on the other hand the experiences of the user.

The human side of the affordance relation we initially conceptualised through concepts inspired by Bedny's [23] strategy concept. Bedny describes strategies as plans for goal achievements that are responsive to external and internal conditions. We distinguished between strategies and routines; strategies have to be consciously developed when human beings are to perform an action. Strategies turn into routines when the predisposition to perform a certain action becomes partly automated or non-conscious. In [Paper I](#) we generalise the human side of the model through Gal'perin's [75] concept of orienting basis (introduced in [Section 2.4](#)). [Table 3.2](#) is the first iteration of what becomes the

Affordance	Aspect	Activity
Need-related		Activity-why
Instrumental	Subject / object directed aspects	Action / strategy-what
Operational		Operation / procedure-how
- Conscious	Handling aspects	
- Adaptive	Physical aspects	

Table 3.1 Mapping between affordance levels, Bødker's aspects and activity

human-artefact model presented in [Paper I](#). In [Section 3.4](#) I explain each field of the final model.

Analytic Layer	Functional Organ	
	Mediator/External	Routines/Internal
Need-related-why?	Need-related aspects	Motivational routines
Instrumental-what?	Instrumental aspects	Actions and strategies
Operational-how?	Operational aspects	Operations and routines
Handling	- Handling aspects	- Learned handling
Adaptive operational	- Adaptive aspects	- Adaption

Table 3.2 A functional organ mapped to the (revised) aspects of the mediator coupled with the internalized routines on the three layers of activity

Misfit examples

To give an idea of the kind of analysis we initially performed in [Paper III](#), I now give two examples of misfits observed and explained through our layered model. For the whole analysis I refer to the paper.

The handling aspects refer to operation of the artefacts that assume to be known – but have to be taught. As a simple example; a **handling** misfit arose from unfamiliarity with the paper map, experienced by particularly the members of one group, who did not recognise the grid- and numbering-system for navigating between adjacent map slices, and who consequently had a difficult time finding the right section of the map to measure a distance across pages. The group, however, had no troubles with the adaptive operational aspects of the paper-map – namely to flip the pages of the book ([Table 3.3](#)).

On the **instrumental** layer (which refers to, on one hand, the functionality of the artefact, and on the other, the goals of the user) we observed during one of the tasks, that the group members used routines for interaction with the phone developed in everyday use of cell phones (e.g. using menus, joystick). The interaction with the cell-phone was fluid and unencumbered and the understanding of the menu structures, and general information hierarchies associated with cell-phones was clear. However, the group experienced a break-

Mediator:		Handling layer	Adaptive operational layer
Paper	E		Leafing pages
	I		
Map	E	No perceived structure	
	I	Grid structure	

Table 3.3 Misfit with the grid structure: While the group can leaf pages they do not expect a support structure for this, and they do not recognize the grid structure even though it is visible

down despite a high degree of familiarity with the interactive artefact and knowledge of maps in general: The group members expected to find a distance measuring tool for the map on the phone and started looking for it by browsing through the menus before being told, that the feature did not exist. Hence they experienced a breakdown on the instrumental layer. The cell-phone afforded computation but did not provide distance measuring functionality.

Mediator:		Instrumental layer
Phone	E	Computational power
	I	
Map	E	Computational distance measuring
	I	No distance measuring available

Table 3.4 Misfit at the instrumental layer: a breakdown is caused by the expectations of a functionality (*expected action possibilities* E) that are blocked or not supported in the actual, implemented functionality (*intended action possibilities* I)

3.3.3 Learning and dialectics

Embedded in our work on conceptualising interaction in complex artefact ecologies is a strong belief that artefacts should and cannot be understood in isolation. In [Paper IV](#) we experimented with applying Gal’perin’s [75] theories of learning (and Bardram and Bertelsen [16] interpretation of Gal’perin) to our observations made from the map study and in retrospect discussing a design case on mobile IT for a wastewater plant [31]. Seen through the lens of Gal’perin’s learning, our map study clearly shows us that learning goes beyond the isolated artefact. The example of the breakdown observed with the lacking distance measuring functionality on the cell-phone is a good example of this. Learning and the use of artefacts shape the perception of others. In both [Paper IV](#) and [Paper I](#) we discuss how learning across artefacts can be embraced in design.

Gal’perin argues that when we learn we establish an orienting basis; a mental scheme for realising learned actions. The quality of this orienting basis depends on how it is established – e.g. if an action is learned by trial and error the quality reusability of the orienting basis is not necessarily as high as if it was taught in general terms (further discussed in [paper IV](#) and [I](#)). In [Paper I](#) we match our four levels of aspects of artefacts with four levels of orienting of the user. The orienting basis side of the human-artefact model accumulates

experiences with other artefacts in the user's artefact ecology, some of these historical and some current, some individual and some with others.

In [Paper I](#) we argue that an exhaustive analysis of interaction with artefacts (especially, but not exclusively, non-PC artefacts) require both understanding the history behind the aspects of the artefacts (where do the assumptions of use stem from?) but also the history behind the user (how is the user's orienting basis established and what does it encompass?).

3.4 The Human-Artefact Model

Why?	Motivational aspects	Motivational orientation
What?	Instrumental aspects	Goal orientation
How?	Operational aspects	Operational orientation
	- Handling aspects	- Learned Handling
	- Adaptive aspects	- Adaptation
	Artifact	Human

Figure 3.2 The Human-Artefact Model

The human-artefact model ([Figure 3.2](#)) collects on the one hand the aspects of an artefact and on the other hand the orientation of the user, broken down by the extended levels of analysis suggested by Bærentsen and Trettvik [8].

To paraphrase [Paper I](#); the human-artefact model provides a form, where each field can be addressed one at the time by summarising empirical or analytical findings, by singling out particularly critical issues or findings to match or contradict those of other fields. If the starting point is in the human side, the human-artefact model can be used to structure the analysis of human practice and orienting basis in terms of the four levels, and a particular current or future artefact may be added on the artefact side. This approach makes it possible to address e.g. quality of actions, and possible new designs. Or to the contrary, the starting point may be the artefact, where the four levels continue to be important for structuring the analysis. Consequently the human side can be analysed with regards to potentials and problems of matches with the artefact, raising e.g. issues regarding the development in use. In the analyses it is possible to move back and forth between these two perspectives, to focus on one level at a time, and to address tensions between fields across sides and levels, e.g. breakdowns. The human-artefact model makes it possible to anal-

use the actual human use of an artefact reflected in the ideal, well-integrated functional organ, in order to find trouble spots and areas for development.

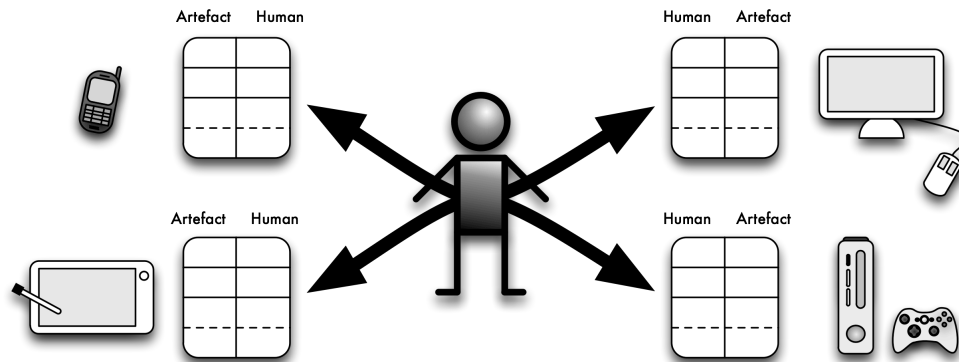


Figure 3.3 A human's orientation is shaped by the artefact ecology

While the human-artefact model is to be applied to analyse the use of a particular artefact, the artefact is not analysed in isolation. The orienting basis side of the human-artefact model, that is the user side, accumulates experiences with other artefacts in the user's artefact ecology, some of these historical and some current, some individual and some with others (Figure 3.3). In other words, the human-artefact model embraces the possibility of change, and as such it is a useful starting point when facing HCI as a design discipline, rather than as something providing static analyses of technologies in use. Similarly the artefact side crystallises and accommodates many uses and many users (Figure 3.4).

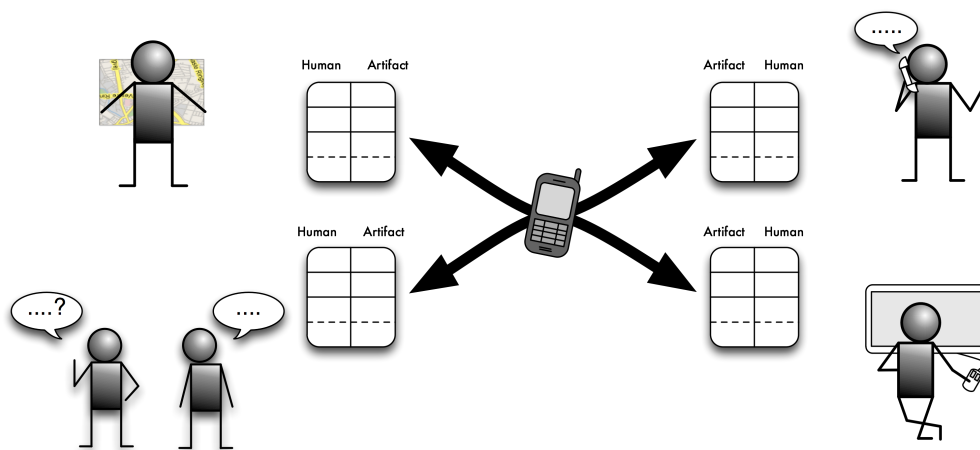


Figure 3.4 An interactive artefact embeds externalisations from a multiple of activities

Before I give an example of its use, I sum up each cell of Figure 3.2. Firstly the artefact side:

- **Motivational aspects:** *Why* is this artefact designed? The assumptions

embedded in the artefact, regarding what drives the user to actually use the artefact. What activity is this artefact supposed to mediate?

- **Instrumental aspects:** *What* can this artefact make the user do? This is what is typically referred to as the functionality of the artefacts. What actions is the artefact designed to help the user realise?
- **Operational aspects:** *How* should the artefact be operated. This aspect is divided into two categories:
 - **Handling aspects:** The handling aspects are part of the operation of an artefact that has to be taught. Moving the slider of the scrollbar to navigate the text of this document while I edit it is part of my editor's handling aspects. The weight distribution in chopsticks is optimised for a certain handling, a handling that has to be taught.
 - **Adaptive aspects:** The assumptions of the low-level Gibsonian affordances. This could encompass assumptions of physical attributes of the user; an icon large enough to be seen, a button large enough to be pressed, or a computer small enough to be carried around. Furthermore assumptions of the low-level operation – e.g. should a screen afford touching.

Secondly the human side:

- **Motivational orientation:** The motivations that drive the users activity. Motivations of use may be subconscious to the user. E.g. in the map-study we saw that many of the participants' actions were driven by a unspoken motive for establishing a common-orienting basis. I.e. agreeing on north, or what street they currently were looking at.
- **Goal orientation:** What are the goals and corresponding actions that the user needs to achieve and execute in order to realise the given activity? When confronted with an artefact, goals previously realised with other artefacts may influence the perception.
- **Operational orientation:**
 - **Handling:** What experiences does the user have in operating similar artefacts. Handling orientation is typically conscious actions that have been operationalised and largely become unconscious.
 - **Adaptive:** This level refers to the user's low-level physical experiences. Does the user have physical restrictions, or are physical restrictions imposed upon the user (Gloves, locomotion, bad lighting, etc.)? What are the experiences with the physicality of the artefact?

Each of the fields in the human-artefact model lends itself to a historical question. On the human side, which artefact has shaped e.g. the motivation or the handling orientation? And on the artefact side, which activities are crystallised into the artefact? And what assumptions of use of previous artefacts do the design draw upon?

3.4.1 Back-of-the-envelope analysis example

In [Paper I](#) we give a detailed walkthrough of analysis with the human-artefact model based on empirical data from the map study. In this section I will provide an example of a quick back-of-the-envelope analysis of the hypothetical use of an artefact structured by the human-artefact model. The findings from the analysis are by no means intended to be profound, but instead they are meant to exemplify the human-artefact model and how it can be applied as a easy tool for structuring analysis.

Imagine that we have to provide ideas for a re-iteration of the Google Map application on the Apple iPhone ([Figure 3.5](#)). Google is interested in providing better support for on-foot urban touring. Hence consider a scenario of a first time tourist in London who has dedicated a day to exploring the city and casually shop for antiquarian books. Our tourist has brought along her iPhone and intends to solely use the Google Map application to navigate around in the city. The iPhone is equipped with internet access, GPS, and a digital compass. I now perform analysis of Google Maps on the iPhone structured with the human-artefact model. In this analysis we will look for hypothetical problems that might arise in this scenario.

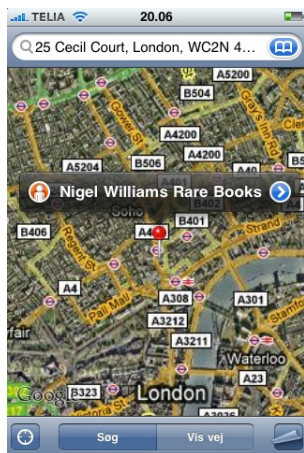


Figure 3.5 Google Maps on iPhone

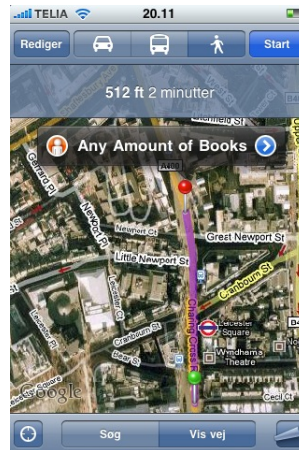


Figure 3.6 Route between two shops

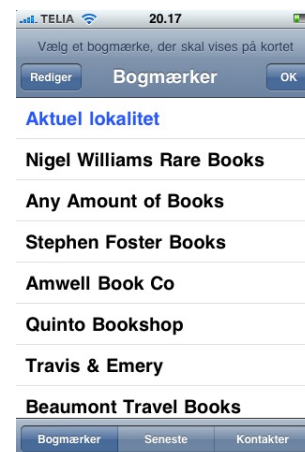


Figure 3.7 List of book-marks

The first step is to provide a rundown of the orienting side of the human-artefact model, meaning the orientation of our tourist. While it is the first time our tourist uses her iPhone to navigate a foreign city she has had a lot of experience with old fold-out tourist-maps, and some experience using different map services on her personal computer.

Motivational orientation

- **Orienting and finding way:** The main motivation for using the map application is to orient her self in the big city, and find her way around.

- **Locating shops:** Secondly she wishes to locate a number of shops she might visit.
- **Planning routes:** Thirdly she would like to find the quickest way to a shop from a given point.

Goal orientation

- **Bookmark shops:** Before venturing out in the city our tourist wants to bookmark potential shops that she might visit.
- **Look up shops:** When moving about in the city she wants to be able to see the shops on the map.
- **Find self:** Furthermore she wants to be able to locate herself if she gets lost.
- **Estimate distances:** Finally to be able to estimate walking distances in the city.

Learned handling

- **Experienced technology user:** Overall our tourist is well experienced with handling the iPhone, PCs, and both PC- and paper-based maps.

Adaption

- **No physical restrictions:** Our tourist has no problems handling the phone in general.
- **One hand interaction:** However, she might need to interact only with one hand when carrying bags of books in the other hand.

Secondly we examine the aspects of the artefact in question. This means the left side of the human-artefact model. For a complete analysis this could include exploring the artefact, consulting documentation and manuals and interviewing developers. However, for this simple example (and simple artefact) I have basically performed an exhaustive exploration. This, however, means that the motivational aspects of the artefact will be based on assumptions and qualified guesses.

Instrumental aspects

- **Look up (place, address, contact, bookmark):** The application supports look-ups on the map either by typing in a location (e.g. "Tower of London"), typing in an address (e.g. "1 Tower Bridge Rd, London"), select-

ing an address from one's contact list, or selecting a bookmarked location (Figure 3.7).

- **Locate your own position:** The GPS enabled iPhone supports pinpointing your (outside) location by the press of a button.
- **Bookmarking:** Locations can be bookmarked for later retrieval one at a time.
- **Get directions, distance, and travel-time:** It is possible to get the shortest path from one place to another, the distance, and the travel-time (either by car, foot, or public transportation) (Figure 3.6). The selection of points is done either through look-up, a placed pin (manually or through a bookmark) or the current GPS position. Directions can be shown either on the map as a path or as a list of turns.
- **Explore different views of the map:** The map can be shown either as a regular geographical map (with markings of roads, sights, hospitals, infrastructure etc.), as a satellite image or as a hybrid between the two.

Motivational aspects

- **Finding your way and orienting in situ:** The GPS and compass afford easy navigation in an unknown area, hence suggesting that the application was designed for in situ use.
- **Planning trips (car, public transportation, or by foot):** The functionality of the applications lends itself well for planning trips – short and long.
- **Visiting a contact:** Integration with the phones address book suggests a design towards easy look-up of known addresses.
- **Exploring the map:** The satellite view of the world and the snappy navigation lends itself well for exploration of the world.

Handling aspects

- **Pan and Zoom:** The navigation on the map consist of traditional panning and zooming.
- **Onscreen typing:** Text-entry is performed through the standard iPhone soft-keyboard.
- **Buttons:** Regular WIMP buttons are used to e.g. shift between modes or locate self via GPS.
- **Modes:** Different views on the map can be selected (regular, satellite, hybrid, and list).
- **Dropping pin:** A pin-dropping tool can be activated and a pin dropped.

- **Drag-and-drop:** The pin can be moved around using drag-and-drop.

Adaptive operational aspects

- **Tapping:** Tapping on the screen is used to activate buttons and double tapping on the map will activate zoom.
- **Dragging:** Dragging a finger across the screen is used to pan around the map. Furthermore dragging can be used to move the pin.
- **Pinching/Reverse Pinching:** Pinching with two fingers can be used to zoom in and reverse pinching to zoom out.
- **Physically moving about:** Given the phone is GPS enabled, the map can be navigated by moving about. Furthermore the map can be rotated with the user. This is supported through the built-in compass.

Figure 3.8 summarises the above in our template for the human-artefact model.

Now assume now that we have an account of the tourist's interaction with the iPhone that day. We will probably see (at least) two issues.

- *Our tourist experiences that one-handed operation is perfectly possible. Interaction is easy, albeit a bit slower, when holding the phone in the right hand and interacting only with the thumb. However, a breakdown occurs when our tourist tries to zoom out to get an overview. Zooming-in works by double-tapping and our tourist tries to triple-tap without luck. Finally she puts her bags of books down and uses two hands.*

The idea is that when a breakdown is observed one should try to explain it, possibly by investigating related artefacts in the user's artefact ecology. Given that our tourist has experience with using the iPhone in general, she has also used it for music. The small button on the head-set cable enables jumping back and forth in an album by double- and triple-pressing respectively. However, this experience was not transferable to the Google Maps application even though afforded by the zooming through double-tapping. This means that a breakdown happened on the handling layer.

- *When wandering around in the city our tourist occasionally wants to know which of her bookmarked shops are close by. She, however, experiences that she cannot have all of her bookmarks shown on the map at a time. She has to load them individually to get them marked on the map. This makes quick estimations of distance cumbersome. Furthermore she cannot quickly estimate a distance on the map without planning a route.*

	Google Maps iPhone	Tourist	
Motivational aspects	Finding your way and orienting in situ Planning trips Visiting a contact Exploring the map	Orienting and finding way Locating shops Planning routes	Motivational orientation
Instrumental aspects	Look up Locate your own position Bookmarking Get directions, distance, and travel-time Explore different views	Bookmark shops Look up shops Find self Estimate distances	Goal orientation
Handling aspects	Pan and Zoom Onscreen typing Buttons Modes Dropping pin Drag-and-drop	Experienced technology user	Handling orientation
Adaptive aspects	Tapping Dragging Pinching/Rev. Pinching Physically moving about	No physical restrictions Potential one hand interaction	Adaptive orientation

Figure 3.8 Google map analysis summarised

Here we see a breakdown on the motivational layer. The artefact is not designed for this particular kind of use. This, however, contrasts the way one normally uses a tourist map, where it is common to annotate the map with multiple sights of interest. Furthermore it contrasts the features available on the PC version of Google Maps where one can easily create annotations on the map. The lack of a quick estimation of bee-line distances mirrors what we observed in the map-study. No scale is provided on the map either. Here the application also exposes internal contradictions; a detailed satellite view of the world without means of estimating distances, seems to conflict with one of the basic features required of a physical atlas.

The design lessons from this very simple analysis would be to:

- Provide zoom-out through triple-tapping to enable complete one-hand operation.
- Enable easy distance estimates. E.g. simply by showing the distance between two fingers on the screen.

- Provide a toggle for showing all bookmarks on the map.

Of course these findings could have been achieved with common sense but the human-artefact model stresses the importance of considering each level and the tensions between the aspect of the artefact and the orientation of the user. Furthermore it frames the investigations of causes of breakdowns. In [chapter I](#) we elaborate on how to perform an extended analysis e.g. encompassing artefact historical analyses.

3.5 Related Work

The work of this chapter (and the associated papers) builds upon and relates to a torrent of work. In [chapter 2](#) I covered the conceptual foundation and in [Section 1.5](#) the research context that the work of this chapter positions itself in. In this section I will more specifically summarise related work through three themes: Conceptualising New HCI, Analysing Artefacts, and Newer Activity Theoretical HCI.

3.5.1 Conceptualising New HCI

One of the central works in conceptualising new trends in HCI that spawned from the move beyond the personal computer is Dourish's 2001 book which is aptly called *"Where the action is"*. Dourish refers to the new types of human-computer interaction stemming from the technological developments in ubi-comp and related fields as *embodied interaction*. According to Dourish embodied interaction is compromised by two interwoven strands of research, which he refers to as tangible computing and social computing. Tangible computing captures the move beyond the desktop computer, the move towards technology that acknowledges us as physical beings. Social computing, in Dourish's words, *"refers to the applications of sociological understanding to the design of interactive systems"*. Hence the acknowledgement of the user as an individual social being, and as a part of a collective. Based on a phenomenological foundation Dourish presents a number of principles for designing embodied interaction. These principles, as Dourish emphasises, are not rules, nor guidelines for designing interactive systems, which would be inappropriate given the diversity of settings where an embodied interaction approach is applied. They are *"things to pay attention to when doing design"* as he writes. I will not cover his principles in details but they are:

- **P1:** Computation is a medium
- **P2:** Meaning arises on multiple levels
- **P4:** Users, not designers, create and communicate meaning
- **P5:** Users, not designers, manage coupling

- **P6:** Embodied technologies participate in the world they represent
- **P7:** Embodied interaction turns action into meaning

The principles echoes many of the insights we have drawn in our work with activity theory, and has previously been drawn in applying activity theory to HCI. One might question why Dourish's book is largely ignorant towards activity theoretical HCI (and other post-cognitivist frameworks), in fact his book does not contain a single reference to either activity theoretical sources or interpretations of activity theory into HCI, even though it had been a player in the field for over 10 years at the time.

To compare Dourish's principles with our work: The work on the human-artefact model addresses directly P2 - P6. We specifically address meaning on multiple levels of activity and the tension between the meaning embodied in artefacts and the user's foundation for making meaning of it. As for P6, embodied technologies not only participate in the world, they dialectically shape the world. Artefacts shape the way we see meaning in other artefacts and objects. This is central in the activity theoretical understanding so I can only agree to P6. Without going too deeply into Dourish's extensive work I can say that our work is quite in line with what Dourish suggests we pay attention to. However, where Dourish has aimed to provide a philosophical foundation for embodied interaction, I have provided means of performing concrete design oriented analysis.

Chalmers and Galani [59] address heterogeneity in theory and in design of interactive systems, leading to a number of guidelines for design. Parallel to my work they question the ideal of seamlessness, invisibility, and embodiment in ubiquitous computing and emphasise that ambiguity and contradictions can be a positive resource for design. They argue that *"Weiser and Dourish both swing from one extreme to the other, focusing almost entirely on design to support embodied or ready-to-hand interaction. They do not fully address the relationship between the two modes. In particular, how does a tool become invisible or ready-to-hand"* [59, pp. 245]. Two of the guidelines they put forth directly back up our work on the human-artefact model. Firstly the notion that that while heterogeneity and ambiguity may be a resource in design it is not an end in itself. Artefacts have to be well designed, while embracing heterogeneity rather than homogeneity. Secondly they point out how past activity is a resource for the current and future action. Hence they need to be explored in design.

To paraphrase [Paper I](#): They conclude their recommendations by stating that *"In the long run, we should consider accommodation and appropriation as a process that designers contribute to by selective revealing system structures, and affordances for their potential use, but it is the users who through their interaction with our system and with each other choose what to use and why"* [59, pp. 241]. Chalmers and Galani do not bring these recommendations any further in general. Through its focus on dialectics, mediation, multiplicity and development the human-artefact model has a contribution to make regarding ambiguity, heterogeneity,

history and seamfulness as a resource in design and in development of use. The human-artefact model brings the focus on use and design further than Chalmers and Galani's guidelines, in particular when it comes to explaining why and addressing how. The dialectics and materialist basis give the foundation for understanding these contradictory resources, whereas with Chalmers and Galani's phenomenological grounding, this is less central. When it comes to how, the four analytic levels of the human-artefact model, its two sides and the unfolding dynamics of these, make it possible to zoom in on the point where designers may make their contributions to a potential change of use, and where and how users may adapt these contributions to their ongoing development of use.

Jung et al. [94] argue in line with me that the increase in use of multiple heterogeneous artefacts calls for considering the ecology of artefacts as a unit of analysis as well as the individual artefact. They empirically show that each interactive artefact in the user's artefact ecology influences how the other artefacts are used. Having both a laptop and a workstation may result in the laptop being used only for work and the workstation entirely for entertainment even though each artefact in isolation is capable of both. The examples from [94] primarily illustrate how artefacts influence one another at the level of actions: What is done with which artefact. With the human-artefact model we propose that it is important to address all levels of activity from the perspective of the artefact ecology, and not only consider the current artefact ecology in use, but also the historical artefact ecology: What artefacts have previously been used to realise the given activity?

Jacob et al. [92] present a framework for what they refer to as *reality-based interaction*. They observe that more and more user interfaces are based on physical metaphors (e.g. BumpTop [3] or the iPhone user interface), or use physicality as a means of interaction (e.g. in tangible computing [91]). Physical-world experiences are broken up in four categories in the paper:

- Naïve physics
- Body Awareness and Skills
- Environment Awareness and Skills
- Social Awareness and Skills

They survey recent research projects for evidence of how these aspects of our real-world experiences are applied in post-WIMP user interfaces. The most interesting part of the paper is their discussion of trade-offs between basing a user interface on physical metaphors and the power of a computer. A discussion, however, that resembles Smith's [156] 20+-year-old discussion of the tension between literal and magic actions in physical simulations.

From an activity theoretical perspective (and an ecological in the light of [8]), it is problematic to distinguish between "real world" or "physical world"

experiences and experiences with for instance handling Finder on the Mac. With the human-artefact model we avoid that clear distinction – instead look at experiences on the different levels of activity. In fact Jacob et al. step into this problem by referring to full-page browsing on the iPhone as more reality based than e.g. column-view browsing found on most other phones with browsers.

One can agree with Jacob et al. that more and more user interfaces base themselves on physical metaphors, especially with the emergence of multi-touch displays and the taxonomy based survey illustrates this trend quite well. However, the human-artefact model is a richer and more suitable tool for analysis, which encompasses both experiences in general and not only from the physical world. The use of the human-artefact model can, however, be supplemented with Jacob et al.'s considerations of trade-offs. The trade-offs are not only relevant in relation to drawing upon physical experiences but on all levels of activity.

3.5.2 Analysing Artefacts

With the task-artefact cycle Carroll et al. [58] present a view on artefacts and human tasks as co-evolving. I.e. a view in clear correspondence with the dialectical understanding of artefacts discussed in [Section 2.3.1](#). The human-artefact model can be seen as a refinement of the task-artefact cycle in three ways: The human-artefact model insists on a dialectical relationship between artefact and user/task, the task-artefact cycle tends to lead to causal explanations; the human-artefact model insists that the human motivation sets a frame for understanding beyond the task; and finally, the levelled approach also gives a useful focus on discrepancies and misfits that eventually lead to a need for change of all levels of the artefact. The task-artefact cycle is defined in a context of personal computing, hence the aspect of other artefacts shaping the user's orienting basis (in my words) is not considered.

Based on the task-artefact cycle, Carroll and Rosson [57] propose letting scenarios drive analysis of artefacts and re-design of artefacts. This approach is highly compatible with the human-artefact model, and in fact the example I gave in this chapter was based on an ad-hoc scenario. This is not an approach we have explored but it is potential future work (perhaps even refined with the use of personas [61]).

As for specific (more application oriented) analytical frameworks and methods Scaife and Rogers [150] analyse graphical representations through a framework based on external cognition. Their work is especially focused on easing the understanding of what the action possibilities of objects on the screen are, and provides a number of guidelines for optimising this aspect.

Beaudouin-Lafon [18] provides (together with the instrumental interaction model, which I build upon in the next chapter) a number of dimensions for analysing user interface instruments. Paraphrasing [Paper I](#): Specifically he develops three ways of assessing such WIMP interface components as scrollbars,

text-fields, buttons etc. in terms of how well-integrated the instrumental, handling and adaptive aspects are. The degree of indirection includes spatial offset and temporal offset. Spatial offset is the distance on the screen from the logical part of an instrument to the object it operates on. This is similar to the relationship between handling aspects and instrumental aspects in our terms; e.g. a dialogue box can be far from the object it acts upon and has a high spatial offset. Temporal offset refers to the time it takes from the point where an action is executed to the visualisation of the effect. Typical scrollbars have low temporal offset while filling out a style-sheet form has high. The degree of integration refers to the compliance between the controlled dimensions of the input device and the dimensions of the logical instrument, the relation between handling aspects and adaptive aspects: 1D scrollbar controlled with a 2D mouse has a degree of $1/2$, while controlling 3D rotation with a mouse has a degree of $1\ 1/2$. Beaudouin-Lafon argues that a degree of 1 is preferable, i.e. that the more direct relationship between handling aspects adaptive aspects the better. Lastly the degree of compatibility is a measure for the similarity between the physical actions of the user on the input devices and the logical actions of the instrument similar to the relationship between adaptive aspects and instrumental aspects. According to Beaudouin-Lafon, dragging an object has a high degree of compatibility while typing in the position of an object to move it on the screen has a low one.

An aspect of analysis which the human-artefact model does not deal with directly is user performance. As for this aspect methods such as the keystroke level analysis [54] should still be applicable although one might have to introduce new constants for e.g. multi-touch interfaces.

Cognitive walkthrough [172] is another classical method that focusses on highlighting misfits between a (hypothetical) user or task and an artefact (typically an application). The human-artefact model could in principle supplement walkthrough based analysis. E.g. supply information for the artefact side of the model. Recently Bertelsen [27] has offered an activity theoretical take on walkthrough methods (the activity walkthrough), which provides a means of contextualising beyond merely a task. This contextualisation could be extended to include the artefact ecology as well.

The activity walkthrough is one of a small handful of HCI methods crystallised from activity theory. The focus-shift analysis [36] is another method used for performing structured tracings of breakdowns in video material of human interaction with computers. In [Paper I](#) we use this method explicitly in conjunction with the human-artefact model to analyse the material from the map-study. Another method is the Activity Checklist [100] which is intended for charting the activity system that a potential interactive system is going to be integrated in.

3.5.3 Newer Activity Theoretical HCI

Besides the more technical work on activity-based computing (e.g. [15]), little work has been done to confront activity theory with the challenges stemming from the new technological developments.

However in recent years a couple of interesting activity theoretical HCI publications have emerged, most of which are covered in my work, but I summarise them here.

Kaptelinin and Nardi's book [98] *Acting with Technology – Activity theory and interaction design* provides a fresh introduction to activity theory in interaction design. Especially their discussion of why theory has a role in HCI is worth reading. However, they do not address the new challenges of ubiquitous computing and all of its derivatives so there is not much new under the sun. For someone unfamiliar with activity theory this book is recommendable.

Bryant et al. [49] provides an exemplary analysis of Wikipedia based on activity theoretical terminology. They analyse the users' development from newcomers to Wikipedians. They characterise the differences between these two groups, and suggest that there are patterns in how users move from being newcomers to Wikipedians. The paper is interesting because it really shows how the motivation for use of an artefact can develop into something quite different. In [Paper I](#) we exemplify the use of the human-artefact model to highlight how the artefact (Wikipedia) supports these two motivations for use.

Finally Bærentsen and Trettvik's paper *An activity theory approach to affordances* [8], which is heavily cited in this dissertation, and on which the human-artefact model is built, provides a quite welcome unravelling of the affordance concept based on an activity theoretical understanding. They conclude by suggesting to explore both the possible and intended uses of an artefact, which is exactly what we do with the human-artefact model.

3.6 Summary

The human-artefact model is a proposed instrument for framing design-oriented analysis based on a solid theoretical foundation. In this chapter I have presented three themes that have pervaded our work leading to the human-artefact model, firstly a reframing of the ideal for design in complex artefact ecologies, secondly a revisit of the affordance concept, and thirdly a focus on learning and dialectics. I have given an introduction to the model and illustrated that it can easily be applied to perform back-of-the-envelope analysis of interaction. Through my overview of related work I emphasised that the human-artefact model is not a model to be used in isolation, but it can co-exist and leverage from a range of methods and approaches. In [Section 5.2](#) I will discuss the potentials of future work in relation to the human-artefact model.

Chapter 4

Complex Artefact Ecologies from Within

This chapter concerns the technological challenges of HCI in complex artefact ecologies. In the light of complex artefact ecologies I question some of the more technical and architectural assumptions embedded in the software executing on our interactive artefacts and present a proposal for an alternative interaction paradigm, *ubiquitous instrumental interaction*. I do not present a complete system realising ubiquitous instrumental interaction – doing that would require an extensive development task. Instead I propose a software architectural model, *VIGO* (Views, Instruments, Governors and Objects), for realising the new interaction paradigm. *VIGO* is a reflection upon what it would require of a software architecture to realise the new interaction paradigm. The potentials of ubiquitous instrumental interaction are elaborated with a concrete study of rethinking paper based laboratory notebooks using the interaction paradigm and interaction model as springboards in participatory design workshops with a group of experimental physicists.

4.1 Chapter overview

This chapter is based on the two papers [V](#) and [VI](#). [Paper V](#) is a technical conference paper introducing the *VIGO* interaction model and demonstrating its potentials. Given the condensed nature of a conference paper I have chosen to use this chapter to elaborate on the premise of the work, the concepts, the implications, and on related work. To create a coherent reading experience it cannot be avoided that some of the body of [Paper V](#) will be repeated in this chapter. [Paper VI](#) will not be covered in detail in this chapter but the results from the study presented in the paper will be used to put light on the potential implications of ubiquitous instrumental interaction and *VIGO* on a concrete application area.

The chapter is structured as follows: In [Section 4.2](#) I problematise the cur-

rent application-centric paradigm for creating interactive software. I introduce the vision of Ubiquitous Instrumental Interaction in [Section 4.3](#) followed by [Section 4.4](#), which introduces the VIGO architectural model for realising Ubiquitous Instrumental Interaction. [Section 4.5](#) provides two examples of implementational considerations when it comes to realising a VIGO based system. [Section 4.6](#) elaborates the potentials of the instrumental approach through the laboratory notebook study. [Section 4.7](#) provides an overview of related work and [Section 4.8](#) summarises the chapter.

4.2 Elaborating the problem

The kind of technology present in Mark Weiser’s ubiquitous computing vision from his seminal paper *The computer for the 21st century* [170] is by and large in our hands today. We have interactive wall-sized displays, tab-sized smartphones and PDAs, and tablet-PCs resembling the pads of Weiser’s visions and prototypes. Wireless and high-speed cellular-networks are becoming ubiquitous, and it can be argued that computing has indeed become ubiquitous – nonetheless not necessarily in the shape Weiser imagined twenty years ago [25]. According to Bell and Dourish [25] the *real* ubicomp is a messy thing characterised by improvisation and appropriation.

Even though our artefacts resemble those of Weiser’s vision and all of our interactive artefacts are interconnected through both local networks and the Internet, a fluent interplay between them is still absent in today’s ubiquitous computing. With a fluent interplay I mean the ability for us to move between artefacts and move work between artefacts in a meaningful and non-problematic way. Interaction with our heterogeneous artefacts today are characterised by continuous reconfiguration and maintenance of synchronisation through considerable effort of the users [134].

I will not enter a lengthy discussion here of whether Weiser’s vision will ever be fulfilled – I believe there are fundamental problems in the notion of the smart or disappearing computer (e.g. addressed by [46, 148]). Weiser’s vision is pervaded by the idea of technology having agency, and fluency is achieved by the technology adapting our surrounding and artefacts to us and our context. Rogers [148] refers to this phenomenon as “hard ubicomp”, which she believes has the same fundamental problems as hard artificial intelligence. Rogers points to Greenfield’s [81] quote: “*We simply don’t do smart very well*”, since it requires solving many hard artificial intelligence problems.

When sympathising with an activity theoretical understanding of technology the notion of agency and computer intelligence is especially problematic. It breaks the asymmetric relationship between human and tools that activity theory insists on. However, it is my belief that interaction with multiple interactive artefacts *can* become more fluent than it is today. I agree with Rogers that it requires us to come down to earth and focus on proactive people rather than proactive computers. Rogers argues that we should enable scientists, doc-

tors and teachers to customise and create mash-ups of technology. Hence we should put tools in the hands of the capable.

A recent study by Oulasvirta and Sumari [135] highlights the ordeals users have to go through when they, as Oulasvirta in [134] puts it, “*do the ubicomp*”. Oulasvirta and Sumari [135] study a group of eleven of what they refer to as “computer jugglers” at Nokia in Finland. For these people computers in various shapes and sizes are truly ubiquitous in their work-life. They all move between workstations, laptops and smart-phones in various combinations through their highly mobile workday. The computer jugglers, however, spend considerable effort in managing their configuration of these interactive artefacts and in synchronisation of data for particular aspects of their mobile work. The intelligent and seamless adaption from the ubiquitous computing vision is far from the reality of the group from Nokia. Oulasvirta and Sumari [135] conclude their study with what can be interpreted as five challenges in future development to multi-device interaction, three of which are relevant in this context:

1. Synchronisation of data between interactive artefacts is so laborious that special discipline of the user is required. The paper identifies a range of strategies for keeping artefacts up to date with the correct and needed data. Common for all these strategies is, however, that they are far from seamless.
2. Communication applications are not designed with migration in mind. Oulasvirta and Sumari give an example of one of their respondents complaining that reading an email on a smart-phone does not necessarily mean that it has been read as properly as if it had been read on a desktop computer, however, the applications do not distinguish. In general users need to be able to inspect data trajectories and histories of how it was modified and used on different artefacts.
3. The resources of co-present artefacts should be combinable. E.g. it should e.g. be possible to use the full-keyboard of a laptop computer to edit an SMS on a smart-phone.

I believe one of the fundamental problems that inhibits fluency and configurability originates from the fact that the software executing on our interactive artefacts is not designed for interaction in complex artefact ecologies.

The tools of interactive artefacts designed today are largely built upon an outmoded assumption of one user – one artefact – one application. This assumption is deeply engrained in modern desktop computing environments and pervades mobile technologies as well. The *read but unread mail* example above quite elegantly highlights this problem. The pervasive concept of *applications* embodies an assumption of fixed work-tasks and work environment, something that is far from the reality of actual multi-device users (as shown in [135]). The dynamism required to realise e.g. the third challenge of above

is contradictory to the static nature of applications. In my own work with re-imagining laboratory notebooks for experimental physicists (elaborated in section 4.6) I saw needs or challenges complementing those observed by Oulasvirta and Sumari. One physicist imagined being able to copy a button e.g. for controlling the force of a pump together with a temperature widget to a mobile artefact so that functionality could be physically brought to the relevant place in the laboratory and not bound to a workstation in the corner of the lab. This kind of interaction seems obviously reasonable when looking at the technical potential of our interactive artefacts, however is far from reasonable in the software architectures *inside* the artefacts.

Beguín and Rabardel [24] conceptualise the dichotomy of *instrumentation* and *instrumentalisation*. The design of today's technology assumes that appropriation happens through instrumentation (the user adapts to the technology) but it rarely supports instrumentalisation (the adaptation of the technology by the user to fit the users' needs). The trend of selling specialised applications for mobile interactive artefacts is a push in the right direction and offers some support of instrumentalisation of one's mobile artefacts, however, it also follows the same paradigm of quite monolithic applications as on desktop computers.

If we take a closer look at the nature of applications as we know them from our desktop computers they are actually quite artificial constructs and have no direct equivalent in the physical world. Applications can be understood as collections of tools dedicated to a certain activity, e.g. the collection of tools of the architect or the graphics designer. However, applications lack the dynamics and flexibility of such collections in the physical world. A graphics designer can freely add or remove brushes from his collection, pass them around, etc., but a brush in a drawing application can rarely be removed and used in another context. Instead, applications typically have a predefined set of tools that is difficult or impossible to reconfigure to one's taste. Applications tend to be inclusive in that they are built to suit every aspect of the activity they embody. Hence they become large and complex with dozens and sometimes hundreds of tools to cover all possible needs. While in the physical world we do complex tasks with simple tools, e.g. draw complex architect drawings with a pen, an eraser and a ruler, in the digital world it is the other way around, we typically do simple tasks with complex tools. It is rarely the case that I use more than a fraction of the functionality of my drawing program or even word processor to do my daily work.

Applications today largely rely on filetypes. If I want to draw an illustration in my text document I typically have to perform a series of exporting and importing from a drawing program to my word processor – and repeat if I need a change in the drawing. Filetypes (especially when proprietary) pose an even larger problem when dealing with multiple different interactive artefacts – it is not always the case that a vendor creates a reader for a given format for a given type of artefact.

Physical tools may have limited properties, but they can be put to unlimited uses and are usually easy to understand and adapt to one's need. For example,

almost any surface, not just paper, affords to be written on with a pen. Collections of physical tools can assist in achieving complex goals and they can be split or recomposed to address specific situations. The architect can bring his drawing, a pencil and an eraser with him to a meeting. He may use the same pencil from the drawing table to write his grocery list or he may use the pen with which he signed the check at the restaurant to draw a sketch on a paper napkin. In other words the tools are not restricted to a specific domain, they can be used across domains, places and people.

Norman [127] argues that in screen based user interfaces affordances in the Gibsonian sense do not exist at all. The user needs to learn the meaning of each function and cannot reason about it without trying or being told what a particular function does; the action possibilities are encoded in the object and not a relation between the subject and object. Physical tools foster affordance-based reasoning and can be appropriated to a degree where the user can “think through” the tool – what in the previous chapter is referred to as the tool becoming a functional organ. Applications can become functional organs to the user e.g. to the skilled Emacs user or the DJ performing live with Ableton Live. However, whether skills acquired in use of an application can be applied in other domains or on other interactive artefacts largely depend on standardisation of the action possibilities offered by similar looking objects. This is not the case for the architect’s pens, rulers and compasses.

It would, however, be overly naïve to let the relationship between tools and objects from the physical world be the ultimate goal for an interaction paradigm. Nonetheless I believe that some of the decoupling between tools and objects can be approximated by relaxing the strong coupling between tools and objects apparent in applications, hereby achieving affordances that are closer to the “real” affordances, which Norman argues are missing.

In this chapter I explore an alternative interaction paradigm named *ubiquitous instrumental interaction* to the current application based on the interaction model *instrumental interaction* [18] (presented in Section 2.6). The aim is to explore the potentials in decoupling interaction logic from domain objects in order to approximate the flexibility of tools in the physical world.

4.2.1 A note on the application paradigm

In this section I establish what I mean with the “application paradigm”. I refer to the notion of the application paradigm as the predominant way we interact with digital objects in computers. In (the yet unpublished paper) [105] we sort out five paradigms for manipulation of digital objects on computers from the perspective of the user based on the perceived relationship between data, data representation and operations.

Data is defined as “symbols inscribed in formalised patterns by human hands or instruments” [65]. In this context data is an abstract term that covers the users’ perception of the data related to a digital object. E.g. the text and

mark-up related to a text document, but not what is concretely stored on the hard-drive or in a database.

Data representation is the result of translations of data to one or several modalities presentable to the user. This is required in order for a user to perceive a digital object. This could for instance be the translation of a text file into a bitmap image shown on a screen.

Operations are defined as the encoded possibilities users have for changing an object.

We argue that from the perspective of the user (in contrast to the developer) these three concepts are always *coupled* in some way or another when the user interacts with a given object. By coupled we mean that if two or more of the concepts are coupled they are inseparable from the perspective of the user. If for instance the user manages to create two simultaneous but different representations of the same data she may infer that in this case *data* and *data representation* are *not* coupled; on the contrary if they always appear “as one”, they are. An example of two simultaneous but different representations of the same data could be a webpage opened both in a text editor and in a browser. What characterises the application paradigm is that operations and data representation are tightly coupled while data is typically decoupled; multiple spreadsheet data can be opened in the same spreadsheet application and multiple webpages can be opened in a browser.

The application paradigm has been dominant since the introduction of the Macintosh in 1984, which introduced the concept of applications to the mainstream consumer. I refrain from talking about the WIMP (Windows, Icons, Menus and Pointing device) paradigm since I am more interested in the relationship between data, data representation and operations, than specific techniques for realising interaction – e.g. whether mouse or touch interaction is applied or applications are presented in windows or not. It can be argued that the iPhone user interface is not following the WIMP paradigm but it certainly is following the application paradigm. The application paradigm, expressed as a relationship between tools, objects and their representation is more general than speaking about the WIMP paradigm as it in principle can also embody post-WIMP interaction as well.

Ubiquitous instrumental interaction is an attempt to conceptually relax the tight coupling between data representation and operations which are predominant in the application paradigm.

4.3 Ubiquitous Instrumental Interaction

Ubiquitous instrumental interaction is a vision of an interaction paradigm alternative to the current application paradigm.

Ubiquitous instrumental interaction extends the instrumental interaction model ([Section 2.6](#)) to an interaction paradigm where instruments are first class

entities to the user when interacting with digital objects. The principal idea is that users should be able to apply their instrument across objects, interactive artefacts and work-domains, instead of moving and converting objects between monolithic applications bound to specific platforms and work domains.

While instrumental interaction was developed in the context of desktop interfaces the concepts are more general and apply particularly well to multi-device interaction going beyond the desktop computer. In ubiquitous instrumental interaction the instrument is an explicit construct rather than a descriptive concept. Ubiquitous instruments should be exchangeable among users, they should work in a similar way on different interactive artefacts, sometimes across multiple artefacts, and they should be independent from the objects they operate on. Instruments should be applicable to domain objects and on interactive artefacts when and where it makes sense for the user, even if they were not designed to do so in the first place. The objects that users interact with through instruments should be able to migrate across artefacts' surfaces, support multiple views, and be manipulatable by an instrument in ways not necessarily anticipated by the object.

The core vision of ubiquitous instrumental interaction can be exemplified by a scenario:

Grace is a graphics designer for a small advertisement bureau. She is currently designing a poster for a political campaign for an upcoming election and heading for a meeting with her client at a café downtown to discuss the poster proposal. To the meeting she brings along an A3 colour-print of the poster. The client is quite pleased with the poster but asks whether it would be possible to create flyers matching the design. Grace thinks for a moment and reaches for her touch screen smart-phone in her purse. Connected to the wireless internet of the café Grace loads the poster object by navigating a list of recently edited objects. Grace already has a configuration of instruments ready for light graphics manipulation on the go. Meanwhile the client is watching and commenting on Grace rearranging, trimming and scaling of the poster into a flyer – the dimensions of which actually match the screen-size of the smart-phone quite well. Given the fast wireless network access of the café, features such as alignment of graphical elements can be computed externally – something the smart-phone does not have the computing capabilities to do without significant latency. Grace's move instrument exploits this and provides alignment feedback to her – Grace has however disabled automatic alignment in the instrument, she hates that. Grace and the client agree on a layout and Grace gives a finishing touch to the flyer at her home workstation in the evening. When she sits down at the workstation that evening she simply "picks up" the flyer from the smart-phone with a finger-gesture and "drops" it on the work-area of the workstation with a keyboard shortcut. In a similar manner she moves a colour-palette with the poster's colour-scheme to the smart-phone and places it left of her keyboard for easy access.

In ubiquitous instrumental interaction interactive artefacts are regarded as opportunities for interaction with one's digital objects. It should be possible for the user to configure artefacts with the objects and instruments needed for a given activity. Grace has a configuration of instruments ready for the kind of work she typically performs on her smart-phone. Objects are accessible from any of the user's artefacts and can be loaded either through search and browsing, or by moving objects between artefacts when they are co-present.

The co-presence of two interactive artefacts should ideally add value to use. If only a simple artefact is at hand a user should be able to do with it what makes sense. It makes sense to make minor adjustments to a poster or annotate an architect drawing on a PDA in order to incorporate the changes thoroughly later on a more capable interactive artefact such as a PC.

Objects are central to ubiquitous instrumental interaction. All HCI in ubiquitous instrumental interaction is directed towards objects and mediated through instruments. Objects are not defined by a specific application or application area, but can evolve through use. This means that objects will change when new instruments are applied and new instruments can be added by the user when needs regarding the object change. Configurations of interactive artefacts and instruments are assembled by users to facilitate manipulation of objects given the nature of the objects.

Given that all interaction is directed towards objects all state should be expressed in objects; anything from text-objects to objects resembling WIMP widgets. Hence it should be possible to create mash-ups of objects whether they be architect drawings, a colour-palette, spread-sheets, a board-game or controls for a pump (see Section 4.6).

Returning to the notion of instruments; to perform specialised operations special instruments are obviously needed but more open-ended instruments that are usable in ways not necessarily anticipated by their designers should be available as well. General instruments like for instance a drawing instrument with colour picking capabilities should be able to pick a colour from anything with a colour attribute and draw on any surface. The same goes for instruments to input text, manipulate geometric properties, cut and paste etc. Furthermore such general instruments have the potential for limiting handling breakdowns as discussed in chapter 3. Hence it should be possible for the user to have a favourite instrument for e.g. text-editing, configured with e.g. custom spellchecker or rules for programming-code auto completion.

4.3.1 Architectural Requirements

I now take a step down from the more abstract vision in order to concretely discuss what it would require from a software architecture to realise ubiquitous instrumental interaction. The road from a vision like the one above to an actual running system is long. And there is not necessarily one road to take. Grace's poster manipulation requires both clever interaction design and a quite differ-

ent software architecture from the ones running beneath the user interface of today's interactive artefacts.

Implementing Ubiquitous Instrumental Interaction requires a software architecture that enables the flexibility that was described above in the context of a distributed object infrastructure. Instruments are central in this line of thought, hence a software architecture based on small-grain components that can be reconfigured according to the users' needs or the available artefacts is the most appropriate. This leads to three overall requirements:

Decoupling

Objects and instruments should be separate components that communicate through a simple protocol allowing instruments to query and modify objects. For example any object that provides a 2D surface can be used by a pen instrument to add annotations. Application logic should be separate from objects in order to allow less-capable interactive artefacts to modify complex objects – or to enable distributed computation as in the scenario above.

Integration

The system should appear as a single, consistent entity to the user transcending the individual interactive artefact. The user should be able to access all her digital belongings from any of her interactive artefacts (what [119] refers to as a *personal information cloud*).

Distribution

Objects should be easily distributable and it should be possible for instruments to communicate across artefacts to support e.g. moving objects between artefacts when co-present. The requirement of decoupling is also a requirement seeded from the need for distribution. Decoupling can facilitate easy distribution and replication of objects across the network.

4.4 VIGO

VIGO (Views, Instruments, Governors and Objects) is an architectural model I have developed in collaboration with Michel Beaudouin-lafon who introduced instrumental interaction in the first place. VIGO is a model along e.g. Model-View-Controller (MVC) [143] or Presentation-Abstraction-Control (PAC) [62], however, it is designed towards realising interactive systems based on the ubiquitous instrumental interaction paradigm. VIGO embodies the basic requirements for achieving the flexibility required by ubiquitous instrumental interaction.

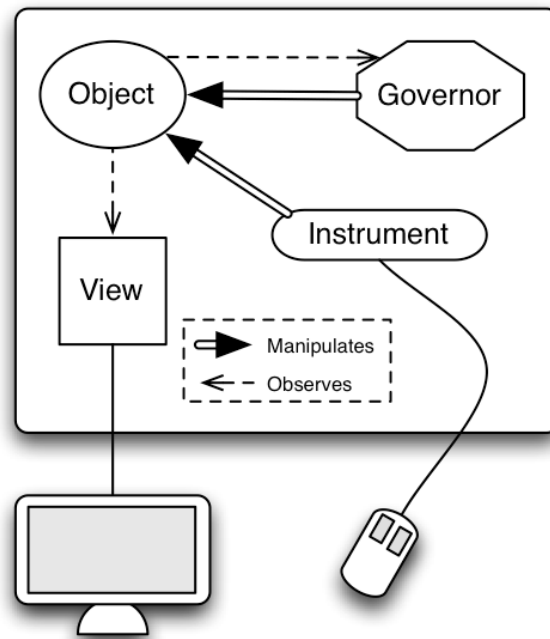


Figure 4.1 The VIGO architectural model

Figure 4.1 presents the VIGO architectural model in pseudo-UML. I give a brief description of the model and then go through each of the components in detail. *Objects* are presented to the user through *views*. Users manipulate objects through *instruments*, which query views to identify the objects being designated. In order to manipulate an object an instrument queries the *governors* attached to that object to validate its manipulations. Governors observe object changes to implement potential side effects. These side effects can both be local to the given object manipulated by an instrument, or be side effects to other objects attached to the governor. Side effects could be anything from objects “physically” affecting each other when moved around, or a time-stamp or a list of recently edited objects being updated when an object is interacted with.

