COMPUTERS
AND
INTERACTION

The Social Organisation of
Human-Computer Interaction
in the Workplace

Paul Luff

Department of Sociology
University of Surrey

PhD Submission:
30th April 1997
ABSTRACT

This thesis is concerned with exploring the nature of human-computer interaction. After examining how this is currently conceived within the field of Human-Computer Interaction (HCI) it suggests that it may be worthwhile utilising methodological and analytic frameworks drawn from the social sciences, rather than from psychology. In particular, it explores how work from an ethnomethodological orientation may be a resource for an analysis of the use of computational artefacts. This framework provides a foundation for a series of analyses of four different settings of human-computer interaction: an open-ended 'experiment' of individuals using a computer application; the work of architects using a Computer-Aided Design package; the talk through a computer-supported phone system on a transport network; and the scheduling of trains through a system located in a control room. In each the details of the activities performed on or through the computer are examined. For the analysis particular innovations are made. These provide for an examination of how computer-based activities shape and are shaped by the interactions of participants in a setting. The thesis concludes with a consideration of the relationship between the analysis of activities in the workplace, particularly those centred on the use of technologies, and the design, development and deployment of novel systems. This discussion focuses on how detailed studies of workplace activities can inform the studies of human-computer interaction, collaborative work and processes for system design. Although it appears that specific and generic implications can be drawn from studies such as those undertaken here, it may be that the more important contribution of this work is with respect to the conceptions which underlie the analysis of computer-related activities, whether these be initially considered as collaborative or as interactions between computers and users.
DEDICATION

TO
MY LOVING MOTHER
AND
MY FATHER WHO I MISS GREATLY
# SUMMARY CONTENTS

## Acknowledgements

<table>
<thead>
<tr>
<th>Chapter 1:</th>
<th><strong>Computers and Interaction</strong></th>
<th>a general introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 2:</td>
<td><strong>Technology mirroring the mind</strong></td>
<td>the uses of Cognitive Science in the analyses of computer use</td>
</tr>
<tr>
<td>Chapter 3:</td>
<td><strong>The Practicalities of Menu Use:</strong></td>
<td>improvisation in screen-based activity</td>
</tr>
<tr>
<td>Chapter 4:</td>
<td><strong>Human-Computer 'Interaction' in an Architecture Practice:</strong></td>
<td>observations on computer use and collaborative work in a workplace</td>
</tr>
<tr>
<td>Chapter 5:</td>
<td><strong>Distributed Collaborative Work</strong></td>
<td>The use of the radio phone system in the Docklands Light Railway Control Room</td>
</tr>
<tr>
<td>Chapter 6:</td>
<td><strong>The Collaborative Production of Computer Commands</strong></td>
<td>the interrelationship between activities and talk in a natural setting</td>
</tr>
<tr>
<td>Chapter 7:</td>
<td><strong>Resources for Technological Design</strong></td>
<td>relationships between workplace studies and methods for system development</td>
</tr>
<tr>
<td>Chapter 8:</td>
<td><strong>The Social Organisation of Human-Computer Interaction</strong></td>
<td>a brief discussion of experiences and possible directions of future work</td>
</tr>
<tr>
<td>Appendix A:</td>
<td><strong>Methodological Background</strong></td>
<td>analytic concerns in the examination of naturally-occurring screen-based activities</td>
</tr>
</tbody>
</table>

## References

| 354 |  |  |  |  |
## Detailed Contents

### Chapter 1: Computers and Interaction:
**a general introduction**

1. **Introduction** 1
1.2 **General Background: Emerging Problems within Studies of Computer Use** 3
1.3 **General Themes and Issues** 7
1.4 **Overview Of Chapters** 13

### Chapter 2: Technology mirroring the mind:
**the uses of Cognitive Science in the analyses of computer use**

2.1 **Introduction** 21
2.2 **Problems of Designing Computers for Use** 26
2.3 **The Cognitive Conception of Activity**
  2.3.1 Turing's Model of Computation as the Foundation to Computational Models of the Mind 33
  2.3.2 The Development of a Computational Model of the Mind 36
2.4 **A Cognitive Approach to HCI**
  2.4.1 The Foundations to GOMS: the Model Human Processor and the Keystroke Level Model 40
  2.4.2 The Symbolic Processor 44
  2.4.3 The Parallel Human Information Processor 48
  2.4.4 A General Discussion of GOMS Models 51
2.5 **Distributed Cognition**
  2.5.1 The Conception of Computational Processes in Distributed Cognition 57
  2.5.2 The Conception of Cognition in Distributed Cognition 60
  2.5.3 Distributed Cognition as a Distinctive Analytic Framework 63
2.6 **A Social Scientific Conception of Cognition**
  2.6.1 Situated Action 65
  2.6.2 An Ethnomethodological Orientation 67
2.7 **Discussion** 71
2.8 **Summary** 75
## Chapter 3: The Practicalities of Menu Use:

improvisation in screen-based activity

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>78</td>
</tr>
<tr>
<td>3.2</td>
<td>Background</td>
<td>81</td>
</tr>
<tr>
<td>3.3</td>
<td>Analyses of Menu Use</td>
<td>84</td>
</tr>
<tr>
<td>3.4</td>
<td>An Initial Exercise for Considering Human-Computer Interaction</td>
<td>87</td>
</tr>
<tr>
<td>3.5</td>
<td>A Provisional Typology of Menu Use</td>
<td>92</td>
</tr>
<tr>
<td>3.5.1</td>
<td>The Selection</td>
<td>93</td>
</tr>
<tr>
<td>3.5.2</td>
<td>The Hold</td>
<td>97</td>
</tr>
<tr>
<td>3.5.3</td>
<td>The Pick</td>
<td>101</td>
</tr>
<tr>
<td>3.5.4</td>
<td>The Straddle</td>
<td>104</td>
</tr>
<tr>
<td>3.5.5</td>
<td>The Chug</td>
<td>107</td>
</tr>
<tr>
<td>3.5.6</td>
<td>The Shuffle</td>
<td>109</td>
</tr>
<tr>
<td>3.5.7</td>
<td>The Ramble</td>
<td>112</td>
</tr>
<tr>
<td>3.5.8</td>
<td>The Meander</td>
<td>114</td>
</tr>
<tr>
<td>3.6</td>
<td>Discussion</td>
<td>118</td>
</tr>
<tr>
<td>3.7</td>
<td>Summary</td>
<td>121</td>
</tr>
</tbody>
</table>

## Chapter 4: Human-Computer 'Interaction' in an Architecture Practice:

observations on computer use and collaborative work in a workplace

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>124</td>
</tr>
<tr>
<td>4.2</td>
<td>Background</td>
<td>125</td>
</tr>
<tr>
<td>4.3</td>
<td>Technologies and Tools for Architectural Work</td>
<td>130</td>
</tr>
<tr>
<td>4.4</td>
<td>The Practice Using Menus</td>
<td>134</td>
</tr>
<tr>
<td>4.5</td>
<td>The Manipulation of Objects</td>
<td>136</td>
</tr>
<tr>
<td>4.6</td>
<td>Navigating around the Design</td>
<td>145</td>
</tr>
<tr>
<td>4.7</td>
<td>The Context of Computer Use: Multiple Activities and Collaborative Viewings</td>
<td>151</td>
</tr>
<tr>
<td>4.7.1</td>
<td>Securing an Alignment to the Objects on the Screen</td>
<td>153</td>
</tr>
<tr>
<td>4.7.2</td>
<td>Managing the Alignment Towards Paper and Screen</td>
<td>158</td>
</tr>
<tr>
<td>4.7.3</td>
<td>Readings of the Screen in Interaction</td>
<td>162</td>
</tr>
<tr>
<td>4.8</td>
<td>The Situated and the Cognitive in HCI: The Interactional Context of Screen-based Activities</td>
<td>166</td>
</tr>
<tr>
<td>4.9</td>
<td>Summary</td>
<td>171</td>
</tr>
</tbody>
</table>
Chapter 5: Distributed Collaborative Work:
The use of the radio phone system in the Docklands Light Railway Control Room