The following part of this section is an elaborated walkthrough of the components of the VIGO architectural model based on the paper *VIGO: Instrumental Interaction in Multi-Surface Environments* (Paper V). In this section I mainly give a general introduction to VIGO whereas more technical details on implementation can be found in the paper. In the paper we especially discuss how to realise Rekimoto’s pick-and-drop technique [144] with VIGO. These details I will not go into here.

4.4.1 Objects

Objects in VIGO are pivotal in all user interaction. An underlying principle in a VIGO based interactive system is that every interaction must target one of several objects and may result in a change to the state of these objects. The idea is that objects in VIGO are distributable across interactive artefacts. This means that the objects and instruments used by a user in a given situation can span multiple interaction artefacts (as it is the case in the scenario above). Interaction with an object on one interactive artefact may have side-effects on another object residing on a different interactive artefact (which I elaborate in [Section 4.4.7](#)). Hence, rather than interacting with the “system” or an “application” running on an interactive artefact, the user interacts with objects that are residing on interactive artefacts configured with appropriate instruments. Given that all interaction must be directed towards objects, it also means that buttons and other traditional widgets are all objects manipulatable on similar conditions as for instance text documents.

Objects in VIGO differ, however, from objects in the classical object-oriented sense. VIGO Objects are passive collections of structured data. This means that they do not provide operations or methods for manipulating them (beyond primitive get/set attribute, append child, etc.). Objects in VIGO expose a set of directly accessible properties rather than hiding their state through encapsulation. Behaviour which is typically provided through private methods¹ is extracted to and encapsulated in *governors*. The decision for passive objects is based on two requirements. Firstly in order to achieve decoupling between operations and data instruments should define how objects are manipulated, not the objects themselves. Secondly, ubiquitous instrumental interaction is a vision of easy access to the users’ objects across multiple interactive artefacts, hence objects should be lightweight. Objects (or parts of objects) can be distributed over multiple artefacts. In our implementation documented in [Paper V](#) the state of objects is kept synchronised across the replicas. Because objects are passive this can be implemented simply and efficiently. Given the structure of objects it is natural to store them in a mark-up language such as XML, which is used for persistence in our prototype implementation.

Objects can be primitive, consisting only of properties, or composite, consisting of properties and other objects. An object can be a part of several objects simultaneously, for instance when a diagram is used in multiple documents: A change in the diagram is then reflected in all the parent documents. This is a different concept than multiple views, which is discussed below. Examples of concrete objects are text documents, graphic canvas, the board and pieces of a board-game or concrete user interface elements such as buttons or menus.

The fact that objects are open gives tremendous power to instruments. On the one hand it makes it possible to “break” an object by putting it into an inconsistent state. Governors are designed to control this potential chaos. On

¹At least this is the case in more traditional Object Oriented languages such as Java. Python, however, relaxes encapsulation and has objects where all attributes can be accessed directly.

the other hand it makes it possible to implement interactions that were not anticipated. For example in a board game the board could be used to play a different game (the rules will be embodied, as we will see, in governors) or it may be possible to annotate the board with arbitrary marks. The traditional approach where objects are only accessed through methods clearly precludes such unanticipated uses.

4.4.2 Views

Views on objects are translations of objects into one or more modalities perceptible by the user. Views in VIGO are very similar to views in MVC. A typical instance of a view is a visualisation of an object (and its subobjects) on a screen. Views are strongly coupled with the objects they represent so that any change of the object is reflected in the view. Views on the other hand do not provide any kind of interaction: any change of the view is the result of a change to the object. Views can, however, be queried by instruments to provide translation of co-ordinates (in the case of a visual display) from its enclosing co-ordinate system to that of the view.

Unlike objects views are device-dependent, i.e. they optimise their representation of the object for the display device at hand. Multiple views can be associated with an object, in which case they are synchronised. Note, however, that multiple views of the same object can also occur when an object is shared among several parents as described in the previous section.

Since objects are passive and views are pure representations they can be implemented efficiently, e.g. through an observer pattern [76].

A view will itself have an object representing it with primitives such as screen position and dimensions for instruments to manipulate. Depending on what kind of view-management is preferred, view-objects can either be local to an interactive artefact or shared. If shared, resizing a view on one artefact will effect the others as well.

4.4.3 Instruments

Beaudouin-Lafon [18] originally defined an instrument as:

... a mediator or two-way transducer between the user and domain objects. The user acts on the instrument, which transforms the user's actions into commands affecting relevant target domain objects.

As a user instruments can be understood as a means of reaching into the computer akin to tools of the physical world: a stick to enhance one's reach, a pen to write on a piece of paper, a hammer to drive a nail. Examples of digital instruments include those to enter text, manipulate graphics, draw, select

objects, etc. The tools in the tool-palette in applications such as Photoshop are typical examples of constructs that can be described as instruments.

As a developer one can think of instruments as event processors. Instruments react to input from the user through input devices or other instruments, and change objects or fire new events for other instruments to process.

Instruments provide feedback to users through objects with associated views. A simple example from our prototype is a gesture instrument using a translucent “glass pane” object covering the whole screen to render gestures to the user. This glass pane object can be used by multiple instruments. Instruments can also have objects associated to them for other instruments to manipulate. E.g. a circle object indicating the brush-size of a drawing instrument, which can be resized by a proper resizing instrument.

Interaction typically occurs through chains of instruments. E.g. an instrument for selecting an object on the screen is chained to an instrument for moving objects. Or a gesture instrument is used to activate other instruments.

Instruments can be distributed and share events and objects across the network. In [Paper V](#) we demonstrate two distributed instruments: A colour picker and a pick-and-drop instrument inspired by [144]. The pick-and-drop instrument enables picking objects on a PDA and dropping them on a PC (in this case attached to a smartboard). One could also imagine creating remote-control instruments treating e.g. a touch-screen of a mobile phone as an input device for controlling the cursor on a PC. The domain objects of the user are oblivious to the distributed nature of instruments applied on them. In [Paper V](#) we show how pick-and-drop can be used on any object, and the colour-picker can be used on anything with a colour.

Furthermore in the prototype implementation of [Paper V](#) we have a range of simpler instruments such as instruments for drawing, moving, resizing, deleting etc. These instruments are usable both on the smartboard and the PDA and the implementation is the same. Furthermore on the smartboard we have a gesture instrument for activating the other instruments.

4.4.4 Governors

Instruments alone are not enough to implement what is required from an interactive system. Consider a board-game such as Checkers or Othello. Implementing a game like this only through instruments, objects and views would require the user to enforce rules manually in the same manner as if played physically. It would be natural to assume that one could implement the rules of the game in an instrument, however, that would be problematic. This would either require implementing a specific instrument for each kind of game, which breaks with the ideal of polymorphic instruments. Or an instrument would have to embody the rules of all known board-games, which would be impossible. This principle applies broadly to many aspects of interactive systems. Another example where this problem is apparent is if one was to implement

a code editor with syntax highlighting or code-completion. A text-edit instrument cannot know every programming and mark-up language.

Solving this problem by implementing the functionality directly *in* objects would hamper the distributability of our otherwise passive objects and overall go against the fundamental principle of decoupling.

Instead we introduce the concept of *governors* to embody the rules and consequences of interaction with objects. Governors embody the cybernetic reactive aspects of a system, hence the name. Cybernetics stem from the Greek *kybernetes* meaning steersman, *governor*, pilot, or rudder [116]. Governors implement the “application logic” or “business rules” commonly found in the MVC Model.

Governors are stateless chunks of code associated to objects at the level of individual properties. Instruments can query governors to validate a change of an attribute and the governor is notified when a governed attribute is changed so it can execute potential consequences. The fundamental idea is that a governor is associated to the parts of an object that is relevant to the given governor – and oblivious to interaction beyond those particular parts. Multiple governors can be associated with a single object.

Now returning to the example of board-games, consider the game of Othello. Othello is a rather simple board-game consisting of an 8x8 board and pieces in two different colours (Figure 4.2).

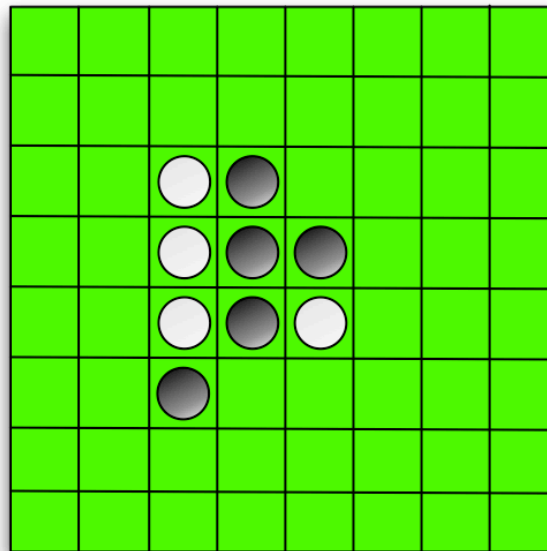


Figure 4.2 Othello

If Othello should be implemented in VIGO the board and pieces would be represented as objects and the game would be played through an instrument capable of moving pieces on the board. The rules of the game would be embodied in a governor associated with the position of the pieces on the board.

This means that if the user manipulates the position of the pieces the governor would be involved. However, the user can freely move e.g. the game board or draw on the surface with a drawing instrument without the Othello governor being involved.

Given that there is not *one* way to implement governors, the interaction between instruments and governors could happen like this: When a piece is moved the moving instrument queries the governor of the piece when it is about to drop the piece on the board. The governor checks that it is a legal move and if so returns the proper position for the piece, i.e. the middle of the destination square. If it is not a legal move the instrument can query for a list of valid moves, which can be used as feedback to the user. Once the piece has been set to the new position the governor is notified of the change and it applies the side effect, i.e. turns over the pieces according to the rules. It should be noted that since it is up to the instrument whether the governor is queried or not, it might decide to bypass it and cheat or even break the game. Finally if the Othello governor is to keep track of turn-taking in the game it must store this state in a separate object or in the board itself.

Governors are not the interfaces to objects that instruments call through to perform manipulations. Nor are governors acting completely behind the scenes only reacting to manipulations performed on objects. Both of these cases would lead to a lack of flexibility. In the first case instruments would not be able to bypass the governors; in the second case instruments would not be able to visualise what the governor proposes, such as the valid moves in the Othello game. Instead governors and instruments negotiate: instruments query the associated governors to validate a manipulation or to get the valid or suggested manipulations (valid moves in the case of the Othello game) which they could ignore in principle (which in Othello would break the game).

To sum up the characteristics and requirements of governors:

- Governors embody rules and consequences for object manipulation.
- Governors are stateless, all state are in objects.
- Governors are only associated to relevant aspects of an object.
- Governors can be bypassed or disabled.
- Governors can be queried for validating a value or for getting a representation of the valid values of an attribute.

As mentioned, multiple governors can be associated to one object. Furthermore multiple governors can work together. Figure 4.3 illustrates an implementation of the Othello game in VIGO. On the first upper left board no governors are associated to the objects, hence the pieces can be placed anywhere on the board. On the upper right board there is a governor making sure that the pieces are only placed in the centre of the cells on the board. When a piece is placed this governor tags the piece with a position on the board in

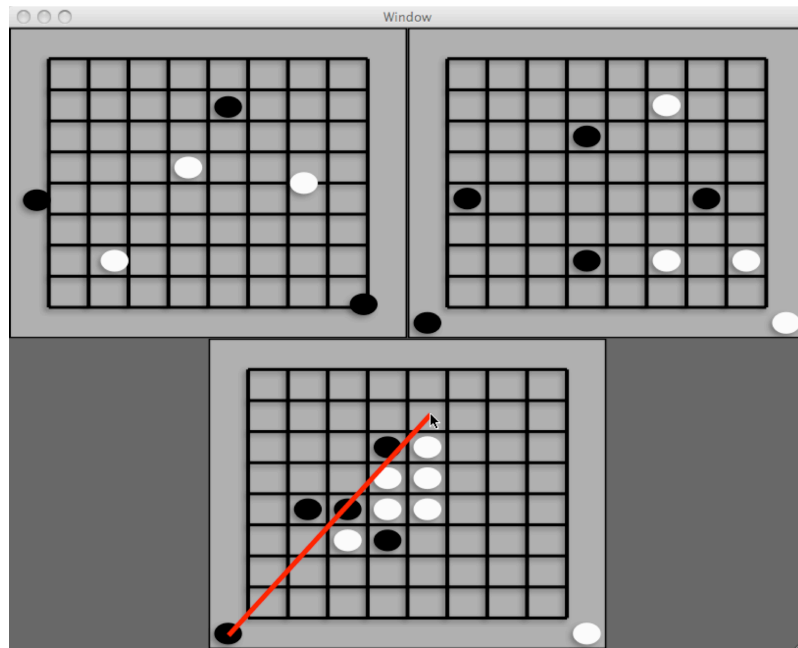


Figure 4.3 Othello governors

the form of e.g. *E5* (ranging from *A1-H8*). It also places a fresh piece in the corner of the board when one has been moved to a cell. On the lower board both a general board governor and an Othello governor are active. The Othello governor is associated to the tagged-on position and when the instrument queries associated governors to a piece, the intersection of valid positions is computed. On the lower board it is shown how the moving instrument uses information from the governors to visualise that the given position is illegal by colouring the line from selection to target red.

Given the fact that governors are stateless, a governor does not necessarily have to be executed the same place as where the interaction with an object is performed. In the prototype implementation of [Paper V](#) all governors reside on a dedicated server and are queried through remote method invocation². However, if an object is shared between two clients each machine could also have an instance of the object's governors or only one of the machines could hold the governors and let others query it. One could also imagine hybrid approaches where some governors reside locally and others are distributed on the network.

In principle views can be regarded as a special case of governors, however they are so frequently used and adhere to a fixed interface towards instruments that they deserve their own concept in the architectural model. In [Section 4.5](#) I will exemplify how it can make sense to let a view be both view and governor for an object.

²This is not necessarily an approach that scales very well. The handling of distributed governors is a subject for future work discussed in [Section 5.2](#)

4.4.5 VIGO and MVC

VIGO is conceptually close to MVC but has a number of key differences. First of all VIGO is designed for distributed interfaces. Secondly VIGO stresses the notion of mediation that is absent in MVC. In MVC the view is both a visualisation of the model and of the “tool” interacting with the model, whereas VIGO puts the visualisation of the model in the view and the visualisation of the tool in the instrument (through the associated objects and views of the instrument). This makes it possible for multiple, independent instruments to interact with the same model through the same view, which is different from MVC’s model-sharing through multiple views.

Instruments have no direct equivalent in, and are somewhat alien to, MVC. Controllers are what resembles instruments the most but they are tied to a certain view. This makes the implementation of traditional widgets, such as scroll-bars, easy. However, a simple interaction such as drag-and-drop is difficult to implement with MVC (in fact, it does not really fit the pattern). In VIGO it is possible both to implement traditional widget-like instruments and instruments that are loosely coupled with their target objects. In section 4.5 I discuss the implementation of two traditional widgets in VIGO.

In other cases the equivalent of instruments are implemented in the controller of an object in MVC. For example an implementation of a text area with MVC typically includes the text editing commands in the controller. This merge of a domain object (the text itself) and the instruments to manipulate it (the text editing commands) in a single MVC entity does not provide the separation of concerns that VIGO encourages. It also limits extensibility, e.g. adding a highlighter tool can be accomplished in a VIGO architecture by adding an independent instrument while it requires changing the code of an existing controller in MVC. Finally, the MVC state-based Model is not explicitly present in VIGO. Instead the Governor represents behavioural aspects of the model that are relevant to interaction. This provides a high degree of flexibility that is difficult to obtain with MVC.

4.4.6 Prototype implementation

In [Paper V](#) we show a prototype implementation of VIGO. The prototype system is a client/server based system implemented in Python. The server side has three facets: A centralised shared object space, distributed event handling and governor handling. All network handling is implemented through the Twisted [73] distributed computing framework.

The client side consists of a Smartboard connected to a Mac with views implemented in Cocoa and a Nokia N810 Internet Tablet with views implemented in GTK (Figure 4.4).

The prototype is not a fully fledged system but implemented to illustrate how the VIGO architecture can support multi-surface interaction. The motivat-

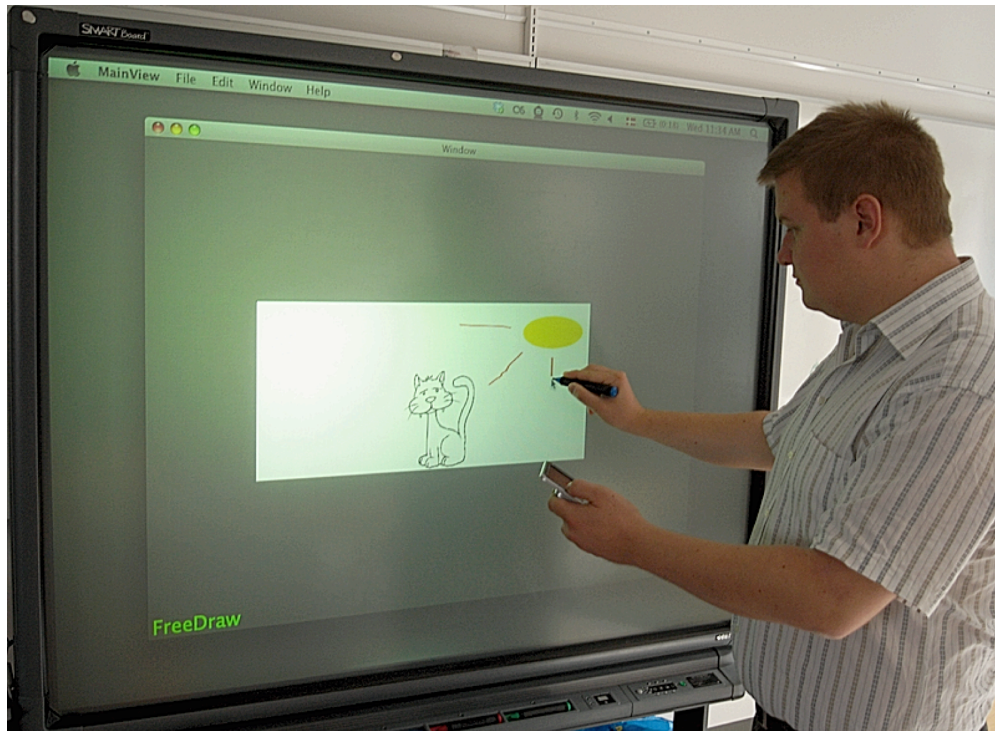


Figure 4.4 Prototype setup

ing example in the paper is Rekimoto's pick-and-drop interaction technique, which is difficult to implement in a contemporary interactive system but, as we show, easy to implement on top of the VIGO architecture as a distributed instrument. The prototype enables us to compose a drawing by picking objects from the PDA (the Internet Tablet) and dropping them on a canvas on the smartboard through the pick-and-drop instrument. It also enables us to draw on the smartboard with colours picked from a palette object on the PDA.

On top of the prototype architecture we have implemented the Othello game through the use of governors. The prototype system enables us to pick up an ongoing Othello game on the PDA with our general pick-and-drop instrument and drop the game on the smartboard. Since the objects are synchronised the game can be continued on both artefacts. Hereby a primitive multi-player version of Othello is created (Figure 4.5). Hence the VIGO architecture enables moving objects that resembles live applications between artefacts.

It should be noted that the pick-and-drop instrument capable of moving e.g. live Othello games around is approximately 150 lines of Python code where a large fraction of this is boilerplate code for the state machine library used to implement instruments.

A video demonstration of the prototype system can be found at the following address³: <http://cs.au.dk/~clemens/dissertation/VIGO-demo.mov>

³Should the link be dead at the time of reading please contact me for a copy of the video.

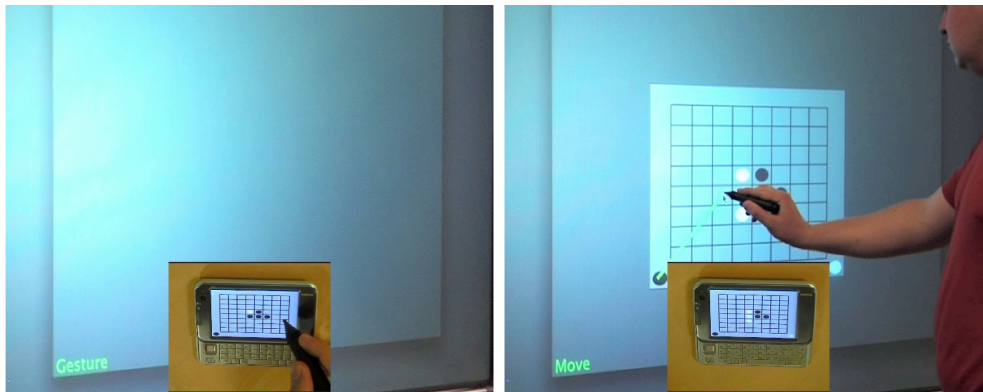


Figure 4.5 Live Othello game picked up from a PDA and dropped on a smartboard

4.4.7 Representation through Governors

Governors embody the rules and consequences for manipulating an object, however, governors can also be used to maintain two representations of an object. Figure 4.6 illustrate how e.g. a programming project could both be represented as code and as a UML diagram. The relationship between the code and the UML diagram would be maintained through a governor. This means that there are two objects, each having their own views and applicable instruments. The governor would react e.g. if the user changed inheritance relationships in the code by updating the UML diagram. However, the governor would only react if interactions on either object affected the other. On the code objects regular text editing instruments could be applied while on the UML diagram instruments for manipulating graphical objects would be applicable.

In the MVC line of thought multiple representations are typically handled by having multiple view-controller pairs to the same model. This is not possible in VIGO. It would not make sense to use instruments on objects where the representation in the view differs fundamentally from what the object represents. It would be exceedingly hard for the user to assess which instruments to apply. It would not make sense to write a view visualising statistical data directly to a pie-chart since instruments otherwise used to manipulate shapes would not be applicable. E.g. if one was to visualise the data as a pie-chart, it first has to be translated into an object consisting of shapes (e.g. circle-slices) and then visualised through view capable of drawing shapes. This transformation would be handled through a governor.

However having the different representations as explicit objects makes unintended use possible. The user could e.g. draw on the UML diagram or if the user needs a snapshot of the diagram at a given time the UML object can be copied and governors detached.

Since objects are easily distributable it would be possible to have two representations on two different interactive artefacts as illustrated in Figure 4.6.

The governor construct also makes it possible to let objects represent phys-

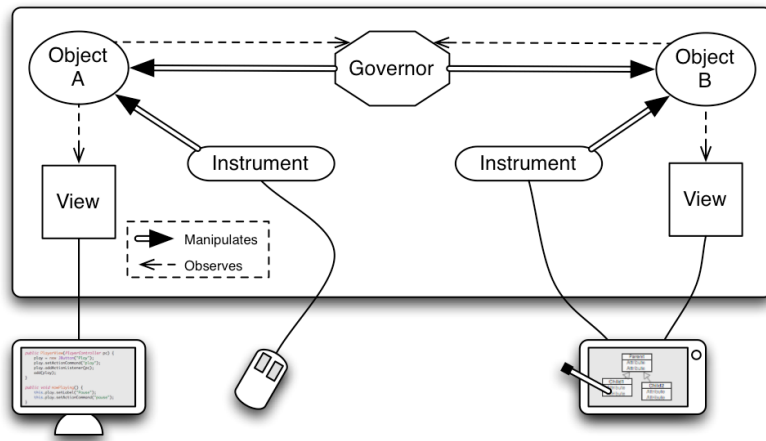


Figure 4.6 Two objects

ical devices, e.g. sensors or servo-motors (Figure 4.7). Certain properties of the object could be mapped to e.g. the position of the servo or the reading of the sensor. It would require a computer connected to the physical device to host the governor and have a replica of the given object. However, it would then be possible to easily move the object between artefacts (as with the Othello game) hence facilitating e.g. remote control of equipment.

In connection with the implementation of the VIGO prototype system, I created a simple object representing a printer through a governor. This printer object enabled the user to print any object (in its XML representation) by dropping the given object on the printer object using the pick-and-drop instrument. The printer object could itself be moved between interactive artefacts, hence a printer could be shared without any setup.

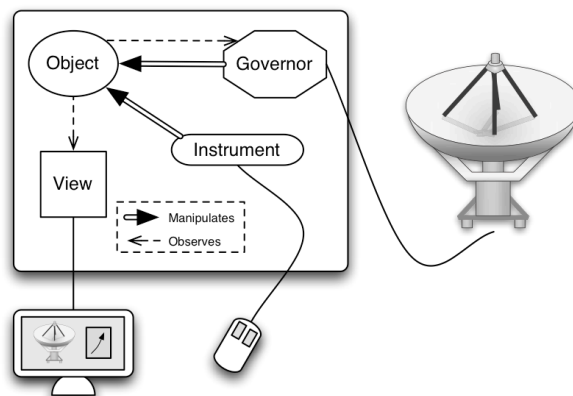


Figure 4.7 External Object

4.4.8 Summing up so far

I have introduced the vision of ubiquitous instrumental interaction and presented a proposal for an architectural model to realise the vision. Table 4.1 summarises the levels of the proposed vision compared with how the world of computing looks today.

	Proposed	Currently Dominant
Interaction Paradigm	Ubiquitous Instrumental Interaction	Application Paradigm
Interaction Model	Instrumental Interaction	Direct Manipulation
Architectural Model	VIGO	MVC

Table 4.1 Comparison of the levels of the two paradigms

Returning to the scenario of our graphics designer Grace. The VIGO architectural model provides a foundation for implementing a system realising the scenario.

If we look at how one might implement the scenario through VIGO the first feature introduced is a recently edited list. Each user has an associated root object accessible on each of the user's interactive artefacts. A list of recently edited objects maintained by a governor would be a child of this root object. Configurations of instruments would likewise be represented by objects. Each artefact – or type of artefact – has a collection of configurations associated to it. This would enable the user to configure an interactive artefact, e.g. the smart-phone, at a workstation (as done in the scenario). With the graphical objects an alignment-governor would be associated. This governor would append alignment information to the objects so the move instrument could use it and provide feedback appropriately. This governor is potentially computationally heavy, hence it is not executed on the smart-phone. Together with the whole of Grace's data-objects the governor could reside on a virtualised server run by a hosting company paid by her advertisement bureau. If no internet connection was present, or only low-bandwidth cell-phone network, the governor would be disabled and the user experience would be gracefully degraded. Distributed governors are demonstrated in [Paper V](#).

There is of course a lot of technical and implementational assumptions in the discussion above. However, with VIGO we have shown that the scenario is realisable in principle. Realising it fully, however, is beyond the extent of this dissertation work. It would require both extensive interaction design and a considerable software engineering task. Furthermore, realising a VIGO based system would require toolkit development from scratch since all modern advanced toolkits are built upon MVC and the assumptions of applications.

4.5 Some implementation considerations in VIGO-based systems

To give an idea of some of the low-level consequences on implementation strategies of the VIGO model I will now discuss how simple widgets such as buttons and text-fields could be implemented. The solutions I propose are not definitive but intended to highlight some of the challenges and potentials that arise.

4.5.1 Example: Buttons

Implementing a regular button in VIGO is simple. All it requires is an object with a boolean stating whether the button is pressed or not. A view associated with the object can then render it and change the visual appearance according to the observed state of the button. The action of the button will be carried out by a governor associated with the boolean attribute. In order to interact with the button an instrument capable of switching the boolean is required. Given that the button is an object it can be replicated on another interactive artefact to e.g. provide remote control.

4.5.2 Example: Textbox

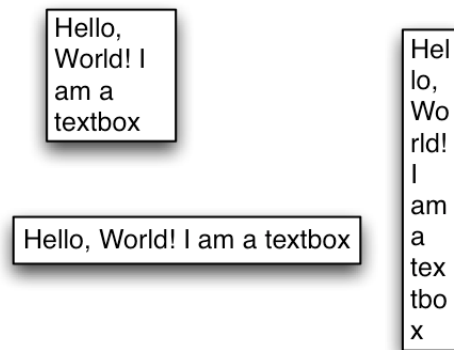


Figure 4.8 Fixed width text-boxes

Consider simple textboxes like the ones found in many graphics editing applications. This is a type of textbox where the width can be resized while the height expands given the content ([Figure 4.8](#)). Such textbox could be implemented in VIGO in the following manner: The object could be as simple as the one in [Figure 4.9](#)⁴. The textbox consists of an object with the properties of the box itself and a child object containing the content. The textbox object has a view type attribute and a reference to the actual instance of the view. This reference is used for instruments to access translation methods of the view. The

⁴The XML syntax in this example is a pseudo-code version of the one used in the prototype of [Paper V](#)

view implementation *expanding_textbox_view* implements the visualisation of the object to the screen. It provides a method instruments can call for translating positions in the co-ordinate system enclosing it to its own co-ordinate system, and furthermore a method for translating a (xy) position to a character position. This is required for instruments to know where the insertion point in the text is given e.g. a cursor position on the screen. Given that the view computes the text-wrapping, the view must provide a *lineUp* and *lineDown* method that takes a character position and returns the position on the line above or below.

```

1 <textbox x="0" y=" 30" width="20" height="10" viewType="exp_textbox"
  viewInstance="iuje28" id="77cc72h">
2   <text id="jjnb2ee">
3     Hello, World! I am a textbox
4   </text>
5 </textbox>

```

Figure 4.9 Text-box object

A cursor is one of the basic prerequisites for text-editing. One of the principles of VIGO is that anything shown on the screen is represented by an object, hence cursors should be objects as well. One approach is to put the cursor directly in the object (Figure 4.10). A text-instrument would then maintain a list of its cursors.

```

1 <textbox x="0" y=" 30" width="20" height="10" viewType="exp_textbox"
  viewInstance="iuje28" id="77cc72h">
2   <text id="jjnb2ee">
3     Hello, World<cursor color="#FF0000" id="kkmn27" />! I am a textbox
4   </text>
5 </textbox>

```

Figure 4.10 Text-box object with cursor

Hereby it is supported to have two text-boxes on two different interactive artefacts where each artefact can have an instrument with its own cursor associated. Without the view dimensions are shared between the artefacts (Figure 4.11). Since objects in the VIGO architecture are replicated multi-user editing through this approach is extremely simple. This would be hard to achieve if either the instrument or the view held the cursor information.

The textbox should increase its height automatically given the content. To be strict to the VIGO principles this should be handled by a governor that observes the content and adjusts the size of the view accordingly. However, the governor can be the view itself. In this case it makes good sense to let it be exactly that. The view holds all information regarding font rendering, which influences the required height of the text-box (Figure 4.11). The view could either be treated like any other governor (Figure 4.12) or this principle could

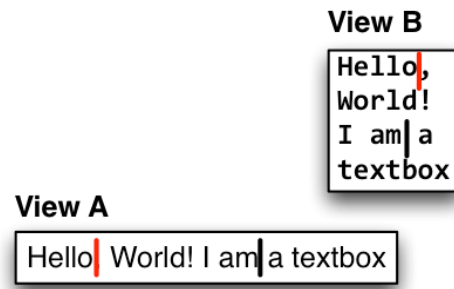


Figure 4.11 Two views on the same text object with two cursors

be relaxed and e.g. a boolean property could define whether it should extend automatically or not.

```

1 <textbox x="0" y=" 30" width="20" height="10" viewType="exp_textbox"
  viewInstance="iuje28" id="77cc72h">
2 <text id="jjnb2ee">
3   Hello, World<cursor color="#FF0000" id="kkmn27" />! I am<cursor color="
  #000000" id="luj28f" /> a textbox
4 <governed by="exp_textbox" refId="iuje28">
5   <governs action="edit" />
6 </governed>
7 <governed by="exp_textbox" refId="fhj329">
8   <governs action="edit" />
9 </governed>
10 <governed by="EN_spellcheck" refId="muj3li">
11   <governs action="edit" />
12 </governed>
13 </text>
14 </textbox>

```

Figure 4.12 Text-box with governors

To finish the example, a simple text-instrument should be able to process e.g. mouse-events for selecting an initial cursor position, inserting the cursor, manipulating the string, and moving the cursor given arrow-key events. The spellcheck-governor in [Figure 4.12](#) could potentially append spelling suggestions in mark-up, which a more advanced text-instrument could use to provide feedback to the user.

4.6 Elaborated vision: Laboratory notebooks

The scenario with the graphics designer of the preceding section is interesting because it shows aspects of computer use that are not easily achievable with our current interaction architectures. However, one could argue that being able to manipulate graphics on a smart-phone is not what would convince the world

of the benefits of ubiquitous instrumental interaction – in principle Grace could just bring along her laptop. In this section I will address an area where the limitations of current interaction architectures seem to hamper the digitalisation of an analogue artefact that seems peculiarly out of place in a definitely complex ecology of digital artefacts. The section has its starting point in a recent pilot study aimed at uncovering the potentials in applying ubiquitous instrumental interaction and the VIGO architectural model in rethinking laboratory notebooks. The study is presented in full in [Paper VI](#). The study consisted of interviews and two future workshops [95] with five experimental physicists.

The research agenda of the study was not to design a new laboratory notebook but to assess whether ubiquitous instrumental interaction could act as a springboard [71] for an expansive solution (as discussed in [Section 1.6](#)).

The study was also motivated by a desire to take a closer look on the role laboratory notebooks play in the physicists' activities. What is it that makes it special? Can we describe it theoretically? How does it suit the principles of ubiquitous instrumental interaction?

In this section I present some highlights from the study and aggregate the ideas from the workshops in three visions. For the theoretical conceptualisation of laboratory notebooks I refer to the paper.

Laboratory notebooks are interesting from an HCI perspective since they are artefacts that seem to be extremely hard to digitalise. In the industry systems electronic laboratory notebooks (ELN) are beginning to become commonplace, especially in the biochemistry and pharmaceutical industry where the patent race make hard demands on what needs to be documented. According to Butler [51] the paper-based laboratory notebook is still the referential source of documentation in academic research laboratories around the globe. However, the paper based laboratory notebook is becoming an anachronism in the increasingly digitalised laboratories and digitalised activities of modern science.

4.6.1 The notebooks

A laboratory notebook is typically the primary record of experimental research. Interviews with the physicists showed that laboratory notebooks are used for quite diverse activities. They were used when actual experiments were executed, in analysis of data, in planning experiments when the equipment of the laboratory was maintained, in paper writing and, last but not least, in documenting the research of the laboratory. Most of the physicists used a shared laboratory notebook in their group resulting in the laboratory notebook also becoming a means of communication, e.g. writing messages between shifts. The notebooks were typically mobile both in the laboratory and between laboratories, offices and meeting rooms. In three of the physicists' laboratories they had notebooks tied to a particular piece of equipment besides their general one. In two labs both shared and personal laboratory notebooks were used.

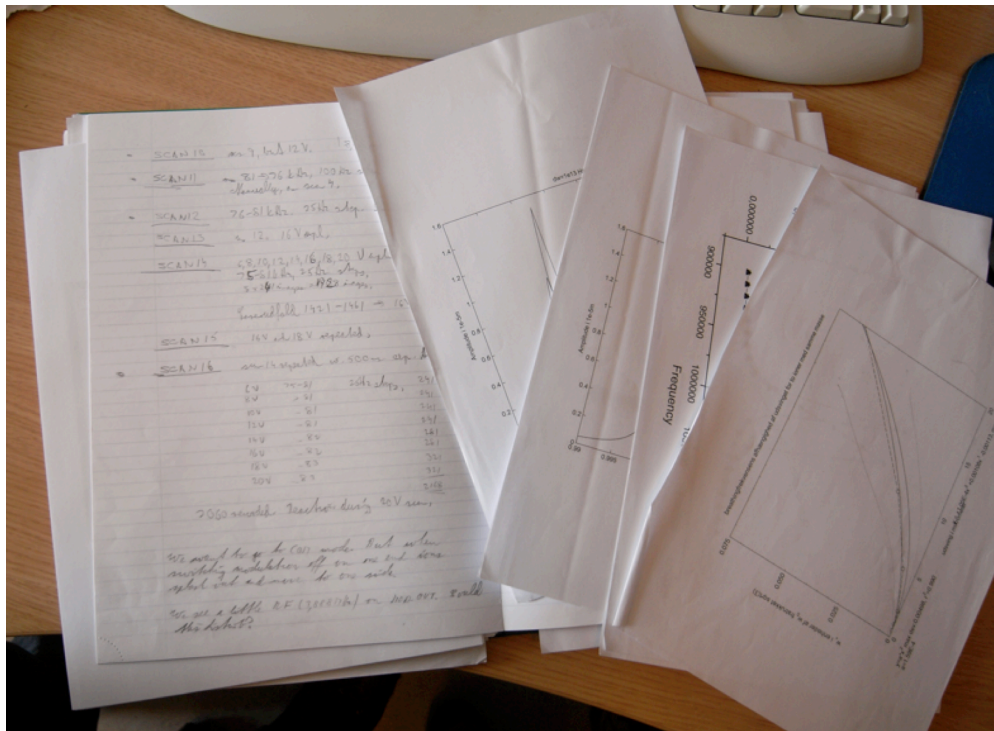


Figure 4.13 One of the physicists' laboratory notebook

The books were all variants of regular A4 sized bound notebooks. All of these books were inscribed chronologically from left to right. The entries in the books consisted of a mix of written text, drawings, handmade tables and loose printouts of graphs and tables of data (Figure 4.13). The most important piece of meta-information inscribed into an entry of all of the respondents' notebooks was timestamps. The timestamps were often coupled with a reference to a data file and sometimes the name of the author of the entry.

All of the physicists used computers for almost all of their activities, for collecting and analysing data, for writing papers, for seeking information, and for controlling equipment. All of which they documented in their analogue notebook as well as they could.

4.6.2 The workshops

As described in Section 1.6 I conducted two future workshops with the physicists. The second workshop was a re-iteration of the fantasy phase of the first where the physicists had been introduced to ubiquitous instrumental interaction and the concepts of VIGO.

In the first future workshop the physicists were asked to criticise their current paper based laboratory notebook. Obviously, much of the critique was aimed at the physical restrictions of a paper notebook – it can be at only one

place at a time, it can be lost or it can burn in a fire, writing space is restricted on a page, etc. They criticised it for creating a lot of redundant information, e.g. graphs both on print taped into the book and on the computer. The physicists also criticised the paper notebook for creating a broken workflow. They missed integration between their different activities and had to do considerable overhead work to maintain their laboratory notebook and later retrieve and process information from it for e.g. paper writing. An obvious limitation of the notebook is the inability to search in it effectively, which was one of the physicists' main critiques. Searching would be an immediate benefit of almost any conceivable electronic laboratory notebook. The second highest ranked point of critique, however, is difficult to mitigate even in an electronic version, namely not knowing what to document in the laboratory notebook before it is too late.

In the following I aggregate the ideas that emerged from the workshops after the physicists had been introduced to ubiquitous instrumental interaction into three short visions. The visions show aspects of a potential system but do not present a vision for a *whole* system. Following the visions I will discuss how they relate to the work on ubiquitous instrumental interaction.

4.6.3 Vision 1: Integration of documentation, analysis, and control

The laboratory notebook envisioned by the physicists is not only an environment for documentation but it is the focal point of their activities. The notebook provides a general laboratory environment integrating both documentation, analysis of data and control of equipment.

In the laboratory notebook there is no distinction between objects for documentation and objects for control. Documenting the position of a laser at a given time involves copying the object that acts as a control to a page in the notebook. Later this control can be reloaded to restore the setting. This idea is illustrated in [Figure 4.14](#). Here a copy of a control pane is stored on a page in the notebook. The controls are changed but reverted back to the original by using the copy stored in the book. The intermediate parameters are documented in the notebook as well and annotated as non-working.

The envisioned laboratory notebook supports automatic pre-analysis of data and live mapping of incoming data to graphs through the concept of governors. The graphs in the laboratory notebook have publication quality as default and hence they are almost immediately usable in follow-up paper writing activities. These graphs can of course be freely annotated through the instruments associated with the notebook.

4.6.4 Vision 2: Access from multiple interactive artefacts

The laboratory notebook is obviously accessible from multiple workstations. However, it can also be accessed by multiple heterogeneous artefacts – artefacts

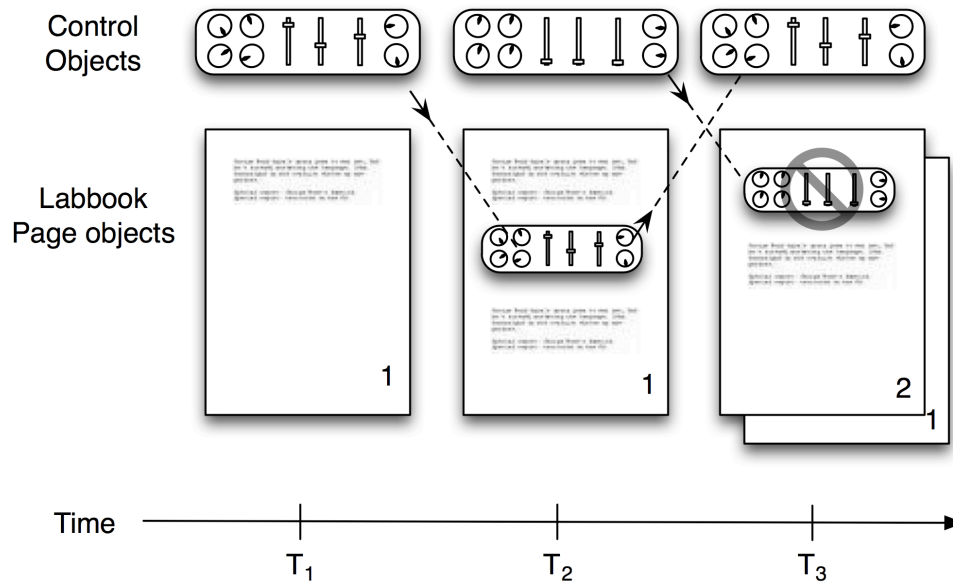


Figure 4.14 Control objects are stored in the notebook for documentation and later reloaded

resembling the three sizes of Weiser's [170] vision quite well.

To not completely abandon the flexibility of hand-writing, a page from the laboratory notebook can be loaded on a tablet sized artefact. Multiple tablets are available in the laboratories and each researcher has her own custom writing instruments. This makes it easy to distinguish who has written what.

Given the quite mobile nature of working in large laboratories, controls and data read-outs can be loaded to PDA sized artefacts easily cartable in a pocket. Previously, one of the physicists had to do rather dangerous acrobatics to read a meter on the screen of the other end of the laboratory while climbing his pressure chamber to control a valve. Now he can just bring the valve control and the pressure read-out with him on a PDA.

The professor in one of the laboratories could enjoy being able to configure his smart-phone with a graph of the latest data-readout and a discussion field in the laboratory notebook. This way he would be able to monitor what happens in the laboratory at a glance and comment directly in the laboratory notebook if necessary – especially when stuck in tedious administrative meetings.

At group meetings, ToDo lists and summary pages from the laboratory notebook are loaded to a large interactive white-board. The ToDo lists are displayed on the white-board in the laboratory as well. When a point on the ToDo list is finished it can be linked to the pages documenting the process. The whiteboard in the laboratory notebook is also configured to show the five latest edited pages from the notebook.

4.6.5 Vision 3: Revision control

To ensure that nothing is ever erased users of paper based notebooks typically only write in permanent ink. In the envisioned electronic laboratory notebook all objects are revision controlled and centrally stored on a university server to ensure nothing is ever lost. High granularity revision control makes it possible to easily browse the history of a given object or page of the laboratory notebook. Various instruments and views are available to visualise and navigate the history of objects.

4.6.6 Discussion

In the workshops with the physicists it quickly became clear that they did not wish for just an application to handle documentation. If the laboratory notebook should be digitalised it should integrate with all of their computer related activities. It was clear from the workshops that being introduced to the vision of ubiquitous instrumental interaction and the concepts of VIGO stimulated the physicists to be visionary. The physicists also fairly easily grasped the concepts of VIGO and were able to conceptualise their ideas in its terminology (see further in [Paper VI](#)). In the following I will briefly discuss the realisability of the three visions in a VIGO based architecture.

The governor concept provides the potential for realising vision 1. Controls in VIGO are just objects with additional functionality associated, hence they can be treated in the same manner as e.g. a piece of text or a graphical object. This is not the case in the application paradigm; I cannot use the font widget to document in this text what font is used by embedding the font widget in the text (given my L^AT_EX editor had a font widget).

The principles of vision 2 are demonstrated in [Paper V](#). In principle a governor of an Othello-game does not differ from the one of a piece of physical equipment. Previously in this chapter I also discussed how an object representing a printer could be moved between artefacts. Both vision 1 and 2 presume an integration between the laboratory notebook software and the equipment of the laboratory. However, one physicist argued that much of their latest equipment has standardised interfaces and can be programmed from many of the major programming languages. I borrowed a small National Instruments USB Data Acquisition box from his laboratory and was quite easily able to program a governor in Python mapping properties of VIGO objects to input from a temperature sensitive resistor and output to a light emitting diode.

Vision 2 and 3 assume functionality that I have not yet discussed as part of a VIGO based system namely hypermedia and revision functionality. VIGO requires a shared object space and an object space where objects are expressed similarly (to ensure the polymorphic nature of instruments). The objects of VIGO are passive and in our implementation implemented in a mark-up language resembling HTML. Hence hypermedia functionality including linking, and even more advanced linking than on the Web, is possible. Likewise, the

passive nature of objects lends itself well to revision control. Revision control is, in my opinion, a central candidate for future work on a shared object space which I elaborate on in [Section 5.2](#). I argue that revision controlled objects might be the only way to handle e.g. undo in a distributed setting.

The laboratory notebook study showed that there are artefacts that seemingly do not translate well to an application dominated interaction paradigm. However the potential is present in our interactive artefacts, and ubiquitous instrumental interaction and the VIGO architectural model could be an approach towards realising it.

4.7 Related work

This section provides a thematic overview of related work to the work of this chapter. The four themes are: Application critique, User interfaces in ubiquitous computing, Architectural models, and Laboratory notebooks.

4.7.1 Application critique

Today we take the concept of applications for granted but in what we regard as the forefather to our modern user interfaces, the Xerox Star 8010 Information System (Star) [\[155\]](#), there were no such things as applications. Applications were introduced by Apple in the early eighties. When opening a document in Star the editing facilities were *there* conceptually as part of the document (what we in [\[105\]](#) refer to as the complete coupling paradigm). Star was monolithic by nature, sold as it is, hence this complete coupling was no problem. Apple, however, was business minded and saw opportunities in having 3rd parties develop software for their system e.g. Microsoft Word was available for the 1984 Mac together with Apple's own word processor MacWrite. Hence arose the idea of having multiple applications potentially for the same type of files.

However, 10 years later Apple was forerunner in an ambitious project of re-thinking applications. Apple's OpenDoc project (see e.g. [\[5\]](#)) was an attempt to provide a platform for working with component based mash-ups rather than monolithic applications. The classical example was how spreadsheet functionality could be embedded in wordprocessing, hence providing a uniform interface for both manipulating text and spreadsheets. When a user selected the spreadsheet, menus and inspectors would be adapted to the type of the component. Apple envisioned this approach to be the future of applications, however, it never really caught on as a general interaction paradigm. This kind of mash-up approach is today becoming more and more common-place on the web.

A modern more direct critique of the application concept comes from David Karger's Haystack project [\[101\]](#). Karger addresses the problems of applications when it comes to personal information management. Karger exemplifies the

problem with a music reporter who has little support for collecting artist information, interview dates, music samples and reviews in the same interface. All of which are relevant for the reporter's activity, but assumed different activities by the calendar, music player and word processor application developers.

Haystack provides an environment for aggregating multiple types of information, and to easily program views for new types of information. Haystack is based on an idea of *one information space*:

There should be no a priori segregation of users' information by "type" or application. Rather all information exist logically in a single space. Users should be able to group and relate any information objects they choose.
[101, pp. 52]

With ubiquitous instrumental interaction we have criticised applications mainly from the perspective of multi-device interaction and the *interaction* with digital objects. These two approaches are two sides of the same coin and it would be interesting to join the two perspectives.

4.7.2 User interfaces in Ubiquitous Computing

Research addressing the more conceptual aspects of user interfaces and interaction paradigms in ubiquitous computing environments is quite limited. There are, however, some exceptions. The work on activity based computing [15] is an example. Activity based computing proposes *activities* as the focal point of user interaction across heterogeneous interactive artefacts. Activities can be suspended, resumed and they can migrate between artefacts. Hence activity based computing is also an attempt to go beyond the isolated personal computer. The current implementations of activity based computing, however, build on top of our traditional application based user interfaces.

An area with some active research is automatic generation of user interfaces adapted to particular interactive artefacts for services on a network. The pebbles project [122] has e.g. investigated how user-interfaces for appliances on a user's home network can be aggregated on a PDA or smart-phone to create a *personal universal controller*.

Some systems have attempted to provide a generic solution for services or applications in a multi-device environment [139, 132, 123]. XWeb [132] and the SpeakEasy Recombinant Computing Framework [123] are probably the closest to the work on VIGO in that respect. XWeb decouples internet services from their user interface to facilitate access from multiple kinds of artefacts by introducing device-independent widgets that specify the possible values of the data items of a service, e.g., a date or time. In SpeakEasy services can provide their own user interfaces to be aggregated on the client. The goal of XWeb, SpeakEasy and related approaches is, however, to automatically generate user interfaces for services while with VIGO we focus mainly on an interaction architecture for manipulating data objects across different artefacts. They take a

bottom-up approach: Starting with services accessible over the network they explore how to provide interfaces to access them from multiple interactive artefacts. With VIGO I have started with a uniform interaction model and shown how the framework provides the flexibility needed to support multi-device interaction. Each of these approaches could probably benefit from each other – it would be interesting to explore how e.g. a VIGO based architecture would handle the streaming of video which is supported in SpeakEasy and generally how services should be handled.

iStuff toolkit [9] and the follow-up work on Patch panel [10] discuss input redirection in ubicomp environments and relate to the discussion of our distributed event handling in VIGO instruments. Our implementation of the event-handling is rather naïve and techniques from the two cited papers could probably be applied to strengthen the implementation. Another approach would be to abandon a dedicated distributed event-handling mechanism and completely integrate it in an advanced distributed VIGO object system.

More generally, software architectures for ubicomp systems have focused mainly on middleware to support system requirements such as distribution, discovery, fault-tolerance or context-awareness, but they do not address the specific needs of interaction (see, for example, [117]).

The BEACH architecture [162] is a rare exception. BEACH is a general ubicomp architecture designed to support interaction with shared objects on heterogeneous interactive artefacts. In [162] Tandler presents a conceptual separation of the components of a system for ubiquitous computing environments or, as he refers to it, *roomware*. He proposes a separation between interaction model, environment model, user-interface model, application model and data model. Furthermore he brings forth a series of requirements from the intersection between HCI, UbiComp and CSCW [162, pp. 269]:

- HCI:
 - H1: Different forms of interaction
 - H2: Different user interface concepts
- UbiComp:
 - U1: Multiple and heterogeneous devices
 - U2: Multiple-computer devices
 - U3: Context and environmental awareness
 - U4: Dynamic configuration changes
- UbiComp and HCI:
 - UH1: Adapted presentation
 - UH2: Multiple-device UI and Interaction
 - UH3: Physical interaction

- CSCW:
 - C1: Multi-device collaboration
 - C2: Modelled collaboration mode and flexible coupling
 - C3: Multiple-user devices
- UbiComp and CSCW:
 - UC1: Collaboration with heterogeneous devices
- Software techniques:
 - S1: Generic functionality – reusability
 - S2: Tailorable functionality – extensibility

The VIGO architectural model is quite compatible with both the separation of concerns and the above requirements of BEACH. VIGO supports different forms of interaction by making instruments explicit and decoupled from the rest of the system. Hence implementing support for a new input device merely requires implementing a new instrument (Addressing H1). U1, UH1, UH2, C1 and UC1 is demonstrated in [Paper V](#). H2 relates to working with different interface concepts than e.g. only the desktop metaphor – this is not directly addressed in our work on VIGO but VIGO is by no means only confined to the desktop line of thought. U2 relates to how something can seem as one computer from the user perspective but underneath multiple computers are in play. Tandler exemplifies with a multi-monitor setup with multiple computers associated. This is addressed in VIGO by governors being capable of receding both locally and externally to the interactive artefact in use. We have not addressed U3 or U4, nor have we addressed C2, CSCW supported through VIGO is future work. C3 is also a candidate for future work, e.g. to create a multi-user VIGO based interface on Microsoft Surface (<http://www.microsoft.com/surface/>) or Diamond Touch table (<http://www.circletwelve.com/>). Our governor concept supports S1; e.g. the general board-governor presented in [Paper V](#) would be usable with multiple games and instruments are reusable across multiple domain objects. S2 is principally supported by VIGO, instruments and governors can be added dynamically – however, *how* this should be realised in an actual system is future work.

If we look at the separation of concerns, BEACH's interaction model is clearly defined by instruments while the application model is defined by governors. The data model of VIGO is however pervading the whole system. Everything in VIGO is represented by objects from sensors and actuators (the environmental model in BEACH) to views on data. Hence a lot of state, e.g. placement of shared cursors in a text-field, which is argued to reside in the application model in BEACH, is actually part of the data or object model in VIGO. Since VIGO does not distinguish between user objects (e.g. a document) and user-interface objects (e.g. a widget) the user-interface model in BEACH terminology would be anything related to presenting any object to the user. The

distinctions between interaction model, application model and data model is actually very clear cut in VIGO.

Therefore VIGO is quite compatible with BEACH and it is actually more compatible with the requirements of ubicomp than MVC (following our argumentation in [Paper V](#)), which is used in Tandler's implementation of BEACH.

4.7.3 Architectural models

The area of architectural models in user interface development is quite interesting since there has been, by and large, only one that has really caught on. This is of course Model-View-Controller (MVC). MVC was originally introduced by Trygve Reenskaug [143] while visiting Xerox PARC's Smalltalk group. Reenskaug never came to publish an academic paper on MVC, hence the reference is a technical note. In the foreword to [137] Reenskaug refers to the model as an obvious solution to letting users have control over their information as seen from multiple perspectives. Perhaps its obviousness explains its huge success and why it is still the predominant approach to user interface structuring today. Granted, MVC has seen a lot of derivatives and interpretations. Each graphical toolkit providing its own little twist on its implementation. Apple's Cocoa e.g. treats the controller as a mediator between the model and the view. Eden and Kazman [70] argue that MVC in Swing [146] is actually used as a design pattern rather than an architectural model per se – hence as a local solution rather than principles pervading the whole system⁵.

However, the basic premise of MVC stays more or less the same. It advocates a separation of the representation and interaction from application logic. Both to support changing requirements to user interfaces and to enable multiple representations of the same model [50]. The MVC line of thought echoes the ideas of user interface management. Creating user interfaces as add-ons to an otherwise well defined application model.

One variant of MVC that is interesting from the perspective of VIGO is the Taligent Model-View-Presenter (MVP) [140]. MVP separates the controller of MVC into a general presenter concept, which orchestrates the traditional main or event loop of an application and interactors that translate events on views into commands and selection on the model. Interactors however conceptually differ from instruments in VIGO in that they are strongly coupled to a presenter – that is the application.

Abstraction-Link-View (ALV) [83] is another variant designed for sharing a common model (the abstraction) among multiple networked clients potentially each with their own view, but it is otherwise quite similar to MVC.

Presentation-Abstraction-Control (PAC) [62] is an architectural model based on co-operating agents in a hierarchy. Each agent is responsible for a specific

⁵Eden and Kazman [70] formally define architecture, design and implementation in their excellent paper bearing the same name. They provide two criterions *intentionality* and *locality* used to formally prove whether a specification is architectural, design or implementation.

aspect of the functionality and presentation of the application. Hence PAC provides both a stronger separation of concerns than MVC and support for change and extension [50] – one agent can easily be exchanged without influencing the rest of the hierarchy.

None of these patterns, however, make instruments explicit constructs to the user. Instead they promote a widget-based type of interaction.

Document-Presentation-Instruments (DPI) [20] is a document-centric software architecture based on instrumental interaction. Like VIGO it decouples instruments from the target objects. However, DPI is a desktop-based framework and does not address the distribution of objects and instruments across multiple machines. VIGO is in many ways a continuation of the work on DPI. One important distinction is that VIGO expresses how business logic of objects should be handled through the concept of governors and therefore it goes beyond interaction with passive documents alone. DPI has later been implemented in the open source toolkit OpenDPI [21].

An approach worth mentioning in this discussion is Richard Pawson's work on Naked Objects [137]. Naked objects is an approach inspired by Alan Kay's original Smalltalk visions. The user interacts directly with real OO objects. The user interfaces are a direct mapping of objects to the user – hence application logic and user interface melt together. VIGO and Naked Objects share the idea that everything on the screen should be an object manipulatable by the user. VIGO and Naked Objects, however, differ in that Naked Objects is based on encapsulation in the spirit of Object Oriented Programming. Naked objects both embody interaction and business logic – whereas VIGO objects embody no logic.

4.7.4 On laboratory notebooks

While the development of an actual laboratory notebook system has not been a main focus of my work, and has mainly served as a driver for discussion, I point out a couple of papers discussing the development of laboratory notebook systems.

For a popular scientific discussion of the current state of laboratory notebooks I refer to Butler's Nature article [51].

schraefel et al. [151] develop a laboratory notebook system for chemists by doing classical work-place studies, although with an inspiring technique⁶. The work results in a system tailored for the work practice of the given chemistry laboratory.

My work on laboratory notebooks was inspired by Aurelien Tabard, Wendy Mackay and Evelyn Eastmond and their work [160] which happened concur-

⁶Given the researchers have little clue about the experiments the chemists are doing, they create an experiment of making tea hence enabling the researchers to understand the processes of a chemistry experiment.

rently in the same lab in Paris where I did most of my research on VIGO. Tabard, Mackay and Eastmond worked with augmented paper through the use of Anoto technology (<http://www.anoto.com/>). They hereby combined the potentials of the computer with the flexibility of paper. While this approach is definitely promising, I believe the kind of flexibility possible with paper is not completely impossible through a computer screen. However, there is a lot of unfulfilled potential in the way we create what is behind the screens today.

Going over the (rather sparse) HCI literature on laboratory notebooks combined with our (granted, quite limited) observations from the laboratory of the physicists, it seems that the literature on laboratory notebooks has focused on somewhat “easy targets” compared to the complexity of experimental physics.

Regarding the theoretical framework applied in the analysis of laboratory notebooks in [Paper VI](#): The notion of syntonic seeds is related to the notion of boundary objects [157]. Laboratory notebooks can of course also be regarded as boundary objects. Boundary objects however stress the collaborative aspects of an object while syntonic seeds stress that an object is part of, and has different roles in, multiple activities – potentially only of a single subject. The notion of syntonic seeds furthermore emphasise how the object oscillates between being mediator and object of activity, which was an integral part of our analysis.

4.8 Summary

In this chapter I have addressed a problem area that has gained little attention despite the development of ubiquitous computing technologies – namely whether applications and their underlying architectural models are suited for the new challenges. The core motivation for the work presented in this chapter comes from the perspective of the user – how can we as developers empower the user to more easily appropriate technology, and not let the technology dictate behaviour. In this endeavour I have not empirically proved that actual use benefit from this new approach, however, I have shown how it is possible to realise the foundation for ubiquitous instrumental interaction through the VIGO architecture and furthermore shown that it provides new possibilities in multi-device interaction. I found that the architecture led to a flexibility that easily enabled e.g. migration of what resembles live applications between different artefacts. This was illustrated through implementing the pick-and-drop interaction technique and showing how it enabled the user to copy a live Othello game between two interactive artefacts. Furthermore I have illustrated how the same simple instruments can be applied across multiple objects and artefacts.

In order to put a perspective on the potentials of ubiquitous instrumental interaction I presented a pilot study with a group of experimental physicists where the paradigm was applied as a thinking tool for envisioning electronic laboratory notebooks. The study especially showed that the non-distinction between domain objects and user interface objects seemed promising in that

context. It also shows that the flexibility embodied in the ubiquitous instrumental interaction vision is compatible with a domain object which is not statically definable – something we characterise as a *syntonic seed* in [Paper VI](#).

Conclusion

This chapter concludes the overview part of this dissertation. In this dissertation I have explored two perspectives on HCI in complex artefact ecologies: One theoretical where I discussed how to analyse the interaction with a single artefact while acknowledging that it is part of an ecology of artefacts; one technological where I proposed an alternative interaction paradigm for our interactive devices breaking with the ingrown assumptions of applications and personal computing and explored the technological realisation of it.

There are multiple common characteristics in the way I have approached the two perspectives. Both approaches share a foundation in activity theory. The theoretical approach explicitly builds upon activity theoretical concepts while the technological approach embodies the activity theoretical understanding of mediation, of externalisation, and of the dialectically evolving relationship between human beings and technology. In both approaches I have been interested in the action possibilities offered by interactive artefacts. On the one hand from an analytical perspective; how can I provide a means of discussing the action possibilities offered by artefacts, and of how they are perceived by the user? On the other hand from a technological perspective; How can we create an interactive environment where the user is empowered to reason about action possibilities beyond the isolated artefact.

In the following I discuss how I have specifically approached the two research objectives stated in the introductory chapter.

5.1 Research objectives revisited

Theoretical Objective: To enrich our analytical vocabulary of artefacts that are, or are to be, part of complex artefact ecologies and provide a means of structuring design-oriented analysis of such artefacts.

The human-artefact model summarises the insights made by revisiting activity theory in the light of complex artefact ecologies. In itself the human-artefact model provides a means of structuring analysis of individual artefacts while maintaining a focus on the ecology of artefacts which the artefact is part of. Multiplicity of artefacts is embraced through a dialectical understanding of human beings' relationship to artefacts; that our perception of an individual artefact is shaped by the ecology of the artefacts it is part of, and artefacts are a crystallisation of a historical development of an activity. The human-artefact model provides a means of systematically exploring the tensions on all levels of activity between, on the one hand, what has been externalised into the artefact, and on the other hand, the activity and orientation basis of the user.

By emphasising the importance of exploring the artefact ecologies, and providing a structure for doing so, I also contribute with a means of talking about context. In [Paper I](#) we argue that there is a tendency in discussions of ubiquitous computing to address primarily the singular artefact, leaving the rest to the ill-specified notion of context. Obviously a focus on the artefact ecology will not render the notion of context unnecessary, as there is always a new context beyond a given situation. However, the human-artefact model makes it possible to focus on the question of which elements of the ecology matter for our understanding of the single artefact and vice versa.

Through the development of the human-artefact model I have revisited a number of concepts: I have argued that the concept of *functional organs* provides a richer ideal for the appropriation of artefacts than transparency or disappearance. I have explained breakdowns in use as tensions between the *aspects* of an artefact and the *orienting basis* of the user. Bødker's original aspects have been refined through Bærentsen and Trettvik's activity theoretical delineation of affordances. Hereby I have also revisited the affordance concept and argued for letting what is the basis for affordances be the units of analysis rather than the affordances themselves.

As an example of analysis I have in [Paper I](#) shown how the model can be applied to frame design-oriented analysis based on a study of maps used on multiple heterogeneous interactive artefacts. The human-artefact model is intended to be applied to frame quick back-of-the-envelope analyses (as illustrated in [chapter 3](#)) as well as more extensive artefact and activity historical analyses (as sketched in [Paper I](#)).

Technological Objective: To reconsider the current predominant interaction paradigm and propose an alternative paradigm that embraces the heterogeneity of complex artefact ecologies. As a consequence show that it is realisable in principle and that it provides new opportunities for HCI.

I have criticised the current application based interaction paradigm for embodying outdated assumptions of interactive artefacts that are solely used in isolation. Based on a critique of the current application centric interaction paradigm, I have proposed *Ubiquitous Instrumental Interaction*. Ubiquitous instru-

mental interaction is a vision of an interaction paradigm where user interfaces dynamically span multiple heterogeneous interactive artefacts, and where each interactive artefact is not seen as an isolated entity. Ubiquitous instrumental interaction extends the interaction model *instrumental interaction* to a general interaction paradigm. Here the interaction possibilities with objects are decoupled from the objects themselves into instruments. Instruments can potentially work across both objects and interactive artefacts.

The VIGO architectural model provides a reflection on the requirements for implementing Ubiquitous Instrumental Interaction. Through my prototype implementation I have shown that the vision is realisable in principle and that it provides new technical opportunities. For instance it makes it possible to move what resembles live applications between two heterogeneous artefacts. This is partly achieved by the introduction of the *governor* construct.

By infusing Ubiquitous Instrumental Interaction and the VIGO model as instruments for change in workshops on rethinking a group of experimental physicists' laboratory notebooks, I have shown that the interaction paradigm has potentials in a concrete application area.

5.2 Future Work

5.2.1 On the human-artefact model

Assessing value in use

The most interesting future work is to observe the human-artefact model in actual use. Currently we are experimenting with letting students use the model. In recent course students used the human-artefact model to generate mood boards (e.g. [133]) for designs in the context of e-Governance (Figure 5.1). In the fall a group of students will apply the human-artefact model to structure the development of, and evaluate on, a social geotagging interface for mobile phones.

Whether the human-artefact model will be appropriated and used by practitioners time only can tell. However by letting students experiment with using the model might provide insight in how to pedagogically present the model to someone who are not activity theory specialists.

Structuring iterative prototyping

Prototypes embody design proposals and ideas and provide hands-on experiences with problems and situations. Prototyping accordingly provides the means that allow users hands-on experience while allowing designers to explore not only the current use situation but also the future. What exactly one might evaluate with a particular prototype depends on purpose of the design



Figure 5.1 Students discussing a mood board structured by the human-artefact model for an e-Governance design case (Picture by Nikolaj Gandrup Borchorst)

situation. It is, however, often left to the common sense of the designers to decide this and structure the process accordingly. They are given very little help in seeking systematic direction for prototyping. The human-artefact model could be a help to systematically explore the different levels of activity in a prototyping process.

Scenarios and personas

In [Section 3.5](#) I related the human-artefact model to the Task-Artefact Cycle [58], and Carroll and Rosson's [57] discussion of scenario driven design based on the Task-Artefact Cycle. It would be interesting to explore this idea further in the light of the human-artefact model. Perhaps in conjunction with the use of personas [61] and an equivalent concept for archetypical technologies and artefacts.

5.2.2 On Ubiquitous Instrumental Interaction and VIGO

Toolkit development

Modern graphical toolkits are almost exclusively built upon MVC and designed towards building desktop applications. Hence implementing something as simple as e.g. text-editing in VIGO requires implementation from scratch (see [Section 4.5](#)). Development of VIGO based user interfaces would be considerably eased if the developers had access to a standard toolkit for creating instru-

ments, views and governors together with a set of standard components e.g. components for text manipulation. We have taken early steps towards a toolkit by implementing a state-machine library (inspired by *SwingStates* [6]) for implementing instruments easily. One could imagine similar abstractions to ease the development of views and governors. A well structured toolkit could also help developers appropriate the VIGO way of thinking – as e.g. *Smalltalk-80* did it and Java’s *Swing* or Apple’s *Cocoa* do with MVC.

Object visualisation

Visualising objects across multiple heterogeneous interactive artefacts is not trivial. Some objects will be too complex to visualise on smaller artefacts. Therefore it would be meaningful to experiment with for instance a stylesheet [2] inspired approach. Here objects could have metadata associated with information on how it could be rendered on a given type of interactive artefact.

Shared object space

The basic premise for VIGO to work is a coherent object space with objects expressed in a uniform way. I, as well as Tandler [162] and Karger [101] have argued that a shared information or object space is fundamental for both post-application user interfaces and user interfaces spanning multiple devices. In our implementation of the VIGO architecture in [Paper V](#) we implemented a simple XML-based object system based on replication. Future work on a shared object space would include revision control on objects. Undo becomes problematic in distributed user interfaces since there is no application environment to store the state. In [26] we argue that objects should keep their own history in order to facilitate e.g. undo. This complements the physicists’ visions of in-built revision handling quite well. Such an approach would also facilitate the implementation of many different types of history handling instruments and support e.g. hierarchical undo. In [26] the governor construct had not yet been conceptualised. Undo becomes slightly more complex when interaction has side-effects. This leads to another area of future work, namely whether a VIGO based system should support transactions and how the interplay between instruments and governors should be handled in that case.

End-user configuration

An essential part of the ubiquitous instrumental interaction vision is to let users configure their devices with views, instruments, governors and objects. A major topic for future work is to design an actual interaction environment on top of a VIGO architecture which enables dynamic configuration. The VIGO architecture and our implementation currently support this flexibility, however, the changes have to be done by hand in XML.

The instrument construct facilitates easy implementation of novel interaction techniques and it would be interesting to support user configuration of input devices with instruments e.g. inspired by [67] and [85]. In the case of developing a laboratory notebook it would be obvious to adopt a dataflow programming language *à la* G for LabView [93] (which most physicists are familiar with).

Interaction design and user involvement

A VIGO based prototype system has to be deployed with users to assess whether the hypotheses of Ubiquitous Instrumental Interaction foster affordance based reasoning holds. This requires design of an interactive system mature enough for users to experiment and evaluate on configurations of instruments and objects. There is, however, not only one way to design this and participatory design involving a specific application domain would be the most sensible approach. The ideas of ubiquitous instrumental interaction are addressing the fundamental basis of interactive systems and it is my hypothesis that it would make most sense, in the first time round, to involve users capable of distinguishing between architectural and design related issues. The workshops with the physicists showed that it was possible to bring subjects that were non-computer scientists but used to abstract thinking and advanced technology, to a level where the abstract concepts of VIGO were constructively applied for envisioning a system. As a follow up it could make sense to stress the usability of the design by applying it e.g. to create a shared family scrapbook by participatory design with actual families.

5.2.3 Combined

Optimally the human-artefact model could be used to structure the development of a VIGO based interactive system. In fact the human-artefact model would lend itself well to evaluate and explore development of individual instruments, and as we argue in [Paper I](#) it can be applied in conjunction with Beaudouin-Lafon's own dimensions for evaluating instruments.

Part II

Papers

The Human-Artefact Model – an Activity Theoretical Approach to Artifact Ecologies

Susanne Bødker

Clemens Nylandsted Klokmose

Abstract

Human computer interaction is currently moving beyond the desktop, towards interaction with multiple and dynamic configurations of smaller and larger computing artifacts. With this paper we present an understanding of artifacts as mediators, addressing multiplicity and dynamic configuration in complex artifact ecologies. We address quality and learning across artifacts. At the core of this is the Human-Artifact Model, which has its roots in activity theoretical HCI. By reinterpretation of the activity theoretical foundation, we present a framework that helps address mediators in plural. We show how the Human-Artifact Model helps structure the understanding of an artifact's action-possibilities in relation to the artifact ecology, which surrounds it, hence addressing multiplicity. Essential to the model is that it provides four interconnected levels of analysis and addresses the possibilities and problems at these four levels of activity.

Practically and in principle, use and artifacts are constantly developing, and we address the ways in which we understand and design for development in and of use. Accordingly, we seek a change-oriented framework. This leads to a methodological approach that focuses on new artifacts to supplement and substitute existing artifacts.

Through a design case we develop the methodological approach and illustrate how the Human-Artifact model can be applied to analyze present artifacts as well as to design future ones. The model is used to structure such analysis and to reason about findings, with focus on mediation.

Submitted: to Human-Computer Interaction December 2008, Revise and resubmit requested from associate editor Jack Carroll April 2009, Revised version submitted August 2009. The revised version is the one presented here.

1 Introduction

The current development of interactive technologies is steadily moving beyond the desktop computer. Computers ranging from hand-held appliances to wall-sized interactive displays are becoming commonplace. All of these computerized artifacts may not necessarily have a well-specified function in daily use; rather, they are used interchangeably in response to context and conditions of use. Cell-phones for instance in many situations give the same access to the Internet as a laptop. Consequently people increasingly blog and use maps interchangeably on cell-phones and laptops. Oulasvirta [134] describes how Ubicomp is facing a big divide, where on the one hand practical applications and interfaces are pluggable and interchangeable only to a very limited extent. On the other hand, much of the research, focusing on the visions for ubiquitous computing, remains too simplistic and abstract. As analysts, practitioners and designers we have to acknowledge that users juggle multiple different interactive artifacts in their daily lives, and each of these artifacts influences the perception and use of the others. In line with Jung et al. [94] we refer to the interactive artifacts in use by a particular user as the user's artifact ecology. Over the last decade users' artifact ecologies have become increasingly complex, as described by Oulasvirta [134], and in our understanding, the boundaries between interactive and other artifacts are blurring, and the artifact ecology consists of all types of artifacts, and not only interactive ones.

The use context of interactive technologies has broadened to embrace almost all aspects of both work and leisure. New elements of human life are being included in human-computer interaction research and practice. Such elements include culture, emotion and experience. Conceptually and theoretically, the most recent wave of HCI has focused on the cultural level ([69, 46], or on experience [115]). Norman [128] connects emotions to experiences and things, and he points out how in addition to relating to things cognitively, human beings relate to them emotionally. Boehner et al. [45] tackle the topic of emotions from a social and interactionist perspective, arguing that meaning in emotion is created in the interaction between people. While emotion accordingly provides a place where it is possible to move beyond the classical purpose-driven understanding of human-computer interaction prevalent to HCI, there is yet quite a distance to something that is conceptually and methodologically useful when working in the void described by Oulasvirta [134]. In our view the entire focus on emotion and experience tend to reinforce Oulasvirta's divide. This is simply because this focus provides very few concepts and methods to help understand what users actually do in the messy ecology of multiple and ubiquitous artifacts.

The dawning complexity of the technological artifacts is approached by Dourish [66]. He argues that we are *embodied* in technology. Such embodiment is not only an issue of culture and emotion, it is equally an issue of how our physical capacities extend and expand through technology. Dourish approaches embodiment from a phenomenological perspective, where individu-

als are not defined by the boundaries of their skin [99]. Individuals are, among other things, defined by the tools they use, whether called ‘equipment’ in Heidegger’s terms or ‘functional organs’ in the terms of Leontiev [111, 110]. Bødker [38] used Heideggerian as well as activity theory, when proposing that human activity is mediated by technological artifacts. She furthermore proposes that this mediation includes both the cultural levels and actual handling of the technology. Beaudouin-Lafon [18] took his starting point in this idea of mediation in his proposal for instrumental interaction as a way of approaching human-computer interaction analytically, as well as designwise. Dourish’s embodiment and Beaudouin-Lafon’s instrumentality point towards an understanding of multiple and ubiquitous interaction, where instruments coexist and replace each other as extensions of the human body.

Chalmers and Galani [59] criticize Dourish’s approach to embodiment for not offering specific design principles and guidelines. They point out that seamless, embodied interaction is mirrored in seamfulness, which reveals differences and limitations to the technology. Seamlessness in this terminology has similarities to what is often called transparency in HCI, the fluid and non-intrusive switch from one artifact to another. Seamful in contrast points out that it is sometimes equally important to be aware, or be made aware where the boundaries are for such as switch. Similarly to Chalmers and Galani we are interested in understanding the dialectics between the interfaces that are designed, and the ways in which people build their understanding and use over time, the central activity theoretical notion of mediation [38]. Both Dourish and Chalmers and Galani, however, seem to be largely unaware of earlier attempts to think Heideggerian phenomenology into HCI. While conceptually clear, both of their approaches fail to be operational in design. Our goal is to make such a conceptual and methodological delivery by revisiting the shared history.

Post-cognitivist thinking brought into HCI e.g. by ethnomethodology, in particular Suchman [158], has given a useful insistence on a focus on the particularities of specific unfolding interaction. Such studies have been an eye-opener to many by their insistence on understanding how human beings, together, cope with messy environments and ill-suited technical solutions. However, Kaptelinin and Nardi [99] point out that a theoretical framework is of importance to HCI to avoid continuously going back to specific, detailed accounts of particular cases. Such needs are even more predominant when we insist on studying use across computing devices, and computing devices across use situations. Hence, we need a theoretical framework that enables us to look at *human interaction with computerized artifacts in plural*. Such a framework, going beyond the singular use situation, can help conceptualize the dynamics of use situations involving complex ecologies comprised of multiple artifacts.

To summarize, the current state is one where we need to address the gap between the practical mess and limited ideals of ubiquitous interaction. We need to address the gap between culture, experience and the practical role of artifacts in embodiment and mediation, between unique examples and gen-

eral theoretical concepts, and even between analysis and design, i.e. between understanding the problems and possibilities of current use, and possible new artifacts infused into such an ecology. To discuss these challenges we present an activity theory-based understanding of how human beings perceive and appropriate artifacts. Where Dourish [66] sees the body and embodiment in contrast to the classical focus of HCI on the mind, activity theory helps us understand the relationship between embodiment and understanding differently, but within the same framework. Although the foundation for the theoretical development is not new, neither as general theory nor in HCI, this paper re-thinks the foundation in new ways in order to address the current situation, and proposes an enriched conceptual focus on the artifact. In addition, we derive the Human-Artifact Model as a framing for analysis of multiple and ubiquitous interaction. We propose that the Human-Artifact Model can be used to frame both comprehensive and casual analysis of observed interaction, and to frame design considerations, e.g. for prototyping.

Just like [82], and for similar reasons, this paper is a conceptual one. The above set of complex understandings and perspectives indicate that a more precise terminology is needed, one that makes us think and act as analysts and designers in the above characterized void. We believe in multiplicity and have no aspirations for a complete and all-encompassing methodology. Nonetheless we present the Human-Artifact Model to deal with the challenges of complex artifact ecologies, and focus on the concepts needed to understand and design for these.

Central to our approach is that activity theory allows for dialectical rather than causal thinking. Dialectics is the method of reasoning that aims to understand things concretely in all their movement, change and interconnection, with their opposite and contradictory sides in unity. There are many approaches to dialectical thinking that we will not scrutinize here (e.g. Hegel, Engels, Marx and many later philosophers). In the following we pursue these challenges by looking deeper at the concepts of mediation, multiplicity and development:

1. Human activity is mediated by artifacts. While we need to address how artifacts support transparency and seamlessness, we need to similarly address seamfulness and breakdown.
2. Human beings use multiple artifacts in multiple, overlapping use activities, i.e. artifact ecologies. We need to address every single artifact on the one hand, while keeping in mind the multiplicity of artifacts on the other.
3. Human understanding and use of artifacts develop over time. Accordingly, we need to address how human use of (particular) artifacts has been historically shaped on the one hand, and is part of shaping possible future use, on the other.

In the remaining of this chapter we provide an overview of activity theoretical HCI. This leads to a chapter, where we develop the new foundational

contribution of artifacts, leading to the human artifact model. We then go on to illustrate the analytic power of the model. The paper is concluded with a discussion of related work and what we have achieved.

1.1 Activity Theoretical HCI

Since the mid 1980's activity theory has been explored as a basic perspective on human-computer interaction. In an attempt to break with cognitive science-based HCI, a theoretical platform has been established, based on dialectical materialism [88, 71] and human beings acting in real-life situations [173, 68, 158]. In the following we revisit this basic perspective with a view towards the current empirical and theoretical situation in HCI.

Fundamentally, activity theoretical HCI [38, 30] led to a focus on extending HCI to focus on analysis and design of artifacts for a particular work practice with concern for qualifications, work environment, division of work, etc. From HCI's classical focus on tasks it moved attention towards actual use and the complexity of multi-user activity, in particular the artifact as mediator of human activity. The development of expertise and of use in general came into focus, and resulted in a further concern for active user participation in design and use as part of design.

Activity theoretical HCI offers a set of concepts and tools that applies in various ways to such analysis and design, but it does not prescribe *one* best approach.

Activity theory addresses more than just individual skills, knowledge and judgment, and is not restricted to the "generic" human being, because it understands human conduct as anchored in collective/shared practice. Activity theoretical HCI focuses on the appropriateness of certain tools for certain practices. It studies how the introduction of new artifacts changes practice, and how practice may change the use of these artifacts. Furthermore, it focuses on individual human development in relation to development of the culture and communities of practice of which it is part (see also [99]). As practice develops over time, concern for the historical context of such artifacts in use is essential to activity theoretical HCI. Learning is not only a matter of how the individual adapts to particular artifacts; it is a matter of how the collective practice develops, in small or larger leaps. In actual use, artifacts most often mediate several work activities, and the contradictions and conflicts arising from this multitude of use activities are essential for activity theoretical artifact analysis and design.

Human activity can be analyzed into a three-level hierarchy of activity, action and operation [111, 110]. Activity deals with the level of motivation; it motivates why a particular set of actions, with particular material or ideal objects, is carried out. Without motive, there is no activity. The subject's reflection (in terms of expectation and evaluation) of this object motivates the activity. Motives are often tacit or unarticulated. With the division of work in society,

activity may be poly-motivated, meaning that a particular activity may have more than one motive, coming from e.g. different areas of life. To describe an activity at the activity level means to focus on the social and personal meaning of activity and its relation to motives.

Human activity is carried out through actions. These actions are governed by the conscious goals of the subject. Goals reflect the outcome of action, and vice versa, quite obviously. This reflection consists of an ongoing evaluation of the actual outcome of actions on objects with the desired outcome. Hence, goals are related to the desired future state of an object and are, thus, different from the motive. Since goals are conscious, it is through goal formulation we immediately meet human activity in an analysis. Actions are realized through series of operations. Actions are recursive structures, consisting of both conscious and non-conscious sub-actions and operations. To describe an activity at the action level means to focus on *what* the subject does and on possible goals, critical goals and particularly relevant sub-goals of subjects [8].

Operations are never fixed, but adapted dynamically to the conditions of the environment. Operations are “triggered” by the conditions for and structure of the action. They are performed without conscious thinking, and oriented by what Kaptelinin [96] calls an unconscious *orienting basis*. Operations are cultural-historically developed or have naturally evolved [8]. They result from appropriated use of tools, educated manners towards other human beings, or movements in the physical world according to concrete physical conditions. To describe the activity at the level of operation means to focus on *how* the activity gets carried out; the concrete way of executing an action in accordance with the specific conditions surrounding the goal. A summary of these levels is presented in Table 1.

The action level is where researchers and designers meet users, when we ask them what they do. Researchers and designers can more easily understand the available action possibilities of the human users when they combine a description of the possible and critical goals and actions of a particular activity, with an understanding of the quality of actions, and the repertoire of operations and conditions for operations for each of those actions.

The activity theoretical tri-partition should not be understood as means for static categorizations of analytical findings. Instead it provides three sets of analytical glasses, each of which focuses on an important aspect of human activity: Motivation (by asking *why?*), goal-orientation (by asking *what?*) and function (by asking *how?*). It is important to sustain the dynamic relationship between these three aspects. Stating that something is an operation is not interesting without questioning what action has been operationalized in the first place? What seamfulness may make the human actor consciously aware of it? Where are its limits? When does it break down?

From HCI the notion of motive may seem abstract. However, it forces the analyst to look beyond the apparent. Motive embodies the unspoken and unconscious. In our design case below we see how a central, albeit unspoken,

Analytic layers	Mental representation	Realises	Level of description	Analytical question
Activity	Motive (need)-not necessarily conscious, but may become conscious	Personality	The social and personal meaning of activity, its relation to motives and needs	Why?
Action	Goal-conscious	Activities (systems of actions organised to achieve goals)	Possible goals, critical goals, particularly relevant sub-goals.	What?
Operation	Condition of actions (structure of activity)-normally not conscious, only limited possibilities of consciousness	Actions (chains of operations organised by goals and concrete conditions)	The concrete way of executing an action in accordance with the specific conditions surrounding the goal.	How?

Table 1 Human activity and its analytical layers – The table is adopted from [7] and [8] and shows the relation between the three levels of activity, as regards their conscious or unconscious mental representation, what elements of human activity that the level realizes through action, the way in which the level may be analytically described, and the analytical key question.

motivation for the actions of two people solving assignments with a geographical map was to establish a common orienting basis. For these actors to successfully work together in solving the given assignments, the artifact needed to support such as common orientation.

This theoretical grounding seeds an exploration of mediation, multiplicity and development. Nonetheless, we are quite far from the detailed concepts and methods needed to provide a revised theoretical basis. Hence, we proceed to pursue the understanding of artifacts in Human-Computer Interaction from the perspective of activity theory.

2 Artifacts

Activity theory was originally introduced in HCI to point out that the relationship between the human being and the computer is not a simple subject-object or subject-subject relationship: Instead of studying the relationship between the user and the computer as something that the user works on, or communicates with, Bødker [38] pointed out how we may more usefully see the computer as something that the user acts through, *on* other objects or *with* other subjects:

Designed artifacts and their role as mediators have been at the core of this way of thinking. A further, overriding concern of the following discussion is to move the focus away from one artifact alone towards *artifacts* in plural. The

move beyond the desktop computer as the sole interactive artifact forces us to be able to conceptualize the interplay between artifacts: How use of one artifact may influence use and perception of another, and how human understanding and use of artifacts at large influence an artifact in use? In this manner the focus is on the interplay between reflection and action, rather than one or the other alone.

2.1 Artifact Ecologies

Artifacts are never isolated, and they cannot be understood as such. All artifacts used by human beings are part of artifact ecologies, whether simple (e.g. pen and paper) or complex (e.g. tools for building a house). Hence, human activity is not just mediated through a single artifact; it is multi-mediated. Bertelsen and Bødker [29], Bødker and Bøgh Andersen [40] exemplify how multiple mediators for specific activities may be connected: *Chains* (e.g. a key-card produces a number that is used to open a door), *meta-instruments* (e.g. the pencil sharpener and the pencil), *levels* (e.g. driving the car by cruise control on the freeway using only the wheel to control the car, versus curb-side parking where several mediators are used to move the car-wheel, rear-view-mirror, clutch, brake, etc.) and *co-occurring mediators* (e.g. the pencil and paper). In environments of many interchangeable artifacts, *substitution* (Paper III) is a similarly relevant relationship, emphasizing how and under which circumstances one artifact may replace another. An artifact ecology often consists of multiple artifacts built for similar purposes, but with slight variations and no clear delineation of when to use which artifact. The specific choice of artifact between these is situated and depending both on the material conditions of the activity and on the specifically intended outcome. Examples are sets of clubs for golf or brushes for artistic painting.

From an activity theoretical perspective, artifact ecologies are defined from the point of view of activity. Primarily we look at ecology as connected to purposeful action of some kind, and not as truly endless action possibilities of the environment. The artifacts used on a regular basis by users in relation to a given activity constitute the *current* artifact ecology of use (see also [94]). In addition, the *historical* artifact ecology plays an important role when appropriating new artifacts, since it is what shapes the user's perception of other artifacts. Such past experiences are elements of the orienting basis (see later). The historical artifact ecology consists of all the artifacts that previously have been applied to realize the activity.

Within artifact ecologies, multiple overlapping activities take place with multiple motivations and purposes. A number of artifacts offer overlapping uses at various levels of activity: A fountain pen and a pencil may both be used for writing a note on a piece of paper. However, they are held and handled quite differently, and whether they should both be considered when designing a new writing device much depends on focus of the design. Jung et al. [94] show how each interactive artifact in the user's artifact ecology influences

how the other artifacts are used. Having both a laptop and a workstation may result in the laptop being used only for work and the workstation entirely for entertainment even though each device in isolation is capable of both.

The examples from Jung et al. [94] primarily illustrate how artifacts influence one another at the level of actions: *What* is done with which artifact. The well-known story of why we use qwerty keyboards today on laptops and cell-phones illustrates that artifacts may similarly influence each other on the *how* level: The qwerty keyboard was originally designed to prevent the arms of a typewriter from getting entangled when the typist went too fast. Hence, frequent key-combinations were spread out on the keyboard. This concern is no longer relevant for computer keyboards, and other layouts of keys have proved more efficient, yet we are somehow stuck with qwerty, in parts because many resources have been put into training typists. With these examples in mind, we propose that it is important to address the artifact ecology from all three levels. We return to this in [Section 2.6](#).

While we propose to address artifact ecologies primarily as collections of artifacts that are determined by looking at empirical situations, artifact ecologies in addition define a wider space of action possibilities. This is because artifact ecologies, in principle, consist of every artifact that the acting subjects have come across in realizing an activity. Taking a dialectical perspective, it is in the field of tension between the specific artifact ecology and the space of experience-based action possibilities we find the *potentials and problems* of relying on *past experiences* and *similarities* from other artifacts.

The focus of artifact ecologies help address the challenges of multiplicity by drawing boundaries and exploring the relationships between artifacts in a certain setting, at all levels of activity. Development is addressed through the historical artifact ecology in relation to the current ecology, and to future action possibilities with new artifacts. We proceed to look at mediation in the following.

2.2 The ideal of the artifact

Evidently, a computing device, like any other thing in our surroundings, can simply be regarded as an object; something that can be looked at, picked up, or even used for something unanticipated, e.g. thrown to break a windowpane. However, this is likely to happen mainly if the “user” has either no past experiences whatsoever with a similar device, or if the user is desperate and caught with no other alternative. One additional situation where computing devices are objects of attention is design, where they constitute the *material*, which gets transformed into some sort of workable computing device.

This role in design points to an important quality, which makes computing devices more than just objects: They are *artifacts*, i.e. they are designed or shaped by human beings with a particular purpose or use in mind. It is this question of artifactness that drives HCI: How do we design computing arti-

facts that work (better) in use? How do we shape artifacts to fulfill particular purposes that are (more or less) well understood?

In the following we leave the focus on multiplicity for a short while to focus on the relationship between an artifact and its user. Beguin and Rabardel [24] introduce the term *instrument* to address this relationship. An artifact becomes an instrument through the activity of the subject. The artifact may be perceived as a hammer given its physical properties, but only when used for hammering does the artifact become the instrument “a hammer”. In other situations the artifact may be the instrument “a weapon”.

Placing the computing device as something that the user acts through, on objects of interest or with other subjects, points to the role of a mediator. The mediator stands between the user and the object of interest, and in this role it helps the user act on the object of interest, in ways she could not act without using the mediator. The hammer helps the user drive a nail as one of the many mediators of house building; the telephone conveys our conversation across distances, and the web browser helps browse web-pages or plan a road trip for the summer. The mediator can be understood as standing between the user and the object of interest, and it is the quality of this standing in between that has been discussed in HCI, often dressed up as transparency or in the Ubicomp setting as seamlessness [66]. A mediator that works well allows the user to focus on the object of interest, when carrying out the necessary acts supported by the capacities of the mediator. A mediator that does not work well causes *breakdowns* and draws the user’s focus towards the artifact as such. Whether or not a breakdown occurs during use depends on the actual material conditions of use. It does however, also depend on the quality of the action repertoire, possessed by the user, as we will return to. Whereas any physical object may ultimately be used to e.g. break a windowpane in case of an emergency, they are not equally well suited to do so. A big hammer is better suited for big nails and a small for small ones. This kind of seamfulness [59] is often present in the physical world, but seems to lack when it comes to computing. We introduce the concept of *mediator* to capture that certain a mediation takes place between a subject and an object through an object or artifact, while the concept of *instrument* is used to capture what is actually done with the artifact or object by the users. Accordingly, mediation is not a simple matter of transparency or making things “go away”. Mediation addresses the dialectics between seamless and seamful, between transparency and breakdown.

The hammering example gives indications of such a dialectical focus on all levels of the activity and leads towards the concept of *functional organs*. The concept of functional organs captures the functionally integrated, goal-oriented configurations of internalized operations and external mediation [97]. An artifact can obtain the form of a functional organ to a user when a certain combination of artifact and internalized routines are present in a given use situation. Hence, it is a dialectical concept encompassing both aspects of the subject and the artifact. The fork and knife become functional organs in the activity of cutting and eating food. Thus, the knife and fork augment the human ca-

capacity with regards to eating. Artifacts that become functional organs become part of human beings, moving the boundaries of the human being “outward”, similarly to how Polanyi [138] describes the blind man “seeing” with his cane. An artifact does not become a functional organ for somebody without development of operations.

In the same manner as Heidegger’s concept of ‘vor handen’ [68], functional organs are both ideal and actual states. An artifact becomes a functional organ the moment it is applied as an instrument to a degree, where the user is able to think through the artifact. Appropriating an artifact as a functional organ enables the human user to perceive the world and reason through it; even when she is not necessarily holding it in her hands. The ideal of functional organs is when we are capable of exhaustively exploiting the action possibilities of an artifact in a given activity. This is indeed an *ideal*, since it is almost impossible to achieve in the fullest extent. However, the blind man and his stick, as well as some virtuous violin players and lifelong Emacs users, actually seem to have this sort of relationship with their artifact.

This section has focused on mediation and the inner dynamics of the relationship between the artifact and its user. Before we proceed to understand the levels of activity, we need to approach development, as stated above.

2.3 Development in artifacts

The dynamics of human activity is essential to activity theory and the Human-Artifact model. As a matter of fact, over the recent years we have seen quite a few examples of the application of activity theory within HCI and CSCW (e.g. [164, 60]) where the theory basically was used to map out the multiplicity of artifacts and activities, whereas the dynamics were ill understood. A noteworthy exception from this is Bryant et al [49], who focus specifically on development of use. Activity cannot be analyzed once and for all. Rather, activity is continuously changing in interplay with other activities, artifacts used, etc.

Bærentsen [7] illustrates the dialectical relationship between activity and tools in an artifact historical analysis of the development of hand-weapons from the time of bow and arrow to current automatic rifles. Bærentsen shows how human operations are gradually incorporated and crystallized into artifacts, and how this development shapes both the goals and motives of human users. The introduction of hand-muskets made re-loading far more complicated than placing an arrow on a string; nevertheless its speed was outweighed by the sheer firepower and the status of using muskets. The operational level of the use of the weapons changed radically. The routine for reloading moved from putting an arrow on a bowstring over stuffing a musket, to inserting a clip into an automatic rifle. While the development of weapons has contributed to changing the society, the central goal of the development of the weapon has not changed; that is to be able to shoot down an enemy from a distance. However, the human action possibilities keep developing in a dialectical relationship with e.g. these weapons, and in this manner the future cannot be fully

anticipated [71].

Historical analyses are significant to the activity theoretical framework. Engeström [71] proposes artifact-historical analyses as part of the methodological cycle of work development. This methodology focuses on developing a particular activity, and identifies the artifacts that are historically central to the focal activity. The historical development of activity implies a development of artifacts and environments. Modes of acting within such activities are historically crystallized into artifacts; in this sense the historical development of activity can, in a non-deterministic manner, be read from the development of artifacts mediating the practice [7].

To summarize, on the one hand, artifacts are fundamentally crystallizations of activity [111, 110]. Activity is crystallized into artifacts in two ways: Firstly, they are externalizations of operations with earlier artifacts, and secondly, they are representations of modes of acting in the given activity. On the other hand, the artifacts shape the activity in which they are used. This way of thinking sets apart activity theoretical HCI from more static or causality-focused approaches.

Activity theory focuses foremost on human action in the material world. In contrast to many cognitivist stances it does not see what is happening in the human mind as separate from what is happening in the material world [88]: Human shooting skills and knowledge are activated when shooting, e.g. understanding of the role of the distance to the target and the manners in which the handgun is held. At the same time human skills and knowledge are reflected in this shooting, and are changed as a result of the specific shooting experience. Norman is one of several researchers (e.g. [141, 142, 86, 58]) attempting to use Gibson's ecological psychology to move away from the idealism of cognitive science, in particular the separation between human cognition on the one hand, and human action on the other. Activity theory, as well, rejects such idealism and combines an interest in actual material conditions of human acting with a focus on motivation of actions. That such a stance has consequences can be seen when looking at e.g. Hartson's [82] attempts to develop different kinds of affordances in interaction design. By separating cognitive affordances from e.g. physical affordances, and putting elements of the interface in similar boxes independent from the actual use situation, he moves Gibsonian thinking "back" in the direction of idealism and causality, and away from reflection and dialectics, as we return to in the next section.

Based on this richer understanding of dynamics and development, we proceed to address how activity is leveled and what that means for our understanding of artifacts.

2.4 Levels of activity and artifacts

To conceptualize what is externalized in an artifact, Bødker [38] proposes the analysis of technological artifacts through three kinds of aspects: *The physical*

aspects; the handling aspects and the subject/object aspects: The physical aspects are the conditions for the physical handling of the artifact. The human adapts to the forms and shapes of the artifact, and a mal-adaptation may prevent the forming of certain operations. The handling aspects constitute the support for operations toward the computer artifact. A breakdown in these operations will make the artifact appear as an object to the human user. In breakdown situations, these kinds of operations can be conceptualized, whereby the user may be forced to conduct actions toward the artifact as an object. The subject/object-directed aspects constitute the conditions for operations directed toward objects or subjects that we deal with in the artifact or through the artifact. Different parts of the subject/object-directed aspects relate to different subjects or objects, but it is also part of these aspects to support the shift between subjects/objects. Let us use the example of a (shared) electronic calendar: The physical aspects relate to e.g. the screen size, and to whether new appointments are added through pointing with a mouse or pen, or by typing on a keyboard, allowing for an analysis of e.g. the limitations to screen size, or the suitability of a pen-based PDA solution. The handling aspects relate to how appointments and meetings are added, whether e.g. multiple, simultaneous appointments are possible, etc. The subject/object-directed aspects address how it is possible to e.g. plan for both a single user and a group, and shift between these foci.

In summary, the aspects help separate out, and summarize elements of the artifact. They address their change and interconnection. By bringing the aspects together as a leveled analysis, it is furthermore possible to focus on their contradictory elements. We move on to refine the understanding of the levels further.

Bærentsen and Trettvik [8] combine and extend the use of Gibson's [79] affordances with activity theory. They identify three types of affordances to match the levels of activity, action and operation: Need-related, instrumental, and operational affordances. The need-related affordances relate to what motivates people. The instrumental relates to the socio-culturally shaped action possibilities in instruments and objects surrounding us. Finally the operational affordances: Gibson's original level of affordances for movement in the four-dimensional physical world are the action possibilities relating to human beings' naturally evolved, ecologically determined patterns of behavior and conditions. Bærentsen and Trettvik identify the lowest level in the operation hierarchy, where human beings get confronted with the operational affordances and describe how, at the operational level, some of these are learned and can be conceptualized, whereas others come out of human adaptation to the environment (the adaptive operational level). This is the difference between the adaptive level, which they define as the level of human low-level response to natural conditions, and conscious operational level, which addresses the human repertoires of cultural-historical ways of interacting with the environment.

Bærentsen and Trettvik are critical to Norman's [127] attempt to solve the confusion of how affordances had hence far been used in HCI: "*As we see it, the problem with affordances stems from attempts to adapt it to the dualistic Procrustes bed*

of cognitivism with the result that it is reduced into something fundamentally foreign to Gibson's use of the concept. In this guise the properties of "affordances" can only exist as mysterious and magic properties of objects" [8, pp. 52]. Hartson [82] falls into a similar trap when he develops his leveled analysis of artifacts based on affordances. While he analyses cork screws using a leveled separation quite similar to that of Bærentsen and Trettvik, the mere use of the term cognitive artifact for one of these levels is puzzling, and seems to indicate that cognition is separate from doing. He similarly discusses an alien corkscrew, indicating that there is a "correct use" that is separate from e.g. the past experience of the human user, which is in contrast to Gibson's ideas (see our further discussion of this example in 2.6). Gibson invented the notion of affordance to talk about action possibilities, which he addressed in terms like stairs being walkable, chairs being sitable, doors being openable and doorknobs being turnable, etc. This strongly emphasized action possibility over property and correct use (see also the discussion of the relationship between affordances and activity theory in [99]).

Instead of pursuing the complex discussions of Gibson's affordances, we follow a path that leads us away from the above Procrustes' bed. Matching Bødker's [38] aspects with the three types of affordances makes it possible to better understand how current mediators are integrated in the functional organs of the human users. We may address the possible and desirable types of uses that are afforded, and which are not. This mapping requires a re-interpretation of Bødker's aspects. Bødker's physical aspects encompass Bærentsen and Trettvik's *adaptive operational* affordances, and we choose to use this term, since 'physical' can be slightly misleading. In Bødker's three aspects, the need-related or activity-related level is missing, and since Bærentsen and Trettvik convincingly argue for its importance, we include this level. Even this aspect deserves a terminological comment: We have previously used the term *motive* as a synonym for need as the term is used here. In order to not confuse matters further, we continue talking about motive and the *motive-related* level. We prefer to use the term *instrumental*, as introduced by Bærentsen and Trettvik, rather than Bødker's subject/object directed aspects. However, we use Bødker's *handling* rather than conscious operational. The matching of activity levels, affordance and aspects is summarized in Figure 1.

Beaudouin-Lafon [18] proposes to address the relationships between the instrumental and operational aspects of artifacts. Specifically, he develops three ways of assessing such WIMP interface components as scrollbars, text-fields, buttons etc. in terms of how well-integrated are the instrumental, handling and adaptive aspects. The *degree of indirection* includes *spatial offset* and *temporal offset*. Spatial offset is the distance on the screen from the logical part of an instrument to the object it operates on. This is similar to the relationship between handling aspects and instrumental aspects in our terms; e.g. a dialog box can be far from the object that it acts upon and has a high spatial offset. Temporal offset refers to the time it takes from an action is initiated to an effect is visible. The *degree of integration* refers to the compliance between the controlled dimensions of the input device and the logical instrumental dimensions, the relation

Affordance	Aspects	Activity level
Need-related		Activity - Why?
Instrumental	Subject-object directed aspects	Action - What?
Operational - Conscious - Adaptive	Handling aspects Physical aspects	Operation - How?

Figure 1 Summary of affordance, aspects and activity levels. Bærentsen and Trettvik's [8] affordances are aligned with Bødker's [38] aspects of the artifact and the activity levels, illustrating the need for introducing motivational aspects, and revising the vocabulary further.

between handling aspects and adaptive aspects. Beaudouin-Lafon argues that when it comes to WIMP interaction, the more direct relationship between handling aspects adaptive aspects the better. Lastly, the *degree of compatibility* is a measure for the similarity between the physical actions of the user on the input devices, and the logical actions of the instrument, similar to the relationship between adaptive aspects and instrumental aspects. According to Beaudouin-Lafon, dragging an object has a high degree of compatibility, while typing in the position of an object to move it on the screen has a low. Beaudouin-Lafon points out ways of working with the four aspects for specific kinds of interfaces, yet the three degrees do not immediately lead to general measures, and cannot be generalized beyond WIMP interaction. We return to this concern for generalizations in 2.5.

The four aspects serve to focus on actual use on the one hand, and on the assumptions made about use in the artifact, the action possibilities and constraints, on the other. The creation of action possibilities is the designers' concern, no matter how the action possibilities are created. Hence we address this in terms of how aspects influence the ways designers may create action possibilities. We have enriched the notion of functional organs by identifying four aspects, related to the levels of activity. We have also pointed out that in specific settings, it is possible to systematically address the relationships between the four aspects. In continuation of the argument of Section 2.3, there is action and understanding at all these levels, and a final and static sorting of interface features into the four aspects will not suffice.

2.5 The orienting basis and quality of actions

Functional organs encompass both artifact and user. Hence, if we are to discuss the quality of this duality, we need to explain the relationship between the artifact, its aspects, and the user's capabilities and experience. This is the focus of our next section, where we turn to the orienting basis of the user and

the qualities of action possibilities.

Users' shared capacities and experiences are not only based on individual acting and learning in the world. Human beings are born into a cultural situation that makes today's human beings capable of acting differently with technology than those born a century ago. In addition, human beings get to share particular practices or praxes through schooling, etc. When getting trained as a carpenter or nurse one gets to share the praxis of carpentry or nursing. At the same time each individual, who is part of a praxis, continues this praxis. He or she changes it as well by developing new ways of doing things, and new artifacts [99]. This praxis allows us to talk about more than just individual skills, knowledge and judgment. When we look in further detail at the relationship between the user and the artifact in the following, it is on this background: There is no user without other users who share experience with artifacts and materials, understanding, etc.

We turn to Gal'perin [75] in order to understand the human capacities and experiences. He presents the *orienting basis* as central to support the application of actions and operations across different artifacts and objects of work. The orienting basis guides actions, and is shaped through experiences. Where the aspects above help address what the artifact is intended or may be used for, the orienting basis similarly addresses the capacities with which the human user meets the artifact. There are three ways in which the orienting basis is constituted. These reflect both how the orienting basis is created and how it may potentially be applied: *Domain independent general actions*, *reusable knowledge within a specific domain*, and *trial and error*.

One example of domain independent general actions is bike riding. Once learned, this may be carried out for work, everyday transportation or leisure activity. Other examples are addition or multiplication of numbers, to be used for many kinds of purposes once it learned in its general form, or writing and sending email messages for work or friendship. Examples of reusable knowledge within a specific domain would in contrast be emailing learned only within a particular work context; hammering and drilling while working with wood and plywood only; playing the bass as rock-and-roll instrument, but not as an instrument for jazz.

Examples of actions learned through trial and error are division of integers by ten without understanding the general principle of the position system; playing "Autumn Leaves" on the double bass by heart after locating one tune at the time on the instrument. If an action has been developed through trial-and-error, there is less chance that it can be reused to develop other quality actions, since fundamental understanding may be missing. Generality and reusability of actions are important factors when it comes to addressing the action possibilities available in the orienting basis of the user. The quality of a specific action may, according to Gal'perin [75], be assessed through the following:

- Generality, i.e. the extent to which recognition of an object's properties

relevant to completing an action is part of the given action.

- Support for abbreviation, i.e. the extent to which the action will help to consciously skip operations in the performance. Operations can be abbreviated either consciously or unconsciously. Conscious abbreviation is natural and non-harmful, e.g. skipping intermediate results in the process of solving a problem. Harmful abbreviation is the opposite unconscious skipping of intermediate results or actions, due to of lacking knowledge or experience with the consequences.
- Mastering, i.e. the independent reproduction of an operation with new material, achieved within the action.

Consider the example of a double bass student who has no formal theoretical music training, and must learn to play the walking bass line of “Autumn Leaves”. One strategy would be to listen to a recording, imitating what is played. However, when asked similarly to play a walking bass line for “Summertime”, the double bass student would have to repeat the process of listening and imitating. Another student, who has trained chords and scales, would immediately be able to play along to any of the two songs just by seeing chords—as well as easily transposing the song to another key—something that would be difficult for the first student. Hence the second student has an orienting basis that is characterized by a larger degree of domain independence (even though it is all about music), and more domain specific knowledge than the first student who had primarily learned through trial and error. The actions of the second student are both more general and mastered. He may be able to do useful abbreviation, whereas the first student is likely to “cut the wrong corners”. At some point, however, the first student may still be able to generalize from all the songs imitated to obtain an equal understanding of walking bass as the formally trained student. As a matter of fact several great jazz musicians are autodidact. Through playing double bass the student builds domain knowledge with playing stringed instruments in general. This would make her capable of applying her orienting basis to new types of string instruments. For instance, a double-bass student might be able to generalize and concretize tone intonation and left hand vibrato and apply these to a violin.

Gal’perin argues that familiarity is what triggers the establishment of an orienting basis. Bardram and Bertelsen [16] refer to this with the concept of initial familiarity, which they argue can be constituted in interface design both through playing on simple affordances or e.g. interface metaphors. The nuanced view of affordances from Bærentsen and Trettvik [8] conceptualizes the dynamics of this familiarity.