5.1 Introduction

5.2 Background

5.3 The General Activities of Controllers on the DLR

5.4 The Organisation of Telephone Calls

5.5 Taking and Making Calls on the Docklands Light Railway

5.6 The Initiation of Radio Phone Calls on the Docklands Light Railway

5.6.1 Summonsing the Other

5.6.2 Producing the 'Reason for a call'

5.6.3 Problematic Cases

5.6.4 Initiating Calls through a Mediating Technology:
    a Summary

5.7 Pauses and Delays - Systematic Resources for the Accomplishment of Activities in Radio Phone Calls

5.7.1 The Second Gap: a Resource for Securing Engagement in the Call and for Prospective Activities in Relation to the Call

5.7.2 The First Gap: a Resource for Distinguishing Calls

5.7.3 Designing Radio Talk through a Mediating Technology:
    a Summary

5.8 The Use of Technology in the Initiation of Phone calls

5.9 Discussion: Interactional Resources for the Production of Distributed Collaborative Activities

5.10 Summary
Chapter 6: The Collaborative Production of Computer Commands:
the interrelationship between activities and talk in a natural setting

6.1 Introduction 231
6.2 Background 234
6.3 The Division of Responsibilities: Scheduling Trains and the Problems on the Line 240
6.4 A Practical Division of Labour: Resources for the Management of Activities within the Control Room 246
6.5 Practices through which Calls are Heard as Relevant 253
6.6 The Interrelationship Between Talk and the Production of Commands 258
   6.6.1 Engagement in Multiple Activities: Cis Monitoring Ciis 258
   6.6.2 Exploring the Scene for Practical Problems: Ciis Monitoring Cis 263
   6.6.3 The Collaborative Accomplishment of a Command 267
6.7 Maintaining the Division of Labour: Explicit Cases of Collaboration 271
6.8 Discussion 276
6.9 Summary 282

Chapter 7: Resources for Technological Design:
relationships between workplace studies and methods for system development 285

7.1 Introduction 286
7.2 Background: Methods for System Design 287
7.3 Naturalistic Approaches: Recent Contributions to the Design Process 291
   7.3.1 Resources for the Assessment of Particular Interfaces 293
   7.3.2 Resources for the Assessment of Deployed Technologies 296
   7.3.3 Particular Resources for the Design of Technologies 298
   7.3.4 Generic Resources for the Design of Technologies 304
   7.3.5 Resources for the Deployment of Technologies 310
   7.3.6 Resources for Supporting the Design Process 313
   7.3.7 Discussion: Naturalistic Analysis and Design Considerations 316
7.5 Conceptual Relevancies 319
7.6 Summary 327
Chapter 8: The Social Organisation of Human-Computer Interaction:
a brief discussion of experiences and possible
directions of future work

8.1 Introduction

8.2 Experiences in the Analysis of Audio-Visual Materials
for the Analysis of Human-Computer Interaction

8.2.1 Data Collection
8.2.2 Data Analysis
8.2.3 Data Presentation

8.3 Future Work

8.4 Summary

Appendix A: Methodological Background:
analytic concerns in the examination of
naturally-occurring screen-based activities

A.1 Introduction

A.2 The Docklands Light Railway

A.3 Problems and Tentative Solutions in the Analysis of
Command Entries Into the ATS
A.3.1 Keystroke Tracking
A.3.2 Train Following

A.4 Example Analysis

A.5 Glossary

References
Acknowledgements

When he was overcome with anxiety he didn't scream. He would sit down at his desk and one by one pull out from his files his research reports, ancient and less ancient, and reprints of his articles, he would rummage in his card index box for the notes he'd been collecting for his lectures or for his planned (and continually postponed) book, he could turn over his ideas and sentences, in order to convince himself, at least of the quantity of his words, by the weight of all that paper that he'd accomplished something.


This thesis is the outcome of several years of part-time work. Because of this it has required the support and tolerance of the various research institutions where I have been employed, the grant-giving bodies for the other research I have been engaged in at the time, and the colleagues with whom I have been collaborating. In this time my research activities have been supported by the Joint Cognitive Science Research Initiative, the EC RACE and ACTS Telecommunication Research Programmes, and Rank Xerox Research Centre, Cambridge (formerly 'Cambridge EuroPARC'). These projects have allowed me the flexibility and time to undertake the work that forms the foundations for this thesis. As well as providing challenging research problems these projects have also brought me in contact with numerous colleagues with whom I could discuss interesting research issues, both related and unrelated to my own concerns.

This is particularly the case with Cambridge EuroPARC where the initial work for this thesis was undertaken. Not only did EuroPARC
provide financial support for the doctoral research, it was an exciting place to work. I would like to thank all those who I collaborated with in the three years I was there, including Victoria Bellotti, Bill Gaver, Richard Harper, Allan Maclean, William Newman, Abi Sellen and Richard Young. I appreciate the various discussions I had with those more directly involved in the project ‘Social Organisation of Human-Computer Interaction’ and the Social Science Group: Bob Anderson, Graham Button, Tom Moran, and Wes Sharrock. Bob and Graham have been particularly helpful and supportive, offering challenging insights concerning the possibilities of relating social science to computer system design, and more generally providing some useful informal tutorials concerning recent arguments and debates in ethnomethodology and conversation analysis.

Early versions of the topics discussed in Chapters 3 and 4 have been published in the ‘Journal of Intelligent Systems’ and ‘Technology in Working Order’ (edited by Graham Button). I thank both the editors of these publications and the anonymous reviewers in their comments on the drafts of these pieces.

Over the last few years I have been fortunate to be associated with a number of academic institutions, including the Department of Sociology at the University of Surrey, the Management Centre at King’s College, London, the Centre for Requirements and Foundations at the University of Oxford and the School of Social Science at the University of Nottingham. These have provided valuable resources and support for my research. Despite these movements, several colleagues in the various guises of the Centre for Work, Interaction and Technology have remained (fairly) constant. These have helped in their discussions of research topics, the consideration of particular instances of data and in collaborating over various of the more mundane work activities. I would particularly like to thank my colleagues David Greatbatch and Jon Hindmarsh of WIT, Agnes McGill at Surrey and Marina Jirotka, Joseph Goguen and Lincoln Wallen at Oxford.

The studies undertaken in this thesis make use of materials gathered in three domains: EuroPARC, an architecture practice and a control room. I would like to thank all those participants either anonymous or anonymised who assisted in these studies. Whilst being asked trivial questions about their activities and being filmed undertaking their work, they showed great patience and tolerance. I have tried to unpick some of
the practices underpinning their activities, but the analyses presented here only touches on the complexity of their skills.