Like the aspects of artifacts, the orientation can be broken down into levels of activity. At the top-level, activity is driven by a motivational orientation. The motivational orientation is partly subconscious, meaning that human beings are not in-situ aware of their motivation. Goal orientation encompasses the user’s repertoire of actions. Applying the understanding of chords

to play “Summertime” *prima vista* is an example of this. The handling orientation refers to the users’ culturally developed operational means for realizing actions that have become partly sub-conscious. Such means include the way of holding the bow for classical bass, according to the German school and the French school. The adaptive operational orientation is the low-level response to physical conditions; e.g. tweaking one’s left-hand intonation for a set of new strings. This leveled analysis helps us provide a symmetrical model where the human side mirrors the artifact and *vice versa*.

Action always takes place as specific constellations of actions combined with a constellation of operations. One particular action may have one high-level intention supported entirely through operations, and in another case the sequence may consist of many conscious actions with low-level and fragmented operations. In between are many possible combinations. Our concern lies within the whole repertoire of action possibilities: To write a short note on a sheet of paper, the writer may pick up the nearest pen and write what is needed, presuming that a functioning pen is nearby, and there is no particular need for color, clarity, etc. In other instances, while the writing with various kinds of pens may still be operationalized, conscious choices regarding for instance color may have to be made. In a third instance, the actual writing may not be operationalized quite as much, and choices regarding the size of the characters must consciously be made, for example if the only available pen is a fountain pen or a carpenter’s pencil.

With the above perspective it is possible to address the quality of the action repertoire, and not only the individual action: A golfer who successfully applies all clubs in the golf bag and who knows, or is able to assess, when to apply which club (and mastering the pitch, the put, etc.) has an overall higher quality set of actions and operations, than the golfer who only masters the 7-iron, no-matter how skillfully he may apply this. While both may still encounter limitations to their technique, and have problems if new types of golf clubs are introduced, there is a difference between their action possibilities.

Bryant et al. [49] describe the development of the use of a particular mediator, Wikipedia. This development is by no means a simple operationalization of actions and Bryant et al. [49] carefully document the development of the orienting basis of the users from newcomer to Wikipedian. They characterize the differences between these two groups, and suggest that there are patterns in how users move from being newcomers to becoming Wikipedians. The users change their perception of Wikipedia, and no longer act upon a random collection of articles. Instead, to the Wikipedians, Wikipedia is primarily a community of co-authors. In becoming Wikipedians, users move from a local focus on individual articles to a concern for the quality of the content as a whole. In summary, the development of use involves new goals, new roles, different tools, and especially, new motives.

The orienting basis helps address users’ past experiences and how these get activated when new artifacts are brought into place through familiarity and more extended use, as in the case of Wikipedia. The three approaches—applying

domain independent general actions; reusable knowledge within a specific domain; trial and error—give indication as to how the human user is prepared to take up the new artifact. Generality, etc. gives yardsticks to compare quality of actions, and hence the possibilities and problems relating to the use of new artifacts or existing artifacts in new situations. From the perspective of designing new artifacts, the orienting basis points towards choices, e.g. whether to target general action possibilities or specific reusable knowledge within a specific domain; or leave the new user entirely to trial and error exploration.

In the preceding sections we have provided a theoretical foundation for a more nuanced understanding of artifacts, their relationship with their user and their mediation of human activity. By emphasizing the dialectical relationship between designed artifacts and user's orienting basis, we have argued that transparency is not simply a matter of artifacts going away. Later, artifacts become functional organs to the user through learning in use, hereby shaping the user's orienting basis. Depending on how the existing orienting basis of the user has come into being, e.g. how general actions are, the user is equipped differently as to face the new artifact, or an existing artifact in a new setting. The theoretical foundation provides concepts for analysis at different levels of interaction and underlines the dynamics between them.

2.6 The Human-Artifact Model

We present the Human-Artifact Model as a summary of the insights of the theoretical foundation, for these insights to be applied in actual analysis and design in a structured form. The Human-Artifact Model is to be used to summarize empirical findings, and to summarize and analytically reason about the dynamics between the levels of interaction and the relationship between the aspects of the artifact and the orienting basis of the user, the possibilities and problems of a functional organ (see [Figure 2](#)).

The Human Artifact Model provides a form, where each field can be addressed one at the time by summarizing empirical findings, and by singling out particularly critical issues or findings to match or contradict those of other fields. If the starting point is in the human side, the Human-Artifact Model can be used to structure the analysis of human practice and orienting basis in terms of the four levels, and a particular current or future artifact may be added on the artifact side. This approach makes it possible to address e.g. quality of actions, and possible new designs. Or on the contrary, the starting point may be the artifact, where the four levels again are important for structuring the analysis. The human side can consequently be analyzed as regards potentials and problems of matches with the artifact, raising e.g. issues regarding the development in use. In the analyses it is possible to move back and forth between these two perspectives, to focus in on one level at a time, and to address tensions between fields, across sides and levels, e.g. in breakdowns. The Human-Artifact Model makes it possible to analyze the actual human use of an artifact reflected in the ideal, well-integrated functional organ in order to find trouble spots and areas

Why?	Motivational aspects	Motivational orientation
	Instrumental aspects	Goal orientation
What?	Operational aspects	Operational orientation
	- Handling aspects	- Learned Handling
How?	- Adaptive aspects	- Adaptation
	Artifact	Human

Figure 2 The Human-Artefact Model. The left hand side of the figure ties to the artifact; the aspects. The right hand side is the human side, and ties to the user's orientation. Both sides are expressed vertically through the levels of activity. We use this figure as a building block in our analyses in the rest of the paper. The two sides are sometimes reversed to create a better overview.

for development.

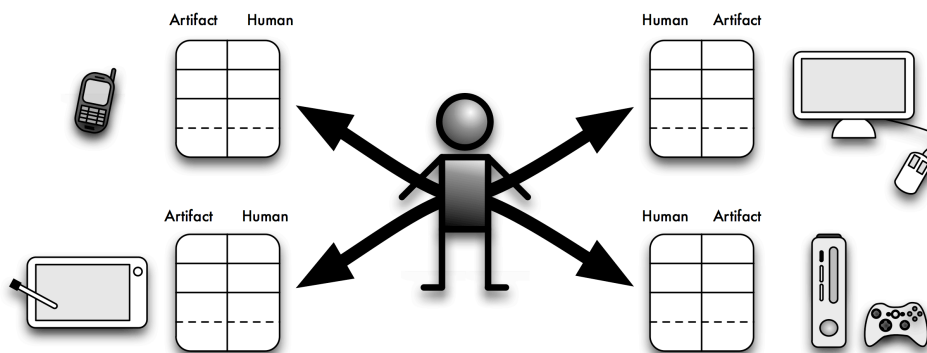


Figure 3 The Human Being surrounded by artifacts. The human being, in the context of other human beings, accumulate experiences with artifacts in the artifact ecology. Some of these artifacts are historical, some current.

While the Human-Artefact Model is to be applied to analyze use of an artifact, the artifact is not analyzed in isolation. The orienting basis side of the Human-Artefact Model, that is the user side, accumulates experiences with other artifacts in the user's artifact ecology; some of these historical and some current, some individual and some with others (Figure 3). In other words, the Human-Artefact Model embraces the possibility of change, and as such it is a useful starting point when facing HCI as a design discipline, rather than as

something providing static analyses of technologies in use. Similarly, the artifact side crystallizes and accommodates many uses and many users (Figure 4).

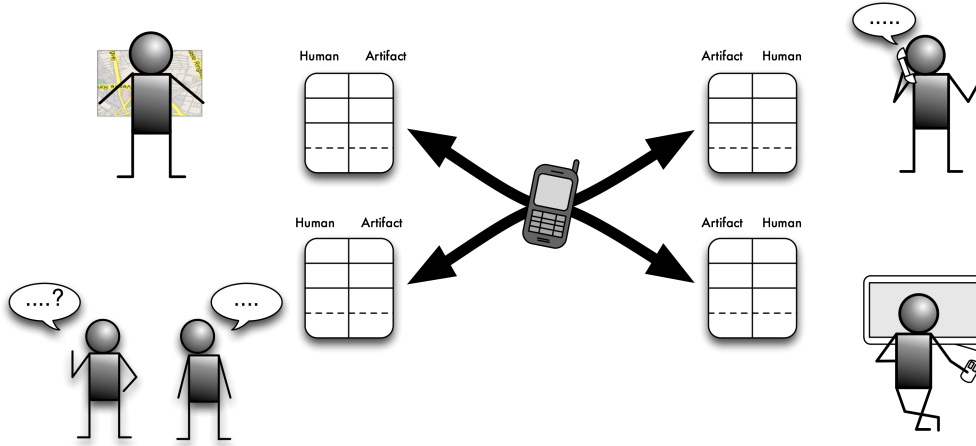


Figure 4 The artifact surrounded by users and uses. The artifact is the meeting place for several activities and actions by the user(s). The aspects influence these activities, and the artifact gets used through repertoires of actions and operations developed across these activities.

Two brief examples

We illustrate the use of the Human-Artifact Model for the analysis of concrete artifacts, using two examples. These examples illustrate what issues and concerns get emphasized through the model. We return to more complete examples in Section 3. We borrow the example of a corkscrew from Hartson [82], and clarify some differences in his approach and ours. We also discuss the use of Wikipedia from Bryant et al. [49], as this adds further dimensions to the Human-Artifact Model analysis of the corkscrew.

Hartson [82] compares the affordances of two corkscrews; a winged corkscrew based on the principle of levers, and a more advanced corkscrew with two modes based on the principle of tightening a nut on a bolt. On the bolt-based corkscrew there is a locking mechanism enabling the user to turn the nut clockwise, so that the bolt (with a pointy spiral in the end) is screwed into the cork. In the second mode the mechanism is unlocked, and the cork is extracted from the bottle by continuing to turn the nut clockwise. By this, the bolt moves up and out of the bottle.

Hartson's findings are based on informally observing a group of people trying to open wine with the corkscrews. He argues that the winged corkscrew has, in his terminology, acceptable cognitive affordances, while the bolt-based corkscrew has non-obvious cognitive affordances. Cognitive affordance in his definition is "... a design feature that helps, aids, supports, facilitates, or enables thinking and/or knowing about something" [82, pp. 319]. Hartson does not, how-

ever, provide a vocabulary to help point out why some affordances are acceptable while others are non-obvious.

Winged Corkscrew		Nut and Bolt Corkscrew		
	Apply little force to uncork bottle Minor damage to cork Cheap	Drinking wine Drinking wine without cork pieces Reusability of cork to close bottle Use little force to open bottle ...	Apply little force to uncork bottle Minor damage to cork (less than the winged) Exclusive	Drinking wine Drinking wine without cork pieces Reusability of cork to close bottle Use little force to open bottle Keeping up appearance ...
	Uncork wine	Uncorking wine	Uncork wine	Uncorking wine
	Screw helix into cork Press down wings	Principle of leverage Fastening a screw	Lock handle to spiral Screw helix into cork Lock releases Continue turning handle to uncork	Physics of bolts and nuts Fastening a screw
	Handle and wings sized for human operation	Using fingers to turn handle and press down on wings	Handle sized for human operation	Using fingers to turn handle

Figure 5 Comparing two corkscrews. The examples are summarized in the form of the Human-Artifact Model. Contrasted with each other, there is a difference between the two at the motivational level, which is more important than indicated in Hartson's analysis. The two corkscrews furthermore assume experiences from two quite different handling principles, that of leverage versus that of bolts and nuts.

Now, consider the two corkscrews analyzed with Human-Artifact Model (Figure 5). Starting with the instrumental layer: Both are designed towards helping the user achieve the action of uncorking a wine-bottle. Hence the instrumental aspects are simple: Uncork a bottle. The adaptive aspects of both corkscrews embody assumptions of human fingers to turn a handle—and in the one case—to press down levers. They both have a physical shape, weight and materiality lending them well to human operation.

The two corkscrews differ most in the handling aspects. Both pullers are based on a basic idea of how to uncork a bottle, that the pointy spiral should be aligned to the centre of the cork. The winged corkscrew assumes that the user is familiar with the principle of leverage; e.g. using a plank to move a large stone. This is something most people learn at an early age; hence it can be expected in all potential users' orienting basis. The bolt-based corkscrew however assumes familiarity with the handling of a bolt and nut, something (assumedly from Hartson's informal studies) that is not as common a knowledge as the principle of leverage. Hence, Hartson's observed users meet breakdowns in the immediate handling of the corkscrew. However, when the user realizes the nature of the handling, the two corkscrews become equally easy to use.

To complete the quick analysis, the motive-related aspects of the two corkscrews embody the assumptions of pulling a cork with ease and avoiding the cork to splinter, along with an underlying idea that some uses may be more exclusive or fancy than others. Ease, technical quality and exclusiveness are reasons why more mechanically advanced corkscrews sell. They also point out why an analysis of a corkscrew in use at a music festival where users focus on being outside

and getting drunk may be entirely different. These motives are not necessarily conscious in actual use. Having learned to use the bolt-based corkscrew, this screw, according to Hartson, is much more efficient and reliable than the winged one, and the user's orienting basis is naturally changed. The next time the user is to fasten a nut on a bolt, although unrelated to wine bottles, she may recall the handling of her bolt-based corkscrew. Applying the Human-Artifact Model as an analytical lens emphasizes the dialectical relationship between subject and artifact; and moves the analysis beyond stating that there seems to be a non-obvious affordance, which is Hartson's conclusion.

The corkscrews illustrate how the Human-Artifact Model may be used to compare artifacts. With the Wikipedia example, we illustrate how the Human-Artifact Model is used to analyze the dynamics and development of use (Figure 6).

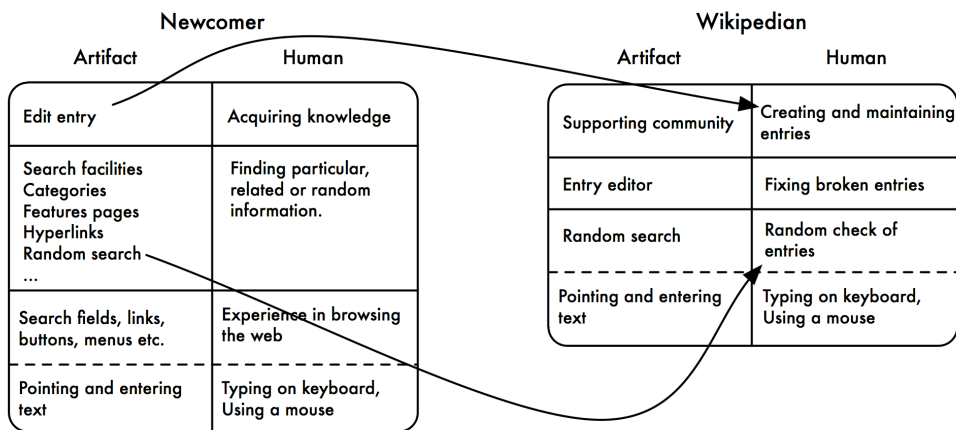


Figure 6 The Human-Artifact Model of newcomers and Wikipedians. The arrows describe the transformation from newcomer to Wikipedian where motivation is changed and operationalization has happened.

Consider e.g. the “edit this page” function for Wikipedia entries. Bryant et al [49] observe two different types of users: Casual readers and active members (called Wikipedians). From the Wikipedian's perspective the “edit this page” functionality is among the instrumental aspects designed for the goal of editing a specific entry in Wikipedia. However, according to Bryant et al., the “edit this page” functionality can be the motivational aspect that is the seed for making the casual reader into a Wikipedian. This is because it helps challenge the reader to start correcting flaws in the text. Hence the transformation from casual reader to Wikipedian is transforming the perception of the artifact and the designed prominence of the “edit this page” function involves both instrumental and motivational aspects.

Wikipedia has different means for navigating the vast amount of entries; e.g. search fields, categories, featured articles and random article search. From a Human-Artifact Model perspective these features can be seen as handling aspects. The experienced Wikipedia user rapidly operationalizes the use of these

means, and they can be used based on experience with navigating web pages in general. Hence they match the learned handling of most web-users. Many of the typical actions related to the use of Wikipedia, are not hardcoded into the artifact; there is no function for bringing forth an interesting entry, however there is a function for bringing forth a random entry. From the perspective of the casual reader this handling aspect can be used to find something interesting to read, while for the Wikipedian it is, according to Bryant et al., used to find incomplete or vandalized articles to be improved or corrected. If the random function only picked completed and well-written entries, this would probably be fine for the casual user, but worthless for the Wikipedian. This would result in a breakdown between the Wikipedian's goal orientation and the instrumental aspects of the artifact.

The examples illustrate how the Human-Artifact Model helps structure and complete an analysis that compares mediators or quality of orienting basis.

2.7 Summary

We have developed the Human-Artifact Model, based on new readings of the activity theoretical basis and existing contributions to activity theoretical HCI. This has provided new insight into mediation, multiplicity and development, and the model sets the framing of analyses of existing and future artifacts. Mediation allows us to focus on the mechanisms in the artifact and the user background that allows the artifact to disappear from the focal awareness of the user in certain situations and re-appear in other, i.e. a dynamic focus on transparency or seamlessness. Furthermore, mediation becomes a concern at all levels of the activity, from motive to handling and adaptation, and in the dynamics between them. Multiplicity is a condition of all artifact use, and in this sense our model does not treat Ubicomp any different from other use of artifacts. Through the focus on multiplicity we draw dynamic boundaries and explore the relationships between artifacts in a certain setting at all levels of activity. Development is addressed through the dynamics between the historical artifact ecology in relation to the current, and to future action possibilities in new artifacts. The functional organ helps address ideal directions to take and identifies trouble along the way to such an ideal use.

In the second part of this article we explore the analytic dimensions of the framework systematically and in greater detail through an example.

3 Design Case

Imagine that we work for a company producing software for palmtops. We are asked to design a new portable artifact for digital geographical maps. The company is open to proposals for new interaction devices, and sees this design as breaking new grounds in terms of interaction and application areas. This new artifact is to smoothly blend into the ecology of map artifacts, whether

paper-based, on personal computers or portable devices, and with other artifacts carried by the user, such as a cell-phone or an iPod. At the same time the design is to provide something new and useful compared to its potential competitors. The artifact is to be used while on location, with visual access to the local geography, and intended for local movement, e.g. walking and navigating as a tourist in a city.

In this fictitious set-up, we step into the design process where the use setting and related artifacts have been studied through observations and interviews. We now exemplify how the Human-Artifact Model is used to structure analysis of the map artifacts and reason about design possibilities. The scenario (Figure 8) describes the map use of the two friends Mary and Kate, and is used to consolidate and illustrate some examples from the study of existing map artifacts in use. These existing map technologies are respectively a paper map, a tablet-PC based map and one running on a cell-phone. While still an alternative, the phonebook map is a historical precursor to cell-phone and PC maps.

While fictitious, this design case is based on a recent empirical study that we were part of (Paper II), where geographical maps were studied while used on three different technologies: A paper map in a telephone book, a web based digital map on a tablet-PC, and a digital map on a GPS enabled Nokia smart phone. The three mediators were different but shared some similarities: All three were small enough to be carried around, and could be shared between users and handed from one person to another while in use.

In the study we asked groups each consisting of two users to carry out simple assignments by looking at the map and out of a window. The groups had no particular qualifications regarding maps. Each group was given eight assignments where they e.g. had to locate north or identify a local school. Furthermore, they were to find the distance to and direction of the target. All assignments were created so that the users could orient themselves through buildings and objects that were visible through a large window providing a good view of the neighborhood of our department (Figure 9). The sessions were recorded, transcribed and analyzed in activity theoretical terms. Detailed analyses are presented in Paper II, Paper III, and Paper IV.

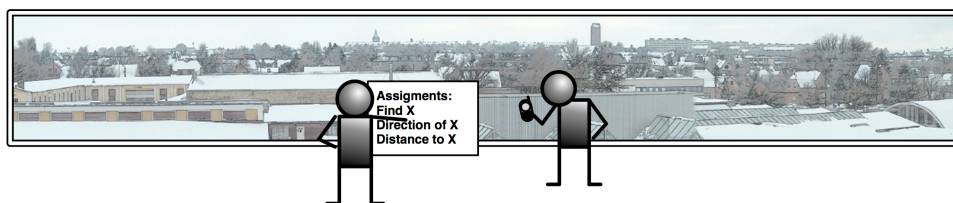


Figure 7 Example setup. Two users are standing in front of a large panorama window, overlooking a local neighbourhood. They are equipped with a set of assignments and a map device, in this picture a cell-phone map.

In the design case we address map use from the expectation that a new

mobile map artifact will be designed, and we focus on the artifact ecology and practices of map use relevant for such mobile map use. As illustrated in the scenario (Figure 7), users' understanding of the general geography of the city and of the local area, learned while riding a bike to school or waiting at a bus stop, are important background experiences. However, we have no systematic studies of such background experience to lean on. In the following sections we use simplified examples from our detailed studies to illustrate the details and dynamics of the Human-Artifact Model in the design case.

Mary and Kate are asked to solve four assignments using a paper map. These includes establishing north, identifying a white villa, and finding it's address. Without further ado they open the paper map on the overview page, and after a short glance at the grid, they turn to the page of the more detailed map. They look up, and have a brief discussion of where, specifically, is north. They are not entirely sure. Kate locates her right hand on their current location (their home), and they look out to find a landmark to locate on the map. A familiar major street crossing at a distance is identified, and Mary puts her left hand on that spot on the map. After some hand-waving and pointing they mentally align the map with the view, even though they do not physically turn the map. They turn to the next assignment, look out to find the villa, and once that is done, Kate, still holding her right hand on home, turns to the map, while Mary keeps looking out. They negotiate the location of the villa, and try laying out the street grid known from the map onto the landscape. Mary, while placing her left hand on the villa spot on the map, compares the direction to what they identified as north, and calls: "North-northeast" without looking at the map. To work out the distance, Kate measures the ruler of the map between two fingers, and moves her fingers as a yardstick to measure the distance on the map.

Mary and Kate now turn to the cell-phone map. One assignment tells them to name and locate a big yellow building in the horizon. Mary picks up the cell-phone, while they both look out. They see the building and discuss which street it may be on. Mary looks at the map and realizes that the school is outside the current map segment. She starts an attempt to use the navigator button of the phone to pan the map. This is slow, and meanwhile the two friends negotiate how they know that street corner from passing it on their bike-rides to school. They point, they move their bodies to illustrate how the turn certain corners, and meanwhile they realize that they need to upload a new map segment on the cell-phone. They turn to the phone, holding it between them, and discuss how they may use the menu structure of the cell-phone to get to the new segment. Even this fails, and they try to answer as many of the questions of distance, direction, etc. without support from the map.

Happily relieved from the small-screen map, they solve a couple of assignments using the tablet-PC. They still struggle with the street address of the school. At first they hope that the school as such will be visible on the map. This is, however, not the case, because the map does not show e.g. public buildings. As they have now remembered the name of the school and hold a computer in their hands, they turn to Google for the address. They get the address, and need to measure the distance to the place. Kate picks up the pen with which the tablet-PC is equipped. Mary identifies an icon that looks like a distance measure. Kate tries to measure the distance by touching the two locations with the pen. However, this means drawing a line between the two locations, rather than getting a distance measure. The two friends are quite confused, and even when they are told that the pen functions as a mouse, and that they have to click at the end-points, they continue to draw lines all over the map.

Figure 8 Scenario describing the map use of the two friends Mary and Kate consolidates and illustrates some examples from the study.

4 Structuring a design-oriented analysis through The Human-Artifact Model

With the Human-Artifact Model as analytical frame we set out to explore mediation in terms of a new artifact that may replace existing map artifacts. We build on the possibilities and problems of existing orienting bases with existing map artifacts. We assume that the new map artifact takes on a role in the artifact ecology that is in some ways similar to existing map artifacts, perhaps in some ways different, and in some ways getting inspiration from entirely different artifacts and use activities. We need to uncover the possibilities and limitations of current artifacts in such an artifact ecology. We address the extent to which such artifacts are or may become instruments for the human user. Structuring the analyses using the Human-Artifact Model is a way of addressing both the horizontal dynamics between the human and the artifact, and the vertical dynamics between the levels. At the same time it is a means to acknowledge that the activity as such may change in parallel with, or as a result of, the introduction of the new artifact.

The levels in the Human-Artifact Model are different analytical foci, and we reduce the complexity of the analysis by systematically shifting the main focus before bringing back together the levels of analysis. In the following we sketch the methodological contents of each step, and the key concerns regarding the artifacts in use. We use the design case to exemplify the kind of analyses that result from this. For each level of the Human-Artifact Model, we focus on both sides of the model: Possibilities in the artifact may be used or not used in the actual repertoire of actions and operations, and actions and operations may be supported or prevented by the artifact, often resulting in workarounds.

The analysis starts with a scrutiny of the current activity. As a vantage point we take the three current kinds of maps. This leads to an analysis of the historical traces of past maps and uses. Finally we look ahead towards the new map artifact.

4.1 Leveled analysis

In the following we present and give examples of what kinds of analyses may be achieved at each level, and how the Human-Artifact Model may be applied to structure the findings in a manner that points towards design. As we have already pointed out the model is based on four different analyses all addressing the entire activity. Hence, using the Human-Artifact Model is not a simple matter of sorting interaction elements into four categories. As pointed out in the introductory examples, when elements play different roles in the analyses, possibilities and problems stand clear.

It is important to note that we use the design case primarily to illustrate conceptual and methodological elements of the design-oriented analysis. Using the case as illustration means that we go into some depth with at least one

finding for each level. However, the case analysis will not be exhaustive or complete in itself.

What? – Goal orientation and the instrumental aspects

We have previously pointed out that the goal orientation is the starting point for further investigation of the motivational level, as well as of the handling level. Therefore the first main focus of the analysis is on the goal orientation and on the instrumental aspects.

Focus

The primary focus of this level is on *what?* What do the users do? What do they say they do? What is the artifact used for? And what may it be used for? The analysis addresses the tension between users' goals in a given activity and expectations towards technology and the assumptions of goals embodied in the artifact, i.e. the instrumental aspects. The analysis places the artifacts of concern in context by focusing on the artifact ecology. Specifically, it is important to understand which artifacts are used together to accomplish the goal.

Method

We use a number of methods at this level, for example in-situ interviews [33], and observations of actual use recorded on video [159]. Taking stock of the artifact ecology through observation and interviews is important, as is an initial understanding of the users' background and related activity. Cultural probes [77] are useful for gathering data for this analysis. The main approach is to use multiple methods to understand what is happening (situated action), and what may happen.

Example

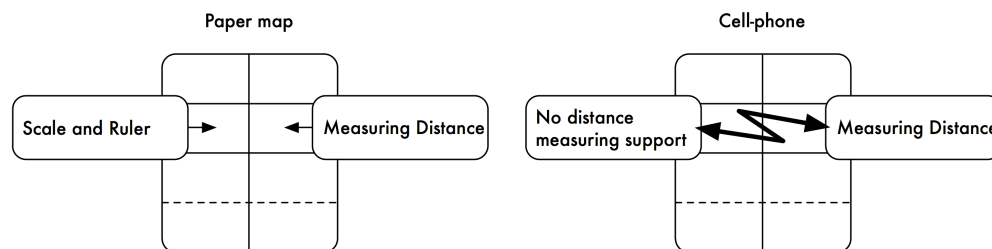


Figure 9 The instrumental level of paper map and cell-phone map. Where the paper map is used without complications at this level, the cell-phone does not support distance measuring as smoothly. Jagged arrow implies mismatch.

In the map case we have identified the goal-orientation as a range of rather clear-cut goals shaped by the assignments given to the participants, namely identifying north, establishing home (where are we?), identifying a target, measuring distance, and finding waypoints and direction. These goals are achieved

by the users through a combination of using the given map device, talking to each other and looking out the window.

All of the three map devices have instrumental aspects that match this goal orientation to some degree. The paper map has a fixed scale with an associated ruler indicating a unit length. Furthermore, on the map are detailed indications of schools, car parks, bus lines etc., as well as an index for looking up streets. The maps on the tablet-pc and cell-phone are slightly less detailed. The tablet-PC map has a dedicated tool for measuring a precise distance, and the cell-phone supports establishing home (i.e. the current location) through its GPS capabilities (Figure 9).

If we take a look at distance measuring, both the tablet-PC and phonebook maps supported the measuring of distance according to the expectations of the participants: On the paper-map a scale was given that could be used to measure distance e.g. with fingers or a ruler. On the tablet-PC distance measuring was externalized into a tool that offered precise measuring of distance (however some complications occurred and will be discussed in the following sections). On the cell-phone no distance-measuring tool was implemented. This conflicted with the expectations of the participants. One group spent considerable time searching the menus of the map application, until they were finally told that such tool was not available. This is an example of a misfit at the instrumental level. The goal of measuring a distance was central to their activity, hence the lacking support on the cell-phone conflicted with the users' experience with similar devices.

The goal of establishing north is interesting because none of the three devices supported it directly. Each of the maps indirectly supported establishing north by following the convention of north being up the uppermost part of the map (the map is "north-up" in navigation terms). However, to establish north, the participants had to rely on juxtaposing landmarks from looking out the window to the map, or on their general feeling of north. Furthermore, they had to convince themselves that the map actually was north up. The participants were often mistaken in their establishing the direction of north, which lead to complications in the later assignments.

Even this "simple" map example illustrates that the artifact ecology and the orienting basis are important for what happens when a new artifact is introduced. Several artifacts are used in combination. The users could not establish north on any of the three maps without an understanding of the geography of the city, and an underlying understanding of north up (or a compass).

Why? – Motivational orientation and motive-related aspects

This level addresses the motivational orientation and the motive-related aspects of the artifact. The motivational orientation, the activity, the *why* question may, as pointed out, not be explicit to users, and as such it may not be addressable directly for example in interviews.

Focus

The focus at this level of analysis is on the (implicit and unspoken) drivers of the activity in question. What motivates the users to perform given actions and how is this supported, or not supported, in the artifact? What motives are supported by the artifact? As we have pointed out the actual use of an artifact is often poly-motivated. Such multiple motives of use need to be explored, along with the multiple motives that the artifact crystallizes and lend itself to.

Method

In-depth interviews may point towards the motivational orientation, as may analyses that contrast conversation with action (see e.g. [39]), and methods that aim at challenging and provoking a current practice such as provotyping [118]. This is because such methods can be used to question unarticulated assumptions. Classical ethnography can often create a good picture of what motivates people in their given activity, whereas artistic expressions of various forms may be used to probe human motivational assumptions about artifacts [69]. This level of analysis is neither addressed by affordance-oriented HCI, Heideggerian HCI, nor by classical cognitive science. We do, however, hope that experience-oriented approaches e.g. that of McCarthy and Wright [115] may eventually lead to methods including this motivational focus.

Example

Kate: The yellow building up there, Mary: So isn't that where you turn? Kate: It is further up? Mary: (sings) Hm, hm, hm, tsk, tsh (..) Mary: Oh what's the name of that street? I ride my bike there every morning Kate: I don't know if we can get much closer Mary: Ahh (sings). I cannot remember at all. The bus stop is called Polytechnics. Kate: Well yes, I pass that too, up that road, it is right after (..) Mary: Well, what is it called? The ice rink is on that street too.	They bend over the phone Mary holds the phone The both look up and on the phone Mary tries to pan Mary is concentrating on the phone, she tries to pan and zoom Mary uses cell-phone menu Kate looks out Kate gestures in the air in front of herself
--	--

Figure 10 Transcription of conversation and action between Kate and Mary.

In the map study we had access to the conversation between users while they carried out the activity, as well as what they actually did.

In the map study participants knew they would be handed a map-device. The motive and purpose were largely those of the assignment. Beyond the particular motivation for solving the assignments, we observed a strong (but unspoken) motivation for creating and aligning a common orienting basis be-

tween the members of the groups. This common orienting basis was established through using the maps, the discussions in the group, the view through the window and references to other activities. All of the groups used anecdotes and stories to make references to other activities where they moved through the landscape (Statements like: “ I ride my bike that way every morning...” (Figure 10) were quite common). The paper map and map on the tablet-PC provided a suitable overview of the city given the assignments, whereas the cell-phone’s limited screen resulted in conflicts between the participants in knowing where and what they were looking at. The paper map supported homing through the participants resting their fingers on the map, once they had established their current location. The small view offered by the cell-phone, which was ill suited for overview, provided an indirect mechanism for homing on a particular slice of the map. Figure 11 illustrates the difference between the quite unproblematic motivational level of the paper map, and the tensions identified in the analysis of the cell-phone.

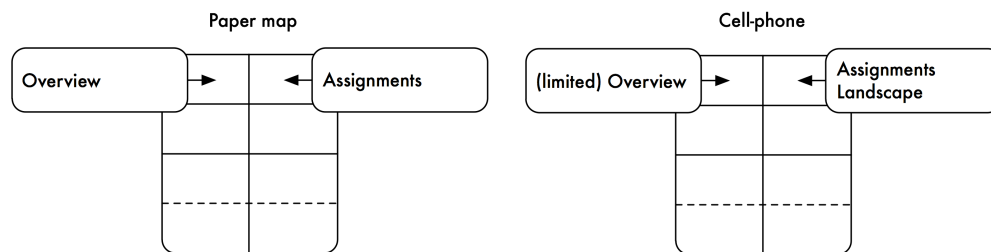


Figure 11 The motivational level illustrates differences between paper and cell-phone

While the main motivation for people participating in the map study was an obligation towards the researchers to carry out the given assignments, maps could also be used in situations where other motivations were present: E.g. a situation where people were lost in a new city and wish to find their way. As a matter of fact it would probably be more “fair” to the cell-phone map if it were analyzed from the perspective of somebody being lost on a street corner in a major city. From the point of view of the design case, the need to create a common orienting basis was more important than solving assignments, and obviously other kinds of motivational issues such as being lost, would be worth considering.

Even though *why* is often tacit and much harder to get to than *what*, the map example illustrates that it is essential to understand both the (multiple) motivational orientation of users, and the motive-related aspects of the artifact. We do not have to go into detail to recognize that how a map is used totally depends on whether the users are solving exercises indoors with a nice overview of a known city, or they are lost, alone, on a rainy night in a narrow alley in a big and foreign city. We proceed with the design case, however, to further scrutinize the handling and the dialectic relationships between levels.

How? – Learned handling

As with motivation, handling is not necessarily articulated. It is essential to study the triggering of operations in the meeting with the actual material conditions of the use situation. Furthermore we address the learned operations that users possess but do not activate in the situation; the quality of operations, and how they are constituted; the crystallization of operations into the handling aspects of the artifact; and the handling aspects that may not be activated in use.

Focus

The focus of this level of analysis is on the tension between the learned routines of the users and the assumptions in the artifact of how it should be handled, i.e. the affordances or action possibilities offered with regards to learned handling.

The immediate analysis focuses on breakdowns [38] as indicators of trouble, when the mediation fails and the artifact draws attention to itself. Breakdowns are necessary for learning, but recurring breakdowns also indicate that there is a mismatch between the artifact and the orientation side as well. Breakdowns regarding the handling indicate whether or not the artifact has become an instrument to its user.

By mapping actions and operations as they unfold, one can address the repertoire of actions and operations further in relation to the action possibilities in the artifact, and the possibility of becoming an instrument: What are the typical and critical operations? What is the quality and constitution of actions and operations (generality, mastering, etc.)? Depending on how experienced users are, there is a further focus on familiarity with the artifact, and on how operations have developed. Furthermore, it is important to identify actions and operations that apply in general across similar artifacts, or independent of the specific, emphasizing the backdrop of the artifact ecology. What are the relationships between multiple artifacts in the ecology? Are they used together in juxtaposition? Chained? As meta-instruments? Etc.

Method

In order to investigate use as it unfolds with particular artifacts we use a version of Focus Shift Analysis [36] adapted to the Human-Artifact Model (example shown at the end of 4.1). This analysis is a starting point for sampling and comparing instances of interaction within and across uses, generalizing patterns based on quantitative analysis (counting the number of instances), and qualitative analysis (focusing on differences between instances) of the data (Paper II). Of interest to the design case, Bouvin et al. (Paper II) proposed patterns of identifying location based on landmarks that were independent of map technology. These were, however, realized through different routines that were dependent of the actual map type.

Example

Each of the map artifacts embodied various assumptions of learned handling. In the design of the paper map it was assumed that the user would be able to connect the grid-based numbering of areas on the map to an adjacent map page, i.e. the paper map afforded grid-based navigability, or navigation. Furthermore, knowledge of how to use the printed yardstick was required in order to perform distance measurements on the map, i.e. yardstick measurability was a possibility with the paper map. All groups typically applied a rather casual transfer from the yardstick to an actual distance, e.g. they were using two fingers to lay out the yardstick length on the total distance. One group was particularly artistic in applying a routine of measuring: The two members measured a known distance using a piece of paper and two fingers as yardstick. This was one example of the kinds of multiplicity of mediation that was noted in the map case. The example also illustrates that the typical action of measuring would not work if distance was critical, simply because it is not precise enough.

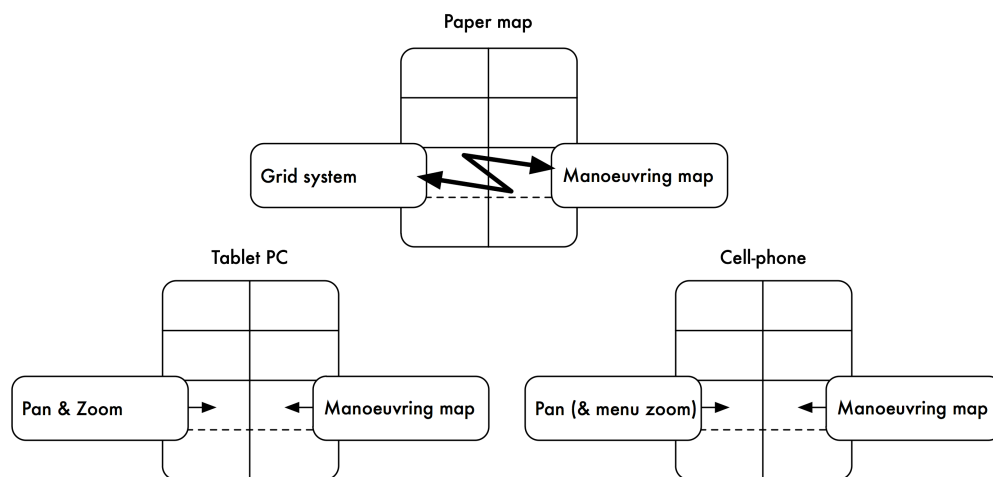


Figure 12 The learned handling level illustrates differences between the three artifacts

We identified how a handling misfit arose from the unfamiliarity of the paper map. This misfit was experienced by a group that did not understand the grid-system for navigation between maps. As a result, the group had a difficult time finding the right section of the map. This was in contrast to most other groups, who identified the relevant map section almost without verbalizing it, and flipped through the map pages while keeping focus on the identification task. The members of the first group had no experience with recognizing the grid structure. Consequently, they had no operations to rely on for getting to the relevant part of the map. There was no adaptive operational conflict (we will return to this concept below), because the group members were well aware of how one would leaf through a book. However, the mere leafing of pages did not support the group members in recognizing the grid structure on the handling level (Figure 12).

Where the paper map applies numbered grids to facilitate orientation on the map and across maps, the cell-phone and the tablet-PC applies panning /- zooming. Panning and zooming on the tablet-PC and the cell-phone was handled in compliance with the general standard interaction for the two devices respectively. Hence participants, experienced with e.g. the PC use, had clearly operationalized its handling in actions general enough to apply for the new map use. Despite this, one interesting problem regarding distance measuring was found: The distance measuring tool assumed a handling that was “mouse-like”; the user had to click at two points between which the distances were to be measured. This conflicted with the participants’ operations for using a pen. Even when told that they should use the pen as a mouse, they tried to draw from point to point instead of clicking. This resulted in panning and misplaced points. The pen was clearly an instrument to most of the participants. However the operations of this instrument were in conflict with the handling aspects of the PC map and as such the ideal of the functional organ is not fully achieved. Or, to use the vocabulary of Beaudouin-Lafon [18], the distance-measuring tool has a low degree of compatibility with the action usually associated with the input device.

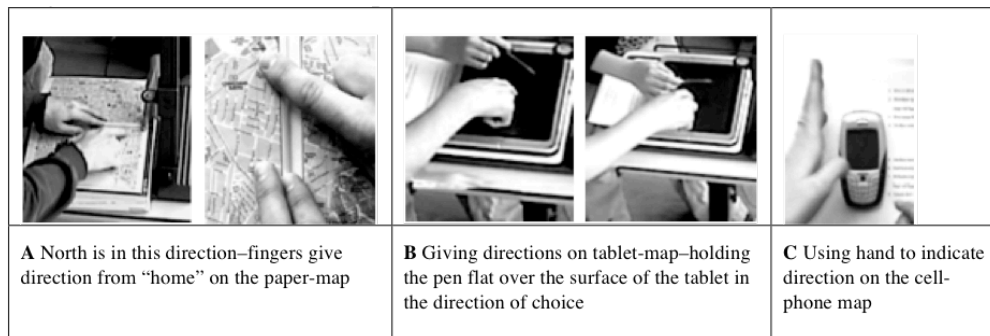
The examples illustrate why it is necessary to look further at the adaptive operational level and across levels. With this kind of setting where new artifacts are introduced into an ecology with several overlapping artifacts, the quality and constitution of actions and operations are important factors for the potential pick-up of the new artifact. The parallel analysis of several artifacts with similar instrumental aspects helps emphasize this focus.

How? – Adaptation

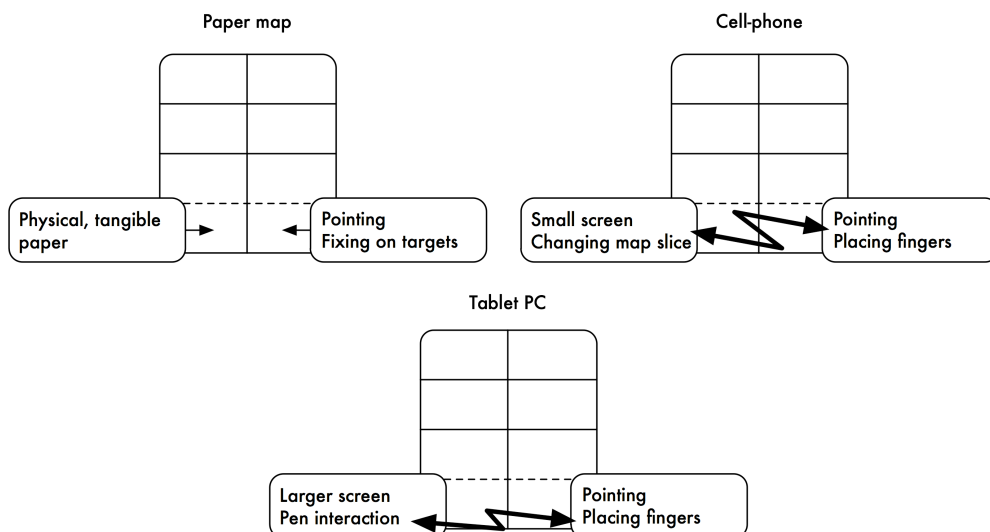
In contrast to the handling level, the adaptive operational level describes human low-level responses to the material conditions of the artifacts and their surroundings. In addition, we focus on the adaptive operations that users possess but cannot or do not use; the quality of operations, in particular how they are mastered; the crystallization of operations into the adaptive aspects of the artifact; and how adaptive operations are activated in use by offering certain affordances. Methodologically, observations are essential as described regarding the learned handling.

Focus

The focus of this level of analysis is on the tension between the adaptive operational routines of the users and the assumptions in the artifact of how it should be handled, i.e. the affordances or action possibilities offered regarding adaptive operational handling. What are the typical and critical operations? What is the quality and constitution of actions and operations (generality, mastering, etc.)? What immediate affordance is offered in terms of e.g. liftability, holdability, turnability (when it comes to maps)?

Example**Figure 13** Direction on the maps

At this level, an important issue is how maps were used when indicating direction, hence co-orienting the two participants. On the paper map it was sufficient to place one finger on home (the users shared current location, as it was commonly known to the participants), and another in the relative direction on the map (Figure 13 A). This pointing was supplemented with the movement of the finger back and forth. This kind of dynamic pointing was also used on the tablet-PC (Figure 13 B), where a strategy unique for the tablet-PC was identified: Holding the pen flat over the surface of the tablet to indicate direction. As with the tablet, hand-waving was the most common way of indicating direction on the cell-phone. This was supplemented with users holding their hand steady and upright, pointing the fingertips as to indicate direction (Figure 13 C). Turning, holding, pointing with hands, and handing over, were operations general to all artifacts.

**Figure 14** The adaptational handling level illustrates differences between the three artifacts

On the paper map, fingers were used, e.g. for marking or remembering a

target. Such marking was an integral part of all the participants' routines of map use. However, on the tablet-PC and cell-phone such a routine was obstructed for different reasons. None of the participants touched the tablet-PC screen in the same manner as they touched the paper map. This may perhaps stem from prior instructions not to touch computer screens. Another explanation is that interacting with the map through a pen led user to perceive that something would happen if they touched the screen, even though the tablet was not touch-based (Figure 14).

On the cell-phone the relatively small screen obstructed the routines of applying fingers on the map. Pointing on the screen simply obscured the map. However the phone was easily handed over between the participants, and turned to match the direction of north.

Gaver [78] and Beaudouin-Lafon [18] have made analyses of the learned handling level regarding multiple instruments existing on one screen, activated through mouse and keyboard. In Beaudouin-Lafon's analysis, the dimensions of movability of a mouse are important, as is the temporal and spatial distance between the action possibility and the objects of attention. The map examples pointed to multiple technologies where the physical and the logical were more mixed, and where holdability, turnability and hand-over-ability were important. Analyses at this level point to problems and explanations arising from the dimensions and physical action range of human beings.

In the above we have presented four analyses, addressing the four levels in the Human-Artifact Model. We have illustrated the dynamics between the human, the orienting basis and the aspects of the artifact at all these levels, and exemplified how the orienting basis gets shaped in relationship with other artifacts in the artifact ecology, and how the artifact similarly affords use with multiple other artifacts. The focus on multiplicity and the artifact ecology is backdrop for the detailed focus on an artifact mediating a particular use offered by the Human-Artifact Model. The Human-Artifact Model also offers the framing of analyses of dynamics between levels. These dynamics will be the concern of the next section.

Bringing the analytic levels together: Cross-field dynamics

Even though the Human-Artifact Model helps separate issues analytically, the levels need to be brought together as well as described above.

Focus

In the above we have exemplified a systematic analysis of each level, and pointed out that there are fits and misfits between the human side and the artifact aspects at each level. With the Wikipedia case we discussed how development in use may make specific parts of the design take on entirely new roles in the artifact aspect hierarchy. We have given examples of how breakdowns shed light on the tensions and movements between levels.

Quality of actions is a concern for this part of the analysis, because this issue cannot be isolated to one level of the human orienting basis. Quality, as addressed through generality of actions, abbreviation by appropriately skipping operations, and mastering of extensive and high-level operations, relate to the entire space of actions and operations available to the user in the use, and need to be addressed in addition to tensions between levels.

In continuation of [18] it is important to systematically address the relationships between the aspects of the artifact. The *degree of indirection* addresses the relationship between handling and instrumental aspects. The *degree of integration* refers to the relation between handling and adaptive aspects. The *degree of compatibility* is a measure for the similarity between adaptive and instrumental aspects. These relationships do not lead directly to general measures. Nonetheless, when it comes to maps we propose that it would be possible to develop measures similar to those of Beaudouin-Lafon: An important issue for compatibility is the turnability of the maps to face north or in the same direction as the view; integration includes issues of homing and finger-placing, and indirection includes the relationship between landmarks in view and on the map, and issues of scale and scalability of the map in relation to the view. These degrees only scratch the surface of the relationships between aspects of the artifact, and it is beyond the scope of this paper to develop them further.

Method

Bødker [44] developed a focus shift analysis aimed at tracing and analyzing the focus of an actual unfolding use when one or more users use an artifact to mediate their relationship with one or more objects of interest. The concept of breakdown was used to understand when and for what reason the artifact as such becomes the object of attention of use: Shifts of focus among objects of activity and between those and the mediators indicate the dynamics of the situation and these focus shifts are the main points of concern in the analysis. The outcome of the focus shift analysis is a trail of foci that are mapped onto the levels of analysis of the Human-Artifact Model. The original work does not include the motivational level, which may, like the operational be unarticulated. The focus shift analysis helps focus on the dynamics of the mediation in unfolding action. It is sequential and does not make the distinction and connection between the human orienting basis and the artifact clear. However, we propose that focus shift analyses can be used to investigate breakdowns leading to the identification of structural elements to be summarized in the Human-Artifact Model (Figure 15).

Example

In Brodersen et al. (Paper III) we illustrate the dynamics across levels with distance measuring as example. We focus on the dedicated tool for measuring distances between two points, or along a series of segments available on the tablet-PC. As described previously, this measuring tool gave very accurate measures of distance in contrast to the routines that users applied on the pa-

		What Instrumental aspects				How Learned handling		How Adaptation	
		School	View	Street	Bus stop	Pan/zoom	Menu	Joy-stick	Hold
Kate: The yellow building up there, Mary: So isn't that where you turn? Kate: It is further up? Mary: (sings) Mm, hm, hm, tsk, tsh (..) Mary: Oh what's the name of that street? I ride my bike there every morning Kate: I don't know if we can get much closer Mary: Ahh (sings). I cannot remember at all. The bus stop is called Polytechnics. Kate: Well yes, I pass that too, up that road, it is right after (..) Mary: Well, what is it called? The ice rink is on that street too.	They bend over the phone Mary holds the phone The both look up and on the phone Mary tries to pan	M&K K M <							

Figure 15 Simplified focus shift analysis. Kate and Mary move in and out of common focus, while Mary struggles with the handling of the cell-phone (transcript from Figure 10). The relevant objects of focus are lined out horizontally, and the unfolding action vertically. The letters M and K are used to indicate who has which focus where and when.

per map. Due to the dynamic map scale, the finger measuring operations of the paper map did not apply well for the tablet-PC. This is because they assumed a fixed scale across the map. As the scale got changed with zooming, the distance-measuring tool provided an immediate and accurate mapping of scale to the map segment in focus, but not one easily measured with fingers. The tool of measuring distance supported what the users were trying to accomplish in terms of motive. It was an externalization of a whole range of actions and operations developed through other artifacts, and, as such, it was a quite complex mediator in itself. However, the tablet-PC map failed to support several groups in measuring distance because of an ambiguity of the pen-like input. The assumption that the pen-shaped device could be recognized as a mouse was problematic, because it forced the users to disassociate the physical shape of the input device from the act of using it. Figure 16 summarizes the artifact ecology of the tablet-PC example, and Figure 17 sketches some of the important dialectical tensions between Human-Artifact Models in the example.

The map study pointed out that the multiple background practices of map

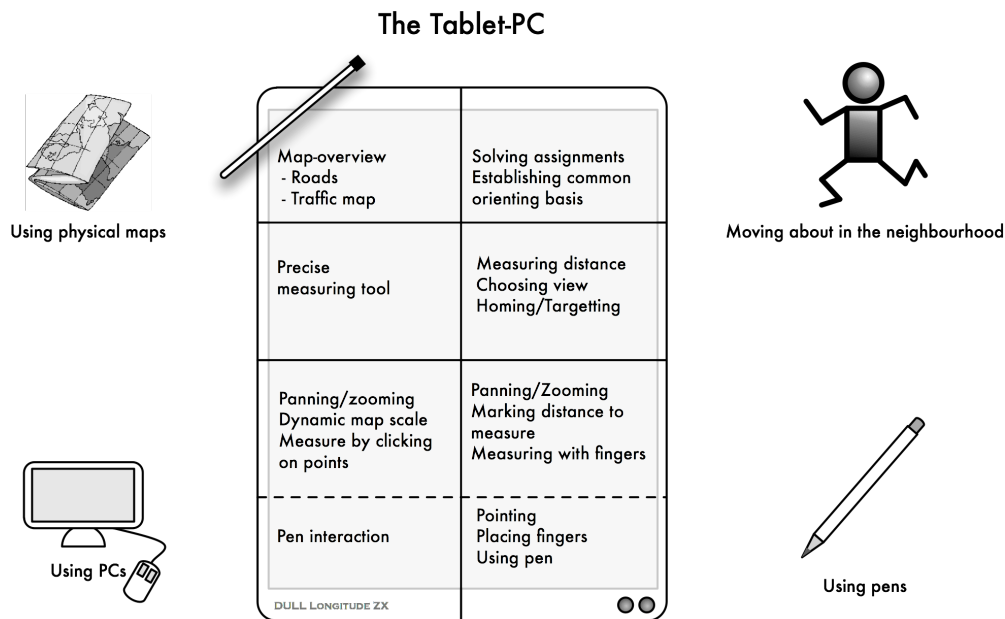


Figure 16 Summary of the artifact ecology and human artifact model of the tablet-PC example.

use needed to be analyzed further. Users understood the local geography that they looked at in different ways, and they had importantly different experiences as cell-phone users. Their understanding of grid-navigation and scale differed significantly, as did their experience as map users in general. These issues illustrate that it is necessary to look for multiplicity of experience and artifact ecology in such analyses. The analyses need to look outside the specific mediation. However, the example also illustrates the difference between the quality of actions when using the three devices. For example, where one group mastered the paper map by leafing back and forth between the correct maps, holding fingers on different pages, improvising ways of measuring distance, etc., even the most skilful users of the cell-phone map ended up in breakdown situations, where they either mastered the map or the cell-phone, but not both (Paper III). In other words, the quality of their actions was neither general, nor mastered enough to skillfully include both map and cell-phone. In this sense, quality is an important concern for design of future mediators, e.g. should a new design address a use that masters both maps and cell-phones, or is it better to assume one and let the new artifact help develop a mobile map use based on this? Which is the most viable to start from?

A design process will need to address these questions, not least if e.g. cell-phone use is to become a necessary prerequisite for the new design, and whether or not it is essential and possible to help users do coarse finger-measuring of distance in some sort of combination with pan and zoom. The latter seems like a design challenge that has not yet been solved.

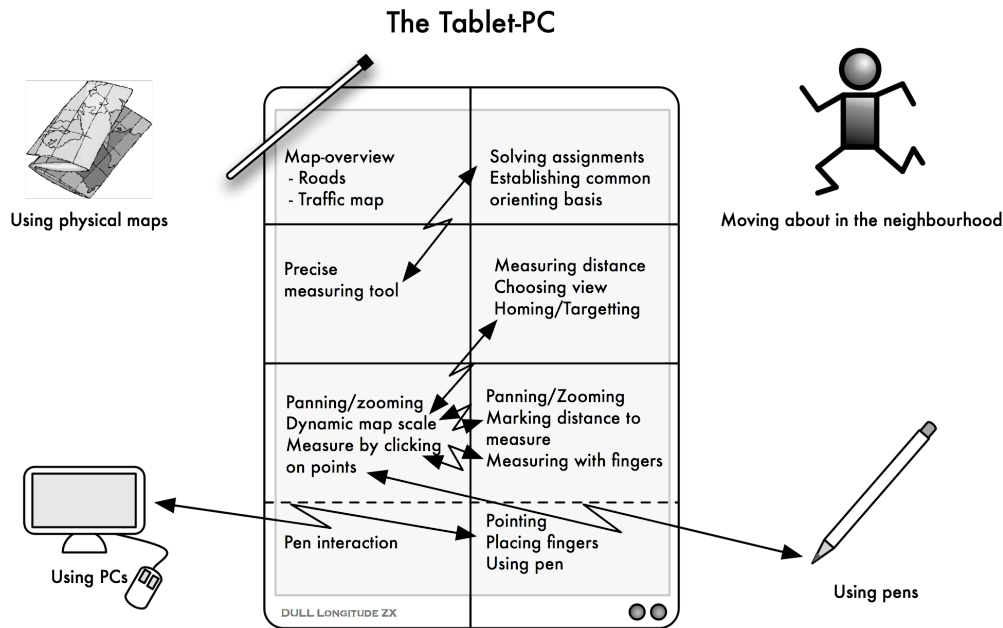


Figure 17 Sketching important tensions in tablet-PC example.

Summary

The questions of *what?*, *why?* and *how?*, with the further specific focus on learned handling and adaptation, respectively help separate out analytical questions and concerns of understanding the levels of artifacts in use, and their potential of becoming functional organs. This leveled analysis is the starting point for applying the Human-Artifact Model to structure empirical and analytical findings, and focus on mediation and multiplicity. Each level has different descriptions and analyses and point to different potentials and problems of artifacts in use, and of functional organs as an ideal and reality. By focusing on cross-field dynamics the model furthermore helps explore the development problems and potentials of the functional organ.

4.2 Tracing explanations looking back on history

As we have seen, there are many different ways in which the history of particular artifacts, and general human development and use of artifacts play into the understanding of current uses and the anticipation of future potential and problems when designing a new artifact. In short, these dynamics need to be addressed to deal with the developmental focus of activity theoretical HCI. Artifacts stem from an ongoing dialectical relationship with previous artifacts and practices. Hence a historical analysis enriches the understanding of observed practice and use of artifacts. Performing a historical analysis can in this context serve two purposes. First, a historical analysis of the goal-oriented activity in question can potentially explain breakdowns and misfits encountered

in the analysis of present interaction. Second, studying the past may serve as valuable inspiration for future design.

A complete historical analysis of the development of a specific goal-oriented activity and the use and development of the related artifacts, can be extremely voluminous and time-consuming. For this reason, we do not give detailed instructions for how to do systematic historical analyses. Instead, we outline approaches that apply either on the back of an envelope, or for more extensive study:

One approach is to identify *historical milestones* in artifacts and practices (as done in Bærentsen's [7] analysis of the development of handguns). This analysis is carried out by asking: How, and with what mediators, has the activity in question been realized before? What are the relationships between those mediators, and how did they leave traces on human practice along the way? How to identify the milestones of the historical development of the activity can differ widely. In some cases, interviews and archaeology of written material [71, 39] may be helpful, while the actual historical artifacts may be available for evaluation in others, such as in the case of maps.

A second approach is *artifact archeology*, which more directly targets tracing where elements of particular use of a particular artifact come from. The starting point of such analysis is typically a misfit in the particular use of the artifact. This may be used to trace on the one hand, where the actions and operations come from (other purposes, or other artifacts)? Their quality and how they have been developed? What in the artifact has triggered that they are applied in a particular situation? What are the resemblances between the particular artifact and those that are the roots of the actions and operations? Bødker [39] focuses in particular on use situations where there are differences or even conflicts between these roots at the four levels of activity.

Example

In the map example it is possible to see the paper map as an older version of the other maps. Here we sketch a simple example of the analysis. We may both find explanations, and identify earlier historical practices to return to with the artifact under design. In the example, we compare distance measuring on the tablet-PC with the paper map as a historical milestone.

From the study of the paper map (Paper II), three observations were clear: Users used fingers resting on the map to indicate home; the gridding mechanism slicing the paper map into smaller sections was neither understood, nor used routinely by many of the users; actual distance measure was often done quite coarsely by measuring out the distance between two fingers, and then moving the fingers to the scale to estimate the ratio.

On the tablet map users never rested their fingers on the map-home. Grid-ding was replaced with panning, and zooming by clicking and dragging. Actual measuring could be done with great precision by mouse clicking on the endpoints of the segment to be measured. The latter was, however, not easily

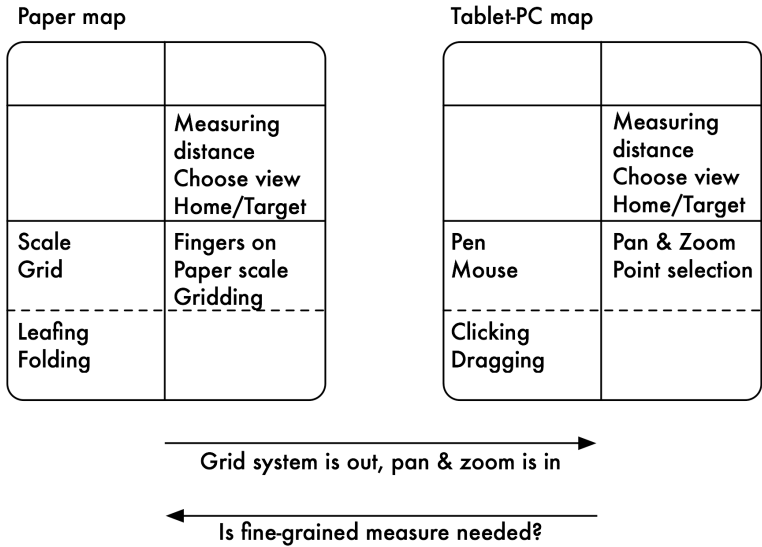


Figure 18 From Paper map to Tablet-PC map

adapted by the users. Barely any of the learned handling or adaptive operational features of the paper map had been considered when designing the tablet-PC map, we suspect. As we have seen, users often had problems activating the distance measure of the tablet-PC. At the same time, for users of the paper map, coarse measure of distance was sufficient, and they did not need anything more fine-grained for this kind of action. Just like in Bærentsen’s [7] discussion of hand weapons, users did not quite identify with the goal of tablet-PC distance measuring, as it was externalized into a tool in the tablet-PC map measure (Figure 18).

Summary

There is no causality stating that because users used the paper map in a particular manner, they now handle the tablet-PC map the way they do; or similarly, should we design a new artifact in a particular manner. The historical analysis of the goal-oriented activity of map use indicated that with this kind of map use, users may have no need for a detailed distance measure that caused breakdowns in the tablet-PC map interaction, at the same time as the detailed measuring was indeed more inaccurate and complicated with the paper map. The choice between the finger-measuring of the paper map and point-and-click measuring of the tablet-PC is probably not the way to proceed in future design. Instead designers need to consider how to make informed choices and come up with something new, given the problems and potentials of both solutions.

4.3 Design decisions

There is no direct way in which the kind of analyses presented here can lead to a design. There is no causality between the future, the present and the past. There are however three ways in which the Human-Artifact Model and the analyses may support design of future artifacts:

1. By applying the functional organ as ideal, the model points to the necessity of designing for all levels.
2. Analyses of current artifacts and the artifact ecology as such help identify from which other mediators to seek inspiration, and on what levels. Similarly, understanding the quality of actions and the construction of the orientating basis can help make design choices.
3. Due to the ongoing dynamics in use, and the lack of causality in anticipation of the use of a future artifact, iteration is essential in design. It is possible to make analyses of prototypes by applying exactly the same foci and methods as proposed here for a finished design. With this perspective on prototyping, the Human-Artifact Model provides a set of filters to focus each prototype, quite in the same manner as described by Lim et al. [112] in their anatomy of prototypes. And the prototypes become support for design exploration, rather than early versions of a final design [112].

We will not go into details with how such design may be carried out systematically and in general, simply because this would lead too far. However, we will return to the design case to summarize the findings of the map study with respect to how these findings may inform design.

Example

Briefly summarized, the findings from the map study as they have been presented here, lead to a number of concerns:

1. The experience of cell-phone handling highly impacts the use of cell-phone maps, yet it is desirable to design for cell-phone savvy as well as cell-phone novices.
2. Zooming/ panning was the most general strategy for map navigation (compared to grids in a phonebook), yet coarse-grained measuring seems important.
3. Landmarks were important for matching map with the actual visual view, yet direction and distance were often dealt with indirectly and implicitly.
4. Users in general want hands/fingers on the map, and have problems with joystick, mouse and pen.

These concerns open for a range of questions for vision-generation on all levels of the Human-Artifact Model.

One vision that focuses on handling in the design case can be formulated as follows: “The map should support continuous zooming and panning without the use of menus”. This vision requires further sub-visions for what adaptive aspects to explore. A motive-related vision for the same map is: “The map should support fast approximation of distances”. Such approximation is supported in historical map use when fingers are used on the scale of a paper map. However, the adaptive aspects of a mobile device prevent this, simply because the size of the screen and the average size of fingers make such measuring impossible. Hence a vision for solving this issue must be developed.

After our map-study was carried out, Apple introduced the iPhone, which has a native implementation of Google Maps. The interface for Google Maps on the iPhone uses pan and zoom through two-finger interaction. Zooming is performed by a pinch gesture with two fingers, and panning by dragging one finger across the screen. Distance measuring is still based on inputting exact locations. Coarse measuring with e.g. two fingers is not supported. One could, however, assume that the iPhone has overcome the problem with users being afraid to rest fingers on the screen, an issue we have not verified empirically.

Aspects of the paper map can be used as triggers of familiarity in the new design. However, the designer should be aware that using e.g. the scale for distance measuring not only seems to trigger familiarity on the instrumental level (something to measure distance with), but potentially also on the lower levels (how to actually carry out the distance measuring). This could lead to breakdowns to be investigated in the next round of analysis of a prototype in use.

The new design must be explored by users through hands-on experience. Designers are facing an iterative process, where the Human-Artifact Model helps structure the findings from prototype evaluation, and consequent new prototypes.

4.4 Summary and perspective of Human-Artifact Model analysis

We have illustrated how the Human-Artifact Model serves to support analyses of specific artifacts in use in particular artifact ecologies. The functional organ serves as an ideal against which to measure the actual mediation. We have further demonstrated how the Human-Artifact Model helps design from existing and past artifacts and helps formulate visions of future artifacts with a concern for all levels of the functional organ. The map example illustrates how the specific insight can be systematically applied to design.

The new insight into mediation, multiplicity and development has been applied to a design case, where the concepts are used to understand the artifact ecology and the actual mediators, as well as to point towards future artifacts. Examples of mediation have been given along with examples of the mechanisms in the artifact and the user background that allow the artifact to disappear from the focal awareness of the user in certain situations and re-appear

in other. It has been illustrated specifically how mediation is a concern for all levels of the activity, from motive to adaptation, and in the dynamics between them. We have addressed multiplicity and the dynamics of artifacts in artifact ecologies. Development has been addressed through the dynamics between the historical artifact ecology, exemplified by the paper map in relation to future action possibilities in a new map artifact.

The Human-Artifact Model is in itself an artifact made for analysis and design. It coexists with other such artifacts, models, methods, etc. in the artifact ecology of HCI analysis and design. It offers itself to be learned, and to be developed in use. Whether or not it becomes a mediator for analysts/designers is as yet an open question. However, it is important to emphasize that the systematic way of addressing levels, aspects versus orientation and dynamics across the fields, is not aimed at helping analysts/designers to put findings in boxes without thinking. Rather, the model is intended to help them stop and reflect as well. The dialectical method and tensions help focus on where, in use and in the model, we find trouble. However, they do not predetermine trouble, and consequently an actual analysis of use is unavoidable.

Recommendations regarding the design case

This paper has been mainly conceptual. As part of its foundation it focuses on the relationships between the past and the future, between analysis and design. We have used a fictitious design case to underline this dimension, and to target the applicability of the model. In the following we briefly summarize our recommendations as regards process and product for such a design case.

In order to design an application for digital tourist maps, we have demonstrated how to analyze the ecology of map artifacts, and the specific alternatives, whether paper-based, or running on personal computers or portable devices. The Human-Artifact Model can be used to map out the ecology focusing on multiplicity. It helps shed analytical light on the four levels of use of specific artifacts, and hence it helps compare artifacts, and understand both general and artifact-specific patterns of actions and operations. We have demonstrated that the motivational level is important to understand artifacts, whether these are maps, corkscrews or wikis.

The Human-Artifact Model helps addressing choices for the future by tracing back actions, operations and aspects to the past. Both general investigations of map-history and specific traces of elements and actions from the past turned out to be of interest in the design case. These do not lead to ways in which the new may causally be derived from the past, but they point to alternatives and choices to be made.

In general the concern for quality of actions has been helpful in singling out which actions and operations to design for, and which to leave out. The Human-Artifact Model points towards the functional organ as ideal. It is necessary to design for all levels, but ultimately this design needs to be evaluated

in use, and hence iterative design is necessary.

Productwise, the example has pointed towards true dilemmas in the design, such as the complications of coarse distance measure with zoomable technologies. Our recommendation to the design case is to pursue exactly such dilemmas.

We propose a prototyping process where the levels of the Human-Artifact Model are used as filters to focus prototypes to be explored in design. We are currently exploring how the Human-Artifact Model may be used to structure also mood-boards, scenarios and personas. These results will be presented in a future paper.

5 Discussion and related work

In our discussion of what we have achieved in this paper, we look at the contribution from three different angles in order to finally address what it takes for the Human-Artifact Model to become a functional organ in interaction analysis and design. The three angles are the current discussions in ubiquitous computing of interchangeable use, seamlessness and context; experience and emotion-based contributions to HCI and the most recent contribution to HCI coming out of activity theory [99]. Furthermore, we end with a discussion of the challenges facing activity theoretical HCI.

Chalmers and Galani [59] have provided a highly inspiring analysis of heterogeneity in theory and design of interactive systems, leading to certain guidelines for design. This reminds us that ambiguity and contradictory information is a resource, not a problem to design and use. They look at heterogeneity and emphasize how this concept also is a positive resource for design. They point out how past activity is a resource for the current and future action. Finally they turn to the issue of limitations, boundaries or seamfulness, as we have mentioned above. They conclude their recommendations by stating that “In the long run, we should consider accommodation and appropriation as a process that designers contribute to by selective revealing system structures, and affordances for their potential use, but it is the users who through their interaction with our system and with each other choose what to use and why” [59, pp. 251]. Chalmers and Galani do not bring these recommendations any further in general. Through its focus on dialectics, mediation, multiplicity and development, the Human-Artifact Model has a contribution to make regarding ambiguity, heterogeneity, history and seamfulness as a resource in design and in development of use. The Human-Artifact Model brings the focus on use and design further than Chalmers and Galani’s guidelines, in particular when it comes to explaining why, and addressing how. Without repeating ourselves too much, the dialectics and materialist basis provides the foundation for understanding these contradictory resources, whereas with Chalmers and Galani’s phenomenological grounding, this is less central. When it comes to how, the four analytic levels of the Human-Artifact Model, its two sides and

the unfolding dynamics of these, make it possible to zoom in on where designers may make their contributions to a potential change of use, and where and how users may adapt these contributions to their ongoing development of use.

As discussed in the introduction, the tensions between seamless and seamless, and between operationalized handling of an artifact and understanding the limitations and possibilities of the artifact, are important to us, as well as to Chalmers and Galani, who takes a direct outset in Dourish [66]. We would like to pick up one element from Dourish that comes out less clearly in the above: In the ubiquitous situation we need to be able to address both the single artifact and the artifact infrastructure or artifact ecology. There is a tendency in discussions of ubiquitous computing to address primarily the singular artifact, leaving the rest to the ill-specified notion of context. [1] is a good example of a paper that has a quite interesting discussion of everyday computing, leading to a number of recommendations, where context is essential. By focusing primarily on the single artifact, the multiplicity and artifact ecology remain reduced to context. Even worse, it seems that anything beyond the direct action level is actually part of the context, including motivation. While context has been in our vocabulary for about as long as post-cognitivist HCI, it has always been quite ill-defined. Unfortunately these recent applications of the notion in Ubicomp have not improved on that. We propose that it is feasible to make the artifact ecology a first-class citizen in HCI, side-by-side with the single artifact. Obviously this will not render the notion of context unnecessary, as there is always a new context beyond a given situation. However, the Human-Artifact Model makes it possible to focus on the question of which elements of the ecology matters for our understanding of the single artifact, and vice versa.

Complex artifact ecologies pose a challenge to our HCI methods because they have often focused on complete replacements of one artifact with another, designed often from one authoritative perspective sengers [152]. When a newly designed artifact has to be designed to co-exist with a range of other interactive artifacts, there is indeed a challenge in addressing this multiplicity from cultural and emotional perspectives, as well as from perspectives of handling and affordance. Such cultural and emotional perspectives have not been in focus to our above development of the Human-Artifact Model, and need to be discussed further here. Zimmerman [175], Stolterman (with various collaborators, e.g. [130]) address human beings' passionate relationships with things. Norman [127] connects emotions to experiences and Boehner et al. [45] tackle the topic of emotion from a social and interactionist point of view, arguing that meaning in emotion is generated through the interaction of people, and that emotion is understood and often modified through interaction with others. Palen and Bødker [136] point to how the multiplicity of experience influences actual interaction with artifacts in emotional situations, and hence how both the direct connection of emotions to things and the entirely social perspective lacks in explanatory power, when it comes to understanding human-computer interaction in emotional situations. They argue that the focus on emotion should not replace that of mediation; rather, a new perspective to include both is what is needed. The levels of activity, we propose, give some

help in analyzing such situations, and the motivational level gives some of the explanatory power to help address passion. The Human-Artifact Model helps see the double nature of artifacts as things that we may be passionate about, and the social relationships between human beings in communities of practice that may create and enforce such passion. McCarthy and Wright [115] apply a pragmatist approach in discussing experience as people's lived lives. Their choice of theoretical anchoring in Dewey and Bakhtin shares foundation with ours. The authors argue that human past experience with technology is a rich and multi-faceted phenomenon, which keeps changing when experiences get recounted to the human being herself and to others. They also point to the fact that these experiences are important in human anticipation of future activity at all activity levels. We have not yet made any attempts of describing such trajectories of technology as experience in the Human-Artifact Model. We would, however, see this as a quite interesting step.

Kaptelinin and Nardi [99] provide a thorough and interesting introduction to activity theory and positions in relation to other recent theoretical and methodological trends in HCI, in particular Actor-Network Theory and distributed cognition. From this, they proceed to discuss how to understand complex human use situations, with emphasis on the social and the psychological. The authors propose that what is unique to activity theory is its focus on development as a phenomenon to be analyzed, and as a tool for analysis and design. Interestingly, despite this, they do not unfold the dialectical perspective very clearly, which makes the study of development difficult, leaning towards a causality that we have previously criticized. By their focus on the social and the psychological, Kaptelinin and Nardi to some extent loose the focus on mediation and artifacts which in our perspective is essential to activity theoretical HCI. Despite their emphasis on the importance of development as phenomenon and as tool, they do not present a very clear design epistemology.

The essence of such a design epistemology is really what it takes for the Human-Artifact Model to become a functional organ in interaction analysis and design. In interesting similarity to most of the references in this section, we have pointed to issues to be addressed, and to methodological elements of analysis and design. When dealing with design for the highly complex, but mono-praxical domain of control rooms, Vicente [166] complains that activity theoretical HCI has failed to give directions for the future artifact. This critique could well be extended to most of the references discussed in this section. More interestingly, however, is to understand whether the Human-Artifact Model changes this picture? Based on the design case of maps we have illustrated that it is possible to give directions for the future map. Through a historical analysis it is possible to choose which specific features of specific aspects of past maps, to add to the new artifact, or to build the new artifact on. An example of this is to choose pan and zoom navigation rather than grid navigation. This example also illustrates that a concern for quality and generality of action repertoires is helpful for design. Beaudouin-Lafon's [18] ways of measuring and qualifying relationships between the levels of aspects for WIMP/post-WIMP interaction is a further place to start. The focus may be generalized, but the dimensions

seem to be quite specific to the technology. In continuation of the quote from Chalmers and Galani [59], we propose that the Human-Artifact Model can help designers decide which structure and affordance to reveal to the users, and when. However, this is only the start. As the design choices made have no simple causalities, it is important to explore the actual consequences of design choices as use unfolds.

Activity theoretical HCI is often criticized for being too complex, and too heavily based on an array of theoretical conceptions that are difficult to understand, unless one really devotes oneself to reading the basic literature. The complexity is empirical, conceptual, and methodological.

Empirically, it is evident that assumptions regarding the importance of the actual use activity, and even more, the multiplicity of interfaces and devices, lead to a more complex empirical situation than assessing a singular device in a toy-situation.

Conceptually, many approaches and lines of thought exist. However, we have worked to reduce this complexity and focused on extracting the essentials as regards understanding of mediation and artifacts. This work has parallels to research on affordances. Affordance is a highly used and much discussed concept in general, and in HCI in particular. While our lines of thought have been inspired by, and are related to the concept of affordance, we have tried to avoid discussions about the nature of affordances. An example of this is whether or not they “are there” without human action, and whether they can be designed or not. By emphasizing the relationship between the artifact and orienting basis throughout, we have tried to avoid discussing one without the other.

Methodologically, it is evident that once one insists that designing the future also involves understanding the past and the present, this adds to the methodological toolbox. However, the understanding of the fact that these elements are needed has been established in post-cognitivist HCI (e.g. [58, 57]), and is not often discussed. Within post-cognitivist HCI itself, however, other frameworks have been proposed, for instance the task-artifact cycle, emphasizing the interplay between the shaping of the artifact on the one hand, and the development of the task on the other. Whereas the task-artifact cycle is in some ways simpler than the Human-Artifact Model, it is more limited in three aspects as well: Whereas the Human-Artifact Model insists on a dialectical relationship between artifact and task, the task-artifact cycle tends to lead to causal explanations; the Human-Artifact Model insists that the human motivation sets a frame for understanding beyond the task; and finally, the leveled approach also gives a useful focus on discrepancies and misfits that eventually lead to a need for change at all levels of the artifact.

6 Conclusion

We have presented the Human-Artifact Model and discussed how the application of the model leads to a more nuanced understanding of mediation, multiplicity and development, and hence interaction in general. The Human-Artifact Model helps address multiple artifacts; how they replace each other or work side by side in changing use settings; and how they may be put together and taken apart in changing configurations. The Human-Artifact Model embraces possibilities and problems at all four levels of activity. The ideal of the functional organ is highly dependent on the multiple background practices of the users, i.e. on other functional organs in the ecology of artifacts surrounding the use activity.

Development and learning are important concerns, even in situations where artifacts are picked up and used without extensive training, so as to avoid simplicity of singular artifacts as the only design goal. Through the Human-Artifact Model it is possible to address, both analytically and design-wise, the purpose of artifacts and their role in the artifact ecology, while maintaining attention to the concern for unanticipated use.

In this article we have thoroughly pointed to the fact that what sets the Human-Artifact Model apart is its insistence on dialectics rather than causality, addressing the tensions between levels and multiple artifacts as resources and potentials and not only as problems.

Acknowledgements

Olav Bertelsen, Nikolaj Gandrup Borchorst, Erik Grönwall, Niels Raabjerg Mathiasen, Marianne Graves Petersen, Pär-Ola Zander for comments. Niels Olof Bouvin, Christina Brodersen, Allan Hansen for cooperation on the map study. Nikolaj Gandrup Borchorst and Dorthe Haagen Nielsen helped improve the language. The UUID project was sponsored by the Danish Technical Research Council.

A Comparative Study of Map Use

Bouvin, N. O. Brodersen, C. Bødker, S. Hansen, A.
Klokmoose, C. N.

Abstract

We present a study comparing the handling of three kinds of maps, each on a physical device: a paper map, a tablet-PC based map, and a cellular phone based one. Six groups of users were asked to locate eight landmarks, looking out a window, and using a particular map. We have begun analyzing video recordings of the situations, and this paper will give examples of the handling of the three kinds of physical devices.

1 The study rational and focus

In an earlier project [47] kids were sent out to do project work in the streets. Some were equipped with annotated maps on a tablet-PC and others with cell-phones. We videotaped this interaction only to realize the difficulty of capturing the detailed handling of the artifacts while the kids were moving about. To better compare map use and handling of the different physical devices we decided to set up the study outlined here.

Published as: Bouvin, N. O., Brodersen, C., Bødker, S., Hansen, A., and Klokmoose, C. N. A comparative study of map use. In *CHI 2006 Extended Abstracts on Human Factors in Computing Systems*, (Montreal, Quebec, Canada, April 22 - 27, 2006). CHI '06. ACM Press, New York, NY, 592-597.

2 The three interaction technologies

In the study, three kinds of maps were used on different devices: a paper map in a book, a digital map on a tablet-PC, and a digital map on a Nokia smart phone running the mobile hypermedia system HyConExplorer [47], developed at the University of Aarhus.

The technologies are characterised by different interaction techniques, but share some similarities: all three devices are relatively small and can be carried around, shared between users, and handed from one person to another while in use. The book format of the paper maps lets the users use their fingers to point to locations on the maps and to quickly “bookmark” related maps on other pages. The tablet-PC supports interaction with the maps through an electromagnetic pen with a single button. This allows users to use the pen much as an analogue pencil but at the same time select objects on the screen using the button like a computer mouse. The smart phone is a one-handed device controlled by a keyboard and a five-way joystick.

The three technologies support roughly the same set of navigation tools. The paper maps support panning from one map to another by a grid index printed at the margins of the pages. Viewing the same area at different scales is accomplished by turning pages. The scale of each map is printed in the corner of the map and can be used to manually measure distances. On the tablet-PC, panning is supported by clicking on a point on the map. Different map scales can be displayed by invoking tools from a small palette, or by selecting a square area on the screen. The application has a dedicated tool for measuring distances between two points or along a series of segments. The HyConExplorer application on the smart phone uses a five-way joystick to control a small pointer on the screen. By clicking with the joystick, the map is panned and centred on the pointer position. Scaling is supported through menu-selectable scales. In the field, these tools can be automatically controlled through GPS [47]. Both the tablet-PC and smart phone retrieve maps from the net, which can cause a delay when panning and zooming.

3 The Experimental Setup

The study was performed with six groups with 2-3 members, all adults. Each group was given eight assignments – e.g., locating north, or identifying a local school and the distance and direction to it – half to be done with one technology, half with another. All assignments relate to objects that are (more or less) visible through a large window (as seen in Figure 2).

The experiments were recorded using the setup shown in figure 1 plus a tripod-mounted digital camera. The top-mounted camera provided a general view of the participants’ posture and gestures, and the table mounted camera provided a close-up of their interactions with the map technology at hand.

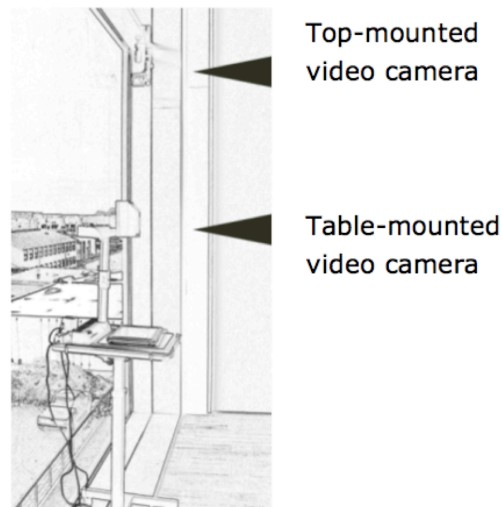


Figure 1 The Experimental Setup



Figure 2 The view with demarcations of the answers to the eight assignments given

4 The Analysis

The video material has been analyzed in three steps, in interaction with the theoretical model (figure 3). First, we logged and transcribed the video, letting the material speak in the manner of interaction analysis [163]. Second, the model below was developed, inspired by activity theory [23], [30] and rooted in the empirical analysis. In the third step, we gathered incidents and compared these across technologies. We have primarily used a qualitative approach, despite the rigid set-up.

4.1 Theoretical Model

All assignments result in two people agreeing on a target, common orientation, distance, etc. (the answer). They possess some knowledge about the landscape that they are looking at and the technology that they are applying (their orienting basis). They create a common orienting basis, mediated by the individual orienting bases as well as by the maps and technologies applied and the view through the window. Based on the assignment, the individual and common orienting bases, and the maps/technology, the groups perform a number of tasks to establish the outcome. These tasks are neither isolated from each other, nor carried out in one particular order. However, the assignments give a main direction of the tasks leading from A to F. Iteration is frequent. Each task is me-

diated by the map, the particular device, the outside view, and the orienting bases of group members, as we shall illustrate.

4.2 Carrying out the analysis

Activity-theoretical HCI provides a number of techniques for carrying out an analysis of use. We used checklists [30] to understand the *context* of the events studied, after which we sampled *instances* of how each task was carried out, and decomposed them further [23]. The tasks were categorized, and further incidents gathered. The orienting basis was traced through speech and bodily references. Fundamental to our understanding of the role of the three technologies are the *breakdowns* that happen.

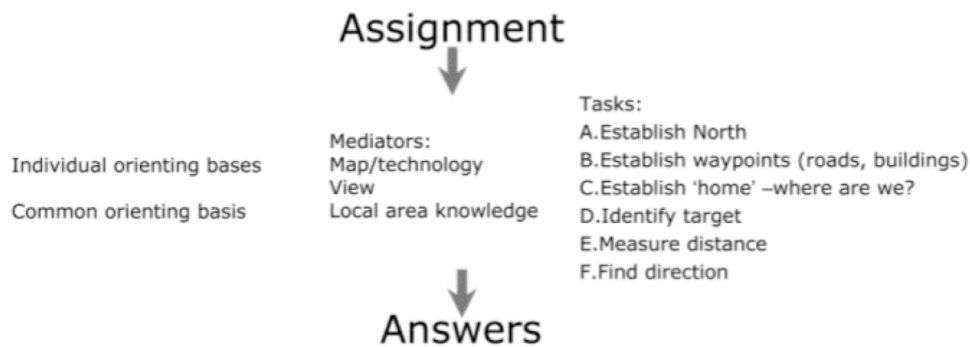


Figure 3 The model

5 Examples

In the following, we will give preliminary examples of the analyses that we are currently carrying out, showing the richness of our material.

5.1 Visible identification of target

This task results in the creation of the initial common orienting basis for the group, with respect to the particular assignment. We see in the video two overall strategies with sub-strategies: A. Both (A0), or one group member can make a visual identification of the target and point it out to the other (A1), or negotiate the location through scanning the landscape, talking out loud and pointing (A2); and B. none of the group members can make an identification of the target and they go through iterations of looking at the map (B1), looking out (B2), and trying to remember from past experience (orienting bases) (B3) until they identify the target and end the iteration in one of the three substrategies. Furthermore, there is a strategy C which is basically a failure to identify the target.

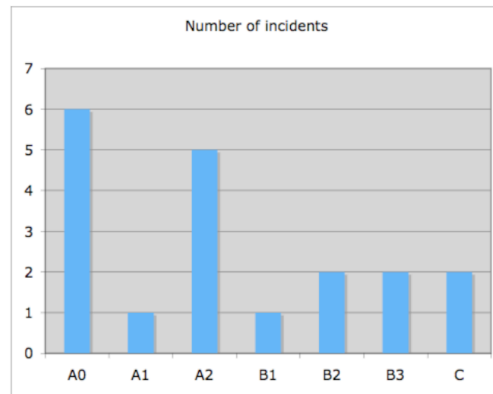


Figure 4 Distribution of incidents

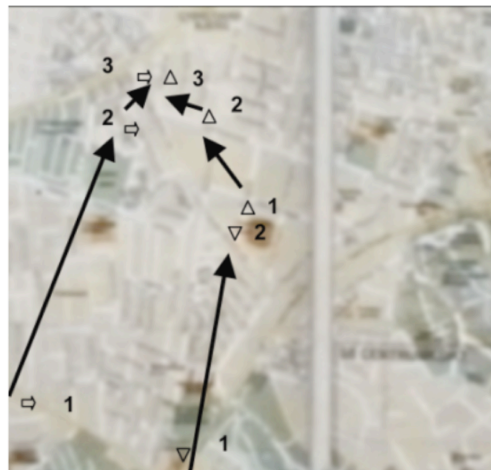


Figure 5 A three-step situation

When counting the number of incidents (figure 4) in each category, we see that only 1 out of 19 incidents leads to the target being identified on the map (B1), in 2 out of 19 (C) there is a failure to identify the target, and in the rest the target is found because the target is known to the group or could easily be identified visibly (“the yellow buildings at a distance”). One assignment focuses on a target that is not easily seen (6—the local school). The B1 group makes the identification by looking at the paper map which shows names of schools and other public buildings. The two groups, who fail (C), do so because neither the cell-phone (group 1) nor the tablet-PC (group 2) gives such support. Several other assignments could have been solved using paper maps in much the same way. Nonetheless, people tend to make a visual identification first. In the future research we will go into further detail with this type of analysis and we aim to get a better understanding of the relationship between what is known by group members, what can be seen, and the maps.

5.2 Using the maps to make detailed identification

The next step in the procedure is from visual identification to identification on the map. We identified, based on the transcript, the elements of the procedures and how they were interconnected. We were particularly concerned with procedures carried out using (or not) the cell-phone and the tablet-PC, respectively.

Two kinds of procedures stand out: A. identification through landmarks, and B. grid-based identification of roads. All participants seem to have a dual purpose of creating a common orienting basis, and of being able to point at something outside, or on the map.

For the cell-phone, we have the four main examples, which each include several instances of both types of navigation: 5 instances of grid-based mapping and 6 landmarks. Landmarks are anchor points in the common orienting basis. In the following example, the group looks only at the map, using the common orienting basis to place the building on the map:

A: This is Skovvejen where I live, and then it is further up Randersvej
 B: but, it is on the corner of Nordre Ringgade, so it is right there//
 A:// Right where we are
 B: So//

Regarding grid-based identification, it is characteristic of the instances we have sampled that the superimposing of the map and street names upon what can be seen is very local and rather depending on people's previous knowledge of the terrain. When discussing which road is which, our users tended to rely on their knowledge of the environment:

E: Then perhaps it is up there
 D: Vågøgade
 C: No it is the one that passes, passes, at it comes from down Defta. What is it called?
 D: There is a road without a name
 C: What is it called? It is called...

These analyses will be supplemented with an analysis of the use of the tablet-PC, in order to focus on the ways in which these technologies support navigation strategies.

5.3 Handling the maps

It is extremely common that both persons rest both hands on the paper map, or hold one or more fingers on particular locations. An example from one group serves to illustrate this: One of the participants (A) rests his right hand on the paper map while identifying where the group is (home) with the right hand. He maintains this position of the hands while looking out the window. The group makes visual identification of the target, and their gazes move back to the map. At the same time, A starts to move his right hand towards the location of the target, while B places his right hand near the target, and A moves his left hand towards home (figure 5 and 6 a). In the third step, A and B both point at the target with their right hands. People do not rest their hands on the tablet-PC. In very few instances a hand or the pen is held over the tablet to e.g., indicate a direction, or the pen and/or finger is used to point on the map (figure 6 b).



Figure 6 Handling (a: paper, b: tablet, c: phone)

There are several, shorter examples where both users have their hands on the phone, e.g., one person holds the phone and both point (figure 6 c). In at least one example, a person tries to activate the joystick while another person is holding. In several instances, the phone gets handed over from one person to another as part of the unfolding action. Holding the phone and resting on the paper map serve some of the same purposes, as a kind of homing. We will analyze this in further detail with particular focus on the lack of hands-on handling of the tablet.

5.4 Finding north and other directions

Throughout the assignments, the participants were asked to identify the direction for their target. Specifically, assignment 1 was an identification of the direction 'north' which we hoped would serve as a good starting point for aligning the view through the window with the view provided by the particular map. We were also curious as to how the different technologies could support their subsequent negotiation about direction, if at all (tasks A and F in the

model). Preliminary analysis strongly suggests that the success of identifying a direction is based on the success of the teams to connect their visual identification of the direction with the identification of the same direction on the map. This is achieved through choosing a visual landmark or a target through the window and comparing this to similar waypoints and/or the knowledge about the map's orientation. More detailed analysis will determine what role the technology plays in this context.

6 Conclusion and lessons for redesign

We have developed a first theoretical model of the participants' use of the different maps and technologies. The participants solved most assignments by first creating a common orienting basis in the model. The common orienting basis is mediated by the individual orienting basis of each person in the group (prior experience and knowledge), the employed technology, the map, and the view (the physical world).

We saw different procedures for going from a visual identification of a target to the identification on the map. Two procedures for aligning the map with the physical view stood out: landmarks and grid-based identification. We will analyze and develop these further: e.g., how we may place waypoints on the electronic maps and what would be the effect of using them; and how we may support the grid-based procedure through design.

The assignments and set-up gave least justice to the cell-phone, because it has its main force when on the move in the field. However, the study gave a substantial amount of immediate feedback to the design of, in particular, our own HyConExplorer.

The paper map affords homing with the hands on the map that neither of the other technologies do. This seems important for both applying local area knowledge in relation to map information, and for relating map and landscape. Especially, when none of the group members could make an immediate identification of a target, homing was used a quick way to iterate between looking at the map, looking at the world, and discussing the correspondence between the two based on the individual's past knowledge and the common orienting basis. The question is then: how can homing be supported on the electronic maps to support these activities?

Ubiquitous Substitution

Christina Brodersen Susanne Bødker
Clemens Nylandsted Klokmose

Abstract

Ubiquitous interaction places the user in the centre of dynamic configurations of technology, where work not necessarily is performed through a single personal computer, but supported by a multiplicity of technologies and physical devices. This paper presents an activity-theoretically based framework for analyzing *ubiquitous substitution*, i.e. a set of mediators that are or can be continuously substituted with the purpose of highlighting expected and indented uses, and the conflicts encountered when attempting substitution between them. The paper develops a four-leveled analysis of such mediators, and point towards a minimalist approach to design of ubiquitous interaction.

1 Introduction

The recent development of computing – whether it is called pervasive, ubiquitous, tangible, or ambient – changes the relationship between the human being and the computer from “one-to-one” to “many-to-many”. This development challenges our understanding of human-computer interaction; where previously the focus of HCI has been on one technology-one application-one user, all packaged into one monolithic unit, ubiquitous interaction is about changing configurations of input and output devices, applications and users.

Published as: Brodersen, C., Bødker, S., and Klokmose, C. N. Ubiquitous Substitution. In *IProceedings of INTERACT 2007.*, (Rio de Janeiro, Brasil, September 10-14, 2007). Springer Verlag.

In classical HCI, there has until now been two ways of understanding and designing to suit users: Either to ignore user competence and design for novices only, as was often done in the first generation of HCI; or to do in-depth analyses of user needs as in the second [13]. In both ways of thinking, however, the idea has been to design a new mediator as a *permanent* substitution of the old.

In first generation HCI, very minimal assumptions about the users were made explicitly. Yet, Buxton [52] derives the underlying model of the human user from an analysis of existing mediators at the motor-operational level. Bannon [12] notes how all mediators contain a theory of the user and the task domain, with reference to “idiot-proof” design philosophies that were implicit in some first generation HCI. With a few exceptions, such as key-stroke level analysis [54], this quite reductionist approach did not provide any methods to help compare mediators and forms of interaction.

In the second generation, the focus has been on context and on understanding current and future use situations to understand the needs of the users. Learning has been important and hence, it was OK for users to spend time learning, and accordingly developing a new repertoire of actions and operations. Bardram and Bertelsen [16] and Bødker and Petersen [42] took early steps towards focusing on learning from one mediator to another, yet the challenges for ubiquitous interaction design are a bit different: The *ubiquitous substitution* that we design for is not a permanent substitution of a full mediator with another (for example a typewriter with a word processor). It is a sort of design that focuses on how to substitute one mediator with another in a way that enriches the human repertoire of actions possibilities, hence, making possible continuous substitutions in terms of switching between mediators. Ubiquitous substitution is not a matter of breaking down and replacing one kind of mediation with another, but providing a larger “toolbox” and a better understanding of which mediator to apply when.

This paper focuses on such ubiquitous substitution. The theoretical account is based on activity theoretical HCI as it has been developed by Bødker [38], Bertelsen [32], Kaptelinin [96], with more recent inspiration from Gibson as developed by Bærentsen and Trettvik [8], and from Bedny et al. [23]. We use the term *mediator* to cover a configuration of physical devices, logical interaction possibilities, functionality, etc., that serves a particular purpose to users [97]. We will return to these theoretical concepts in the following.

To illustrate our theoretical elaborations, we use examples from a recent empirical study (Paper II), where geographical maps were used on three different technologies: A paper map in a telephone book, a web based digital map on a tablet-PC, and a digital map on a GPS enabled Nokia smart phone. The mediators are different but share some similarities: All three devices are relatively small and can be carried around, shared between users, and handed from one person to another while in use. In the study, we asked groups of two users to solve simple assignments looking at the map and out the window. Each group was given eight assignments that they had to solve, e.g., locate north, identify a local school and find the distance and direction to it. All assignments were cre-

ated to relate to buildings and objects that were visible through a large window. The sessions were recorded with two videocameras covering different angles; still-shots were taken with high frequency, and all sessions were transcribed and analyzed in activity theoretical terms [30].

The paper is structured as follows: First, we develop the theoretical framework with emphasis on functional organs. Through an analysis of the use of different map technologies, we demonstrate elements of the framework and discuss how the results can be used to understand substitution between mediators. We conclude by describing the challenges they pose for ubiquitous substitution, and the formation of functional organs in ubiquitous interaction.

2 Human Activity and Functional Organs

Human activity can be analyzed according to three layers: Activity, action, and operation [30]. Activity is directed towards satisfying a need through a material or ideal object and the activity layer focuses on motivation and analytically addresses the question of *why* something takes place. Human activity is carried out through actions. These actions are governed by the conscious goals of the subject, and accordingly the analytical layer focuses of *what* takes place. Goals are different from the motive. They belong to the level where we immediately meet human activity in an analysis, because they are conscious. Goals reflect the objective results of action. Actions are realized through series of operations; each “triggered” by the conditions and structure of the action. They are performed without conscious thinking, but are oriented in the world by what Kaptelinin [96] calls an unconscious orienting basis. According to Bærentsen and Trettvik [8], operations may be cultural-historically developed or naturally evolved, ecologically determined. Accordingly, they may realize internalized cultural-historical patterns of conduct or inborn species-specific patterns of behavior. Thus they result from appropriated use of tools, educated manners towards other human beings or movements in the physical world according to concrete physical conditions. Actions are dynamic structures where only the goal can be assumed to be conscious. Operations are never fixed, but adapted dynamically to the conditions of the environment. Actions are thought as a recursive structure, consisting of both conscious and non-conscious sub-actions and operations.

Bedny et al. [23] introduce the concept of strategies described as plans for goal achievements that are responsive to external and internal conditions. We distinguish between strategies and routines; strategies have to be consciously developed when human beings are to perform an action. Strategies turn into routines when the predisposition to perform a certain action becomes partly automated, or into non-conscious – meaning we become familiar with while performing the action and similar actions. Table 1 is adopted from Bertelsen and Bødker [30], Bærentsen [7] and Bærentsen and Trettvik [8] and shows the relation between the three analytic layers.

Analytic layers	Mental representation	Realises	Level of description	Analytical question
Activity	Motive (need)-not necessarily conscious, but may become conscious	Personality	The social and personal meaning of activity, its relation to motives and needs	Why?
Action	Goal-conscious	Activities (systems of actions organised to achieve goals)	Possible goals, critical goals, particularly relevant sub-goals. Strategies	What?
Operation	Condition of actions (structure of activity)-normally not conscious, only limited possibilities of consciousness	Actions (chains of operations organised by goals and concrete conditions)	The concrete way of executing an action in accordance with the specific conditions surrounding the goal. Routines	How?

Table 1 Human activity and its analytical layers

Activity theory understands human beings as dialectically recreating their own environment Bertelsen and Bødker [30]. Subjects are not merely choosing from possibilities in the environment, but actively creating the environment through activity; “*new tools shape new goals*”, as Kaptelinin [96] puts it. Whereas any specific human activity can be understood as activity, action or operation, none of these are fixed; an action can become an operation through internalization, and an operation becomes an action through conceptualization in breakdown situations [38].

The historical development of activity implies a development of artifacts and environments. Modes of acting within an activity system are historically crystallized into artifacts. In this sense, the historical development of activity can be read from the development of artifacts mediating the practice [7, 13]. Activity is crystallized into artifacts in two ways. Firstly, they are externalizations of operations with earlier artifacts, and secondly, they are representations of modes of acting in the given activity. Kaptelinin [97] captures the functionally integrated, goal-oriented configurations of internalized operations and procedures (routines and strategies), and external mediation in the term *functional organ*. When a mediator is well integrated into the functional organ, it augments the human capacity, whereas if not, the mediator is outside the human user.

This means that when we study mediators, tools and devices, it is important to understand the extent to which they can be integrated into a functional organ. What efforts are needed to provide such integration? How may exploration of the intended use be supported, and how may some of the possible, yet less desirable uses be blocked [8]?

3 Mediation and Functional Organs

Bødker [38] proposes to analyze mediation through three aspects of the mediator: *The physical aspects* are the support for human operations towards the computer application as a physical object; *the handling aspects* support operations towards the computer application. The handling aspects are the conditions that allow the user to focus on the objects of the activity. *The subject/object directed aspects* constitute the conditions for operations directed towards objects or subjects dealt with in or through the artifact.

Bærentsen and Trettvik [8] combine and extend the use of Gibson's [79] affordances with activity theory. They talk about the "*network of technologies and praxis that are the basis of (...) affordances*", and identify three types of affordances to match activity, action and operation in activity theory: Need-related, instrumental, and operational affordances. The need-related affordances relate to what motivates people, the instrumental to the socio-culturally shaped action possibilities in instruments and objects surrounding us. The operational affordances, Gibson's original type of affordances for movement in the four-dimensional physical world, describe action possibilities relating to our "*naturally evolved, ecologically determined patterns of behavior and conditions*." By distinguishing between the adaptive operations that are our low-level response to natural conditions, and conscious operations that are our repertoires of cultural-historical "training", we are able to identify the lowest level in the operation hierarchy, where we are confronted with the operational affordances. Affordance is a concept focusing on the relationship between the human being and the environment. In order to grasp how devices or artifacts are and become part of such relationships, we need concepts that help us focus on the device as well as on human action. Bødker [38] three kinds of aspects match these three types of affordances in a manner that makes it possible to understand how current mediators are integrates in the functional organ of the human users; which possible and desirable uses they afford and how, and which they do not. The relationship between affordances, aspects and activity can be seen in table 2. This mapping requires a re-interpretation of Bødker's aspects and integration with the concept of functional organs. Physical aspects encompass adaptive operational affordances, and we will choose to use this term, since "physical" is slightly misleading. In Bødker's three aspects, the need-related or activity-related aspect is missing, and since Bærentsen and Trettvik convincingly argue that it needs to be there, we include it. We prefer to use the term "instrumental" to subject/object directed aspects, whereas we use handling rather than conscious operational, as it indicates to us a focus on how mediators become transparent in use, rather than on how they have once been conscious. The resulting four analytical layers of a functional organ can be seen in table 3.

By mapping the way the mediator supports *intended* action possibilities and blocks other uses on the one hand, and the routines that the user is equipped with when exploring the mediator on the other (the *expected* action possibilities)

Affordance	Aspect	Activity
Need-related		Activity-why
Instrumental	Subject/object directed aspects	Action/strategy-what
Operational - Conscious - Adaptive	Handling aspects Physical aspects	Operation/procedure-how

Table 2 Mapping between affordance levels, Bødker's aspects and activity

Analytic Layer	Functional Organ	
	Mediator/External	Routines/Internal
Need-related-why?	Need-related aspects	Motivational routines
Instrumental-what?	Instrumental aspects	Actions and strategies
Operational-how?	Operational aspects	Operations and routines
Handling	- Handling aspects	- Learned handling
Adaptive operational	- Adaptive aspects	- Adaption

Table 3 A functional organ mapped to the (revised) aspects of the mediator coupled with the internalized routines on the three layers of activity

on the four analytic layers of the functional organ, we explore and compare mediators in connection with the repertoires of actions and operations connected to them. Hence, we provide a better understanding of how one mediator may substitute another, and how well the substituting device may be integrated into the functional organs of the users (e.g. which effort it takes to do so) (table 4). A well-integrated functional organ is one where there is a fit as regards all four layers, whereas understanding problems of the integration focus on discrepancies at either layer (table 4), or, as we shall see later, across layers. We talk about *breakdowns* whenever the integration stops working, and non-conscious routines become conscious actions [38]. In the following, we will use our map examples to illustrate each layer of analysis of a functional organ.

4 Analyzing Functional Organs

Integration of a mediator into a functional organ clearly played an important role in the success or failure of solving a task on multiple levels of interaction: Groups who were familiar with a particular device had achieved a type of transparency in the interaction which allowed them to focus on the task/map specific issues, rather than on how to make the device perform in a certain way. Similarly, groups who were familiar with maps could focus on e.g. identifying different relevant waypoints rather than on interpreting the map; groups who were familiar with the local area visible through the window, could focus on verifying distance and direction by contrasting multiple visible waypoints and/or information on the map.

Functional organ		Functional organ	
Mediator/external		Mediator/external	Routines/internal
Need-related aspects	↔	Need-related aspects	Motivational routines
Instrumental aspects	↔	Instrumental aspects	Actions and strategies
Operational aspects	↔	Operational aspects	Operation and routines
- Handling aspects		- Handling aspects	- learned handling
- Adaptive aspects	↔	- Adaptive aspects	- adaptation

Table 4 Well-integrated and non-integrated functional organ

In the following, we will discuss the four analytical layers with examples from the map study, and for each level we will give examples of principal conflicts in the integration of mediators in functional organs. Such conflicts illustrate the fits and misfits of the functional organs on and between layers, and point to where misfits may occur when aiming to make ubiquitous substitution happen.

4.1 Need-related Fits and Misfits

Given a specific task or work domain, the need-related layer defines and sets the stage for the analysis. The need-related aspects refer to the question of *why* is there use for a given mediator. Our examples reflect the groups' attempts to solve the given assignments. We note that the need-related possibilities of a map also support e.g. being lost in a new city, looking for the way back to your hotel, where the motivational routines lie outside a very specific assignment.

In the map study examples, all assignments started with two people who had a joint assignment, access to a map-device and to the view out the window. They resulted in the two people agreeing on a target, common orientation, distance, etc. (the answer as it was defined by the assignment). The size of the map, and the way it was shareable affected how the two users would orient each other to the direction outside the window, on the map and in their common orientation.

Need-related misfits can be trivial; does the device match the needs of the user at all? But they can also be subtler; does the given implementation of the functionality needed match the actual needs? In our study, the tablet-pc provides tools for fine-grained distance measuring, but this does not match the need of the users, who were quite satisfied with a more coarse measure.

4.2 Instrumental Fits and Misfits

The instrumental layer relates to the question of what *is*, what *seems* or what *should* be possible to do with a mediator. Understanding these aspects of a mediator is fundamental for appropriating it as a functional organ. When ana-

lyzing ubiquitous substitution, this appropriation gets more complex, as there may be more mediators in play as background experience – mediators more or less appropriated as functional organs. It is necessary to know these individual mediators and understand what is normally possible, or assumed to be possible with them. To ascertain such assumptions, we require the mediator to clearly indicate the extent of its capabilities at the instrumental level.

We exemplify how the combination of functional organs can create instrumental conflicts by focusing on one group using the cell-phone map. During one task, the group members used routines for interaction with the phone developed in everyday use of cell phones (e.g. using menus, joystick). They focused on the task (in this case finding a specific street name) and the response they got from the phone map. The interaction with the cell-phone was fluid and unencumbered, and the understanding of the menu structures, and general information hierarchies associated with cell-phones was clear. However, the group experienced a breakdown despite a high degree of familiarity with the technical device and knowledge of maps in general: The group members expected to find a measuring tool for the map on the phone, and started looking for it by browsing through the menus before the researchers stopped them by telling them that feature did not exist.

Drawing upon the description of a functional organ in table 4, we focus on the relevant layer of analysis (in this case the instrumental) to be able compare the different mediators and highlight any misfits found. The result is seen in table 5.

Mediator:		Instrumental layer
Phone	E	Computational power
	I	
Map	E	Computational distance measuring
	I	No distance measuring available

Table 5 Misfit at the instrumental layer: a breakdown is caused by the expectations of a functionality (*expected action possibilities* E) that is blocked or not supported in the actual, implemented functionality (*intended action possibilities* I)

So, even though both the cell-phone and the map had become functional organs for the group members, the combination of the two had not. This was caused by their expectations of the functionality exceeded what was actually offered by the cell-phone map, given their current need for providing a distance measure (table 5).