The contributions and assistance of my two supervisors cannot be understated. Whilst I was a member of the Social and Computer Sciences Research Group Nigel Gilbert transformed my initial scepticism for the social sciences into an enthusiasm. Throughout the course of this work he has constantly been available to discuss the issues that have emerged, to direct my focus of attention and to provoke me to address particular arguments. I must also thank him for his unerring attention to detail, particularly with respect to the drafts of the chapters which I presented him in a range of states.

I am particularly indebted to Christian Heath for encouraging me to commence a Ph.D. Throughout this exercise I have maintained my motivation for engaging in this line of research. There is no doubt that my interest has derived from his enthusiasm for all matters in association with these activities, from examining the details of particular instances of data, exploring the implications for new technologies and through to discussing more general debates within the social sciences. He has been a constant source of support, both in helping with the mundane practicalities of getting it done and in offering academic and scholarly insights and encouragement.

The duration of the work could have provided me numerous opportunities to bore a good many friends with its contents. Fortunately, they have been sensitive enough not to allow for this. I am particularly grateful for their continuing tolerance, particularly in not frequently asking about the progress of the thesis. Nevertheless, I am appreciative of those who have been subjected to discussions about my work, including Tom Routen, Michael Smyth, Patrick, Paul, Tim, Chris, Graham, Rebecca, Sarah and Jonathan. Above all I thank my parents, my brother and sister who have been unfailingly patient and supportive, and the only person who has known the trouble this work has been, Marina - a really good friend.
Chapter 1

Computers and Interaction

a general introduction

Some discouragement, some faintness of heart at the new real future which replaces the imaginary, is not unusual, and we do not expect people to be deeply moved by what is not unusual...If we had a keen vision and feeling of all ordinary human life it would be like hearing the grass grow and the squirrel's heart beat and we should die of that roar which lies on the other side of silence.


1.1 INTRODUCTION

This thesis is concerned with exploring the details of the very ordinary activities surrounding the use of computers. Although this domain has been an area of considerable research interest with potentially wide-ranging practical implications, most studies have been constrained by the conceptions they adopt. This is particularly so for the principal field concerned with these studies: Human-Computer Interaction (HCI). Utilising conceptions drawn from psychology, particularly from Cognitive Science, researchers in this field principally consider the activity of using a computer in terms of the mental processing underpinning the
individual's 'interaction' with the technology. In this thesis, drawing from resources from recent developments in the social sciences, the way in which individuals 'interact' with computers is reconsidered.

The studies reported in this thesis examine not only the nature of the activities surrounding an individual's use of a computer system in an experimental domain, but also how technologies are utilised to support activities in natural settings, in workplaces. Even in the analysis of an individual user carrying out a circumscribed task the current conceptions of the activity in HCI may to be too restrictive, and neglect qualities of the activity. When exploring the uses of technology in work settings, the focus of HCI appears to be more problematic. Individuals collaborate with colleagues on and through the technologies, and the systems become a resource for collaborative activities. Such uses have become the interest of the field of research called Computer-Supported Cooperative Work (CSCW). However, characterising the uses of a computational artefact as 'collaborative' may be as vague and unclear as characterising these in terms of interactions. The principal concern of this thesis is in unpicking the ways in which individuals produce and render intelligible activities accomplished on and through computer systems. In doing this, it will address issues of concern to both the fields of CSCW and HCI. However, its principal focus is on details of the use of the computational artefacts, in other words, reconsidering the nature of 'interactions' with computers.

An objective of the analyses is to examine the nature of computer use within the domains under scrutiny. This requires the use of naturalistic materials, principally audio-visual recordings and field notes. Although materials such as video-recordings are frequently utilised in experimental and assessment work within HCI, their use in natural settings is potentially problematic. Indeed the analysis of activities in the workplace can be considered too messy, or too 'noisy', to examine within existing methodological frameworks drawn from psychology and Cognitive Science. Whereas within the social sciences there has been a long-standing concern with the analysis of naturally occurring activities, particularly drawing on fieldwork carried out in work settings. However, even given recent developments concerning the analysis of talk and interaction, resources are not readily available to begin an analysis of naturally-occurring activities in general, let alone human-computer interactions. In undertaking analyses of computer-related activities analytic and methodological innovations are required. Constant concerns throughout
the studies reported here are the nature of the activities examined, the ways in which the analyses can be warranted and the methods through which these analyses are accomplished.

A remaining concern of this thesis is the relationship between the analyses and the design, development and deployment of novel technologies. As is noted in the forthcoming chapters it is becoming recognised that the relationship between studies of human-computer interactions and design is far from straightforward. It may be that more fundamental innovations are required before a more systematic relationship can be developed. Such developments, including novel methods and approaches to design and deployment, will have to rest on an understanding of the ways technologies are used in practice. It is as a contribution to this understanding that the work in this thesis is primarily directed.

1.2 GENERAL BACKGROUND: EMERGING PROBLEMS WITHIN STUDIES OF COMPUTER USE

The fields of HCI and CSCW are multi-disciplinary, involving both academics and practitioners who are concerned with technical and theoretical matters. Although interests in human-computer interaction predate those in CSCW, and even, for some, subsume them, a division has emerged between the two fields: HCI being concerned with the individual computer user; and CSCW with computer systems for groups, or numbers, of individuals.¹ There are many overlaps between the two fields; each member of a group is also engaged in interacting with a computer, and individual computer use often occurs in a collaborative setting. Nevertheless, issues of principal concern to the two fields can be distinguished, HCI focuses on matters of the interface with the user and CSCW with problems such as how items are shared and displayed across various devices in different locations. One concern raised in this thesis is that the distinction between the two areas of interest may not be so straightforward.

Amongst the broad interests of practitioners and researchers within the HCI community are not only those concerned with developing more

¹ For example, see the distinctions made in Dix, et al., (1993) where the two fields are also distinguished by the principal disciplines that inform them: psychology for HCI and sociology for CSCW (p. 423-4).
sophisticated interfaces and devices, but also of understanding the nature of human-computer interaction. Indeed, it is often considered that the design of systems requires a scientific foundation (Barnard, 1991b; Card, et al., 1983). Although a wide range of analytic orientations are utilised within the field, the dominant one is drawn from psychology, and more particularly, Cognitive Science. This focuses on the empirical, where attention can be paid to how activities are accomplished on computer systems, principally through experiments. Cognitive Science also provides an analytic apparatus to consider the results of this empirical work, so that such phenomena as the motor activities of the user, the effects of displaying particular items on the screen in different ways and the use for different kinds of devices can be subjected to scrutiny.

The interaction under consideration within HCI is that between the computer and the individual user. Unlike prior programmes of study concerned with computer use, such as Human Factors and that of the Man-Machine Interface, HCI explicitly aims to address the dynamics through which the system components influence, and are influenced by, the capabilities of the user. Researchers within HCI also attend to the details of computer use, the ways activities on systems are structured, and the ways in which the appearances on the screen and how devices can be manipulated, transform the activity at hand. The Cognitive Science orientation infuses these studies. The tasks an individual accomplishes through the technology are considered in terms of an individual's goals and mental processes, so that, for example a search for an item on the screen is related to an a priori goal; this goal being matched to the displayed objects through the application of particular kinds of knowledge. One objective is to map out the kinds of knowledge that a user has when interacting with a system. This knowledge, at least for the occasion of use, is considered as static and propositional. The resources relied on by users then are principally cognitive ones, specified with respect to mental processes.