To avoid instrumental conflicts, it is necessary to understand the needs catalyzing the interaction: The need-related level frames what we do with the mediators on the instrumental level, and consequently plays an active part in deciding what to support and what to block. If the users had focused on the phone as a communicative device, the computational power might not have been the centre of the users' expectations as was the case in this example.

4.3 Handling Fits and Misfits

The handling and adaptive operational layers both refer to *how* an action is performed through a mediator. The handling layer covers acquired, once-conscious operations. Our examples show a variety of device-type dependent action-possibilities: E.g. leafing over the paper map using the number grid for orientation is an alternative to zooming and panning. None of those are natural. They have to be taught and trained; yet, they both lead to the same set of overall strategies. We have identified a high-level landmark-based identification strategy that is supported on all three devices, while the repertoires of operations used to carry the strategy through are clearly different. On the paper map, landmark-based identification is achieved through homing by placing hands and fingers on the map, indication of direction with hands on or over the map, adjusting the map either north-up or towards the target, and moving between map views by means of the number grid and leafing through the pages. In addition, several visual landmarks, e.g. churches and schools are marked on the map. On the tablet, landmark-based identification is achieved through homing, which in turn is achieved by panning/zooming (albeit not as precise as on the paper map or on the mobile phone), and indication of direction with pen or fingers over the screen. On the cell-phone, landmark-based identification is achieved through homing by means of the GPS-unit, panning/zooming for a wider or more detailed view, adjusting the map by moving the phone and one's own physical orientation, and handing over the device to support sharing. The well-integrated functional organ is, handling-wise, a matter of supporting routines for moving around on the map, panning, zooming, scaling, and for connecting the map to visible landmarks and streets as they appear in the view. With the existing map-devices, there is very little overlap between the necessary routines at the handling layer.

Our example of a handling misfit arised from the unfamiliarity of the paper map, experienced by particularly one group who did not recognize the grid-system for moving through maps, and consequently had a difficult time finding the right section of the map. This is in contrast to most other groups, who identified the relevant map section almost without verbalizing it, and flipped through the map pages while keeping focus on the identification task.

Mediator:		Handling layer	Adaptive operational layer
Paper	E		Leafing pages
	I		
Map	E	No perceived structure	
	I	Grid structure	

Table 6 Misfit with the grid structure: While the group can leaf pages they do not expect a support structure for this, and they do not recognize the grid structure even though it is visible

In this example (table 6), the group members had no experience in recognizing the grid structure, and consequently no routines to rely on for getting to

the part of the map they needed. There is no adaptive operational conflict; the group members were well aware of how one leafs through a book. However, the mere leafing of pages did not support the group members in recognizing the grid structure on the handling level. Thus, the group members experienced a handling breakdown that was overcome by getting clues from bystanders about how to decipher the encoding of the grid. On closer scrutiny, the grid structure provided its own clues on how to proceed, and, as such, the grid structure was not difficult to appropriate.

Handling conflicts originate from a lack of recognizability on an instrumental level; in this case, the inability to understand the grid structure of the map. Taken to the extreme, the lack of recognizability at the handling layer will make users unable even to create a new strategy to help them proceed, because strategies are constructed through fragments of routines based on recognizable elements on the instrumental level [23].

The grid-structure misfit is an example of the importance of making the handling of a mediator explicit, in order to support the conscious appropriation of it, if the structure is unfamiliar to the users. This, however, must be done without compromising or blocking the automated use by those who recognize it.

4.4 Adaptive operational Fits and Misfits

In contrast to the handling layer, the adaptive operational layer describes our low-level responses to the physical conditions of and surrounding a mediator.

At this layer, an important issue is how maps are used when indicating direction, hence co-orienting the two participants. On the paper map it is sufficient to place one finger on home (their current location which is commonly known to the participants), and another in the relative direction on the map (figure 1). This pointing is supplemented with the movement of the finger back and forth.



Figure 1 North is in this direction – fingers give direction from “home” on the paper-map

This kind of dynamic pointing is also used on the tablet (figure 2), where



Figure 2 Giving directions on tablet-map

also a strategy unique for the tablet-PC is shown – holding the pen flat over the surface of the tablet in the direction of choice. As with the tablet, hand-waving is the most common way of indicating direction on the phone. This is supplemented with people holding their hand steady, upright pointing the fingertips in the direction of choice (figure 3). One group turns the cell-phone to align the map with the view in order to indicate direction. At this layer, the well-integrated map-device is a matter of turning, holding hands on, and handing over.

To illustrate the type of adaptive operational conflict, we use an example of how the groups measured distance on the tablet-map (table 3). The tablet-map used a pen-like device that actually functions as a mouse. Even though all users knew how a mouse worked, and were told that the pen worked like a mouse, they ended up in situations where they tried to use the pen as a pen: They drew a line from point to point when wanting to measure distance, rather than what was supported by the map: To click on the end-points.

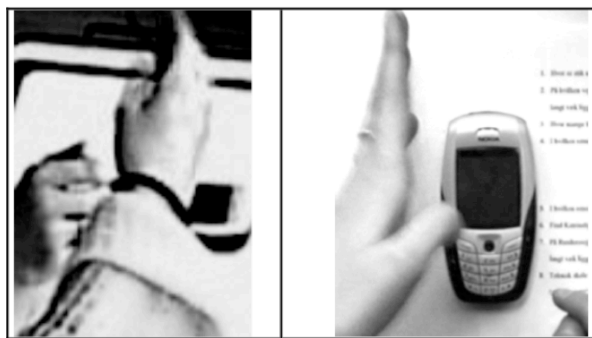


Figure 3 Using hand to indicate direction on the cell-phone map

In this example, the pen suggested action possibilities that were blocked, because the tablet-map only supported measuring distance by clicking at end-points, and not through “drawing” by dragging a line from end-point to end-point.

		Adaptive operational layer
Tablet	E	Drawing with pen
	I	Drawing and clicking with pen
Map	E	Drawing with pen
	I	Clicking with pen

Table 7 Conflict with the pen device

The pen example clearly shows how using the pen as a mouse cannot be taught by instruction, but must be appropriated through use (by changing routines) because these belong to the adaptive operational aspects of the mediator.

Adaptive operational conflicts force the user to change routines, possibly with the consequence of causing a ripple effect, shaking the basic foundations of other functional organs involving the same or similar mediators.

4.5 All together now!

To fully spell out the analytic power of the functional organ, it is essential to bring the layered analysis back together again. The example of distance measuring on the tablet-map shows how connections between layers need to be addressed: The tablet map has a dedicated tool for measuring distances between two points or along a series of segments. This gives a very accurate measure of distance in contrast to the routines that users apply on the paper map, and helps compensate for the dynamic map scale. As the scale changed with zooming, the distance-measuring tool provides an immediate and accurate mapping of scale to the map segment in focus. The measuring distance tool supports what the users are trying to accomplish in terms of needs; it is an externalization of a whole range of routines developed through other mediators and collapsing routines from several different levels, and, as such, it is a quite complex mediator in it self. However, the tablet map fails to support several groups in distance measuring because of the ambiguity of the pen-like input. One of the underlying assumptions of the tablet map is that the pen-shaped device is recognized as a mouse rather than a pen (which, for other tasks are unproblematic). This is a problematic assumption because it forces the users to disassociate the physical shape of the input device from the act of using it.

To target what mechanisms in the mediator that help and prevent development of the functional organs, we turn to Bærentsen and Trettvik [8] question of how exploration of the intended use may be supported, and how some of the possible, yet less desirable uses may be blocked? Evidently, the possible uses are numerous, which is a methodological problem. The best indicator to us in our map study of possible and desirable uses, is what uses the groups make

of the other map devices. Obviously, there are many more, some of which are outside our limited setup. But even within this framing, the possibilities are virtually limitless. In table 8 we summarize examples of the supporting and blocking mechanisms we have seen in the three map-devices.

	Paper-map	Tablet-map	Cell-phone map
Support of intended uses	Homing with fingers Coarse-grained distance measure Exploration by leafing Grid-based navigation	Fine-grained distance measuring Exploration of map through zooming and panning by pointing and clicking	Homing on home by GPS
Blocking of possible uses	Distance measure over large distances	Homing–no fingers on	Homing on target–no fingers on Distance measure

Table 8 Example mechanisms for supporting and blocking

5 Ubiquitous substitution

We now address the challenge of ubiquitous substitution – substitution that does not abandon the old mediator, but supports the ongoing re-substitution of the mediators. There are two motivations for doing so: a desire to perform direct comparison of mediators vis-a-vis a certain purposeful activity, and an analysis of a design space of existing mediators with the focus of building a new one. In terms of our examples, we are interested both in what it take to become fluent users of several existing maps (creating more complex functional organs), and in approaching the design of a new map-device (to be appropriated as a functional organ).

As we have shown, many elements must be in place for creating a well-integrated functional organ: There must be a match between the mediator and the routines at all four layers. It takes little to disturb this harmony – we have seen that a paper map is a well-integrated functional organ to most groups, but that it only takes a breakdown at one layer, e.g. the lack of experience with the grid structure, to disturb this. Even worse, the cell-phone map is neither map, nor PC, which causes a lot of misfits. When functional organs become integrated into a new one that supports ubiquitous substitution, and e.g. lets the user apply both a paper-map and a tablet-map fluently for a task, the case becomes even more complex. Here, it is important that the use of one mediator does not disturb the understanding and routines connected to another. Again, we see this at all layers - learning to use the pen as a mouse – may not only be difficult, it may actually also prevent the future use of the mouse; getting used to scrolling in one manner may be disturbed by the introduction of an application that scrolls differently. In Bødker and Andersen [40], we see an example where historical generations of machine telegraphs are used on a ship bridge, as supplements to one another. Thus, it becomes essential to study the current repertoires of actions, operations and handling aspects as the starting

point for new design. This is not to directly aim to replicate these in the new mediator, but rather to understand what minimal sets of action possibilities are necessary, which to be supported, and which to be blocked.

In a direct evaluation of how the tablet-map as new mediator may substitute the paper-map which is part of an existing functional organ, we found that such substitution becomes a matter of

- fitting into the needs of the user (though not as a direct substitution of one set of motivational elements with another).
- being understandable in terms of what it does or does not offers to the user; a minimal set includes visible scale, visible direction, and means for navigation
- supporting (minimal) sets of routines at the handling level, so that e.g. the necessary strategies can be planned and carried out; at the same time, it may block other action possibilities.
- supporting adaptability (e.g. that all map-devices need to be turn-able to face north-up).

When looking at how navigation is supported on paper versus on tablet – a direct comparison on the handling level – we compare the grid-structure with panning and zooming through pointing and clicking. Those who have been taught the grid structure handle this well, and those, who were not, encounter problems. Panning through pointing and clicking instantly became routine (through instant feedback from the tablet). Thus, as an example of the minimal sets of routines to support, we would go for panning and zooming without making use of the grid structure known from paper maps. Yet, a new map-device would benefit from offering itself more to putting and holding hands on the map, and borrowing visible building-landmarks as they are marked out on the paper map.

In our design-oriented analyses, we see a move of focus away from the handling, which previously in HCI has been the predominant focus (e.g. in GOMS and key-stroke level analyses [54]), and towards the instrumental one on the one hand, and the adaptive operational one on the other. This obviously does not mean that the handling will or should disappear from consideration. If that happened, every user would be faced with a device where every step, except the basic holding and turning, would need to be planned and consciously carried out, i.e. the total breakdown. At the other extreme, handling could be highly standardized, just like car manufactures apply strictly standardized support for handling, which means that anybody can basically jump into – and securely drive – any car. We do not argue for any of those solutions. Instead, we argue that it is essential to identify and support minimal sets of routines at the handling level (e.g. the ones needed for using a map, together with the ones of using a cell-phone), so that the necessary strategies can be planned and carried out; at the same time other action possibilities may be blocked. It is equally

essential that the action possibilities are recognizable as action possibilities in order to be included in strategies that get transformed into new routines through use.

This way of thinking gives indicators as to how to design a new functional organ, e.g. a map-device, yet the actual implementation of e.g. pan and zoom on a new technology depends on the capabilities of the technology in question, and must ultimately be evaluated in use like any other interaction design.

With this paper, we have taken the first step in confronting the challenges of ubiquitous interaction; we have provided a theoretically based foundation for describing and understanding *ubiquitous substitution* by analyzing interaction across different technologies in terms of mediators and functional organs intended and expected use.

Breaking mediators down into four analytical levels has provided us with a strong analytical tool that, through examples from a recent case study, enabled us to analyze interaction, and pinpoint exactly from where trouble stem. This points towards development of new mediators from the perspective of minimalist handling.

6 Discussion

We have chosen to focus entirely on the topic of substitution of mediators, while in many ways ignoring the many kinds of complexity surrounding the use of the particular mediators. We have for instance largely ignored the juxtaposition of mediators [30], i.e. the combination of address register with the paper map, and even the view out the window, or the binoculars that one might use to enhance the view. We have ignored the web of activities and mediators [48] that surrounds the substituting mediators in general, and the way-finding situations in particular. Nevertheless, with the proposed framework, we provide a concrete basis for designing within “webs of technology” [124]. We have not ignored these elements because they are irrelevant, or because the theoretical framework does not deal with them. On the contrary, the framework reflects those exact ideas, and our only reason is to reduce the complexity of the analyses.

The mediator and the functional organ are the centrepieces of our analyses. The mediator is everything, from a simple cursor shape applied for a particular purpose in a design, a scrollbar, the Windows desktop, to other complex arrangements. The framework helps us spot misfits, breakdown in any of these. This is indeed a strength and a weakness. A strength because we escape the separation of hardware from software, which to the best of our belief matters little to the user, while maintaining a comparison across designs. Also, because we have seen that a tiny problem in one element may jeopardize the entire functional organ. The weakness is obviously that we inherit the problem from activity theoretical HCI of limiting the analysis [38]. By taking up the challenge

of ubiquitous interaction, we have, however, gone a long way towards analysis and designs that cross interaction styles.

7 Acknowledgements

Allan Hansen and Niels Olof Bouvin are kindly acknowledged for their contribution to the map study, Olav Bertelsen, Steven Harris and Johan Trettvik for many fruitful theoretical discussions, and Dorte Haagen Nielsen for language improvements.

Quality of Learning in Ubiquitous Interaction

Christina Brodersen Susanne Bødker
Clemens Nylandsted Klokmose

Abstract

Ubiquitous computing places the user in dynamic configurations of technology. As a result, learning in use has new complexities. With this paper, we aim to move beyond understanding and designing one-off interfaces by developing concepts to understand and design for learning in ubiquitous settings based on empirical examples and a foundation in activity theory. Specifically, we point to core concepts: Quality of an action; functional organs; and routines and strategies as being pivotal in analysing and designing for learning in ubiquitous settings. We argue that by focusing on general, high-quality routines we can better support learning in dynamically changing webs of technology.

1 Introduction

Previously, the focus of HCI has been on “one technology – one application – one user”, all packaged into a monolithic unit with a fixed set of input and output devices, and a standardized software platform, from which any task that the user would want to carry out, can be done so. Standardization, integration, and consistency framed design of such user interfaces (see e.g. Shneiderman and Plaisant [154]). With the recent trends of ubiquitous interaction the user is placed in dynamic configurations of technology, where human activity

Published as: Brodersen, C., Bødker, S., and Klokmose, C. N. Quality of Learning in Ubiquitous Interaction. In *Proceedings of ECCE 2007*, (London, UK, 28-31 August, 2007), ACM Press.

is not necessarily performed through a single personal computer application, but supported by a range of technologies that supplement each other and are dynamically changing over time or through mobility. In ubiquitous interaction, the user may conduct a task using a sequence of devices, e.g. starting to work on a palmtop while being on the move; moving the task to a large screen in a meeting room, and later continuing work on a laptop. This kind of multiplicity is well-known from everyday artefacts in general: When a pilot is manoeuvring an aircraft, or the captain a cargo ship, there is not necessarily one single unified interface between the user and whatever is the object of the activity. Instead, one will see one-to-many or many-to-many relationships between the (collaborating) users, the instruments they are applying, and the objects of their activity at various levels. When looking towards learning in ubiquitous interaction, the multiplicity and diversity entail a different approach. When devices and tools are continuously and repeatedly substituted (Paper III), the user should not start the learning process from scratch with each new device. At the same time, the old mantra of *“keep it simple, stupid”* is neither the goal. Users should be able to apply their domain knowledge, their knowledge of use with similar devices, and their understanding of human-computer interaction in general to the device.

When designing ubiquitous interaction it is essential to understand how the multiple devices supplement each other, while making the best possible use of the capacities of each and every device. Accordingly, it is no longer evident that standardisation, integration and consistency are the main design criteria. We urge designers to assume and utilize the competences of the users, and not think their designs in isolation. In this paper, we suggest an understanding of learning in ubiquitous interaction focussing on general high quality routines to support interaction in dynamically changing webs of technology better.

2 Theoretical Background

Learning, in general, is the development of operations, actions and activities (and transitions between them) as they take place in use, or in dedicated learning activities. Along with Bødker and Petersen [42], we focus on the learning that takes place in use of instruments, between users of instruments and in the transfer of experiences that users make between artefacts. In more classical human-computer interaction settings, we would look at the learning that takes place in a well-established community of users, surrounded by manuals, training courses etc. Hence, the learnable instrument may encompass the models, references to other instruments, etc. that are part of the web of activities of the users. This is less so in design for ubiquitous interaction qua its changeability. In Paper III we combine the analysis of computer-mediated human activity ([38], and [29]), with an extended notion of affordance [8]; Kaptelinin’s development of the notion of functional organ [97], and Bedny et al.’s [23] concept of strategies as plans for goal achievement. Table 1 shows the relation between the three levels of activity and the descriptive focus on

strategies and routines (see further in [Paper III](#)).

Analytic layers	Mental representation	Realises	Level of description	Analytical question
Activity	Motive (need)-not necessarily conscious, but may become conscious	Personality	The social and personal meaning of activity, its relation to motives and needs	Why?
Action	Goal-conscious	Activities (systems of actions organised to achieve goals)	Possible goals, critical goals, particularly relevant sub-goals. Strategies	What?
Operation	Condition of actions (structure of activity)-normally not conscious, only limited possibilities of consciousness	Actions (chains of operations organised by goals and concrete conditions)	The concrete way of executing an action in accordance with the specific conditions surrounding the goal. Routines	How?

Table 1 The hierarchy of human activity

Kaptelinin [97] captures the functionally integrated, goal-oriented configurations of internalized operations and procedures (routines and strategies), and external mediation in the term *functional organ*. When a mediator is well integrated into the functional organ it augments the human capacity, whereas if not, the mediator is outside the human user. In [Paper III](#) we conclude that a well-integrated functional organ fits on four analytical layers (see table 2), whereas understanding problems of the integration sheds focus on discrepancies at either layer, or across layers.

Analytic Layer	Functional Organ	
	Mediator/External	Routines/Internal
Need-related-why?	Need-related aspects	Motivational routines
Instrumental-what?	Instrumental aspects	Actions and strategies
Operational-how?	Operational aspects	Operations and routines
Handling	- Handling aspects	- Learned handling
Adaptive operational	- Adaptive aspects	- Adaption

Table 2 Functional organ

We distinguish between *strategies* that are consciously developed to perform an action, and *routines* where the predisposition to perform a certain action is partly automated or non-conscious. Human beings initially deploy strategies. Through execution and repetition of these actions under various material conditions, routines get developed, and the repertoire of action possibilities becomes richer [23, 38]. In ubiquitous interaction, the user cannot rely completely on the specific development of routines for each artefact, as the different con-

figurations of artefacts are virtually endless. Accordingly, we need to focus on the *quality* of the actions and their connected routines. In addition, the adaptation of routines to sound strategies for interaction across artefacts need to be explored further.

2.1 The Quality of Ubiquitous Interaction

Gal'perin [75] judges the quality of an action through the levels of

- generality (i.e. to recognize properties of an object relevant to complete an act),
- abbreviation (i.e. to consciously skip operations in the performance of an act) and
- mastering (i.e. to independently repeat an act with new material) achieved within the action.

Central to supporting transfer of actions and routines between material conditions is the type of plan or orienting basis. Gal'perin [75] distinguishes between three types of orientation:

- Trial and error
- Reusable knowledge within a specific domain
- Domain independent general methods

In the trial-and-error type, the user searches, without help from established routines, between the possible conditions, and learns to distinguish between the correct operations as the different possibilities are tested. Through trial and error, the user learns to perform one specific task (where the setting is stable), but transfer to other tasks in other settings is highly limited because the routine lacks the necessary generality.

Re-usable knowledge within a specific domain is characterized by a much higher degree of guidance of the learning process, and a deconstruction of the problem into separate operations that support the users' ability to analyze and handle the situation. Transfer between tasks within the domain is significant, and relies on the similarity of material conditions, purposes and artefacts. However, routines do not necessarily transfer to domains, and the performance is therefore restricted to the specific domain. Thus, each new domain requires the development of a new orienting basis.

General methods allow the user/learner to develop an orienting basis independently of any specific act or assignment. General methods may be learned through a more complex type of instruction that go beyond the concrete phenomena, at the same time as they allow experimentation with this. Such learning allows the user to gain an understanding of the underlying structure of

the problem (a theoretical generalization), in addition to working with concrete instances of it. This type of orienting basis is very stable under changed conditions, meaning that transfer is virtually unlimited within the given field. Furthermore, the user can direct his attention towards the model for keeping track of the progress of the concrete task, and this supports abbreviation, i.e. consciously condensing the task by “skipping” ahead in the line of operations necessary for achieving the goal.

The ability to conceptualize a routine for performing a task into a strategy for performing another relies heavily on the orienting basis, and *how* the routine has been developed in the first place. If the routine has been developed through trial-and-error, there is less chance that it can be reused to develop sound strategies for other tasks, since fundamental understanding might be missing. In Gal’perin’s terminology, routines can be abbreviated either consciously or unconsciously. Conscious abbreviation is natural and non-harmful whereas unconscious skipping of intermediate results or actions because of lacking knowledge can be extremely harmful.

Bardram and Bertelsen [16] use Gal’perin’s work to focus on learning in use, and Bødker and Petersen [42] extend this focus on the learnable artefact. None of these papers deal specifically with the challenges of ubiquitous interaction. Bardram and Bertelsen [16] introduce three dimensions of the learnable/transparent interface that are developed further in [42]:

- initial familiarity;
- formation of new operations; and
- development in use.

When we enter the area of ubiquitous interaction that is characterized by dynamically changing conditions, it becomes necessary to move beyond trial-and-error. We need to concern ourselves with how we may heighten the quality of the interaction through the support of mastery, abbreviation and generalization in the design of the devices and interfaces used in the materialized stage. Using *initial familiarity*, *formation of new operations*, and *development in use* as key concepts for analysis and design presents us with strong tools for understanding the learnable, ubiquitous interface.

3 Unfolding the learnable interface

We move on to unfold the three dimensions of the learnable interface. In doing so, we use empirical examples to illustrate and discuss our concerns for quality of learning. Hence, they shape the theoretical concepts and discussions, not vice versa. The empirical material pre-cedes these theoretical discussions. We use a study of interaction with three map-devices to explore the dimensions of learning in a ubiquitous setting. The map-study (Table 3) is particularly useful

for this purpose, because it focuses on what happens when people are asked to use several map-devices after each other. Furthermore, the map study helps us address differences and similarities resulting from expertise in map reading and knowledge of the local area. The wastewater example (Table 4) supplements this insight with a focus on design of multiple, ubiquitous interfaces. Wastewater treatment is a highly explorative process, and the section discusses how ubiquitous inter-action may be designed to reflect this.

In [Paper II](#), we presented a study of the use of maps on three different technologies: A paper map in a book, a digital map on a tablet-PC, and a digital map on a Nokia smart phone. The study was performed with six groups with 2-3 members, all adults. Each group was given eight assignments – e.g., locating north, or identifying a local school and the distance and direction to it – half to be done with one technology, half with another. All assignments relate to objects that are (more or less) visible through a large window in our building. The focus for the study was to analyse and understand how users solve similar tasks across heterogeneous devices.

Table 3 Map study

In this study (presented in [[28](#), [30](#), [125](#), [31](#)]), field studies and interventionist approaches were carried out to develop new computer support for the running and optimisation of a wastewater plant. The study was part of a long-term research co-operation in the areas of HCI and CSCW involving Danfoss and the Computer Science Department at Aarhus University, as well as several other partners. The purpose is to explore the theoretical notion and practical design of common information spaces. A prototyping experiment was conducted, focusing on how to get and maintain a local overview of parts of the plant while moving about, and on compiling and interpreting information that is massively distributed on meters and dials [[28](#)].

Table 4 Waste-water study

3.1 Initial Familiarity

With the perspective presented above, initial familiarity is obviously very important. Initial familiarity of an artefact sets the stage for the users' ability to recognize the potential of integrating it as a functional organ, and curtail the level of trial-and-error the users have to resort to in order to appropriate the artefact. Our hypothesis is that this is done by recognition on one or more of the analytic layers of the new artefact, and triggering in the user a set of routines and procedures to match these. Bearing in mind the four aspects of functional organs, initial familiarity includes recognisability of all four analytical layers of activity: Motive and purpose (why and what) as well as handling and adaptation possibilities. The initial familiarity can trigger strategies bound to the

specific domain or from more general knowledge, and formation of new operations can be based upon these. Accordingly, the initial familiarity is both de-pendent on immediate similarity with other artefacts, and on the quality of the actions of the users before they enter the situation.

The map study sets a scene where participants knew that they were handed a map-device, and hence, the motive and purpose were largely given by the assignment. As pointed out in [Paper III](#), the users drew upon routines and knowledge stemming from three sources or domains: The local neighbourhood that they looked at, where some group members walked or biked daily and made reference to; maps, e.g. dealing with scale, road names, direction, etc.; and the device, e.g. leafing pages or using the pen on the tablet.

<i>B: the big red thing that's a church</i> <i>A: ... it is a school</i> <i>B: aaah... there must be a church there</i> <i>A: Let's see. Christian- skirken</i>	<i>B looks out</i> <i>A looks at the map</i> <i>They both search the map</i> <i>A: points at the map</i>	
--	---	---

Table 5 Using a paper-map together

All groups recognized the paper-map immediately. None of the groups showed any hesitation in recognizing they were given a map of the neighbourhood that they were located in, and all groups proceeded to identify their own location (table 5). At this point, however, there were clear differences between the groups. Most of them understood the grid¹ structure common to paper maps, and used it to identify which page on the map to leaf to. At least one group, however, had no understanding of the grid system, and resorted to a trial and error-exploration of the map pages.

In [Paper III](#) we point out how several groups use general strategies for finding a location, independently of grid or zoom. One group was particularly artistic, however, applying a routine of measuring across two devices: They measured a known distance using a piece of paper and two fingers as yard-stick.

When handed the tablet-map, all groups were told that the pen used for interaction functioned like a mouse. This was enough for most of the groups to start using the map, and to pan and zoom to find their own location. However, not all groups found it immediately easy to measure distance on the tablet; when these groups were asked to measure distance, the 'pen-as-mouse'-interaction caused problems: The pen was seen as a pen ([Paper III](#)), and some groups tried to measure distance by drawing, rather than clicking at end-points. At least one group did not manage to use the tablet for measuring distance. This 'false' initial familiarity triggered routines not usable on the tablet.

¹The division of a map into numbered zones using a grid.

Of the three devices, the cell-phone map is the most challenging in terms of understanding how it supports initial familiarity. The map itself is the least useful of the three maps, because it is very limited in size, and is slow in panning/loading new map fragments. For almost all groups this meant that they gave up looking at the map, and relied entirely of their area knowledge and view out the window in answering questions. In one example, the group had been quite good at establishing a mapping between the paper map and the view through the window. However, after switching to the phone-based map, they immediately stopped any attempt to use the map as an aid in finding the direction.

The cell-phone map is interesting because our user groups seem to range from users who already have well-established general procedures as regards cell-phones, to users who have barely held a cell-phone before. The initial use of the phone was therefore based on either an understanding of map-use and only weak understanding of a cell-phone, or understanding of both. During one task, the group members used routines for interaction with the phone developed in everyday use of cell-phones. They focused on the task and the feedback. The interaction with the cell-phone was fluid and unencumbered and the understanding of the menu structures and general information hierarchies associated with cell-phones was clear. The group members expected to find a distance-measuring tool on the phone and started looking for it by browsing through the menus. However, this feature did not exist and, even though both the cell-phone and the map had become functional organs for the group members, the combination of the two caused problems and what seemed to be initial familiarity with the phone-map turned out to be a dead end.

Initial familiarity can derive from multiple aspects of a mediator. In the map study, we saw how the familiarity can be constructive, e.g. most of the groups' unhindered use of the paper map, and most groups' ability to easily pan and zoom on the tablet-PC, however counter-constructive when false. Designing for initial familiarity in the case of the kind of map use addressed here, would require providing a decent overview with the ability for zooming and panning, interaction true to the provided input device, and considerations in the design of what kind of functionality will be expected given the users familiarity with the device.

3.2 Formation of Operations

Understanding and designing for formation of new operations addresses two further elements of use: To support the development of more elaborate routines/operations, and the development of a richer set of alternative routines. Formation of operations also means to form quality actions at more elaborate levels. Hence, generality includes a concern for how the appropriate operations among alternatives get triggered. Abbreviation must be allowed for, or facilitated, in terms of users understanding redundant steps to be skipped, while not being led into skipping significant ones. Mastering independently

with new material relates to the generality of the operations being formed, and to the ways in which experimentation and trial-and-error finds its place in the repertoire of operations – as controlled ways of taking in e.g. new materials, rather than as total break-down based on inappropriate routines.

The 'false' pen caused extensive trial and error interaction by one group, specifically resulting in drawing numerous lines on the map when attempting to measure distance. The group also panned in rather uncontrolled ways while still in the process of measuring the distance. However, at some point, and seemingly at random, they clicked with the pen and got feedback to indicate that they were measuring, and from then on measuring stopped being a problem.

Furthermore, the same group made an early wrong identification of the orientation of the map; north was off by approximately 45 degrees. For a number of assignments, group members used this flawed identification because they trusted their initial determined direction, and quickly abbreviated without checking it again. It was not until their very last assignment, and after having changed map device, that they realized they had made a visible identification that would have to be located behind them, according to their own previous identification of north. While this abbreviation was not directly related to the map, it seemed to be connected to how they used, or rather gave up using, the cell-phone map. In general, all groups 'using' the cell-phone map did not use it unless they absolutely had to: They guessed, tried to remember, looked out the window, and used experiences from previous exercises as much as possible. We see this as a result of the very limited size of the cell-phone map view ([Paper III](#)).

In this setting, almost all of the users had never used maps on tablet-PC and cell-phone before. Their formations of new operations were therefore heavily based upon their degree of initial familiarity. Hence, their problems in forming new operations were mainly connected to problems with initial familiarity. Abbreviation of operations where users used fingers as bookmarks was extensively used on the paper map, and to some degree on the tablet-PC. On the paper map, previous distance measurements were used as the basis for performing new ones. If we were to redesign the map on e.g. the tablet-PC based on the experience from the map study, some implementation of a bookmarking technique would be appropriate, hereby making the users capable of abbreviating repetitive task, such as re-finding home or target.

3.3 Development in Use

Where formation of operations is about the development at the operational layer, development of use is a matter of setting up conditions in the artefact that make users apply them beyond their current (conscious) needs and motivation, i.e. the instrumental and need-related layers. Hence, artefacts that aim to support development in use must be challenging the taken-for-granted motivational routines of the users (while, obviously, still supporting initial fa-

miliarity).

Evidently, the map study contains limited examples of development in use, since the groups use the map-devices only briefly. Nonetheless, we do see what we could call *seeds* [71] of such development in some of the situations. One group recognized the tablet-PC as a PC as well as a map. When they realized that they could not identify buildings on the tablet-map, they started to look for solutions outside the map. They realized that they could get access to a search engine in a web-browser. However, they lacked input-possibilities, and proceeded to look for a soft keyboard. Once that was found, they looked up the address of the school they were looking for, and as they had 'un-locked' new action possibilities, they proceeded to look up more precise answers to several of the assignments they had already completed. Another group, as mentioned earlier, failed to identify a way of measuring distance on the cell-phone. At one point, they decided to investigate if the menu system of the cell-phone would help them, and they explored the phone, rather than the map. The group experienced a breakdown despite their high degree of familiarity with the technical device, and knowledge of maps in general. Their expectations of the functionality simply exceeded what was actually offered by the cell-phone map. Accordingly, a further development of this particular kind of use was prevented due to a clash between functional organs: On one hand, being used to measure distance on the map, and on the other using the computational power of the cell-phone.

Whether successful or not, the kind of development in use found in the map study stemmed from the users combining their knowledge of the map, the assignment and the capabilities of the device to realize new possibilities. And, though the study was limited, it seems clear that a redesign should allow users to move from the map to other functionalities available on the particular device (and as such, these other functionalities should be visible to the users). At the same time, however, it should be possible for users to 'return' to the map, and bridge between the uses and experiences. To avoid the clash of models we saw with the cell-phone, it is an important part of a redesign to aim for a wider, expandable model of the new map-device.

4 Designing the learnable artefact

We will now move on to the wastewater case, to show how concerns for learning across interfaces can shape the design of prototypes. Oversimplified, the process of wastewater treatment deals with the extraction of unwanted elements from sewer water, in order to produce water clean enough for human consumption. Work at a wastewater treatment plant includes an act of balancing elements of mechanical, chemical and biological purification, while keeping down the cost of the purification process. The wastewater operators were constantly optimizing the process to produce cleaner water, increase the throughput of the plant, etc. Furthermore, they were constantly engaged in tasks

to maintain the different components in the plant. In the following, we will identify key elements and routines of the waste-water operators' work, and give examples of how the high-level goals (optimization and trouble-shooting) were translated into the design.

A typical use scenario looks as follows: The wastewater treatment process is managed by the central control and regulation system, located in the control room. However, none of the workers spend more than five or ten consecutive minutes at a time in that control room. They work outside the room, outdoors, or in other buildings of the plant where their daily routines depend on having direct access to the local environment and the components in it. The specific locations at the plant, like the primary clarifier, provide the workers with sensory information not measurable by the automatic sensors, e.g. the way the water look and smell. As part of his daily round, Bob is checking the area around the sludge tanks, and notices that the surface water in the tanks is brownish. He immediately proceeds to check a pump, which he, rightfully, suspects to be malfunctioning. The sensory clues prompt Bob to alter his routine.

Analytic layer	Routines/internal
Need-related-why?	Optimizing, trouble shooting
Instrumental-what?	Sampling automated sensor readings, manual readings, e.g. lab results. Monitoring control system
Operational-how? Handling	Operation and routines - interpreting smell, looks, etc. - adding sensor readings to control system, reading controls - smelling, looking, moving about
Adaptive operational	- interacting with control system and individual components

Table 6 Existing routines to be accommodated

This scenario gave rise to a number of design challenges: Local sensory information along with information from the control system, formed the basis for how work was performed, and in most cases, any one source was not sufficient for deciding how to progress. However, the information obtained from walking around in the area was completely detached from the information from e.g. the sensors, which were only available through the control system. This made the task of getting an overview of the 'state' of the process extremely complicated. Table 6 provides an overview of the tasks and routines connected to the high-level goals of optimisation and trouble-shooting. We proposed two prototypes, on a PC and a PDA respectively, to support optimisation and trouble-shooting by providing a better overview of the purification process at different levels of detail. The intention of the two prototypes is to support users in being out in the plant, walking and using ones senses, while having some access to the control system. At the same time they are able to work in the control room with the extended functionality there.

Figure 1 shows the two topmost levels of detail. First is the plant overview, where the entire purification process and the corresponding sludge and gas

processes are represented in the same abstract form in both prototypes. We chose a simple geometric shape for both platforms, without completely overloading the PDA. At the same time, we wanted to avoid designing for the lowest common denominator by extending the PC-version with a “linking space” to the left, where the user could set up relevant links to other parts of the system, in combination with the control system overview. Second is an overview of a local area, in this case, the inlet, with available sensor data. At this level, focus is on getting an overview of the different components available to the wastewater operator. The circumference of the circle represents the max limit for each sensor, which is easily interpretable, despite the fact that the units for the individual sensors differ significantly.

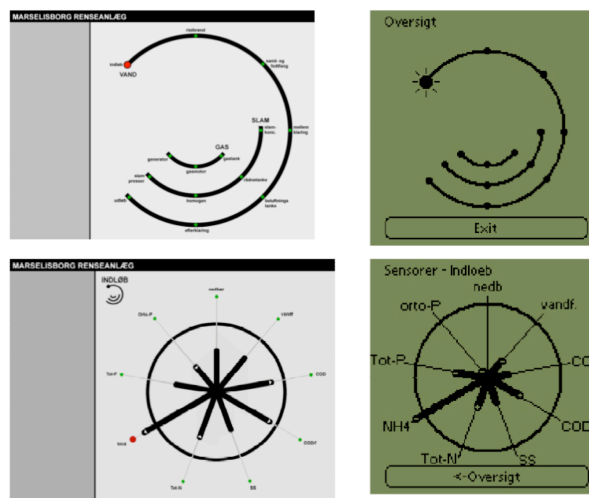


Figure 1 Plant overview. First and second level

The design idea is that using the same graphical representation for a collection of different devices allows the wastewater operators to bring along a ‘window’ into the control system when they leave the control room, and combines the two hitherto independent information spaces without compromising the overall security of the control system. The strong graphical similarity across platforms aims to ensure that the wastewater treatment process is recognized as the overarching structure within which tasks are performed, and information relating to the same parts of the plant is recognizable because of their graphical similarities. These prototypes were evaluated with wastewater operators in workshops as described in Nielsen and Søndergaard [125]. These experiences along with the design as such will form the basis for our exploration.

4.1 Designing for Initial Familiarity

Designing for *initial familiarity* means creating a design that supports and takes advantage of the strategies and routines possessed by the user, with respect to the tasks and domain. We drew upon the wastewater operators’ vast under-

standing of the wastewater treatment process, in the creation of the prototypes. By focussing on over-view of the plant processes, and possibilities of zooming in on information, we made use of the understanding of the wastewater workers in terms of why, what and how.

In the prototypes, we used the terminology known from the current control system to support recognition despite the radically different interface; the elements used to represent key points in the wastewater purification process were taken from their 'old' system, but re-structured to represent an abstract version of the process. Furthermore, using a minimalist approach to building the interface on both devices – the simple geometric shapes, the colour and the sparse textual information – helped the wastewater operators sort between the available properties. However, on the PDA, the lack of textual cues on the first level proved to be *too* minimalistic; the wastewater operators had to rely on their knowledge of the purification process, and what they remembered from the PC-version to find the right information on the PDA. Hence, they quickly reverted into a trial-and-error pattern of use. In a sense, wastewater operators were forced to guess how we had interpreted the wastewater purification process and embodied it in the design.

Designing for initial familiarity, in a case like this, means to take the existing routines seriously. Despite the fact that the graphical user interface was radically different from what the wastewater operators were used to work with in the control system, they had no problems recognising the structure and different elements of the wastewater treatment process in its abstract form. They quickly learned to navigate through the different prototype levels, and asked questions about where to find additional information about key elements in the process. Their feedback strongly indicated that they found this structure to be superior to the existing presentation of the wastewater treatment process with respect to maintaining an overview, as well as being very useful for explaining the purification process to others, for example visiting school children. This latter point is an important indication that they recognised the prototypes as something that would help them beyond their own sampling of data, towards an understanding of waste-water processes that was more general. This is well in line with how the wastewater operators see themselves as part of an ongoing experiment to optimise the running of the wastewater plant [28]. This perspective is discussed further below.

4.2 Designing for the Formation of New Operations

Designing for the *formation of new operations* means creating a design that supports the development of more elaborate routines and operations, as well as a richer set of alternative routines. To achieve this, the design must support optimisation and trouble-shooting in a way that does not prescribe a restricted set of routines bound to the different analytical layers of the functional organ. The work of wastewater operators is characterised by being highly unpredictable, because they often have to respond to unexpected events (trouble-

shooting). Thus, the interaction with the control system and the individual components must be flexible enough to adapt to the changing circumstances, and their repertoire of routines and operations must be general and mastered over a wide set of conditions. However, despite our efforts to support a high degree of flexibility, we realised the structure of the prototypes was flawed; the design allowed users to move one level up or down, but it was not possible to jump across several levels in the information hierarchy. This restriction made it difficult for the wastewater operators to create navigational short-cuts, and thus to consciously skip operations.

This is an example of a lack of possibilities for abbreviation in the prototypes on the instrumental level of the prototypes, which would affect the need-related aspects of the use of the prototypes. It presented a hindrance of the formation of new operations on the two devices, and must, naturally, be supported better in both prototypes. Without it, it is doubtful that the prototypes would become functional organs to the wastewater operators. Thus, designing for the formation of new operations in this context puts focus on the importance of combining an understanding of the work practice and tasks, with an understanding of the technological possibilities because we, to some degree, design for unanticipated use. This example also illustrates how ubiquitous devices are rarely singular. Obviously, we chose to design a family of interfaces, but in doing so we also brought out a focus on the differences and boundaries between the two: The similar graphical layout appeals to the similarities of use between the two, but nonetheless, it is important that the users understand and operationalize the differences, so that they have a set of routines for both. A set of routines that gets appropriately applied when facing the “right” interface, and that includes the operationalized choice between the two. We aimed to support “telling the difference” through the different degree of sophistication, in terms of use of colour and resolution, and through the different screen use. We have no real evidence to show if this worked in the long run, but we find this an interesting design challenge when designing multiple interfaces in general.

4.3 Designing for Development in Use

We have already pointed out that wastewater treatment is by its very nature highly experimental and unpredictable [28]. To design ubiquitous artefacts that support *development in use* seems well in line with this unpredictability, and focuses on ways of challenging the taken-for-granted motivational routines of the users while obviously still supporting the initial familiarity at all levels of activity.

In the PC prototype, the ‘linking space’ encourages users to mix information from other parts of the system as the need arises. The wastewater operators can build their own information space, in relation to the view provided of the wastewater treatment process. The PDA version does not have this linking space access, because the tasks it was to be used for are highly local, dealing pri-

marily with maintenance and trouble-shooting at specific components. With the prototype design we tried to present an explicit model of the wastewater treatment process at the same time as we allowed the wastewater operators to work on specific components. They would maintain the overview of the process and the individual components. The similarities across the user interfaces were to support an orienting basis that was general enough to include both interfaces, and a high degree of transfer of routines between the two.

The focus on development in use and challenging of existing routines made us aware of more areas not originally thought of as relevant for the prototypes, and hence were not supported. For example, it should be possible on both devices to add and remove sensors from the different points of interest in the process. This would allow users to re-use knowledge about, and repertoires of, operations on sensors while applying this understanding to new conditions.

The example of how wastewater operators were imagining using versions of the new design with visiting school children, represents the seeds of an entirely new type of use: Visualizing the processes of wastewater in a pedagogical setting, while moving about the plant.

Since optimization and trouble-shooting are such essential parts of wastewater treatment, it is particularly challenging to design instruments that allows for experimentation, and unanticipated use as well. The overall design that runs across the two sorts of devices has all the way through been focusing on generality of actions, possibilities of formation of new operations, while emphasizing the mastering across material conditions and across the two interfaces. The simple visual design was an important part of emphasizing learning and exploration across interfaces. Unanticipated by designers, it gave rise to possibilities of adding and removing e.g. sensors in a manner that supported transfer of handling and understanding across plant components. How such adding (and removal) of sensors would become a part of the everyday handling of the two overview-devices is a question that could well be explored as an issue of tailoring or end-user programming. Addressing end-user programming from the learning perspective presented here is a further challenge for design.

4.4 The learnable artefact in summary

In this paper, we have addressed the quality of learning in the context of ubiquitous interaction. Ubiquitous interaction differs from classical one user-one device-one application interaction, as being interaction with multiple, and changing configurations of devices and settings. Hence, long-term learning processes with a single device in isolation cannot be assumed. Learning should transcend the isolated device, and rely on the use and development of high-level high-quality routines. High-level high-quality routines are routines applicable across devices, and perhaps even across domains. Though imprecise as measuring tool, the measuring of distance between two fingers, or using a sheet of paper to mark a ruler, are examples of this.

Central for designing learnable artefacts in the context of ubiquitous interaction, is supporting of transfer of use between artefacts – not designing for use of the artefact in isolation. In this paper we have argued how it is important to support reuse of high-level routines, rather than relying on trial-and-error alone.

Trial-and-error however is a natural part of any learning process. It would be naïve to assume that trial-and-error could be completely avoided. What is import though, is that either domain-specific routines or general routines frame the trial-and-error. Blind trial-and-error occurs either when the user has no clue of how to approach the interaction (they may not understand what the artefact serves for or why), or when the general routines, or the domain specific routines drawn from experience, are not applicable to the specific device. Central for the learnable artefact is that trial-and-error is limited to the low-level adaptation of a high-level routine to a specific artefact; e.g. being able to adapt the routines of navigating to a map, or the knowledge of the structure of a wastewater plant to a new artefact, without breaking with these routines and resorting to blind trial-and-error.

One of the challenges in designing the learnable artefact, is making the users initial familiarity trigger the appropriate high-level routines, and avoiding to trigger 'false' familiarity as e.g. in the case of expected but lacking distance measuring on the cell-phone. The learnable artefact should further exude its capabilities, so a user in a ubiquitous interaction setting will be able to chose what device to use when, and for what.

Just like trial and error is a natural part of learning, so is the focus and understanding of the domain. While we argue for simple and general artefacts, we would not get far in modern society if we only designed for totally generic artefacts. Central to the learnable artefact is that it takes its starting point in a purpose, a motivation and the reusable knowledge and routines within the domain. This sets the stage for trial-and-error exploration as well as for identifying the possible general routines to be supported in design. At the same time ubiquitous interaction takes place in dynamically changing webs of technology and activities. These webs means that domains of knowledge and understanding in the ubiquitous interface design stem from many overlapping arenas of life – from wastewater optimization, to cell-phone use – which can be 'activated' trough the appropriate design choices. Accordingly, it is essential to design for sustainability and adaptability of routines beyond the specific use situation. We propose that it is in this space that we look for quality of actions.

VIGO: Instrumental Interaction in Multi-Surface Environments

Clemens Nylandsted Klokmose Michel Beaudouin-Lafon

Abstract

This paper addresses interaction in multi-surface environments and questions whether the current application-centric approaches to user interfaces are adequate in this context, and presents an alternative approach based on instrumental interaction. The paper presents the VIGO (Views, Instruments, Governors and Objects) architecture and describes a prototype implementation. It then illustrates how to apply VIGO to support distributed interaction. Finally, it demonstrates how a classical Ubicomp interaction technique, Pick-and-Drop, can be easily implemented using VIGO.

1 Introduction

In his seminal paper on ubiquitous computing [170], Mark Weiser envisioned how computers would take on multiple sizes, from the small tab to the notebook-sized pad and the large interactive wall. These devices would be used interchangeably and in combination: pads as “sheets” of paper and tabs as, e.g., input devices for an interactive wall. Part of this vision has been realized today: we now have myriads of small devices, similar in size to the envisioned

Published as: Klokmose, C. N. and Beaudouin-Lafon, M. 2009. VIGO: instrumental interaction in multi-surface environments. In *Proceedings of the 27th international Conference on Human Factors in Computing Systems* (Boston, MA, USA, April 04 - 09, 2009). CHI '09. ACM, New York, NY, 869-878.

tabs and pads, and interactive walls are on their way to becoming consumer level products. But the seamless interplay between the multiple device surfaces that Weiser imagined is still far from reality: the simple act of exchanging data among devices typically requires complex configuration or the use of a physical USB key rather than Rekimoto's simple and intuitive *pick-and-drop* [144]. Other examples of multi-surface interactions include the use of a PDA as a personal and portable tool palette [145] or as a remote control [122] when interacting with an interactive whiteboard. While such techniques have been prototyped in the lab, they are still not available to the users at large. We believe this is due to the lack of adequate software support to develop such interactions.

In this paper we specifically address *multi-surface interaction*, i.e. interaction spanning the surfaces of multiple devices. We question the adequacy of the current predominant user interface paradigm, the application-based WIMP interaction model, and its underlying architectural models such as MVC [143] for building user interfaces going beyond a single desktop computer. We examine the requirements for a user interface software architecture that supports multi-surface interaction. We argue that instrumental interaction [18] provides an appropriate framework for interaction in multi-surface environments and introduce *ubiquitous instrumental interaction*. We then present VIGO (Views, Instruments, Governors and Objects), an architecture that supports ubiquitous instrumental interaction and show how it is used to create a generalized version of pick-and-drop.

2 Ubiquitous Instrumental Interaction

The vision of ubiquitous computing includes the idea that multiple users can interact with multiple devices through a variety of interfaces, including interfaces spanning multiple surfaces. This vision challenges the traditional assumption of one surface / one interface / one application that is very deeply engrained in today's desktop computing environments. This assumption is also pervasive in the software tools used today for creating interfaces – tools which are tightly bound to the platform hosting them and to the WIMP interaction style. These tools typically do not support the multiplicity, dynamism, heterogeneity and distribution that characterize the ideal of multi-surface interfaces, making it difficult in particular to develop multi-surface interfaces.

Two central goals in creating user interfaces for multi-surface environments are to provide users with distributed interfaces that support fluid interaction across stationary and mobile devices and the ability to dynamically configure those interfaces according to the available devices and users' needs. Two major challenges in this context are: Supporting reuse and the quality of learning across different devices (Paper IV) (User perspective), and technically supporting the continuity and distribution of work across multiple devices (Developer perspective). We argue that one approach to address these problems is to *deconstruct* applications and distribute the interface across multiple surfaces rather

than simply create scaled-down versions of PC applications to run on, e.g., Personal Digital Assistants (PDAs).

Beaudouin-Lafon [18] introduced *instrumental interaction* to model WIMP and post-WIMP interfaces on desktop computers. The key idea is a conceptual separation between instruments and domain objects. Instruments consist of a physical part (the input devices) and a logical part (the software component). Instruments are *mediators* [38] between the user and domain objects: The user acts on the instrument, which transforms the user's actions into commands that affect the relevant domain objects and provides feedback to the instrument and the user.

Instrumental interaction was inspired by the way we use physical tools: A painter can freely add or remove brushes from his collection, pass them around, etc.; Brushes are not bound to painting on a canvas, they can also be used to paint on the wall or on the hand of the painter. Computer applications do not currently support this level of flexibility: a brush in a drawing application can rarely be removed and used in another context, such as a text editor. Applications typically have a predefined set of tools that is difficult or impossible to adapt to one's needs or taste. This lack of flexibility limits the mobility, distribution and customizability of interfaces. It also typically results in complex monolithic applications, built for general-purpose personal computers, with dozens or even hundreds of tools to cover all possible needs.

While instrumental interaction was developed in the context of desktop interfaces, the concepts are more general and apply particularly well to multi-surface interaction. We call *Ubiquitous Instrumental Interaction* our extension of instrumental interaction to this context. In Ubiquitous Instrumental Interaction, the instrument is an explicit construct rather than a descriptive concept. Ubiquitous instruments should be exchangeable among users, they should work in a similar way on different devices, sometimes across multiple devices, and they should be technically separated from the objects they operate on. Instruments should be applicable to domain objects and on device surfaces when and where it makes sense, even if they were not designed to do so in the first place. The objects that users interact with through instruments should be able to migrate across device surfaces, support multiple views, and be manipulatable by an instrument in ways not necessarily anticipated by the object. Note that this does not preclude an instrument from "breaking" an object, i.e. to make it unusable by other instruments¹. Finally, instruments should themselves be objects and therefore be manipulatable by other instruments.

While we recognize the need for specific instruments that work only with certain objects on certain surfaces for performing specialized operations, we also want to support the ability for instruments to be more open-ended and usable in ways not anticipated by their designers. For example, pick-and-drop is a very flexible and generic instrument capable of picking up an object of any type on any of the surfaces available to the user and dropping it onto another

¹This may be useful, for example in the board game that we describe later, where a player may want to create art with the pieces rather than play the game.

object on another surface. A color picker is another example of a generic instrument that can be used in many contexts. It typically works with a color palette that displays the set of available colors, but can also be generalized to pick the color of any object with a color attribute. A last example is an annotation instrument that can add annotations to any object, e.g. by supporting hand drawing on any surface or adding electronic Post-it Notes to any object. Such flexibility supports what Illich [89] calls *convivial use*: “Tools foster *conviviality* to the extent to which they can easily be used, by anybody, as often or as seldom as desired, for the accomplishment of a purpose chosen by the user. The use of such tools by one person does not restrain another from using them equally.”

2.1 Architectural Requirements

Implementing Ubiquitous Instrumental Interaction requires a software architecture that enables the flexibility that we have described above in the context of a distributed infrastructure. We believe that a software architecture based on small-grain components that can be reconfigured according to the users’ needs or the available devices is the most appropriate. We identify two main requirements for this architecture:

Decoupling: Objects and instruments should be separate components that communicate through a simple protocol allowing instruments to query and modify objects. For example, any object that provides a 2D surface can be used by a pen instrument to add annotations. Such decoupling will facilitate the distribution and replication of objects, or parts of objects, across multiple devices.

Integration: Despite the fact that interaction may involve multiple surfaces, multiple processes and multiple machines, the system should appear as a single consistent entity from the user perspective. The ability to use the same instrument with objects of different types should be seamless and in general the system should support a seamless user experience.

3 Related work

While most research on multi-device interaction has focused on *migrating* applications across devices, especially through model-based approaches [120], little research has addressed true multi-surface interaction, i.e. interactions that involve multiple surfaces. Notable exceptions include the Pebbles project [122] and Demeure et al.’s work on distributed user interfaces [64].

Some systems have attempted to provide a generic solution to interact with applications in a multi-device computing environment. XWeb [132] and the SpeakEasy Recombinant Computing Framework [123] are probably the closest to our work in that respect. XWeb decouples internet services from their user interface to facilitate access from multiple kinds of devices by introducing device-independent widgets that specify the possible values of a service’s data

items, e.g., a date or time. In SpeakEasy services can provide their own user interfaces to be aggregated on the client. The goal of XWeb and SpeakEasy however is to automatically transform user interfaces for different devices, while we focus instead on a uniform interaction model to create custom interfaces that leverage the capabilities of the environment.

Other systems, such as the iStuff toolkit [9], are designed to explore multi-device interaction but do not embody a specific interaction model. Cameleon-RT [11] on the other hand is a reference model rather than an implementation framework. It focuses on the automatic adaptation of plastic interfaces while we focus on adaptability by the users.

More generally, architectures for Ubicomp systems have focused mainly on middleware to support system requirements such as distribution, discovery, fault-tolerance or context-awareness, but do not address the specific needs of interaction (see, for example, [117]). The BEACH architecture [161] is a rare exception as it addresses interaction explicitly although it seems limited to classical interaction techniques based on mouse and gesture input.

Architecture models for user interfaces have a long history, with the MVC (Model-View-Controller) design pattern [107] being by far the most widespread solution. A well-known problem with MVC is the strong dependency between the view and controller that limits reusability. Abstraction-Link-View (ALV) [83] was designed for sharing a common model (the abstraction) among multiple networked clients potentially each with their own view, but is otherwise quite similar to MVC. None of the existing patterns however make instruments explicit, instead they promote a widget-based type of interaction. Finally Document-Presentation-Instruments (DPI) [20] is a document-centric software architecture based on instrumental interaction. Like our approach the goal it decouples instruments from the target objects, however DPI is a desktop-based framework and does not address the distribution of objects and instruments across multiple machines.

4 The VIGO architecture

We now present *VIGO* (Views, Instruments, Governors and Objects), the architecture that we have designed to implement Ubiquitous Instrumental Interaction. *VIGO* is an alternative to MVC designed to create distributed interfaces based on the principles described in the previous section.

Figure 1 presents the *VIGO* architecture. *Objects* are presented to the user through *views*. Users manipulate objects through *instruments*, which query views to identify the objects being designated. In order to manipulate an object, an instrument queries the *governors* attached to that object to validate its manipulations. Governors, on the other hand, observe object changes to implement potential side effects. Finally governors can manipulate their attached objects if the user's actions on the object have side effects beyond that object.

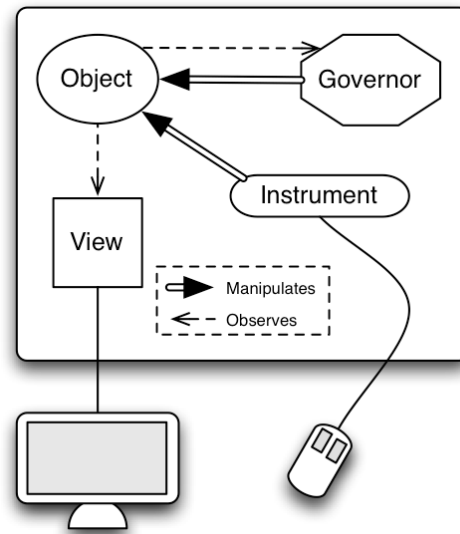


Figure 1 VIGO in pseudo-UML

The following description shows that this design ensures a strong separation of concerns, provides great flexibility and supports distribution among multiple devices and machines.

4.1 Objects

VIGO objects are different from objects in the classical object-oriented sense. In order to achieve the required decoupling between objects and the instruments manipulating them, objects are passive, i.e. they do not provide operations or methods applicable to them. Rather than hiding their state through encapsulation, they expose it as a set of directly accessible properties, while the behavior that is usually provided through methods is encapsulated in *governors* (see below). Compared with MVC, the Model is separated in VIGO into Objects, implementing state, and Governors, implementing behavior.

Objects can be primitive, consisting only of properties, or composite, consisting of properties and other objects. An object can be a part of several objects simultaneously, such as when a diagram is used in multiple documents: a change in the diagram is then reflected in all the parent documents. This is a different concept than multiple views, which is discussed below. This object structure lends itself to a natural description in XML, which is the format we use for persistence. Examples of concrete objects are text documents, graphic canvas, the board and pieces of a board-game (which we will see later) or concrete user interface elements.

Objects (or parts of objects) can be distributed over multiple computers. In our implementation their state is kept synchronized across the replicas. Because objects are passive, this can be implemented simply and efficiently.

The underlying principle behind our notion of object is that every interaction must target one or several objects and may result in a change to the state of these objects. Rather than interacting with a “system” or “application” as with traditional interfaces, the user interacts with objects, by means of instruments. The fact that objects are open gives tremendous power to instruments. On the one hand, it makes it possible to “break” an object by putting it into an inconsistent state. Governors are designed to control this potential chaos. On the other hand, it makes it possible to implement interactions that were not anticipated. For example, in a board game, the board could be used to play a different game (the rules will be embodied, as we will see, in governors), or it may be possible to annotate the board with arbitrary marks. The traditional approach where objects are only accessed through methods clearly precludes such unanticipated uses.

4.2 Views

Views on objects are translations of the objects into one or more modalities perceptible by the user. They are very similar to views in MVC. A typical instance of a view is a visualisation of an object (and its subobjects) on a screen. Views are strongly coupled with the objects they represent, so that any change to the object is reflected in the view. Views on the other hand do not provide any kind of interaction: any change to the view is the result of a change to the object. Views provide a service to translate coordinates (in the case of a visual display) between the view of an object and that of its subobjects or vice-versa.

Unlike objects, views are device-dependent, i.e. they optimize their representation of the object for the display device at hand. Multiple views can be associated with an object, in which case they are synchronized. Note however that multiple views of the same object can also occur when an object is shared among several parents, as described in the previous section. Since objects are passive and views are pure representations, they can be implemented efficiently, e.g. through an observer pattern [76]. On the other hand, for efficiency reason, the object should be available on the machine that holds the view. This is easily achieved using the ability to distribute an object, as described in the previous section.

4.3 Instruments

Beaudouin-Lafon [18] defines an instrument as:

*... a mediator or two-way transducer between the user and domain objects.
The user acts on the instrument, which transforms the user's actions into
commands affecting relevant target domain objects.*

The concept of instrument is inspired by the real world: a stick to enhance one's reach, a pen to write on a piece of paper, a hammer to drive a nail. Ex-

amples of digital instruments include those to enter text, manipulate graphics, draw, select objects, etc. The tools in the tool-palette in applications such as Photoshop are typical examples of instruments.

An instrument may need to provide feedback to the user by presenting some information. This is achieved by associating instruments with objects and creating views for these objects. This means, in addition, that an instrument can manipulate another instrument through the object associated with it and the proper governors.

One way of concretely thinking of instruments is as event processors. Instruments react to input from the user or from other instruments, change objects and fire new events for other instruments to react to. Interaction occurs through chains of instruments, e.g. an instrument for selecting an object on the screen is chained with an instrument for moving objects.

An instrument can be distributed, i.e. it may require input or output from/to multiple devices connected to different machines and yet function as a single instrument. For example, a PDA used as a remote control will require some feedback on the controlled device, while the pick-and-drop technique [144] uses input from two devices, the source and destination. Since instruments are event processors, this requires that the event system must be distributed, i.e. that events produced by a source on one machine are transmitted to a destination on a different machine. Since events are asynchronous, this can work with any network transport system.

Instruments have no direct equivalent in MVC. Instruments that correspond to traditional widgets, such as a scrollbar, can be implemented with an MVC triplet where the controller is the instrument itself while the model and the view implement its associated object. But MVC forces such an instrument to be linked to a target object, such as a text area, through a parent MVC controller. This is one of the reasons why a simple interaction such as drag-and-drop is difficult to implement with MVC (in fact, it does not really fit the pattern). By contrast VIGO instruments are loosely coupled with their targets.

In other cases, the equivalent of instruments are implemented in the controller of an object in MVC. For example, an implementation of a text area with MVC typically include the text editing commands in the controller. This merge of a domain object (the text itself) and the instruments to manipulate it (the text editing commands) in a single MVC entity does not provide the separation of concerns that VIGO encourages. It also limits extensibility, e.g. adding a highlighter tool can be easily accomplished with VIGO with an independent instrument while it requires changing the code of an existing controller in MVC.

4.4 Governors

We have now described how objects are passive constructs visualized through views and manipulated by the user with instruments. The manipulations issued by an instrument consist in changing the state of the targeted object.

Specific rules governing these state changes or the consequences of these state changes are not the responsibility of the instrument, otherwise it would be very difficult to create polymorphic instruments [19] that are independent from the objects they manipulate. Consider for example a board game such as checkers or Othello. If the instrument used to move the pieces of the game implements the rules of the game, then it cannot be used for other games, or indeed for anything else. Another solution is to implement the rules in the board object, but this breaks our object model.

To solve this problem we introduce *governors*² to embody the rules and consequences of interactions with objects. Governors implement the “application logic” or “business rules” commonly found in the MVC Model.

Governors are associated to objects at the level of individual properties: a client, typically an instrument, that needs access to a property of an object asks the governors associated with that property whether that change is acceptable and what an acceptable change would be. Once the change is made it notifies the governors so they can take any additional action. The idea is that a governor controls certain aspects of an object and only the properties that are relevant to that governor are associated with it. Several governors may be associated with one object and several objects may be associated with one governor. Finally governors are stateless, i.e. all state that governors need to manage must be stored in a separate object or added to an existing object.

Let us illustrate this with the Othello³ game. The positions of the pieces on the game board are associated with a governor handling the game rules and the consequences of manipulating the pieces. The move instrument queries the piece’s governor when it is about to drop the piece on the board. The governor checks that it is a legal move and if so returns the proper position for the piece, i.e. the middle of the destination square. If it is not a legal move, it returns the list of valid moves, which the instrument may decide to highlight. Once the piece has been set to the new position, the governor is again notified of the change, and it applies the side effect, i.e. turns over the pieces according to the rules. Note that since it is up to the instrument to query the governor, it could decide to bypass it and cheat or even break the game. Note also that the same move instrument used to move the pieces can be used to move the whole board. One could imagine a governor for the board that “shakes” the pieces when the board is moved, as often happens with a real board. Finally if the piece governor is to keep track of turn-taking in the game, it must store this state in a separate object or in the board itself. This ensures that if another instrument manipulates the board, it can access that state as well.

Governors should not be seen as direct mediators between instruments and objects: the governors are not an interface to the objects, neither are they transparent to the instrument and just react to the manipulations of the objects. Both

²The name *governor* was chosen since this construct adds the cybernetic reactive aspects of the system. The word cybernetics stem from the Greek *kybernetes* meaning steersman, governor, pilot, or rudder [116].

³Also known as *Reversi*.

cases would lead to a lack of flexibility. In the first case, instruments would not be able to bypass the governors; in the second case instruments would not be able to visualize what the governor proposes, such as the valid moves in the Othello game. Instead, governors and instruments negotiate: instruments query the associated governors to validate a manipulation or to get the valid or suggested manipulations (valid moves in the case of the Othello game) which they could in principle ignore (which in Othello would break the game).

Since governors are stateless there are various ways to handle governed distributed objects. Each machine could have an instance of a shared object's governors, only one of the machines could hold the governors and let the others query it, or the governors could reside on a central server. Note that governors can be attached and detached from objects dynamically. In the case of the Othello game, detaching the pieces' governor allows to use the move instrument to move the pieces freely, while attaching another governor would allow to play a different game.

4.5 Using governors to synchronize objects

Multiple representations of a single object are common in computer applications. For example, a UML specification may be represented as a text or as a diagram, a chess game may be represented by an animated board or a list of moves, etc. Representations are also often used to create computer renderings of physical objects, such as the reading of a sensor. Such representations involve a fairly large semantic distance between the object being represented and the representation, to the point where, in the user's mind, the different representations and the original object are separate (but related) objects. In VIGO, this notion of representation does *not* correspond to the notion of view. Since instruments manipulate objects through their views, the mapping between a view and its object must be fairly straightforward.

In order to support the kind of representation that involves a large semantic distance, we use objects and governors instead. Consider the case of a mapping between object-oriented code and an interactive UML diagram: The code and diagram are edited through text-editing instruments and diagram instruments respectively. The code has a text view and associated governors for handling syntax highlighting, indentation, etc. while the UML diagram has a graphical view and governors for aligning boxes and snapping edges to boxes. The two objects, however, share a governor handling the mapping between code and UML (Figure 2) so that adding a new box in the UML class diagram creates the associated code and visa versa. Controlling physical devices is analogous: The external device is represented by an object with an associated governor handling the synchronization of state between object and machine.

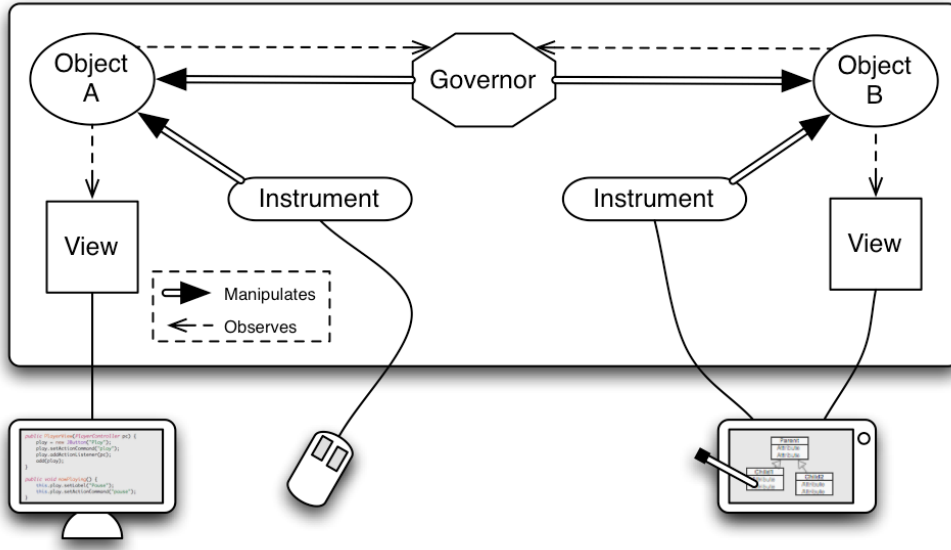


Figure 2 Two objects representing each other through a governor

5 Using VIGO: Realizing Pick-and-Drop

In this section we describe our prototype implementation of the VIGO architecture. Our hardware setup consists of a SMART BoardTM connected to a Mac together with a Nokia N810 Internet Tablet. The implementation is a client-server system developed in Python using the Twisted [73] distributed computing framework. For visualization we use the Apple Cocoa framework on the Mac and PyGTK + Cairo on the Nokia N810. The server has three facets, an object-server, an event-server and a governor-server.

5.1 Objects

Objects are defined in a simple XML language with primitives such as basic shapes and text. Using graphical objects simplifies the mapping to views and is sufficient for our experiments. Graphical objects are laid out in a Canvas, which is itself contained in a view-object. View-objects are the only objects that are not replicated, they are created locally on a device when a canvas is loaded.

Objects have a unique id and are replicated across clients by the object-server. Each client manipulates a local object and the changes are propagated to other clients sharing the same object through the object-server.

5.2 Views

Views are device-specific components that display objects and provide methods for translating between screen and view coordinates and. In our implementation, they also implement *picking*⁴. Our implementation is naive in that it redraws the whole view when the corresponding object changes. This could be improved by observing the changes in the object and optimizing redrawing. Multiple views on a single object are supported. Since views are the only device-dependent construct and have a fairly small interface, it is easy to port VIGO to another platform using a different graphical toolkit. This is what we have done with Cocoa on one device and PyGTK and Cairo on another.

5.3 Instruments

Instruments take input events such as button presses and transform them into object manipulations. This transformation is described by a state machine. We have implemented a Python library for state machines similar in its principle to SwingStates [6]. We have noticed that state machines not only reduce the traditional “spaghetti of callbacks” problem of user interfaces [121], they also provide a good hint of the complexity of the instrument being implemented.

An example of a simple instrument is an instrument for moving objects on the screen. This instrument actually consists of two instruments: An instrument to select an object on the screen, and an instrument to move the selected object. The selection instrument processes input events and fires a `select` event when something has been selected and can now be moved. The move instrument is triggered by the `select` event and tracks the mouse until the button is released (Figure 3). From the user’s perspective, these two instruments act as a single, integrated one. Separating them has the advantage that they can be reused more easily to create more complex instruments.

Instruments can share events through the event-server hence they can receive local as well as remote events. The latter are sent by other clients and automatically dispatched by the event server.

We can now show the implementation of the pick-and-drop instrument. We use two state-machines (Figure 4). Selecting an object when in the `idle` state triggers the transition to the `selected` state and a `select` event is fired with the `id` of the picked object. The remote pick-and-drop instrument receives this event and transitions to the `selected` state. The picked object can now be dropped with the remote pick-and-drop instrument by loading the object whose `id` is in the `select` event from the object-server. When one of the state machines drops the object, it fires a `drop` event that reverts all instruments to the `idle` state. Figure 5 gives the Python pseudo-code for the pick-and-drop state machine. It

⁴Picking could be implemented in a dedicated instrument, however it is more efficient to take advantage of the point-shape intersection computation of the underlying graphical toolkit. Advanced selection, e.g. of hidden objects, and selection based on queries are implemented in instruments.

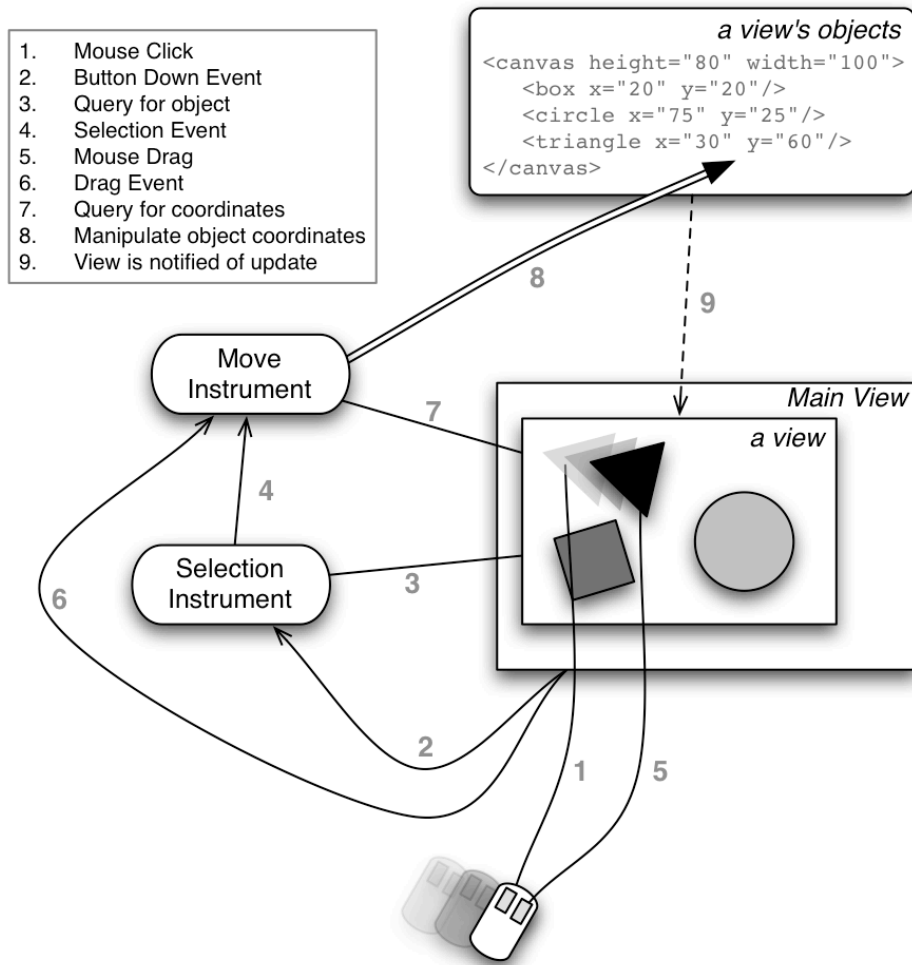
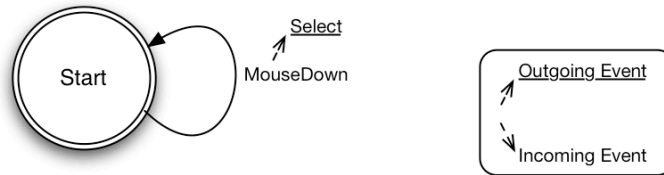
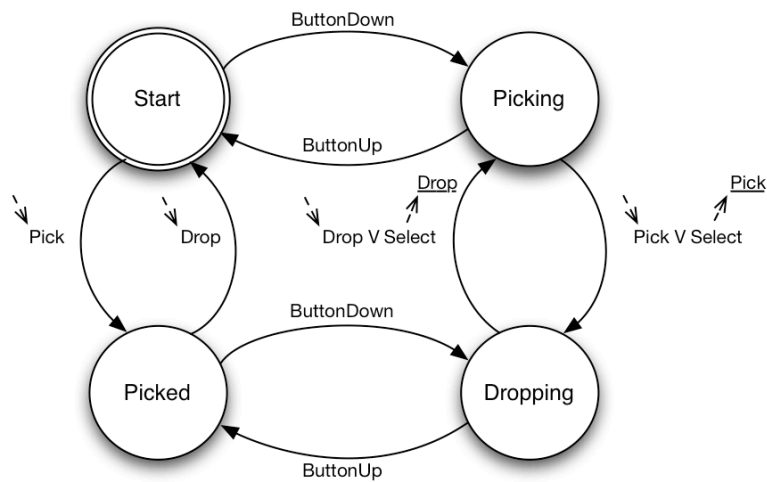


Figure 3 Ungoverned interaction between move instrument and object

uses our syntax for defining states and transitions in instruments with Python decorators (@ declares a state while @ declares a transition).

The color-picker instrument is implemented in a similar way. On the N810 PDA we have implemented an instrument that fires events indicating that a color has been picked, and on the SmartBoard a drawing instrument that reacts to these events and changes its color (Figure 10). The color-picker can work on anything that has a color attribute, hence a specific palette object is not required. We have implemented other instruments for resizing shapes, terminating views and activating other instruments. The latter uses gesture input on the SmartBoard whereas on the PDA it uses the keyboard.

Instruments can provide visual feedback by creating an associated object and using the same vector-graphics language as other objects. The view for this object is a transparent overlay above the other views. This is used, e.g., by

Selection Instrument:**Pick-and-Drop Instrument:****Figure 4** Selection and pick-and-drop state machines

the gesture instrument to provide feedback about the gesture being made and whether the gesture has been recognized or not.

5.4 Governors

Governors provide a simple interface for instruments to validate changes and to retrieve sets of valid changes. In the Othello example a move instrument can ask the governor of a piece whether a position is valid and can query for the set of currently valid positions. When a change to a governed object has been made, the governors are notified and can react accordingly, e.g. to change the color of one or more pieces on the Othello board.

The Othello game uses two governors: The board-governor makes sure that pieces can only be placed within the squares of the board, and adds a new property to the pieces on the board to hold their grid coordinates (-). It also places a fresh piece in the corner of the board when one has been moved to a cell. The second governor is attached to the pieces and implements the Othello rules: checking that a move is correct and turning over the proper pieces after a piece has been put down. It uses the grid coordinate properties added to the pieces by the board governor. Figure 6 illustrates three boards, the first board

```

1  @state  # defining state "start"
2  def start(self):
3      # transition on button press to state "picking"
4      @transition(event=ButtonDown, to=self.picking)
5      def action(event)
6          pass # no action
7
8      @transition(event=Pick, to=self.picking)
9      def action(event)
10         self.picked = event.picked # record picked object
11
12 @state  # defining state "picking"
13 def picking(self):
14     @transition(event=Select, guard=self.picktest, to=self.dropping)
15     def action(event):
16         self.picked = objectHandler.getObj(event.element)
17         eventHandler.fireEvent(Pick(self.picked)) # fire Pick event
18
19     @transition(event=Pick, to=self.dropping)
20     def action(event):
21         self.picked = event.picked
22
23     @transition(event=ButtonUp, to=self.start)
24     def action(event):
25         pass
26
27 @state  # defining state "dropping"
28 def dropping(self):
29     @transition(event=Select, guard=self.picktest, to=self.picking)
30     def action(event):
31         if self.picked is not None:
32             <insert picked object into selected object>
33             eventHandler.fireEvent(Drop(self.picked)) # fire Drop event
34
35     @transition(event=Drop, to=self.picking)
36     def action(event):
37         pass
38
39 @state  # defining state "picked"
40 def picked(self):
41     @transition(event=ButtonDown, to=self.dropping)
42     def action(event):
43         pass
44
45     @transition(event=Drop, to=self.start)
46     def action(event):
47         pass

```

Figure 5 Pseudo-code for pick-and-drop instrument

has no governors (pieces can be placed anywhere), the second board only has the board-governor (pieces are in the middle of the squares) and the third board

has both governors (it shows a correct game). Figure 7 shows an XML object attached to both governors.

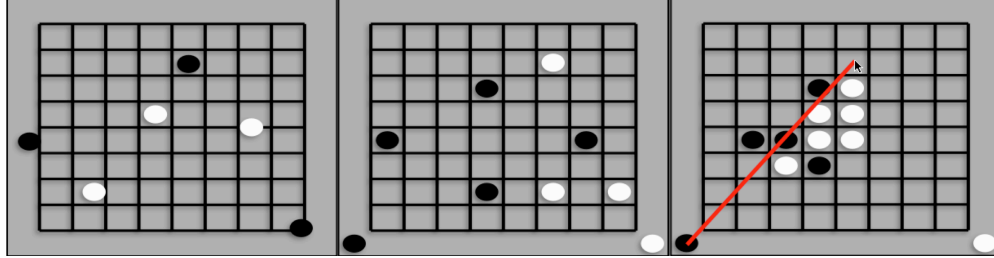


Figure 6 Othello Governors

```

1 <circle x=" 10" y=" 10" radius="7" boardPlacement="D4" fillColor="#000000" id
  ="6bd0">
2   <governed by="boardGovernor" instance="d27d">
3     <governs attribute="x"/>
4     <governs attribute="y"/>
5   </governed>
6   <governed by="othelloGovernor" instance="d09e">
7     <governs attribute="boardPlacement"/>
8   </governed>
9 </circle>

```

Figure 7 Object with governors

The interaction between the move instrument and governors is as follows: The move instrument queries the governor associated with the - position of the piece being moved for valid values. The board governor computes the mapping from the grid coordinates stored in the piece (-) to the - position on the board. It then asks the governors associated with the grid coordinate properties whether their position is a valid move. It collects, say, the list (,). The board governor now takes the intersection between its valid grid coordinates and those returned by the governors and returns their translation into - positions. The instrument can now use the returned positions as feedback to the user or to place a piece. When placing the piece, the board-governor is notified, sets the new grid coordinates for the piece and notifies the Othello governor which exchanges the pieces on the board according to the game rules. This shows that the board governor can be used not only for playing Othello, but also for any game that relies on a board with a square grid.

Instruments can query governors for valid values in order to provide feedback to the user about the rules associated with an object. In the Othello game this is used by the move instrument to show whether a position is legal: in Figure 6, the move instrument turns the line red if the user points at an illegal position.

Instruments interact with governors through a client-side governor-handler (Figure 8). In our implementation governors reside on a governor-server, and

```
1 # Retrieve the governors associated with the object's x and y properties
2 governors = governorHandler.getGovernors(obj, ["x", "y"])
3 ...
4 # The instrument checks whether the new object position is valid
5 if governors.validVal(["x", "y"], x, y):
6     object.set("x", x)
7     object.set("y", y)
8     # It notifies the governors that the object has changed
9     governors.changed(obj)
```

Figure 8 Pseudo-code illustrating the interaction with governors

are instantiated when a client loads an object for the first time. Governors could however be cached locally, e.g. in case the network connection is temporarily lost. If multiple governors are associated with a set of properties the governor-handler computes the intersection of the valid values that they return. If an instrument wants to change the properties of an object that is not governed, e.g. drawing annotations on the board of the Othello game, it can do so without having to negotiate with the governors. An instrument can also simply ignore the governors and change object properties directly.

5.5 Supporting multi-surface interaction

Now that we have described the components, we can illustrate some multi-surface interactions supported by the architecture. We start with the pick-and-drop instrument. Since any object can be used with that instrument, we can apply it to the Othello board itself. This works with no extra code and allows to instantly share live applications across multiple devices and surfaces: picking an ongoing game on the SmartBoard and dropping it on the PDA, we instantly have a networked multiplayer version of Othello (Figure 9).

Another application is to put the pieces of the Othello game on the PDA of each player. They just use pick-and-drop to move the pieces from their PDA to the SmartBoard, again without modifying anything to the existing game. We have also implemented a printer object associated to a print governor. The print governor simply prints the object that is dropped on the printer. The printer can be picked-and-dropped like any other object and therefore can easily be shared across multiple devices and taken away on one's PDA.

Using these components, we have created a scenario where the PDA is used as both a color and object palette for creating graphics on the SmartBoard. Images and shapes are picked on the PDA and dropped on the SmartBoard, while colors for freehand drawing are selected on the board (Figure 10). Finally the canvas on the SmartBoard can be picked and dropped on the PDA, or dropped on the printer for printing.

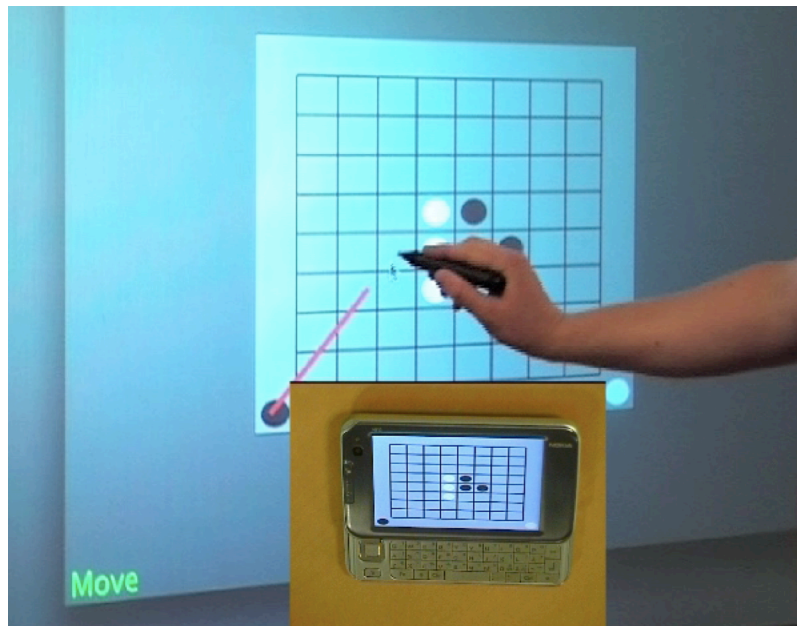


Figure 9 Shared game between the SmartBoard and the N810 PDA



Figure 10 Image created on the board using palettes on the PDA

6 Discussion and qualitative evaluation

VIGO was designed specifically to address multi-surface interaction and support a flexible interaction style where users can use any instrument on any object. We first compare it with MVC and then provide some formative evaluation.

VIGO is conceptually close to MVC but has a number of key differences. First, VIGO was designed from the ground up for distributed interfaces. While there are some implementations of distributed MVC [80], in particular in the context of Rich Internet Applications, they usually have a number of limitations with respect to the original pattern, and to our knowledge none of them natively support multi-surface interactions such as pick-and-drop. Second, VIGO stresses the notion of mediation that is absent in MVC. In MVC the View is both a visualization of the model and of the “tool” to interact with the model, whereas VIGO puts the visualization of the model in the View and the visualization of the tool in the Instrument. This makes it possible for multiple, independent instruments to interact with the same model through the same view, which is different from MVC’s model-sharing through multiple views. MVC makes it difficult, or at least non-natural, to create non widget-like interactions. For example, drag-and-drop types of interaction are not well supported by MVC, but they are a key interaction in multi-surface environments. Finally, the MVC state-based Model is not explicitly present in VIGO. Instead, the Governor represents behavioral aspects of the model that are relevant to interaction. This provides a high degree of flexibility that is difficult to obtain with MVC.

We now summarize a qualitative evaluation of VIGO based on some of the criteria defined by Olsen [131] to assess user interface systems research:

Generality of the solution: Instrumental interaction makes the assumption that interaction is mediated by an instrument. Whether all interaction with a computer is mediated or not is beyond the scope of this paper⁵, however it does cover a wide range of interaction styles, including traditional WIMP interaction, tangible interaction and pen-based interaction [18] that are relevant to multi-surface environments. Our experience so far is that VIGO provides appropriate support to implement ubiquitous instrumental interaction, and we have yet to find an instrumental technique that does not fit the pattern.

To give a perspective on how ubiquitous instrumental interaction could change some basic uses of the computers, consider instant messaging: with VIGO, a conversation would be an object that two or more users can share (a generic instrument would support a directory and a means to share objects). The governors associated with the conversation would add a user name and time stamp to the entries added to it. Entries could be images, text, drawings or even a live game like Othello, added with pick-and-drop. The conversation itself could be moved around between the surfaces available to the user and manipulated with the instruments available on the devices at hand, e.g. freehand writing on a PDA and keyboard text entry on a laptop.

Viscosity: According to Olsen viscosity includes flexibility, expressive leverage and expressive match. The goal of VIGO is to clearly separate concerns among its strongly decoupled components. We have illustrated the level of flexibility this provides. In particular, adding or removing governors radi-

⁵Speech-based interaction, for example, does not seem to be mediated, unless one considers the speech recognizer itself as an instrument.

cally changes the behavior of objects, and instruments can operate on objects they know very little about. New instruments can easily be implemented and tested with existing objects, instruments and governors. Expressive leverage is demonstrated by the polymorphic aspect of instruments and the ability to reuse the various components. For example, the draw instrument that was created for the paint application can be used to annotate the Othello board, and the color picker can pick the color of any other object. The flexibility of the architecture made it easy to create a pick-and-drop instrument that can move pieces of a board game as well as the complete live game from one device to another. Expressive match is supported by the use of concepts familiar to developers, such as a (distributed) event system, state machines and XML-like objects, and the “concrete” character of the components that makes it easy to decide what should be in the objects, the governors and the instruments.

Power and Scale: The power of instrumental interaction is the ability to apply any instrument to any object, and the power of VIGO is to provide developers with simple means to achieve this interoperability. Combining objects, combining governors and combining instruments is simple, what may prove more challenging is controlling which interactions are desirable and which are not. While we see evidence that the approach is scalable, we need to validate it by implementing a larger system. In case the flexibility of the model proves a weakness rather than a strength, for example if we loose control over the combinatorial explosion of interactions between instruments, governors and objects, we will consider adding appropriate control mechanisms.

7 Conclusion and future work

This paper addresses a problem area that has gained little attention despite the development of ubiquitous computing technologies. We have shown the potential of multi-surface interaction and presented an extension of instrumental interaction called Ubiquitous Instrumental interaction that supports distributed interaction among multiple devices and computers. It has then presented VIGO, a software architecture pattern designed for the implementation of ubiquitous instrumental interaction and illustrated it with several examples. VIGO supports reusability of instruments by users, since the same instrument can be used with different objects. It also supports reusability of components by the developer, in particular through the flexibility provided by the dynamic management of governors attached to objects. Finally we have shown how to implement the classical Ubicomp interaction technique pick-and-drop and discussed a number of evaluation criteria for the proposed pattern. We plan to continue the development of VIGO to further explore how to best support multi-surface interaction and address issues such as the configuration of instruments, the scalability of the architecture and the design of novel multi-surface interaction techniques.

8 Acknowledgements

We thank Pär-Ola Zander, Susanne Bødker, and Olav Bertelsen for fruitful discussions. We thank Pierre Dragicevic, Rob Jacob, Allan Hansen and Wendy Mackay for insightful inputs on the paper, Rasmus Berlin for work on an initial implementation and Jonas Petersen for AV assistance.

Rethinking Laboratory Notebooks with Ubiquitous Instrumental Interaction

Clemens Nylandsted Klokmoose

Pär-Ola Zander

Abstract

This paper presents a pilot study where a newly proposed interaction paradigm is applied as a thinking tool in workshops with a group of experimental physicists to envision a future laboratory notebook system. Writing laboratory notebooks are one of few complex inscription-based activities where a digital solution has not really caught on, especially not in academic laboratory settings. We argue that in order to successfully design a laboratory notebook system it is beneficial to divert from traditional ways of designing information systems, e.g. diverting from the application centric approach. We show how introducing the physicists to ubiquitous instrumental interaction empower them to generate visions that go beyond what is possible in applications today – and beyond what they were capable of envisioning before they were introduced to the thinking tool. The results from the workshops enable us to conceptualise the role laboratory notebooks play to the physicists as *syntonic seeds*; inscribed objects that oscillate between being instrument and object of activity, and is intended to sublate a range of contradictions rather than becoming a final product – a type of inscriptions we believe have been neglected in the design of contemporary information systems. We show that ubiquitous instrumental interaction fits characteristics of syntonic seeds especially to support rapid oscillation between instrument and object, something, which was emphasised as important by the physicists.

Published as: Klokmoose, C. N. and Zander, P.-O. Rethinking Laboratory Notebooks with Ubiquitous Instrumental Interaction. *Proceedings of IRIS 2009* (Molde, Norway, 9-12 August, 2009)

1 Introduction

In many aspects of our everyday life information technology gives an obvious advantage over the analogue counterpart for our inscription-based activities. We write our papers, manipulate images, create music and build detailed three-dimensional models with the hassles of doing it by hand long forgotten.

However, there are some inscription-based activities where the flexibility of paper has continued to outweigh the benefit of digital solutions. Patient records are one example of paper-based inscriptions that has been exceedingly hard to translate to information systems. Another example is laboratory notebooks, which is the focus of this paper.

Systems for electronic laboratory notebooks (ELN) are beginning to emerge especially in the biochemistry and pharmaceutical industry, but in the academic research laboratories around the globe the paper-based laboratory notebook is still the referential source of documentation [51]. However, given the increasing complexity of modern research, and the computerisation of the majority of the experimental equipment the traditional paper-based laboratory notebook is beginning to age. But seemingly the paper-based bound notebooks have qualities that are difficult to translate computer application [160].

Tabard et al. [160] augment paper-based laboratory notebooks through the use of Anoto technology (<http://www.anoto.com/>), hence bridging the gap between the analogue and digital. This appears to be a promising approach, however we believe some of the inherent problems in creating a purely digital laboratory notebook system can be addressed by diverting from traditions and fundamental assumptions in how we design information systems.

In this paper we present a pilot study where a newly proposed interaction paradigm, *ubiquitous instrumental interaction* (Paper V), is applied as a thinking tool in workshops with a group of experimental physicists to envision a future laboratory notebook system. We illustrate how introducing this interaction paradigm to the physicists, has a constructive impact on their visions for a future laboratory notebook system.

Furthermore, we take a theoretical approach towards answering the question of what it is that makes laboratory notebooks different from other inscribed objects, and how it can be conceptualised. The analysis is based on the concept of *syntonic seeds* [174].

The paper is structured as follows: Section 2 presents our method for collecting empirical material and the participant of the study. Section 3 introduces ubiquitous instrumental interaction, the thinking tool we presented to the physicists. In section 4 we give a brief overview of the respondents' laboratory notebook, followed by findings and discussions of the workshops in section 5. In section 6 we discuss the role of the laboratory notebooks through the lens of syntonic seeds. Section 7 concludes the paper.

2 Method and Participants

During fall 2008, a group of researchers from the Department of Physics and Astronomy at Aarhus University were invited to participate in a series of workshops and interviews discussing the future of laboratory notebooks.

The research objectives were to assess whether the ubiquitous instrumental interaction line of thought could serve as a springboard for creative thinking for the physicists. Furthermore we wanted to understand the role laboratory notebooks play to the physicists, and whether it suited the principles of ubiquitous instrumental interaction.

The group consisted of five PhD students in experimental physics but with different fields of study: Photoemission, semiconducting, high-energy physics, laser physics and mass spectrometry. Four of the five participants had extensive experience in using paper-based collaborative laboratory notebooks. The fifth had discarded the use of paper-based notebooks for a partly digital solution – a choice, however, leaving him with some amount of bad conscience. One of the participants had been involved in larger experiments at CERN where both paper-based and makeshift web-based laboratory notebooks were used.

The participants all agreed to be interviewed and to participate in two half-day workshops. Each participant was formally interviewed about his or her use and understanding of laboratory notebooks and asked to give a tour of his or her laboratory where more informal discussion took place. Four of the participants were able to show us around their laboratories. The workshops were structured as future workshops [95]. Future workshops were originally used as methods for democratic brainstorming on social and community issues (ibid). It has however shown itself useful from a design and HCI perspective [104]. Future workshops consist of three phases:

A *critique phase* aiming at elucidating all problems and issues regarding, in this case, the researchers use of paper based laboratory notebooks.

The *fantasy phase* where ideas, no matter how far fetched, for solutions to the problems in the previous phase are stimulated and brought to the table.

And finally the *realisation phase* where the potentials in realising the ideas of the last phase are discussed.

The first future workshop was held as prescribed by [95]. For the second workshop the physicists were given a one-hour introduction to the concept of ubiquitous instrumental interaction and the VIGO architecture (introduced in the next section). They were presented example scenarios and demonstrations illustrating its technological potentials. Furthermore, the participants were given small assignments during the presentation. E.g. they were asked to characterise word processing with the concepts of VIGO. It should be noted that this presentation was purposely devoid of examples taken from the context of laboratory notebooks. The presentation was followed by a re-iteration

of the fantasy phase of the future workshop. Audio from both interviews and workshops were recorded. All topics from the workshops were documented, either by participants or the researchers, on large paper posters.

Ubiquitous instrumental interaction was applied as a springboard [71] for design; “a springboard is a facilitative image, technique or socio-conversational constellation ... misplaced or transplanted from some previous context into a new...” [71, chap. 4]. Bødker and Christiansen [41] characterise springboards with the following words: “Springboards do not come about smoothly or automatically, and they are not as such solutions to the problem one is facing. They are starters that may lead to an expansive solution.” The introduction to ubiquitous instrumental interaction was intended to serve as a mean abductive or hypothesis-driven thinking [41]. Hence in this paper we will not present a new design of electronic laboratory notebooks but illustrate how the concept of ubiquitous instrumental interaction can foster thinking that goes beyond the tradition bounds of how we design information systems.

3 Thinking tool: Ubiquitous Instrumental Interaction

In this section we will introduce the vision of ubiquitous instrumental interaction, which we presented to the physicists as a thinking tool to envision a future laboratory notebook.

Klokmoose and Beaudouin-Lafon (Paper V) introduce *ubiquitous instrumental interaction* as an alternative to the current application centric interaction paradigm. It is an extension of upon Beaudouin-Lafon’s interaction model named *Instrumental Interaction*.

Instrumental interaction generalised and augmented the direct manipulation interaction model [153] for supporting design and evaluation of post-WIMP user interfaces.

The core of instrumental interaction was a conceptual separation between instruments and the domain objects worked upon. Instruments consist of a physical part (the input devices) and a logical part (the software component). Instruments are mediators [38] between the user and domain objects: The user acts on the instrument, which transforms the user’s actions into commands that affect the relevant domain objects and provide feedback to both the instrument and the user.

Instrumental interaction was inspired by the way we use physical tools: A painter can freely add or remove brushes from his collection, pass them around, etc.; Brushes are not bound to paint on a canvas, they can also be used to paint on the wall or on the hand of the painter.

Computer applications do not currently support this level of flexibility: a brush in a drawing application can rarely be removed and used in another context, such as a text editor. Applications typically have a predefined set of tools that are difficult or impossible to adapt to one’s needs or taste. This lack

of flexibility limits the mobility, distribution and customizability of interfaces. It also typically results in complex monolithic applications, built for general-purpose personal computers, with dozens or even hundreds of tools to cover all possible needs ([Paper V](#)).

Ubiquitous instrumental interaction extends the instrumental interaction model to a general interaction paradigm – an interaction paradigm where instruments are first class entities to the user when interacting with digital objects. The principle idea is that users should apply their instrument across objects, interactive devices and work-domains, instead of moving and converting objects between monolithic applications bound to specific platforms and work-domains.

Klokmoose and Beaudouin-Lafon argue that while instrumental interaction was developed in the context of desktop interfaces, the concepts are more general and apply particularly well to multi-device interaction going beyond the desktop computer. In ubiquitous instrumental interaction, the instrument is an explicit construct rather than a descriptive concept.

Ubiquitous instruments should be exchangeable among users, they should work in a similar way on different devices, sometimes across multiple devices, and they should be technically separated from the objects they operate on. Instruments should be applicable to domain objects and on device surfaces when and where it makes sense, even if they were not designed to do so in the first place. The objects that users interact with through instruments should be able to migrate across device surfaces, support multiple views, and be manipulatable by an instrument in ways not necessarily anticipated by the object.

The core vision of ubiquitous instrumental interaction can be exemplified by a small scenario: A graphics designer is capable of doing rudimentary modifications of a poster on a PDA using a subset of the instruments she applied in the first case to create the poster on a desktop computer. Hence overcoming the restrictions of having to have specialized applications for graphic editing on each device.

However, in the context of this paper the object-centrism brought forth in the ubiquitous instrumental interaction vision is more interesting than the flexibility of configuring instruments in itself. Klokmoose and Beaudouin-Lafon argue that in order to realise ubiquitous instrumental interaction a unified object model is a pre-requisite. But an object model where objects dynamically can evolve and be deconstructed.

In order to implement ubiquitous instrumental interaction, Klokmoose and Beaudouin-Lafon present the VIGO (Views, Instruments, Governors, Objects) architecture.

VIGO embodies the two concepts from instrumental interaction, namely instruments and objects. Views is a naturally added concept, and embody the representation of objects, e.g. to the screen. However, in order to enable that interaction with objects can have consequences they introduce *governors*. Klokmoose and Beaudouin-Lafon give an example of the board game Othello; In Othello

ello the game and pieces would naturally be objects, an instrument would be used to move pieces, but in order to benefit from playing on a computer, the player should not uphold rules and consequences by hand. The instrument cannot uphold the rules, hereby the idea of polymorphic instruments would be broken, and instead they introduce the concept of governors. Governors embody the rules and consequences of interacting with objects – particular aspects of objects. An Othello governor does not care if the user draws on the board with a drawing instrument, only if the user changes the placement of the pieces on the board.

In the VIGO architecture the core components are objects. However, objects in VIGO differ from objects in the object-oriented thinking. Objects in VIGO are passive collections of structured data. They do not contain methods for self-manipulation or any application specific logic.

Instead of calling methods on the object, the user manipulates objects through *instruments* and indirectly through *governors*. Instruments in VIGO are basically input processors; they translate user actions on input devices to manipulations or queries on objects. Governors on the other hand embody the cybernetic aspects of an information system, namely the rules and consequences of manipulation of objects in the system.

The distinction between instruments and governors echoes the activity theoretical distinction of passive and active externalisations of tools [102]. Passive externalisations are simple extensions of the user, e.g., a hammer to drive a nail, while active externalisations encapsulate the performance of actions on behalf of the user, e.g., as in an automated toaster, or the cybernetically enforced rules of the Othello game above.

Governors are loosely coupled with objects and can dynamically be associated and disassociated with an object. The governors of the Othello game could be removed so the board and pieces could be used e.g. digital art.

VIGO emphasizes an understanding of the system as a reactive rather than a proactive system. Governors should not be understood as actors, or in anyway resemble agents with artificial intelligence – rather they are externalizations of otherwise human actions. E.g. the action of bringing a board of Othello to an acceptable state after a move has been performed.

4 The laboratory notebooks

When the physicists were asked about the role of their laboratory notebook they responded:

- To ensure scientific probity.
- To enable reproduction of experiments.
- The keep track of what happens in the laboratory.

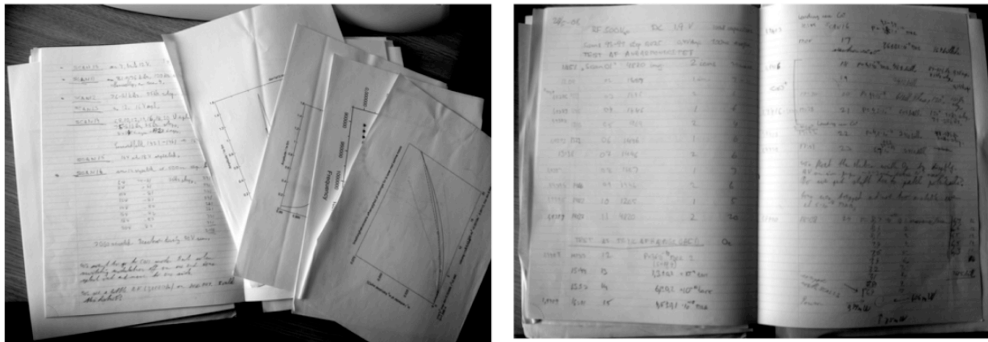


Figure 1 Examples of the physicists' laboratory notebooks

- To log and structure experimental results.
- To analyse intermediate results.
- To document considerations for future experiments.
- To document the mistakes of the past.
- To act as a shared memory for the researchers in the lab.

The books were all variants of regular A4 sized bound notebooks. All of these books were inscribed chronologically from left to right. The entries in the books consisted of a mix of written text, drawings, handmade tables and loose print-outs of graphs and tables of data (Figure 1). The most important piece of meta-information inscribed into an entry of all of the respondents' notebooks was timestamps. The timestamps were often coupled with a reference to a data file and sometimes the name of the author of the entry.

Four out of five of the respondents used a laboratory notebook shared in a group. Three of the respondents furthermore used notebooks tied to a particular piece of equipment in the laboratory, e.g. for a certain laser. Such equipment notebooks could be inherited from the previous owners of the equipment. In two of the participants' laboratories both shared and personal laboratory notebooks were used, which sometimes were source of conflict and confusion. The shared notebooks were used to more or less explicit degrees as means for communications, e.g. for writing messages between researchers measuring in shifts.

One of the physicists explained that her laboratory notebook had three general types of inscriptions: At the right page of the bound notebook, the data and the instrument settings were noted (1). Sometimes this data were coupled with visualizations (2), e.g. printed or hand drawn graphs or drawings of the equipment. The left page was reserved to comments and analysis (3). This particular physicist navigated her book partly by judging if there were many entries on the left page – if so, that experiment probably was important.

5 Rethinking the analogue laboratory notebook

We will now present and discuss the findings from our two workshops. The first workshop was held as a regular future workshop. However, given the pilot nature of these workshops, the realisation phase turned out not to be of much interest in this context. It was mainly a general discussion of what it would require politically and economically in order to get some development going.

In the second workshop the physicists were presented with the ubiquitous instrumental interaction vision and asked to re-iterate on the fantasy phase of the first workshop.

5.1 The physicists' critique and initial visions

During the first phase of the initial workshop, the physicists identified 15 problems in their use of their current paper-based laboratory notebooks.

A number of the problems raised were obvious problems regarding analogue inscriptions. The notebooks are vulnerable to being lost for instance if the laboratory should catch fire. Furthermore, analogue information stored on paper is hard to search through. Another problem raised was that the physical notebook produced too much redundant information, especially when computers for data collection were used together with the paper-based notebook. The paper-based notebook is naturally restricted to being one place at a time which conflicts with it being part of many different activities at one time.

The physical properties of the bound notebook led to a series of problems as well. A notebook has finite space, both in the area and number of pages. Furthermore, a bound notebook enforces chronology (both for good and bad). Important notices can easily disappear and cannot, as one of the physicists stated, act like stickies on a web forum. Categorisation of entries in the notebook is a cumbersome task, and will require third part objects, e.g. index pages or post-it notes.

The physicists furthermore listed a range of problems, where the solutions not necessarily have to be electronic. Namely that there was a lack of standardisation of what to document – a blank page gives no hints of what parameters should be documented. The physicists missed tool support for project management, and another (sometimes critical) issue was that the laboratory notebook in no regard enforced a proper handwriting.

One of the heavy weighing problems lay in what was *not* written in the laboratory notebook. The physicists argued that often they missed important readings (e.g. when they had to be manually read from a display) or didn't know until it was too late what to look for and inscribe in the book.

A final problem stated by the physicists was that all the different objects of inscription created a broken workflow. There was no direct connection be-

tween data collection, control of equipment and analysis. They moved between performing hand-written analysis in the notebook, to data analysis and equipment control on the computer and then again to document in the notebook e.g. with printouts of graphs (Figure 1).

In the second phase of the future workshop the physicists were asked how they envisioned a future electronic laboratory notebook system. This was before they had been injected with the ubiquitous instrumental interaction vision.

The overall vision brought forth by the physicists was of a laboratory notebook system, that integrated documentation, equipment control and data analysis in the same user interface. They envisioned a laboratory where all information from the equipment was logged by a computer.

The notebook should be configurable to document relevant information automatically and to support (or force) the researchers to document what had to be observed manually. They envisioned that through the automatic logging the system could help with error recognition through comparing with old entries in the notebook. Furthermore, automatic pre-analysis of the data could be configured, e.g. to produce intermediate graphs directly in the notebook.

Regarding interaction with the notebook the physicists envisioned being able to access the notebook from multiple devices – including mobile devices. They envisioned multi-modal interaction, and the support of interaction through handwriting e.g. on a tablet-PC.

As for the organization of the notebook, a structuring resembling a Wiki was suggested, where all data could be interlinked. One of the researchers suggested abandoning the forced chronology of the bound notebooks and having something resembling a loose-leaf system, where multiple structures could be put on top the pages.

5.2 Discussion

From the critique and visions brought forth by the respondents it is clear that integration between the different aspects of the activities in the laboratory is one of the key aspects that is missed in their current use of laboratory notebooks.