It is hoped that these studies will have a practical import, revealing implications for the design of new technologies, so that these systems may be, for example, easier or quicker to use or perhaps more importantly, the

---

2 In the thesis the relationship 'human-computer interaction' will be distinguished from the field of research called Human-Computer Interaction (or HCI) by the use of capital letters.
studies could suggest tools and techniques for system design. However, it is unclear whether the research and proposed methods devised within HCI are having such an influence. Not only do novel computer systems and interfaces appear to have been developed with little attention being paid to studies by HCI practitioners, but HCI methods do not appear to be widely utilised by system designers (Bellotti, 1988). When analyses of particular human-computer interfaces are undertaken these often have little influence on the design of the technology, being performed too late within the development process or not yielding results which are of concern to designers (Barnard, 1991a).

These problems may be due, as Barnard suggests, to the precise nature of the findings that are drawn from HCI studies, or it may be that the cognitive orientation circumscribes too narrowly the domain of study. Recently, several researchers have begun to question the analytic orientation underpinning HCI, suggesting that it should be extended in various ways (e.g. the collection of critiques and proposals in Carroll, 1991a). One common suggestion is that studies in HCI are extended to take account of the social aspects of computer use, particularly the real-world contexts in which it takes place.

A similar set of concerns have emerged from the more applied interests within computer science and software engineering. It is becoming increasingly recognised that there are limitations to current approaches and methods within the system design process, particularly in the early phases of design (see contributions to the collection edited by Jirotka and Goguen, 1994). Although it is essential to be able to implement a system which correctly meets a pre-defined specification, what is supplied in the specification and how this is derived is problematic. These difficulties have become the principal concern of the emerging field of Requirements Engineering. Here too, it appears important to explore the social context of a potential technology (e.g. Goguen, 1994). The current concerns of HCI may suggest some areas of concern for the requirements of computer

---

3 Even when the achievements of applying psychological findings and approaches to human-computer interaction are examined critically, and alternatives are suggestions from within the discipline are proposed, these suggestions are often in terms of transformations to the kinds of relevant results that the approach can generate. For example, after questioning the relevance of particular studies of the use of interfaces and individual competencies for HCI, Barnard (1991a, 1991b) proposes more indirect ways, through intermediary 'bridging representations', by which psychological theories can inform design.
systems, like what can be appropriately presented on a screen, however these appear to be rather limited when the organisational setting in which a technology is to be deployed is considered. Hence, requirements engineers have looked to the possibility of utilising studies of naturalistic settings to inform system design (e.g. Randall, et al., 1994).

There are immediate problems in attempting to extend the domain of interest within HCI to the social and the naturalistic. The analytic framework drawn from Cognitive Science focuses on the details of individual activities in circumscribed domains. It may be possible to gather quantitative data concerning naturally-occurring computer use (e.g. Gray, et al., 1993), and use these to verify theoretical analysis, but in doing this the details concerning the activity are lost. It is also unclear how the social nature of activities, rather than just the dynamic, can be commensurate with a framework drawn from Cognitive Science (cf. Vera and Simon, 1993a). Hence, several researchers have begun to explore possible developments to an individualistic and psychologistic orientation for HCI, drawing from innovations from Russian Psychology like Activity Theory (Nardi, 1996a) or from novel conceptions of cognition, such as those which have become known as Distributed Cognition, for example (Rogers, 1992). What appears to be distinctive about these developments is their utilisation of fieldwork and ethnographies for exploring the use of artefacts in natural settings.

These moves towards the 'social' in general, and ethnographic approaches, in particular, have not only been due to practical and pragmatic difficulties with the application of findings of HCI. Recently, the conceptual underpinnings of the field, those drawn from Cognitive Science, have been subject to a series of critiques. The particular conceptions utilised within the field, such as those of planned and rational action, for example, have been called into question (Suchman, 1987; Winograd and Flores, 1986). These critiques have suggested that not only the scope of Cognitive Science is problematic, but also its theoretical framework.

As a response to these critiques, and also because of the practical concerns already outlined, a body of work has emerged which focuses on the collaborative nature of activities with respect to the use of computers. Although there are other interests within the field, researchers within the field of CSCW have begun to draw on a range of studies of natural settings in order to suggest both the implications for technology and the ways in
which activities in complex technological domains are accomplished. These studies principally utilise ethnographic approaches.

The studies outlined in this thesis seek to contribute to the growing corpus of work within CSCW. However, the focus of these studies remains on the details of computer-related activities, how these are accomplished in natural settings. This focus of interest raises a series of practical and conceptual problems for analysis. These are discussed in the remaining chapters and outlined in the following section.

1.3 GENERAL THEMES AND ISSUES

It hardly needs to be noted that the use of fieldwork for data collection and the adoption of an ethnographic orientation towards the analysis of materials have a long-standing tradition within the social sciences. These have been utilised to address a wide range of empirical concerns within the social sciences and their use has been discussed with respect to a range of theoretical and conceptual issues. Critical amongst these is the framework within which observations made are warranted as relevant for analysis. For researchers in CSCW there are additional problems concerned with how warranted observations can be seen to be relevant for the purpose of design. Needless to say, these problems have up to now been of little concern to social scientists.

The studies in this thesis utilise an analytic orientation drawn from ethnomethodology and conversation analysis. Ethnomethodology takes as a topic the methods by which members accomplish and display their practical reasoning through their actions in social settings (Garfinkel, 1967). As a consequence, the everyday talk and conversation through which members produce and render intelligible these accomplishment and which constitute these methods have been a concern of researchers adopting an ethnomethodological orientation. A body of work, conversation analysis, has emerged that explores the ways in which turns of talk are heard as, and accomplish, understandings through their production in conversation. Conversation Analysts have revealed how turns of talk are sequentially organised, displaying through their production an understanding of a prior turn, and providing resources for a next (Sacks, 1992). The preliminary studies in conversation analysis revealed organisations to naturally occurring talk, between particular adjacent turns of talk (Sacks, 1992), in openings and closings of
conversation (Schegloff, 1968) and in the management of turns (Sacks, et al., 1974). Through these sequential organisations participants display and recognise the appropriateness and relevance of a co-interactants’ contributions. They provide a systematic resource for participants to analyse the conduct of co-participants in the conversation – members’ own methods for displaying an understanding of social actions. These methods, through displayed and hearable talk, are also available to analysts and, hence, sequential analyses can provide a warrant for the relevance of observations concerning social actions in interaction.

Recently, concern has been expressed about the trajectory of these studies and other work in ethnomethodology, because despite their interest in naturally occurring activities, these studies appear to be proposing context-free procedures, rules or methods (Lynch, 1993). They appear to neglect what Garfinkel calls the ‘unique adequacy requirement’ – the indexical and indigenous qualities of particular activities (Garfinkel and Wieder, 1992).