The physicists use a range of different applications for analyses and control; the obvious approach would be to integrate documentation into these applications. However, applications as they are constructed today embody a static understanding by a group of designers and developers of a given activity. An application is a monolithic collection of tools designed for a specific work domain. This is for instance to produce printable documents, creating PowerPoint presentations or constructing CAD models of buildings. Microsoft Word embodies a specific understanding of how we produce written documents, if we want to do something out of the ordinary, e.g. make free-hand annotations with a tablet pen, we need to look within Word for a tool to do so. If it is not

possible we can hope it will in the next release.

Applications put a tool centric perspective of interaction with digital objects. The applications static collection of tools defines interaction, and define the shape of objects, rather than letting the objects define interaction.

Furthermore, applications used for inscription-based activities are constrained to specific file type or set of types. The file type for e.g. Microsoft Word, like the application it self, embodies a range of assumptions about the activity of writing; the kind of formatting possible, and what kind of objects can be embedded in the text. This makes it difficult to build a laboratory notebook application as general for lab work, as Microsoft Word is for typing printable documents. Development of use and changing requirements are the norm rather than the exception in the use of laboratory notebooks. But in word it is not possible to go beyond what is prescribed by the file-type. If the file-type does not support embedding of web-cam feed, it is not possible to embed it.

The technological requirements put forth by the physicists, e.g. that they should be able to access the laboratory notebook from multiple different devices, stress the application concept even further. Especially this perspective is difficult to achieve through the application line of thought, which one of the main foci of [Paper V](#).

Central for the respondents' activities is that they are in constant development – academic research is rarely routine work. This is apparent when browsing through one of their laboratory notebooks. The inscriptions change over time given changing requirements due to changing understanding, equipment, people and experimental foci. However the documentation environment, the bound notebook, does not change.

The bound notebook is extremely flexible; it embodies only very low-level assumptions of the activity it is part of. You can write what ever you like wherever you like with whatever type of pen you like. The bound notebook is easily extendable with printouts of digital graphs, objects that did not exist in the point of time where the first bound notebook was created. But it has some fundamental limitations as brought forth by the physicists.

5.3 Injecting ubiquitous instrumental interaction

A whole range of visions for a laboratory notebook system was catalyzed from the introduction to ubiquitous instrumental interaction. Some were a direct product of the physicists' introduction to the ubiquitous instrumental interaction vision, while others where probably a product of the participants being even further stimulated to be visionary.

The overarching vision continued to be a fusion of documentation, control and analysis. The physicists envisioned being able to blend documentation and control in a manner so e.g. controls could be annotated the same way annota-

tions were made in the laboratory notebook. They envisioned the possibility of taking equipment control objects and store copies or snapshots of them in the notebook – and potentially use these objects to reload settings.

The physicists envisioned an environment where instruments and governors easily could be programmed and shared. One physicist suggested that she wanted to be able to easily write an instrument that could multiply selected settings with a given constant. This should be enabled through a dynamic scripting language or a visual programming language.

The physicists suggested a range of governors. Governors could be implemented to support analysis, e.g. governors that plotted data directly to a graph on the page of the notebook – a graph that should be annotatable with the proper instruments. Governors to perform error recognition of the data, e.g. if the data express a known deviation we also envisioned. Another type suggested was governors that wrapped around existing software such as LabView (<http://www.ni.com/labview/>) programs and brought those tools into the VIGO architecture. One physicist suggested a LaTeX (<http://www.latex-project.org/>) governor that could translate code for mathematical expressions into nicely readable equations.

The physicists envisioned having the possibility to bring parts of the laboratory notebook's user interface to mobile devices. For instance support for loading a page on a tablet-PC to have the similar feel of a notebook or to move a control object or readout of data to a PDA to support mobility in the laboratory. As one participant remarked, he sometimes had to do rather dangerous acrobatics to read a meter on the screen in the other end of the lab while climbing his pressure chamber to control a valve.

It was envisioned that the notebook supported hypermedia functionality such as linking and tagging of objects. They wanted to be able to link intermediate conclusions with graphs and data to help the process of paper writing. Through this hypermedia functionality it should be possible to create ToDo lists and set up milestones that could be linked up with entries in the laboratory notebook as they are created. Furthermore, it should be possible to create links to the web, e.g. to papers.

Another participant suggested that the system should support the, either manual or automatic, generation of summary pages that could be loaded on the interactive whiteboard in the professor's office.

It should be possible to structure and categorise entries in the laboratory notebooks in multiple ways. E.g. sorted by activity or by chronology, but doing away with the forced chronology of the bound notebooks.

One participant imagined that the system supported personal versions of entries to the laboratory notebook, where notes and annotations could be made that were not intended to be read by others – but could be made public through interaction with a simple instrument.

It should be possible to generate templates both comprising fields and lay-

out for documentation, comments etc. but templates could also encompass controls relevant for a given experiment.

The physicists wanted to be able to have multiple views on the same information, e.g. a view presenting everything relevant to a given piece of equipment – hereby avoiding to have separate equipment notebooks.

An important aspect of a future laboratory notebook system is persistence. The physicists suggested that if interaction was decoupled from a unified object system, a general revision system could be built in. The physicists suggested that it would make sense that all data was immutable, and that the history of all objects could be navigated, but otherwise hidden to avoid cluttering. This revision data should be stored on a central and secure data central, for instance governed by the university.

5.4 Discussion

Our experience with introducing the physicists to the ubiquitous instrumental interaction paradigm was that it indeed stimulated them to think beyond their previous visions for a laboratory notebook.

The physicists emphasised already before the introduction to the ubiquitous instrumental interaction paradigm that *integration* was the central requirement for transitioning to a digital laboratory notebook. Each activity should not have its own stand-alone tool.

With the introduction to the ubiquitous instrumental interaction paradigm and the concepts of VIGO the physicists were able to conceptualise how this integration could be realised. They envisioned how the object centric approach, with application logic loosely coupled to objects, could support objects shifting between having the role of a widget to being a piece of documentation in the form of e.g. letting a copy of a control object both document and store settings for future re-uses.

Overall the object centric nature of ubiquitous instrumental interaction seemed to fit well with the physicists' use of laboratory notebooks. The physicists talked about the system in terms of the laboratory notebook as an object rather than the laboratory notebook system as an application.

The concept of instruments did not play a particularly large role in the physicists' visions. Except for one physicist's vision of being able to create macro like instruments. However, the kind of multi-device interaction envisioned by the physicists, was a product of their introduction to the idea of decoupling interaction from objects as illustrated in [Paper V](#).

Ubiquitous instrumental interaction and the VIGO architecture may not be the only approach for successfully designing a laboratory notebook system. But the workshops showed that there is potential in an object centric approach with a strong focus on a dynamic relationship between inscriptions and the tools used for performing the inscriptions. Letting re-configuring be the norm

rather than the exception can avoid clashes with the surrounding activities that are in constant development.

Furthermore, treating objects used for control, such as dials, meters and buttons, as inscriptions as well seem to appeal to the physicists, who plead for a stronger integration between documentation, control and analysis.

The kind of system the physicists talked about had a lot of similarities with what the World Wide Web (WWW) offers. E.g. the ability to load objects on various devices from a referential source, and support various hypermedia concept like link. However, in some key aspects it goes beyond the traditional WWW.

The technologies of WWW, the mark-up language HTML and protocol HTTP are server centric in nature. The user asks for a page, and input from the user goes through queries defined by the server. Hence what the user can do with an object is restricted by the interface to the server. Therefore embodies some of the same restrictions as applications.

In instrumental interaction the user would load a proxy of an object, and manipulate it directly on the device, and not explicitly communicate with a server through a specific protocol. Given the visions of the physicists this object system should also keep track of the history of objects.

The object system could probably be realised with modern WWW technologies such as AJAX and SOAP, however it would conceptually differ from the typical interaction with and implementation of modern interactive web pages.

6 What is a laboratory notebook?

It becomes clear from the workshops with the physicists that inscribing a laboratory notebook does not transfer to a typical application as easy as writing a letter does. While letters and laboratory notebooks both are inscribed artefacts, the laboratory notebook appears to have distinct characteristics. The laboratory notebook is used by many people for many activities, and changes between being used to document and as a reference, and the inscriptions do not seem targeted towards creating a final product, like the inscriptions of a letter or a paper are.

In order to conceptualise the role laboratory notebook plays to the physicists and hereby potentially making the findings of the presented study more generalisable we apply the concept of *syntonic seeds* [174] as a lens for understanding. Syntonic seed is a concept informed by activity theory and associated approaches [71, 38, 168] suited for understanding development in activities characterised by collaborative inscription of objects which purpose is to *sublate contradictions*.

To understand the nature of syntonic seeds it is necessary to understand the concept of contradictions, which are central concepts in activity theory and

in dialectical thinking. A contradiction is an entity, which is defined at a fundamental level of existence, i.e. not as something that is necessarily immediately and in an unproblematic manner presented to the senses. Activity theory insists that reality in the last instance is contradictory, not only in our perception of it, but also in reality itself [90, 114]. It is sometimes hidden; sometimes shown at the surface. It is not metaphorical, i.e. as if we could “see” something as a conflict/contradictory situation. Of course we can do that, but this is not the sense in which the term is used. The conflict situation is only the surface structure, which may be used to identify the contradiction. It is the form of the contradiction, not its substance.

Bhaskar has eloquently nailed the difficult concept of contradiction in human activity as a “situation, which permits the satisfaction of one end or more generally result only at the expense of another: that is, a bind or a constraint.” [17, pp. 56].

To sublate (from german: *Aufhebung*) a contradiction means that the contradiction is at the same time negated yet preserved [174]. E.g. using post-its as memory aids negates the contradiction between what should be remembered and what can be remembered, but preserves the indication of an insufficient memory.

Syntonic means that something is in harmony or synergy with its environment. A syntonic seed is an inscribed object that is pivotal in a web of activities [30], inscribed to bring relative harmony in the implied activities.

An example of syntonic seeds is architects’ sketches. They are typically pivotal in a number of activities; they are used for negotiating ideas with colleagues, experimenting with own ideas, used as a reference when making new design and added to a portfolio to further the architect’s career. The sketches sublate contradictions between the architect’s mental capabilities and the requirements in an architectural activity, but they may also sublate contradictions between what other stakeholders in a project believe is possible to build and what the architect does [174].

More formally a syntonic seed is characterised by the following:

- It is persistently representing contradictions, and is inscribed to sublate these contradictions.
- It is simultaneously and/or sequentially a mediator in one or many activities and an object of one or many activities.
- It can at any given moment oscillate between being a mediator and object of activity.

Over time contradictions represented by the syntonic seed may change and so may the activities part of the syntonic seed’s web of activities.

To return to the previous example; The architect sketches oscillate between being objects of the architect’s drawing activity and being mediators of e.g.

planning activities.

6.1 Laboratory notebooks as syntonik seeds

The practice of the respondents' involving laboratory notebooks does not consist of one activity only. Rather, it consists of a web of activities. There are six main activities, which are distinctly different from each other in terms of e.g. object, their histories, rules and subjects but all involve the laboratory notebook:

- Experimentation
- Analysis
- Planning
- Lab maintenance
- Paper writing
- Documentation

There are also several types of educational activities going on in the lab: undergraduate, PhD student work and postdoctoral education.

Documentation activity and the educational activities stand out from the others. They are largely unconscious and routinised, and much of the time in an unconscious and co-ordinated state [169], whereas the experimentation activities, but also the other activities, are often collectively co-ordinated and reflected upon. The documentation for ensuring scientific probity is the only activity where the inscribed notebook is the actual objective product of the activity.

In interviews, the lab book is nonetheless referred to as a documentation device, but in daily operation, that function is accomplished by routines rather than by conscious effort. The same goes for educational activities.

The lab book plays an important role in this web of activities. It is acting as an unfinished object in which data are inscribed in the experiment activity, and at the same time used for reading explanations of results in the paper-writing activity. In several activities, it was found to oscillate between being an instrument and being a tool. For instance, a physicist may first enter some data into it, and then read in the lab book one hour later in order to change the settings in some machinery. Another example of this is when the notebook frequently oscillates between being used for documentation, analysis and paper writing.

A central characteristic of syntonik seeds is that they are objects inscribed to sublimate contradictions in activities. We identified four overall contradictions, which the paper-based laboratory notebook successfully sublates:

1. Contradictions **between two persons' plans of the experiment**, due to their different understandings. In a new experiment, the physicists come to the scene with different plans about prioritized things to investigate. These plans have to be negotiated, and the lab book is one of the important arenas for this.
2. Typically, the physicists deliberately try to make the data in the lab books stand in **productive opposition to previous knowledge** in their research community. They try to show that previous papers did not take important phenomena into account or that previous research cannot be corroborated by empirical data from the laboratory – in order to generate interesting research.
3. Doing scientific research is a complex intellectual activity. It is not possible to remember the result from the experiment; there may be extremely large data sets. It is necessary with a memory aid in order to make the planned calculations and to construct connections between data. The memory skills clashes with the demands from the research community of doing physics at a contemporary level – and the lab book serves as **memory aid** and resolves the contradiction.
4. Experimental physics typically require a lot of instruments and equipment. These artefacts take time to master and at the same time the technological development is rapid. Therefore, no individual can master all instruments at the same time. The organization has solved this by **specializing** physicists and training technicians. Rarely a single individual understand all details in a specific experiment, but only their part. The lab book can hide the complexity each specialist brings to the experiment by letting them enter short configuration notes, understandable in a superficial way for other physicists.

The contradictions stated above are targeted by a number of inscriptions in the current practice. The key inscription is of course the physical lab book. Our observations and interviews reveal, however, that a large number of inscriptions are active and necessary constituents of the lab work. From an atemporal viewpoint, they are clearly distinct from each other, but in delineating the final outcome (a well-documented research process), all these inscriptions are parts of the unfinished documentation object in various stages (Table 1).

Given the analysis of above we can see that the praxis of inscribing a laboratory notebook differs from inscribing a letter. Writing a letter is an extremely product-oriented activity clearly oriented toward creating a finished letter product – which is in stark contrast to praxis of inscribing a syntonic seed, which is oriented to ongoing processes. Hence building an information system for laboratory notebooks on the same premises as we create a word processor would not work. In the workshops with the physicists we have seen that introducing them to ubiquitous instrumental interaction had a positive impact on their ability to conceptualise visions for a future laboratory notebook sys-

Object of inscription	Elaboration
The laboratory notebooks	The main objects of inscription
Electronic files of data	Produced by the experimental equipment
The physical configuration of instruments	Everything is not kept in the notebook; during an experiment looking at the setup in the laboratory can support the researchers understanding
Software for data collection and analysis	Software is used to collect and analyse data. Comments appear both directly in the data and e.g. in the program code for performing analysis. Code which itself also acts as a meaning bearing inscription.
Private notes	One physicist stated that he had co-workers who did not want to show their possibly erroneous analysis.
Email correspondence	The physicists discuss interpretations with each other and collaborators.
Whiteboards	Both in the researchers laboratories and offices there are whiteboards used for notices and analysis.
Specialized lab books	In addition to the main lab book, some instruments have their own notebook, e.g. used for maintenance.

Table 1 Objects of inscription

tem. But what is relationship between ubiquitous instrumental interaction and syntonic seeds?

6.2 Syntonic seeds and ubiquitous instrumental interaction

Instrumental interaction and syntonic seeds share the same theoretical heritage in activity theory, but furthermore, there are a number of qualities in the ubiquitous instrumental interaction vision that match the nature of syntonic seeds.

Ironically *ubiquitous instrumental interaction* is *object centric* whereas the application paradigm is *tool centric*. In ubiquitous instrumental interaction, interaction evolves around objects and interaction evolves objects. How objects can be manipulated is not defined by a specific set of tools fused together in an application. Objects evolve with the instruments used on them – using a drawing instrument on a surface creates a drawing. Likewise new instruments can be added to the users configuration when new requirements arise. Hence ubiquitous instrumental interaction gives the user an ability to sublate contra-

dictions as they arise in an activity in ongoing development. The dialectics between tools and objects are present at a micro level – the user can adapt tools to changing needs, and objects change nature with the introduction of new tools. In the application paradigm the dialectic relationship between tools and objects is only present on a macro level – tools are not changed before a large group of users demonstrate a new need, which then is appropriated by the application developers in a new version of the application.

The concept of governors gives the potential for objects to support *explicit oscillation* in a richer sense than in the application paradigm. Using a copy of a equipment control object to store the setting both for documentation and for later re-use, is a prime example of such explicit oscillation. Given the object centric nature of ubiquitous instrumental interaction, controls are objects in the same manner as a piece of text is. This is not the case in the application paradigm; I cannot use the font widget to document in this text what font is used by embedding the font widget in the text.

7 Conclusion

In this paper we have addressed the design space of laboratory notebooks. Through workshops with the physicists we uncovered how integration between the main activities of their work was a key vision for developing an electronic laboratory notebook system. We showed that introducing the physicists to the ubiquitous instrumental interaction paradigm helped them to conceptualise how this integration could and should be realised.

Applying ubiquitous instrumental interaction as a thinking tool to the physicists without us proposing how it could realise a laboratory notebook system was an experiment. The concepts of VIGO are rather abstract and software engineering-like, however the physicists easily grasped them and were able to express a laboratory notebook with the concepts of VIGO.

The physicists were however used to abstract thinking and were all, to a high degree, computer literates. Hence their use of paper-based laboratory notebooks cannot be blamed on computer illiteracy.

Our experience was that a general vision like ubiquitous instrumental interaction, serves well as a springboard for envisioning an expansive solution to a complex practical problem. However, the physicists are used to abstract thinking, and if the respondents had not had the kind of experience with manipulating abstract concepts, the level of abstractness of the vision presented should have been adjusted accordingly.

In order to put the discussion of the paper in a theoretical perspective we analysed the role of laboratory notebooks through the lens of syntonetic seeds. We have argued that ubiquitous instrumental interaction has qualities that suit the characteristics of syntonetic seeds particularly well. Laboratory notebooks are

however not the only inscribed objects characterisable as syntonic seeds; Zander [174] e.g. discuss inscriptions in regional collaborative writing processes as syntonic seeds. Hence the findings of this paper might be applicable beyond laboratory notebooks, for instance to mundane activities as keeping a shared family scrapbook or diary.

Realising the kind of system with the features envisioned by the physicists is a major development task. The design and construction of application-based systems can lean on toolkits that have been refined over the last 20 years or more. Ubiquitous instrumental interaction tinkers with the underlying assumptions that these toolkits are build upon; hence they potentially have to be rethought and reconstructed. The sum of the visions brought forth by the physicists in previous system would require extensive research, design and implementation – research on the technical realisation and of course comprehensive workplace studies and longitudinal participatory design.

The pilot study in this paper however has indicated that in order to create an electronic laboratory notebook system that maintains some of the qualities of the paper-based notebook, some rethinking of the traditional ways of building information systems is required. Especially a level of flexibility is required which currently is absent from both our desktop and web applications. One approach could be ubiquitous instrumental interaction.

8 Acknowledgement

We thank the physicists participating in this study for their willingness to contribute time to research where the practical implications on their daily work lie some days ahead. We thank Susanne Bødker and Olav Bertelsen for insightful comments on earlier drafts of the paper. We thank Marianne Dammand Iversen for proofreading the paper.

Bibliography

- [1] G. Abowd and E. Mynatt. Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(1):29–58, 2000. [144](#)
- [2] S. Adler, A. Berglund, J. Caruso, S. Deach, T. Graham, P. Grosso, E. Gutentag, A. Milowski, S. Parnell, J. Richman, et al. Extensible stylesheet language (XSL) version 1.0. *W3C Recommendation*, 15:2001, 2001. [93](#)
- [3] A. Agarawala and R. Balakrishnan. Keepin’it real: pushing the desktop metaphor with physics, piles and the pen. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1283–1292. ACM New York, NY, USA, 2006. [47](#)
- [4] H. Albrechtsen, H. Andersen, S. Bødker, and A. Pejtersen. Affordances in activity theory and cognitive systems engineering. *Riso National Laboratory*, 2001. [9](#), [23](#)
- [5] J. Alfke. Opening Up OpenDoc. *MacTech Magazine*, 11(1). [80](#)
- [6] C. Appert and M. Beaudouin-Lafon. Swingstates: adding state machines to the swing toolkit. In *Proc. ACM User Interface Software Technology (UIST’2006)*, pages 319–322, New York, NY, USA, 2006. ACM. [93](#), [200](#)
- [7] K. Bærentsen. Mennesker og maskiner. *Et virksomt liv: Udforskning af virksomhedsteoriens praksis*, pages 142–187, 1989. [20](#), [21](#), [103](#), [107](#), [108](#), [138](#), [139](#), [159](#), [160](#)
- [8] K. Bærentsen and J. Trettvik. An activity theory approach to affordance. In *Proceedings of the second Nordic conference on Human-computer interaction*, pages 51–60. ACM New York, NY, USA, 2002. [9](#), [12](#), [19](#), [20](#), [23](#), [34](#), [37](#), [47](#), [50](#), [102](#), [103](#), [109](#), [110](#), [111](#), [113](#), [158](#), [159](#), [160](#), [161](#), [168](#), [174](#)
- [9] R. Ballagas, M. Ringel, M. Stone, and J. Borchers. istuff: a physical user interface toolkit for ubiquitous computing environments. In *Proc. ACM conference on Human factors in computing systems (CHI’03)*, pages 537–544, New York, NY, USA, 2003. ACM. [82](#), [193](#)
- [10] R. Ballagas, A. Szybalski, and A. Fox. Patch panel: Enabling control-flow interoperability in ubicomp environments. In *Pervasive Computing and*

- Communications*, 2004. *PerCom 2004. Proceedings of the Second IEEE Annual Conference on*, pages 241–252, 2004. [82](#)
- [11] L. Balme, A. Demeure, N. Barralon, J. Coutaz, and G. Calvary. CAMELEON-RT: a software architecture reference model for distributed, migratable, and plastic user interfaces. In *EUSAI 2004, LNCS 3295*, pages 291–302, 2004. [193](#)
 - [12] L. J. Bannon. From human factors to human actors: The role of psychology and human-computer interaction studies in system design. *Design at work: Cooperative design of computer systems*, pages 25–44, 1991. [4](#), [158](#)
 - [13] L. J. Bannon and S. Bødker. Beyond the interface: encountering artifacts in use. In *Designing interaction: psychology at the human-computer interface*, pages 227–253. Cambridge University Press, 1991. [158](#), [160](#)
 - [14] J. Bardram, H. Christensen, and K. Hansen. Architectural prototyping: an approach for grounding architectural design and learning. In *Proc. 4th Working IEEE/IFIP Conference on Software Architecture*, pages 15–24, 2004. [12](#)
 - [15] J. E. Bardram. Activity-Based Computing: Support for Mobility and Collaboration in Ubiquitous Computing. *Personal and Ubiquitous Computing*, 9(5):312–322, July 2005. [9](#), [50](#), [81](#)
 - [16] J. E. Bardram and O. W. Bertelsen. Supporting the development of transparent interaction. In Blumenthal, Gornostaev, and Unger, editors, *Human-Computer Interaction. International Conference, EWHCI '94. Selected Papers.*, pages 79–90. Springer Verlag, 1995. [36](#), [113](#), [158](#), [177](#)
 - [17] R. Bashkar. *Dialectic: the pulse of freedom*. Routledge, 2008. [224](#)
 - [18] M. Beaudouin-Lafon. Instrumental interaction: an interaction model for designing post-WIMP user interfaces. In *Proc. ACM Conference on Human Factors in Computing Systems (CHI'2000)*, pages 446–453. ACM Press, 2000. [3](#), [6](#), [8](#), [15](#), [28](#), [29](#), [48](#), [55](#), [62](#), [99](#), [110](#), [131](#), [133](#), [134](#), [145](#), [190](#), [191](#), [195](#), [207](#)
 - [19] M. Beaudouin-Lafon and W. E. Mackay. Reification, polymorphism and reuse: three principles for designing visual interfaces. In *Proc. ACM Conference on Advanced Visual Interfaces (AVI'2000)*, pages 102–109, New York, NY, USA, 2000. ACM. [197](#)
 - [20] O. Beaudoux and M. Beaudouin-Lafon. Dpi: A conceptual model based on documents and interaction instruments. In *Proc. Computers XV Interaction without frontier (HCI 2001 and IHM 2001)*. Springer Verlag, 2001. [85](#), [193](#)
 - [21] O. Beaudoux and M. Beaudouin-Lafon. OpenDPI: A toolkit for developing documentcentered environments. *Enterprise Information Systems VII*, pages 231–239, 2006. [85](#)

- [22] B. Bederson and J. Hollan. Pad++: a zooming graphical interface for exploring alternate interface physics. In *Proceedings of the 7th annual ACM symposium on User interface software and technology*, pages 17–26. ACM New York, NY, USA, 1994. [28](#)
- [23] G. Z. Bedny, M. H. Seglin, and D. Meister. Activity theory: history, research and application. *Theoretical Issues in Ergonomics Science*, 1(2): 168–206, April 2000. [20](#), [21](#), [34](#), [151](#), [152](#), [158](#), [159](#), [166](#), [174](#), [175](#)
- [24] P. Beguin and P. Rabardel. Designing for instrument-mediated activity. *Scandinavian Journal of Information Systems*, 12:173–190, 2000. [54](#), [106](#)
- [25] G. Bell and P. Dourish. Yesterday’s tomorrows: notes on ubiquitous computing’s dominant vision. *Personal Ubiquitous Comput.*, 11(2):133–143, 2007. [52](#)
- [26] R. Berlin and C. N. Klokmoose. Undo in dynamic and distributed user interfaces. In *Proceedings of the Sixth Danish Human-Computer Interaction Research Symposium.*, Copenhagen, Denmark, November 2007. [93](#)
- [27] O. Bertelsen. The activity walkthrough: an expert review method based on activity theory. In *Proceedings of the third Nordic conference on Human-computer interaction*, pages 251–254. ACM New York, NY, USA, 2004. [49](#)
- [28] O. Bertelsen and S. Bødker. Cooperation in massively distributed information spaces. In *Proceedings of ECSCW*, volume 1, pages 1–18. Kluwer Academic Publishers, 2001. [178](#), [185](#), [186](#)
- [29] O. Bertelsen and S. Bødker. Interaction through multi-artifacts. In *Cognition, culture and design: Eleventh European Conference on Cognitive Ergonomics*, pages 103–111, 2002. [104](#), [174](#)
- [30] O. Bertelsen and S. Bødker. Activity theory. *HCI models, theories, and frameworks: Toward a multidisciplinary science*, pages 291–324, 2003. [17](#), [18](#), [19](#), [20](#), [21](#), [101](#), [151](#), [152](#), [159](#), [160](#), [171](#), [178](#), [224](#)
- [31] O. Bertelsen and C. Nielsen. Augmenting reality in mobile substrates On the design of computer support for process control. In *Proceedings of the 22nd Information Systems Research Seminar in Scandinavia*. Jyväskylä University Printing House, Jyväskylä, Finland, pages 123–134, 1999. [7](#), [36](#), [178](#)
- [32] O. W. Bertelsen. Design artefacts - towards a design-oriented epistemology. *Scandinavian Journal of Information Systems*, 12:15–17, 2000. [158](#)
- [33] H. Beyer and K. Holtzblatt. Contextual design: A customer-centered approach to systems designs. *CITY: Morgan Kaufman*, 1997. [125](#)
- [34] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 73–80, New York, NY, USA, 1993. ACM. [28](#)

- [35] G. Bjerknes, P. Ehn, M. Kyng, and K. Nygaard. *Computers and democracy: A Scandinavian challenge*. Avebury, 1987. [10](#)
- [36] S. Bødker. Applying activity theory to video analysis: how to make sense of video data in human-computer interaction. In B. Nardi, editor, *Context and consciousness: Activity theory and human-computer interaction*. Massachusetts Institute of Technology Cambridge, MA, USA, 1995. [16](#), [49](#), [129](#)
- [37] S. Bødker. When second wave hci meets third wave challenges. In *NordiCHI '06: Proceedings of the 4th Nordic conference on Human-computer interaction*, pages 1–8, New York, NY, USA, 2006. ACM. [4](#), [31](#)
- [38] S. Bødker. *Through the interface: A human activity approach to user interface design*. L. Erlbaum Associates Inc. Hillsdale, NJ, USA, 1991. [6](#), [8](#), [11](#), [17](#), [18](#), [21](#), [22](#), [23](#), [26](#), [99](#), [101](#), [103](#), [108](#), [111](#), [129](#), [158](#), [160](#), [161](#), [162](#), [171](#), [174](#), [175](#), [191](#), [214](#), [223](#)
- [39] S. Bødker. *Historical analysis and conflicting perspectives-contextualizing HCI*. Springer, 1993. [127](#), [138](#)
- [40] S. Bødker and P. Andersen. Complex mediation. *Human-Computer Interaction*, 20(4):353–402, 2005. [104](#), [169](#)
- [41] S. Bødker and E. Christiansen. Scenarios as springboards in design. In G. L. S. S. T. W. Bowker, G., editor, *Social Science Research, Technical Systems and Cooperative Work*, pages 217–234. Erlbaum, 1997. [214](#)
- [42] S. Bødker and M. Petersen. Design for learning in use. *Scandinavian Journal of Information Systems*, 12:61–80, 2000. [158](#), [174](#), [177](#)
- [43] S. Bødker, P. Ehn, J. Kammersgaard, M. Kyng, and Y. Sundblad. A UTOPIAN experience: On design of powerful computer-based tools for skilled graphic workers. *Computers and democracy*, pages 251–278, 1987. [10](#)
- [44] S. Bødker, K. Grønæk, and M. Kyng. Cooperative design: Techniques and experiences from the Scandinavian scene. 1995. [134](#)
- [45] K. Boehner, R. DePaula, P. Dourish, and P. Sengers. Affect: from information to interaction. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility*, pages 59–68. ACM New York, NY, USA, 2005. [9](#), [98](#), [144](#)
- [46] J. Bolter and D. Gromala. *Windows and mirrors: interaction design, digital art, and the myth of transparency*. MIT Press Cambridge, MA, USA, 2003. [9](#), [52](#), [98](#)
- [47] N. Bouvin, C. Brodersen, F. Hansen, O. Iversen, and P. Nørregaard. Tools of contextualization: Extending the classroom to the field. In *Proceedings of the 2005 conference on Interaction design and children*, pages 24–31. ACM New York, NY, USA, 2005. [32](#), [149](#), [150](#)

- [48] C. Brodersen and J. F. Kristensen. Interaction through negotiation. In *NordiCHI '04: Proceedings of the third Nordic conference on Human-computer interaction*, pages 259–268, New York, NY, USA, 2004. ACM. 171
- [49] S. Bryant, A. Forte, and A. Bruckman. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Conference on Supporting Group Work: Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. ACM, 2005. 8, 50, 107, 114, 117, 119
- [50] F. Buschmann, R. Meunier, H. Rohnert, P. Sommerlad, and M. Stal. Pattern-oriented software architecture: a system of patterns. 1996. *West Sussex, England: John Wiley & Sons Ltd.* 84, 85
- [51] D. Butler. Electronic notebooks: A new leaf. *Nature*, 436(7047):20–21, 2005. 75, 85, 212
- [52] W. Buxton. There's more to interaction than meets the eye: Some issues in manual input. 1990. 158
- [53] M. Callon. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. *Power, action and belief: a new sociology of knowledge*, 32:196–233, 1986. 27
- [54] S. Card, T. Moran, and A. Newell. The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, 23(7):396–410, 1980. 15, 49, 158, 170
- [55] S. Card, T. Moran, and A. Newell. *The psychology of human-computer interaction*. Lawrence Erlbaum Associates, 1983. 8, 15, 16
- [56] J. Carroll. *HCI models, theories, and frameworks: Toward a multidisciplinary science*. Morgan Kaufmann, 2003. 4
- [57] J. Carroll and M. Rosson. Getting around the task-artifact cycle: how to make claims and design by scenario. *ACM Transactions on Information Systems (TOIS)*, 10(2):181–212, 1992. 48, 92, 146
- [58] J. Carroll, W. Kellogg, and M. Rosson. *The task-artifact cycle*. Cambridge University Press New York, NY, USA, 1991. 48, 92, 108, 146
- [59] M. Chalmers and A. Galani. Seamful interweaving: heterogeneity in the theory and design of interactive systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 243–252. ACM New York, NY, USA, 2004. 46, 99, 106, 143, 146
- [60] M. Cluts. The evolution of artifacts in cooperative work: constructing meaning through activity. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work*, pages 144–152. ACM New York, NY, USA, 2003. 107

- [61] A. Cooper. *The inmates are running the asylum*. SAMS publishing, 2004. [48](#), [92](#)
- [62] J. Coutaz. Pac: An object oriented model for implementing user interfaces. *SIGCHI Bull.*, 19(2):37–41, 1987. [9](#), [59](#), [84](#)
- [63] P. Dalsgaard. *Designing Engaging Interactive Experiences: A Pragmatist Perspective*. PhD thesis, Dept. of Information and Media Studies, Aarhus University, 2009. [9](#), [27](#)
- [64] A. Demeure, J. Sottet, G. Calvary, J. Coutaz, V. Ganneau, and J. Vanderdonckt. The 4C reference model for distributed user interfaces. In *Int. Conf. on Autonomic and Autonomous Systems*, pages 61–69, 2008. [192](#)
- [65] P. Denning. The profession of IT: who are we? *Communications of the ACM*, 44(4), 2001. [55](#)
- [66] P. Dourish. *Where the Action Is : The Foundations of Embodied Interaction*. MIT Press, 2001. [6](#), [8](#), [26](#), [27](#), [31](#), [98](#), [100](#), [106](#), [144](#)
- [67] P. Dragicevic and J. Fekete. Input device selection and interaction configuration with icon. In *Proc. Computers XV Interaction without frontier (HCI 2001 and IHM 2001)*, pages 543–548. Springer Verlag, 2001. [94](#)
- [68] H. Dreyfus, S. Drey-fus, and L. Zadeh. Mind over machine: The power of human intuition and expertise in the era of the computer. *IEEE Expert [see also IEEE Intelligent Systems and Their Applications]*, 2(2):110–111, 1987. [101](#), [107](#)
- [69] A. Dunne and F. Raby. *Design noir: The secret life of electronic objects*. Birkhauser, 2001. [98](#), [127](#)
- [70] A. Eden and R. Kazman. Architecture, design, implementation. In *Software Engineering, 2003. Proceedings. 25th International Conference on*, pages 149–159, 2003. [84](#)
- [71] Y. Engeström. Learning by expansion: An activity-theoretical approach to developmental research. *Oriente Konsultit, Helsinki*, 1987. [10](#), [75](#), [101](#), [108](#), [138](#), [182](#), [214](#), [223](#)
- [72] Y. Engeström, R. Punamäki-Gitai, and R. Miettinen. Perspectives on activity theory. Cambridge University Press, 1999. [27](#)
- [73] A. Fettig. *Twisted Network Programming Essentials*. O'Reilly Media, Inc., 2005. [67](#), [199](#)
- [74] C. Floyd. A systematic look at prototyping. *Approaches to prototyping*, pages 1–18, 1984. [12](#)
- [75] P. Gal'perin. Ya.(1969). Stages in the development of mental acts. *A handbook of contemporary Soviet psychology*, pages 249–273. [7](#), [25](#), [34](#), [36](#), [112](#), [176](#)

- [76] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reuseable Object-Oriented Software*. Addison-Wesley, 1994. [62](#), [195](#)
- [77] B. Gaver, T. Dunne, and E. Pacenti. Cultural probes. *Interactions*, 6(1): 21–29, 1999. [4](#), [125](#)
- [78] W. Gaver. Technology affordances. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, pages 79–84. ACM New York, NY, USA, 1991. [9](#), [23](#), [133](#)
- [79] J. Gibson. *The ecological approach to visual perception*. Lawrence Erlbaum Associates, 1986. [5](#), [8](#), [12](#), [23](#), [27](#), [109](#), [161](#)
- [80] T. C. N. Graham, T. Urnes, and R. Nejabi. Efficient distributed implementation of semi-replicated synchronous groupware. In *Proc. ACM symposium on User interface software and technology (UIST'1998)*, pages 1–10, New York, NY, USA, 1996. ACM. [207](#)
- [81] A. Greenfield. *Everyware: The dawning age of ubiquitous computing*. Peachpit Press Berkeley, CA, USA, 2006. [52](#)
- [82] R. Hartson. Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology*, 22(5):315–338, 2003. [9](#), [100](#), [108](#), [110](#), [117](#)
- [83] R. D. Hill. The abstraction-link-view paradigm: using constraints to connect user interfaces to applications. In *Proc. ACM Human factors in computing systems (CHI'92)*, pages 335–342, New York, NY, USA, 1992. ACM. [9](#), [84](#), [193](#)
- [84] J. Hollan, E. Hutchins, and D. Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2):174–196, 2000. [27](#)
- [85] S. Huot, C. Dumas, P. Dragicevic, J. Fekete, and G. Hégron. The MaggLite post-wimp toolkit: Draw it, connect it and run it. In *Proceedings of the 17th ACM Symposium on User Interface and Software Technology (UIST 2004)*, pages 257–266, New York, NY, USA, 24–27 Oct. 2004. ACM/SIGCHI, ACM Press. [94](#)
- [86] E. Hutchins and T. Klausen. Distributed cognition in an airline cockpit. *Cognition and communication at work*, pages 15–34, 1996. [108](#)
- [87] H. Hutchinson, W. Mackay, B. Westerlund, B. B. Bederson, A. Druin, C. Plaisant, M. Beaudouin-Lafon, S. Conversy, H. Evans, H. Hansen, N. Roussel, and B. Eiderbäck. Technology probes: inspiring design for and with families. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 17–24, New York, NY, USA, 2003. ACM. [13](#)

- [88] L. Hydén. *Psykologi och materialism. Introduktion till den materialistiska psykologin*, 1981. [101](#), [108](#)
- [89] I. Illich. *Tools for Conviviality*. Fontana/Collins, 1973. [192](#)
- [90] E. Ilyenkov. *The Dialectics of Abstract and Concrete in Marx's Capital. Moscow, Progress (translated into English by S. Syrovatkin)*. [224](#)
- [91] H. Ishii and B. Ullmer. Tangible bits: towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241, New York, NY, USA, 1997. ACM. [47](#)
- [92] R. J. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum. Reality-based interaction: a framework for post-wimp interfaces. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 201–210, New York, NY, USA, 2008. ACM. [28](#), [47](#)
- [93] G. Johnson and R. Jennings. *LabVIEW graphical programming*. McGraw-Hill Professional, 2006. [94](#)
- [94] H. Jung, E. Stolterman, W. Ryan, T. Thompson, and M. Siegel. Toward a framework for ecologies of artifacts: how are digital artifacts interconnected within a personal life? In *NordiCHI '08: Proceedings of the 5th Nordic conference on Human-computer interaction*, pages 201–210, New York, NY, USA, 2008. ACM. [5](#), [47](#), [98](#), [104](#), [105](#)
- [95] R. Jungk and N. Müllert. *Future workshops*. Institute for Social Inventions, 1987. [13](#), [75](#), [213](#)
- [96] V. Kaptelinin. Activity theory: implications for human-computer interaction. In B. Nardi, editor, *Context and consciousness: activity theory and human-computer interaction*. MIT Press. [20](#), [102](#), [158](#), [159](#), [160](#)
- [97] V. Kaptelinin. Computer-mediated activity: Functional organs in social and developmental contexts. *Context and consciousness: Activity theory and human-computer interaction*, pages 45–68, 1996. [22](#), [25](#), [106](#), [158](#), [160](#), [174](#), [175](#)
- [98] V. Kaptelinin and M. Czerwinski, editors. *Beyond the desktop metaphor: designing integrated digital work environments*. The MIT Press, Cambridge, Mass., 2007. [9](#), [50](#)
- [99] V. Kaptelinin and B. Nardi. *Acting with technology: Activity theory and interaction design*. MIT Press, 2006. [8](#), [15](#), [16](#), [17](#), [26](#), [27](#), [99](#), [101](#), [110](#), [112](#), [143](#), [145](#)
- [100] V. Kaptelinin, B. A. Nardi, and C. Macaulay. Methods & tools: The activity checklist: a tool for representing the “space” of context. *interactions*, 6(4):27–39, 1999. [49](#)

- [101] D. Karger. Haystack: Per-User Information Environments Based on Semistructured Data. *Beyond the desktop metaphor: designing integrated digital work environments*, page 49, 2007. [9](#), [80](#), [81](#), [93](#)
- [102] B. Karpatschof. Grænsen for automatisering. *En virksomhedsteoretisk bestemmelse af informationsteknologi*, 5:201–220, 1984. [22](#), [216](#)
- [103] A. Kay and A. Goldberg. Personal dynamic media. *Computer*, 10(3): 31–41, 1977. [9](#)
- [104] F. Kensing and K. Madsen. Generating visions: future workshops and metaphorical design. In J. Greenbaum and M. Kyng, editors, *Design at work: cooperative design of computer systems*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, 1992. [13](#), [213](#)
- [105] C. Klokmoose and P. Zander. Five paradigms for manipulation of digital objects. Working paper. [55](#), [80](#)
- [106] C. N. Klokmoose. An instrumental paradigm for ubiquitous interaction. In *DHRS 2006 Proceedings of the Sixth Danish Human-Computer Interaction Research Symposium*, Aarhus, Denmark, November 2006. [12](#)
- [107] G. E. Krasner and S. T. Pope. A cookbook for using the model-view controller user interface paradigm in smalltalk-80. *J. Object Oriented Program.*, 1(3):26–49, 1988. [193](#)
- [108] G. Kurtenbach and W. Buxton. User learning and performance with marking menus. In *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, pages 258–264. ACM New York, NY, USA, 1994. [28](#)
- [109] B. Latour. One more turn after the social turn. *The science studies reader*, pages 276–289, 1999. [27](#)
- [110] A. Leont’ev. The problem of activity in psychology. *The concept of activity in Soviet psychology*, 1981. [99](#), [101](#), [108](#)
- [111] A. Leont’ev and M. Hall. *Activity, consciousness, and personality*. Prentice-Hall Englewood Cliffs, NJ, 1978. [17](#), [19](#), [21](#), [99](#), [101](#), [108](#)
- [112] Y. Lim, E. Stolterman, and J. Tenenbergh. The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. 2008. [12](#), [140](#)
- [113] W. Mace. James J. Gibson’s strategy for perceiving: Ask not what’s inside your head, but what your head’s inside of. *Perceiving, acting, and knowing*, pages 43–65, 1977. [5](#)
- [114] K. Marx. *Grundrisse: Foundations of the critique of political economy*. Vintage Books, New York, 1973. [224](#)

- [115] J. McCarthy and P. Wright. Technology as experience. *interactions*, 11(5): 42–43, 2004. 9, 27, 98, 127, 145
- [116] Merriam-Webster. *Dictionary and Thesaurus*. 2008. <http://www.merriam-webster.com/>. 64, 197
- [117] M. Modahl, B. Agarwalla, G. Abowd, U. Ramachandran, and T. S. Saponas. Toward a standard ubiquitous computing framework. In *Proc. ACM Workshop on Middleware for pervasive and ad-hoc computing (MPAC'2004)*, pages 135–139, New York, NY, USA, 2004. ACM. 82, 193
- [118] P. Mogensen. *Towards a provotyping approach in systems development*. Aarhus University, Computer Science Department. 127
- [119] T. Moran and S. Zhai. Beyond the desktop metaphor in seven dimensions. *Beyond the desktop metaphor: designing integrated digital work environments*, page 335, 2007. 59
- [120] G. Mori, F. Paternò, and C. Santoro. Design and development of multi-device user interfaces through multiple logical descriptions. *IEEE Trans. on Soft. Eng.*, 30(8):507–520, 2004. 192
- [121] B. A. Myers. Separating application code from toolkits: eliminating the spaghetti of call-backs. In *Proc. ACM Symposium on User interface software and technology (UIST'1991)*, pages 211–220, New York, NY, USA, 1991. ACM. 200
- [122] B. A. Myers, J. Nichols, J. O. Wobbrock, and R. C. Miller. Taking handheld devices to the next level. *Computer*, 37(12):36–43, 2004. 81, 190, 192
- [123] M. W. Newman, S. Izadi, W. K. Edwards, J. Z. Sedivy, and T. F. Smith. User interfaces when and where they are needed: an infrastructure for recombinant computing. In *Proc. ACM symposium on User interface software and technology (UIST'2002)*, pages 171–180, New York, NY, USA, 2002. ACM. 9, 81, 192
- [124] C. Nielsen. *Designing to support mobile work with mobile devices*. Aarhus University, Computer Science Department. 171
- [125] C. Nielsen and A. Søndergaard. Designing for mobility—an integration approach supporting multiple technologies. In *Proceedings of NordiCHI*, pages 23–25, 2000. 178, 184
- [126] D. Norman. *The psychology of everyday things*. Basic books New York, 1988. 21, 23
- [127] D. Norman. Affordance, conventions, and design. *interactions*, 6(3): 38–41, 1999. 34, 55, 109, 144
- [128] D. Norman. Emotion & design: attractive things work better. *Interactions*, 9(4):36–42, 2002. 9, 98

- [129] D. A. Norman. Signifiers, not affordances. *interactions*, 15(6):18–19, 2008. ISSN 1072-5520. [9](#), [23](#)
- [130] W. Odom, J. Pierce, E. Stolterman, and E. Blevis. Understanding why we preserve some things and discard others in the context of interaction design. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 1053–1062. ACM New York, NY, USA, 2009. [144](#)
- [131] D. R. Olsen. Evaluating user interface systems research. In *Proc. ACM Symposium on User interface software and technology (UIST'2007)*, pages 251–258, New York, NY, USA, 2007. ACM. [207](#)
- [132] D. R. Olsen, S. Jefferies, T. Nielsen, W. Moyes, and P. Fredrickson. Cross-modal interaction using xweb. In *Proc. ACM Symposium on User interface software and technology (UIST'2000)*, pages 191–200, New York, NY, USA, 2000. ACM. [9](#), [81](#), [192](#)
- [133] T. A. Øritsland and J. Buur. Taking the best from a company history - designing with interaction styles. In *DIS '00: Proceedings of the 3rd conference on Designing interactive systems*, pages 27–38, New York, NY, USA, 2000. ACM. [91](#)
- [134] A. Oulasvirta. When users “do” the ubicomp. *interactions*, 15(2):6–9, 2008. ISSN 1072-5520. [31](#), [52](#), [53](#), [98](#)
- [135] A. Oulasvirta and L. Sumari. Mobile kits and laptop trays: managing multiple devices in mobile information work. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1127–1136, New York, NY, USA, 2007. ACM. [4](#), [53](#)
- [136] L. Palen and S. Bødker. Don't Get Emotional. *Affect and Emotion in Human-Computer Interaction*. LNCS, 4868:12–22, 2008. [144](#)
- [137] R. Pawson. *Naked objects*. PhD thesis, University of Dublin, Trinity college, June 2004. [84](#), [85](#)
- [138] M. Polanyi and M. Grene. *Knowing and being*. JSTOR, 1969. [22](#), [107](#)
- [139] S. Ponnekanti, B. Lee, A. Fox, P. Hanrahan, and T. Winograd. ICrafter: A service framework for ubiquitous computing environments. *Lecture Notes in Computer Science*, pages 56–75, 2001. [9](#), [81](#)
- [140] M. Potel. MVP: Model-view-presenter, the taligent programming model for C++ and Java. *Taligent Inc*, 1996. [9](#), [84](#)
- [141] J. Rasmussen. *Information processing and human-machine interaction: An approach to cognitive engineering*. Elsevier Science Inc. New York, NY, USA, 1986. [108](#)
- [142] J. Rasmussen, A. Pejtersen, and L. Goodstein. *Cognitive systems engineering*. 1994. [108](#)

- [143] T. Reenskaug. Models - views - controllers. Technical report, Xerox Parc, 1979. URL <http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html>. 8, 9, 59, 84, 190
- [144] J. Rekimoto. Pick-and-drop: A direct manipulation technique for multiple computer environments. In *Proc. ACM Symposium on User Interface Software and Technology (UIST'97)*, pages 31–39, 1997. 8, 13, 60, 63, 190, 196
- [145] J. Rekimoto. A multiple device approach for supporting whiteboard-based interactions. In *Proceedings of the 1998 ACM Conference on Human Factors in Computing Systems (CHI'98)*, 1998. 190
- [146] M. Robinson and P. Vorobiev. *Swing*. Manning Publications Co., Greenwich, CT, USA, 1999. 84
- [147] Y. Rogers. New theoretical approaches for human-computer interaction. *Annual review of information science and technology*, 38(1):87–143, 2004. 12, 17, 26, 31
- [148] Y. Rogers. Moving on from Weiser's vision of calm computing: Engaging UbiComp experiences. *Lecture Notes in Computer Science*, 4206:404, 2006. 52
- [149] Y. Rogers and J. Ellis. Distributed cognition: an alternative framework for analysing and explaining collaborative working. *Journal of information technology*, 9:119–119, 1994. 27
- [150] M. Scaife and Y. Rogers. External cognition: how do graphical representations work? *Int. J. Human-Computer Studies*, 45:185–213, 1996. 27, 48
- [151] m. schraefel, G. Hughes, H. Mills, G. Smith, T. Payne, and J. Frey. Breaking the book: translating the chemistry lab book into a pervasive computing lab environment. CHI, 2004. 85
- [152] P. Sengers and B. Gaver. Staying open to interpretation: engaging multiple meanings in design and evaluation. In *Proceedings of the 6th conference on Designing Interactive systems*, pages 99–108. ACM New York, NY, USA, 2006. 144
- [153] B. Shneiderman. Direct manipulation: A step beyond programming languages. *Computer*, 16(8):57–69, 1983. 6, 214
- [154] B. Shneiderman and C. Plaisant. Designing the user interface: strategies for effective human-computer interface. Addison Wesley, New York, 2005. 173
- [155] D. Smith, C. Irby, R. Kimball, W. Verplank, and E. Harslem. Designing the Star user interface. *Human-computer interaction*, pages 237–259, 1990. 9, 80

- [156] R. Smith. Experiences with the alternate reality kit: an example of the tension between literalism and magic. *ACM SIGCHI Bulletin*, 18(4):61–67, 1987. [47](#)
- [157] S. Star. The structure of ill-structured solutions: boundary objects and heterogeneous distributed problem solving. *Morgan Kaufmann Series In Research Notes In Artificial Intelligence*, pages 37–54, 1990. [86](#)
- [158] L. Suchman and L. Suchman. *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press, 1987. [16](#), [27](#), [99](#), [101](#)
- [159] L. Suchman and R. Trigg. Understanding practice: Video as a medium for reflection and design. *Design at work: Cooperative design of computer systems*, pages 65–89, 1991. [125](#)
- [160] A. Tabard, W. Mackay, and E. Eastmond. From individual to collaborative: the evolution of prism, a hybrid laboratory notebook. In *Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 569–578. ACM New York, NY, USA, 2008. [85](#), [212](#)
- [161] P. Tandler. Software infrastructure for ubiquitous computing environments: Supporting synchronous collaboration with heterogeneous devices. In *Proc. International conference on Ubiquitous Computing (UbiComp'2001)*, pages 96–115, London, UK, 2001. Springer-Verlag. [193](#)
- [162] P. Tandler. The BEACH application model and software framework for synchronous collaboration in ubiquitous computing environments. *The Journal of Systems & Software*, 69(3):267–296, 2004. [9](#), [82](#), [93](#)
- [163] R. Trigg, S. Bødker, and K. Grønæk. A video-based analysis of the cooperative prototyping process. *Scandinavian Journal of Information Systems*, 3:63–86, 1991. [151](#)
- [164] P. Turner, S. Turner, and J. Horton. From description to requirements: an activity theoretic perspective. In *Proceedings of the international ACM SIGGROUP conference on Supporting group work*, pages 286–295. ACM New York, NY, USA, 1999. [107](#)
- [165] A. Van Dam. Post-WIMP user interfaces. *Communications of the ACM*, 40(2):63–67, February 1997. [28](#)
- [166] K. Vicente. *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Lawrence Erlbaum Associates Mahwah, NJ, 1999. [145](#)
- [167] L. Vygotskij, A. Kozulin, E. Hanfmann, and G. Vakar. *Thought and language*. MIT Press, 1962. [17](#), [19](#)
- [168] M. Wartofsky. *Models: Representation and the scientific understanding*. Springer, 1979. [223](#)

- [169] T. Wehner, A. Raeithel, C. Clases, and E. Endres. Von der Mühe und den Wegen der Zusammenarbeit. Theorie und Empirie eines arbeitspsychologischen Kooperationsmodells. *Zwischenbetriebliche Kooperation. Die Gestaltung von Lieferbeziehungen*, pages 39–58, 1996. [225](#)
- [170] M. Weiser. The computer for the 21st century. *Scientific American*, 265(3): 66–75, Feb. 1991. [4](#), [9](#), [52](#), [78](#), [189](#)
- [171] M. Weiser. Some computer science issues in ubiquitous computing. *Communications of the ACM*, 36(7):75–84, 1993. [9](#)
- [172] C. Wharton, J. Rieman, C. Lewis, and P. Polson. The cognitive walk-through method: a practitioner’s guide. pages 105–140, 1994. [49](#)
- [173] T. Winograd and F. Flores. *Understanding computers and cognition*. Ablex Publishing Corp. Norwood, NJ, USA, 1986. [8](#), [15](#), [26](#), [27](#), [34](#), [101](#)
- [174] P. Zander. *Collaborative process change by inscription: a contested terrain for interaction designers*. Department of Informatics, Lund University, 2007. [8](#), [17](#), [212](#), [223](#), [224](#), [229](#)
- [175] J. Zimmerman. Designing for the self: making products that help people become the person they desire to be. In *Proceedings of the 27th international conference on Human factors in computing systems*, pages 395–404. ACM New York, NY, USA, 2009. [144](#)