The limits of a too constrained focus on talk-in-conversation has also been a concern of many conversation, or interaction, analysts. Hence, following the early interests of activities in and through conversation, studies have emerged of a range of other domains, particularly in other institutional settings (Boden and Zimmerman, 1991; Drew and Heritage, 1992b). Here, there has been a particular concern in exploring how the sequential organisations of talk-in-conversation permeate these other domains. There has also been an interest in developing analyses in settings where participants have available more than just hearable talk of their co-interactants – in so called face-to-face interaction, interaction where the participants are co-present, for example. Thus, Goodwin (1981) and Heath (1986) have begun to explore how, through the use of audio-visual materials, visual conduct - body orientations, gestures and the use of artefacts - is interwoven with the sequential organisation of talk. These are preliminary investigations in systematically considering a broader analysis of activities-in-interaction.

In seeking to move away from too constrained a focus on talk, there are broadly similar concerns in the attempts to explore the uses of artefacts in interaction and analyses of the uniqueness of settings. There are no doubt other demands of the ‘unique adequacy requirements’ of Garfinkel and the ‘post analytic ethnomethodology’ of Lynch that would appear to be harder
to meet (Garfinkel and Wieder, 1992; Lynch, 1993). However, these arguments may suggest some directions of interest for studies of naturally-occurring activities in technological settings.

Recent work within CSCW has begun to explore how activities are accomplished in a range of domains including: control centres of airports (Goodwin and Goodwin, 1996) urban transportation (Heath and Luff, 1992a), general medical practices (Greatbatch, et al., 1993), financial dealing rooms (Heath, et al., 1994-5) and emergency call centres (Whalen, 1995b). Although such studies have been concerned with the talk and interaction within these domains, they have also paid attention to how artefacts feature in the production of collaborative work, whether these are the public displays in transport control centres (Goodwin and Goodwin, 1996; Heath and Luff, 1992a), or the ‘tickets’ recording deals done in dealing rooms (Heath, et al., 1994-5). These studies make use of audio-visual materials collected in the setting, supplemented by other ethnographic materials; field notes and other documentary materials. These provide an invaluable resource for subjecting to repeated scrutiny the moment-to-moment production of activities in the setting.

Collaborative work in these settings rely on a set of tacit, seen but unnoticed practices through which activities are produced, and seen to be produced, by others. The audio-visual materials make these practices available to analysis. With such materials it is possible to begin to investigate whether participants' display and produce an analysis of the conduct with regard to the artefacts in the domain, and thus whether something akin to the sequential analysis of talk can be available for analysts. To explore this domain requires some innovations in the methodological orientation adopted. It is unclear, for example, in what ways, if any, some artefact-based activity can be said to display an understanding of another. By examining activities in relation to a range of technologies and in a variety of settings the empirical chapters in this thesis explore ways of developing such an analysis of the artefacts-interaction.

These analyses draw on previous work investigating the sequential production of activities in technological settings, principally those outlined

---

4 Particularly those concerning 'ordinary language descriptions' (Lynch and Bogen, 1996 p. 287), whether findings should address social scientific concerns (Lynch, 1993) and the 'just that' in a setting of particular activities (Garfinkel and Wieder, 1992). These concerns are considered in more detail in Chapter 7.
in Greatbatch, et al., (1993), Heath and Luff (1996a) and Whalen (1995b). For example, the study of the activities of members of staff in a London Underground Control Room reveals how the activities of another individual, although apparently distinct, can be recognised by others to be relevant and consequential to their own concerns, and thus can be made use of (Heath and Luff, 1996a). Moreover, participants through their next actions can display an understanding not only of a prior request, but also of the prior conduct of another individual, for example, in relation to the problems they are facing. Such analyses suggest a potential direction for an analysis of activities associated with the use of computers.

In these complex environments, with a range of technologies available to the participants, it is at first surprising that the activities on the computer systems have not provided a domain for considering a sequential analysis of activities. However, in each there are particular asymmetries that make such an analysis problematic. In Whalen's (1995b) analysis of a Computer-Aided Dispatch system the remote caller has little access to the activities undertaken by the dispatcher on the computer; just, on occasions, the sound of typing. Similarly, patients in a medical surgery have little access to the screen of the doctors' computer, hence, they appear to coordinate particular activities with the doctors hand movements and typing on the keyboard (Greatbatch, et al., 1993). In the London Underground Control Room the information assistant and the controller have different responsibilities and different technologies available to them. Not only may it be difficult to see the details of the activities of another, it may be problematic to assess the consequences of the particular details. Hence, though attending to the details of the moment-to-moment use of technologies, in these studies the use of the computer system, particularly screen-based activities are not subjected to scrutiny. In Chapter 7, a particular domain is considered, the use of a scheduling system in the Docklands Light Railway Control Room, where a pair of controllers can both can operate and the view the same computer system. The extent to which this allows for a sequential analysis of activities is discussed.

Although the studies in the thesis do not aim to satisfy the unique adequacy requirement of Garfinkel and his associates (Garfinkel and Wieder, 1992; Lynch, 1993; Lynch and Bogen, 1996), they do seek to explore the details of naturally occurring activities produced on and through computer systems. An ethnomethodological orientation can
provide useful resources for commencing an analysis of the materials relating to computer use. The domains considered vary quite widely with regard to the setting, the technologies under consideration and the contributions participants make with regard to those technologies. However, in all the settings an analytic orientation is adopted that considers the practices through which participants make sense of their own activities and the activities of others. These understandings are accomplishments and rely on everyday social competencies. As such, the socially displayed methods for common-sense understanding and practical reasoning are incommensurate with the private propositions of knowledge and mental processing within the Cognitive Science orientation. These conceptual distinctions are considered further in Chapter 2, in the light of recent developments in Cognitive Science and in Social Science.

The analyses in Chapters 3, 4, 5, and 6 each consider a different set of resources through which participants produce and render intelligible activities in relation to computer systems. These include:

- the readings of menus on the screen and the improvisatory development of courses of activities when individuals use a system in an open-ended exercise involving a computer application to construct presentations (Chapter 3);
- the skilled practices through which architects navigate through and transform drawings using a Computer-Assisted Design Package (Chapter 4);
- the ways in which, through talk and visual conduct, architects coordinate their own activities with those of their colleagues and the happenings on a screen in order to produce intelligible activities, and make sense of actions, with respect to changes being made to a building (Chapter 4);
- how a controller of a transportation system and a party who is out on the network utilise interactional resources to make sense of the contributions of the other through a computer-supported radio phone system (Chapter 5);
- how a pair of controllers make sense of the ongoing activities on the network, in the control room and of their colleagues with respect to the interactions to remote parties, the visual conduct of local participants and the activities on a 'shared' computer system (Chapter 6).
These analyses explore naturally-occurring activities on and through computers as social actions, how they are shaped by the contributions of others and also are shaped by them. Through the studies the ways in which collaborative activities unfold are unpicked and subjected to detailed scrutiny. It would appear that such an analysis could have practical implications for developers of systems to support cooperative work. In Chapter 7 the various resources that the naturalistic analysis of the uses of technologies in workplaces could provide for different design activities are considered. The activities include the assessment of interfaces, the deployment of technologies, and the design of particular and generic systems. The studies undertaken in the thesis are utilised to illustrate these potential resources. It may be, however, that although such implications can be given, studies of the work practices surrounding technologies in natural settings may suggest more fundamental concerns for HCI and CSCW. These studies, and others that explore the details of workplace activities, suggest that certain key conceptions underpinning the study of computer use may need to be respecified.

By focusing on the interaction between a computer and an individual studies in HCI can explore the activities surrounding the use of computers through empirical, and predominantly experimental, means. The focus on interaction suggests a dynamic nature to the relationship between individual and computer, and a conceptual framework drawn from Cognitive Science has provided methods for looking at the details of particular computer-based activities in particular circumstances. Despite the efforts of researchers seeking to develop the framework, it may be difficult for the orientation to provide resources for the consideration of the activities of individuals on and through computer systems in naturalistic settings. In the chapters that follow some suggestions are made concerning an approach that explores the social production of computer-based activities. These activities are irremediably social and collaborative. However, as well as contributing to the study of technology-related collaborative activities, it may be that an analytic orientation that seeks to reveal the order and organisation through which social actions are produced and recognised may also provide the foundation to a systematic study of human-computer interaction.
1.4 Overview of Chapters

Chapter 2 Technology Mirroring the Mind: the uses of Cognitive Science in the analyses of computer use

This chapter explores the background to the Cognitive Science programme and why this orientation appears to be relevant to the analysis of human conduct on computer systems. Drawing upon discussions by Gilbert Ryle (1949) it argues that in using a framework developed with a Cognitive Science orientation, HCI characterises the use of artefacts, the knowledge how, in terms of mental processes, the knowledge that. For Ryle such a characterisation is a 'category mistake'. This chapter outlines how this conception of knowledge and mental processes infuses much work within HCI and suggests how an alternative orientation of human activity could inform the analysis of activities in relation to computer systems.

Cognitive Science utilises a conception of human activity as information processing, and thus has drawn from various work concerned with computation and Computer Science. These are briefly considered. Cognitive Scientists such as David Marr (1982) have distinguished their work from that in, say, Artificial Intelligence, by trying to divide a computational theory of human information processing from the representations and the algorithms that implement that processing. Work within HCI is considered in the light of these requirements. Particular attention is paid to the GOMS family of methods which examine in detail the conduct of individuals using a computer system. It is unclear how these meet Marr's requirements or, indeed, in what sense, in their implementation, they adopt a cognitive model of human conduct. Although limitations of GOMS methods have been discussed in the HCI literature, and several extensions have been proposed, the aim persists of drawing from a Cognitive Science framework to analyse the mental processing which underlies the use of computational artefacts.

Recent work has sought to maintain Marr's requirements whilst extending the scope of study to naturally-occurring activities in organisational contexts. These motivations are typified by Ed Hutchins' (1995) radical proposals for a study of 'distributed cognition'. In a range of studies, Hutchins and associated researchers have explored the use of artefacts in natural settings. It would appear that Distributed Cognition could be relevant as a framework within which to consider the use of
computers. Chapter 2 considers the proposals for a study of Distributed Cognition, both with respect to its original motivations and with regard to how analyses of social actions are accomplished. From this discussion, the conceptions of computation and cognition, utilised within the framework appear to be problematic.

It may be that a more radical approach is required. Rather than extending Cognitive Science to take account of the social, orientations drawn from the social sciences could be utilised to explore human activities related to the use of artefacts. Indeed, drawing from ethnomethodology and conversation analysis, several researchers such as Lucy Suchman (1987) have begun to suggest how such an orientation may be relevant to the analysis of ‘human-machine communication’. Social scientists, drawing from an ethnomethodological framework, and prior work concerned with the philosophy of the mind (Ryle, 1949; Wittgenstein, 1953), have also explored the ways cognition and related notions are conceptualised in psychology, and Cognitive Science (Button, et al., 1995; Coulter, 1979; Coulter, 1989). These considerations may provide a way of respecifying some of the key conceptions within Cognitive Science and HCI. They also begin to outline how an alternative framework could be developed for the analysis of human-computer interaction in natural settings.

The development of such a framework is considered in the central chapters of the thesis. These provide analyses of different kinds of uses of various technologies in several settings. Nevertheless, a common orientation to the resources utilised by participants in their uses of the technology is taken in these four chapters; one which seeks to unpick the practices through which computer-based activities are accomplished.

In their examination of different technologies the four empirical chapters can be seen to address issues of concern to HCI and CSCW: Chapter 3, in what has many similarities to an experimental setting, explores the uses of a menus in a presentation application; Chapter 4 looks at the uses of a Computer-Aided Design (CAD) application in an architectural practice; Chapter 5 considers a computer-supported radio phone system within a transportation control room; and Chapter 6 analyses the use of a scheduling system by two controllers in the same control room. Hence, the chapters can be seen to move from the individual to the collaborative, from an individual user moving through the menus appearing on the screen to the interactions, gestures and other activities
through which a transportation system is managed within a control room. However, the issues that emerge may not be that distinct: the users of the presentation package and the CAD application make use of practical methods of reasoning and skills about what appears on the screen; the activities of controllers are shaped by what appears on the screens in front of them. The distinction between individual and collaborative, between the concerns of HCI and CSCW, may thus be hard to draw.

Chapter 3  The Practicalities of Menu Use: improvisation in screen-based activity

Chapter 3 considers a domain with many resonances with previous work in HCI - an open-ended exercise where an individual is asked to perform some circumscribed task on a computer system. It also looks at a screen-based activity which has been a topic of many studies within HCI - the use of menus. This chapter takes a distinctive approach to this analysis. Rather than stipulate categories of the conduct, such as in terms of searches, browses and goal satisfaction, instances of naturally-occurring menu use are considered. When examined, even in this constrained domain, these uses reveal a complexity that is ignored within current studies within HCI. To explore the nature of these activities, two resources are utilised: the development of a transcription system to support the analysis of the conduct; and a typology of menu uses. This typology is not intended as an account of different user behaviours, but rather to open up for analysis the ways in which individuals make use of the menus. Rather than the activity being characterised in terms of individuals using the menus to match a pre-specified goal, there appears to be a more improvisatory nature to the conduct. Thus, a 'search' or a 'browse' may gloss a range of practices engaged in by the participants. The chapter discusses this analysis in the light of previous studies of menu use, in particular, and those within HCI, in general. Although there have been innovatory attempts to recharacterise the knowledge utilised by the users of systems within HCI, considering these in terms of the knowledge that individuals possess and in terms of mental processing may ignore how these activities are accomplished within natural settings.
Chapter 4: Human-Computer 'Interaction' in an Architecture Practice: observations on computer use and collaborative work in a workplace

The analysis of the uses of menus outlines some of the resources that appear to be utilised by individuals when they 'interact' with a computer system. Chapter 4 explores these in a natural setting – an architectural practice. It appears that architects utilise a range of resources in order to accomplish screen-based activities, including practical methods for making use of the objects and space on the screen, for timing an activity with the pace of the system, and for making use of the organisation of the building. If the focus is widened then screen-based activities can be seen to be immersed within the local interactional context. By analysing one particular instance in detail, a simple movement around the system by one architect, it can be seen that this is shaped by the operation of the technology, the architect's conduct with a colleague, and with his activity on other artefacts. The screen-based activity not only is shaped through an interaction between the participants, but also shapes this activity. Such an activity not only appears to be a case similar to those considered as human-computer interaction, it also appears to be a typical case of computer-supported cooperative work. Examining the in situ accomplishment of naturally-occurring screen-based activities appears to make the distinctions between HCI and CSCW problematic.

In Chapters 5 and 6 attention is focused on activities that have been the concern of studies in CSCW: those which are distributed between individuals who are geographically dispersed and those where the individuals are co-located. These studies are undertaken within the domain of a control room of an urban transportation system, the Docklands Light Railway (DLR).

Chapter 5 Distributed Collaborative Work: The use of the radio phone system in the Docklands Light Railway Control Room

The phone system of the DLR may not be a typical technology considered within CSCW, but it is a technology through which social actions are accomplished in the setting, these activities being critical to the safe running of the service. This system is the focus of the analysis in Chapter 5. When examined in detail the use of the phone system is not straightforward. The system appears to transform the resources utilised
by participants in other interactional domains. The communications technology introduces particular problems for the participants to be managed. These communications are mediated by a computer-supported radio phone system. This system provides additional resources to the participants, however it provides these asymmetrically: the controller being given information about who is calling, and also a choice of when to call the remote party back. The remote party does not have this information or this option. By utilising previous analyses of telephone conversations the activities undertaken through the radio phone conversations are examined in detail. These reveal a sensitivity to the resources available to the other and their circumstances. It also appears that through the delicate management of their talk, the pacing and features like the intonation of the voice, the participants can display this sensitivity. Hence, not only are collaborative activities achieved through the talk on the phone, but the organisation of the activity is managed through the self-same talk. The Chapter concludes with a brief discussion on how the organisation of such distributed collaborative activities may provide resources for those involved within CSCW.

Chapter 6  The Collaborative Production of Computer Commands: the interrelationship between activities and talk in a natural setting

The focus of Chapter 6 is on a locale which has been attended to in previous work within CSCW: the collaborative activities of individuals who are co-located in control rooms of various transportation systems (e.g. Filippi and Theureau, 1993; Goodwin, 1992b; Goodwin and Goodwin, 1996; Harper, et al., 1989; Heath and Luff, 1992a; Heath and Luff, 1996a; Hughes, et al., 1988; Suchman, 1996; Watts, et al., 1996). These studies have been particularly concerned with what has been called the informal work practices of participants in the settings on which their activities appear to rely. The studies have revealed a range of practices including: the peripheral participation in activities, the monitoring of a colleague's activities and the coordination between apparently disjoint activities. Chapter 6 considers such practices in some detail with particular interest in the use of computer systems as a resource for collaborative activity. It appears that the technology, particularly the phone system and the scheduler system displays, can be utilised to make sense of the activities of others, including the happenings on the phone. Moreover, the phone call can provide resources for making sense of what is on the displays. When
the activities of the pair of controllers are considered in more detail it appears that their uses of the various technologies and their interactions with each other and with parties both in the control room and outside may be tightly interrelated. In particular, it appears that the commands entered into the system by one controller are related to the talk and visual conduct of the other controller who is on the phone, and the specific happenings in the phone conversation. Analysis of instances of activities on the scheduler system reveal them to be closely tied to the happenings on the line and this coordination is relied upon by colleagues. Indeed, it is not just the typing that is coordinated, but the particular activity the command is accomplishing. It even appears in some cases that the moment-to-moment production of these commands can serve to shape the talk between the controller and remote party to a call. Hence, the talk to a remote party, the visual conduct within the control room and the activities on the computer system are interwoven; a critical resource to the participants' accomplishments being the sequential organisation of the turns of talk and the demands this places on activities to be produced by the controllers in the control room. Computer-based activities can be seen to be relevant to an ongoing course of action, they can also serve to display that a relevant and appropriate next activity is being attended to.

Chapter 7  

Resources for Technological Design: relationships between workplace studies and methods for system development

The detailed analysis of the use of technologies in several domains suggest a range of implications for the design and deployment of new technologies. In Chapter 7 these are outlined with relation to particular activities in the design process. These implications are outlined in relation to other recent work within CSCW and allied fields of research which have sought to utilise ethnographic materials for design. It is argued that paying attention to the details of the organisation of social actions can reveal issues of interest relating to the design and assessment of specific technologies, generic design issues and the deployment of novel computer systems. Of particular utility appears to be the use of audio-visual materials to support an ethnographic orientation. Together these can provide a resource for a variety of activities within a design process. However, what appears to be of critical importance is the orientation through which observations and analyses are warranted. Although it remains to be seen whether an analytic orientation can provide a warrant for particular design ideas, it may be that by being concerned with the
status given to particular observations, issues of relevance to members in
a setting can be raised. Audio-visual materials make talk and visual
conduct available for analysis. A focus on the interactional
accomplishment of these activities suggests a way in which a systematic
analysis of particular workplace activities could begin to be undertaken,
an analysis that could uncover some of the unique features of a setting.
Although the thesis has not been directed towards design concerns, the
resources that could be derived from such analyses could be relevant to
particular design activities. To illustrate these some examples drawn
from the materials are given. These suggest how specific design activities
could be undertaken which make use of a distinctive approach to the
analysis of activities in the workplace. The orientation through which
these studies are accomplished also suggests some methodological
resources for the analysis of computer-related activities. These resources
could provide the foundation for a novel approach to design which is not
cumbered by the constrained conceptions utilised within HCI or the
broad notions adopted by researchers in CSCW.

**Chapter 8   The Social Organisation of Human-Computer Interaction: a
brief discussion of experiences and possible directions of
future work**

The thesis concludes with a few remarks concerning the approach taken to
the materials and the problems which emerged when these materials were
subjected to scrutiny. The work suggests several issues, both practical
and academic, which arise from the analyses. These relate to how the
data are collected, analysed and presented, particularly in relation to the
techniques and technologies that are currently available. Consideration of
these suggest some areas of future research, both with respect to the
particular studies undertaken, and the ways in which the analyses were
accomplished, and with regard to more general issues that relate the
studies of workplaces activities to the design, development and
deployment of new technologies.

**Appendix A   Methodological Background: analytic concerns in the
examination of naturally-occurring screen-based activities**

Given the complexity of the materials in the analysis of the use of the
systems in the DLR Control Room, an appendix is provided that outlines
some further background to the activities in the setting. In particular, to
assist with a deeper understanding of the occurrences in the fragments of
naturally-occurring screen-based activity, a glossary of terms, a list of abbreviations and some more details concerning the operation of the service are given. In addition to these, two examples of ways in which the audio-visual materials have been analysed in order to provide an account of the happenings on the scheduler are outlined. These are illustrated by an example breakdown of the activities on the line in a fragment considered in Chapter 6.

In sum, this thesis seeks to explore the social organisations underpinning human-computer interaction; the methods through which members make sense and display an understanding of the use of computational artefacts. Through detailed examination of particular instances of computer use it explores how these activities are thoroughly embedded within, and constitute, the context in which they emerge. Subjecting such activities to scrutiny forces a re-examination of how the uses of technologies are studied, in particular it offers an alternative way of considering how individuals interact with computers.
Chapter 2

Technology mirroring the mind:

the uses of Cognitive Science in the analyses of computer use

The picture which holds traditional philosophy captive is that of the mind as a great mirror, containing various representations — some accurate, some not — and capable of being studied by pure, non empirical methods.

(Rorty, 1980, p. 12)

Mental models seem a pervasive property of humans. I believe that people form internal, mental models of themselves and of the things and people with whom they interact. These models provide predictive and explanatory power for understanding the interaction... These models are highly affected by the nature of the interaction, coupled with the person’s prior knowledge and understanding. The models are neither complete or accurate..., but nonetheless they function to guide much human behavior.

(Norman, 1986, p. 46)

2.1 INTRODUCTION

Associated with recent innovations in computer technology, and their widespread deployment, have been a range of novel problems for system
designers. Although concerns that the developed systems match their specification, and that the systems' operations can be validated and verified, remain, attention has begun to focus on how to make the technology relevant and appropriate for its users. One particular field has emerged which has centred on these concerns. With the active involvement of academic and commercial participants, Human-Computer Interaction (HCI) has both sought to develop theories of computer use and to provide practical guidance for system design and evaluation. Under the rubric of HCI a large body of work has emerged by participants from many disciplines, exploring such disparate interests as the use of computers within organisations (March, 1991), the semiotics of computer interfaces (Keeler and Denning, 1991) and the mathematical specification of interaction with computers (e.g. Dix, 1995). However, psychology, and in particular, Cognitive Psychology and Cognitive Science, has formed the foundation to most of the work in the field.¹

It has been an objective of these studies of individual competencies to provide explanatory models of computer use from which procedures, methods, or guidelines for design could be drawn. Thus, considerable attention has been devoted to developing a cognitive framework for understanding human-computer interaction. Unfortunately, this approach has recently been subject to severe criticism, which has not only questioned the practical results it has offered designers (e.g. Carroll, 1991a), but also its theoretical and methodological underpinnings (e.g. Suchman, 1987). Thus, it has become unclear how the field of HCI should develop, whether existing approaches should be enhanced or extended, for example, or whether a novel conceptual framework for the study is required. These difficulties have been exacerbated by further innovations in technology, the emergence of fields within system design exploring the relationship between individuals and technologies in differing ways, and the interest in alternative frameworks and methodologies for the study of computer-related activities. In particular, the approach drawn from

¹ There are other programmes of research that examine 'technologies-in-use' which will not be considered in this chapter. These have been less concerned with the design of systems and more on exploring such aspects as the organisational change induced by the introduction of technology, the 'discourses' in which technologies are part, and the various ways in which technology are 'socially constructed'. For brief reviews of these diverse approaches to the study of computer systems see Cooper (1991), Cooper and Bowers (1995), Hine, et al. (1995), Murray and Woolgar (1991) and other contributions to Thomas (1995).
Cognitive Science has been criticised because it concentrates on the performance of a limited set of activities by individuals in experimental settings. This account of human-computer interaction has been considered to be unduly constraining.

This chapter examines the cognitive foundations of the study of human-computer interaction, outlining why this approach appears to be appropriate for the study of computer-related activities, and relevant for the design of novel systems (section 2.2). For this, some background concerning the particular account of cognitive activities embodied within Cognitive Science is required (section 2.3). Drawing on the programme set out by Marr (1982), the 'information-processing' model of human conduct is set out, where within different levels of analysis, the computational level is given primacy, for it provides the potential to provide explanatory theories of competence.

The information-processing approach, and the symbol-processing approach, both take a particular account of mental processes; that these are inner processes which can be modelled principally by techniques developed first in mathematics and then in what became known as Computer Science. Though this view is attractive, it is not an uncontroversial account of the nature of mental processes. Within philosophy there have been long-standing debates concerning the nature of the mind and mental processes, and recently, social scientists have drawn upon these discussions to consider the analysis of social actions and human conduct. The consideration of models underpinning HCI in undertaken in the light of these concerns.

In this chapter, the information processing approach is considered with respect to two widely regarded developments concerned, in different ways, with developing cognitive accounts of the use of artefacts. Both have attracted attention of those seeking to understand how individuals use computers within HCI and CSCW. The GOMS family of models within HCI seeks to explicitly model an individual's use of a computer system. Distributed Cognition aims to provide an account of cognitive behaviour as a social or group process, a particular focus being on communicative activities surrounding the use of artefacts.

The GOMS family of models explicitly adopt a cognitive orientation to behaviour based on information processing. GOMS has evoked considerable interest within the HCI community as, for example,
the best established HCI user model, at least in terms of the body of supporting research and is illustrative of what can be done with user models

(Brooks 1991, p.55)

and the most 'notable example' of theoretical concepts being exploited within the HCI research community (Barnard 1991a, p. 103).

Not only are the GOMS models examples of models developed within the Cognitive Science paradigm to HCI, but they continue to be focus of research. Hence, they are frequently utilised within HCI textbooks as examples of theories of human-computer interaction (e.g. Dix, et al., 1993; Schneiderman, 1992). Given their longevity within the HCI literature, it is not surprising that the GOMS models have also been subject to criticism, review and proposals for extensions (e.g. Carroll, 1991; Kieras, 1988; Olson and Olson, 1990; Gray, et al., 1993). However, despite criticisms of their scope and how they utilise particular findings from Cognitive Science they continue to be regarded as a significant framework for considering computer use as a cognitive activity. In particular, they provide a way of circumscribing the tasks individuals accomplish on a computer system which allows for a mentalistic account of the activity. Section 2.4 considers the GOMS models suggesting some of the strengths and limits of the approach. Such modelling exercises allow for the detailed analysis of a computer user's conduct which has aided the evaluation of computer systems. However, in these exercises the extent to which the derived model can be claimed to be cognitive is unclear. By focusing on the observed activities of individuals, it is often unclear how their 'mental' nature can be accounted for. The models, hence, fall short of being explanatory theories of competence - a requirement of Marr's models of information processing.

Recent work has sought to address the problem of developing an explanatory theory of the competencies underpinning artefact use both by returning to the programme set out by Marr and by extending the domain of interest usually considered by Cognitive Scientists. Distributed Cognition, principally developed from the work of Hutchins (e.g. Hutchins, 1995), aims to provide a computational, explanatory theory, but where cognition is conceived of as a social process. Because of these developments, Distributed Cognition has been of recent interest to those involved in the development of computer systems. It seems to offer a broader framework to conventional HCI approaches whilst maintaining a concern with individual cognitive processes. As it has attracted the
attention of researchers in both HCI and CSCW, its innovatory approach to the study of artefacts is considered in Section 2.5.

Distributed Cognition, whilst being concerned with social and collaborative activities, and with understanding the cognitive properties of groups, maintains a conception of individual mental activity, or of individual 'minds'. Distributed Cognition being a single framework for considering symbolic, and thus cognitive, manipulations both within and between minds. This places considerable emphasis on the nature of symbol processing embodied in the framework. Section 2.5 considers how the possibility of Distributed Cognition providing a framework for understanding the social organisation to the use of artefacts and whether social conceptions of cognition and the individual conceptions of mind can be reconciled in this way.

In section 2.6, the account of cognition, and of individual 'minds', which imbue Cognitive Science and also are maintained within Distributed Cognition are called into question. Utilising critiques of researchers from an ethnomethodological orientation, informed by the work of Wittgenstein and Ryle, the status of such a categorisation of activities as 'mental processes' and 'of the mind' is challenged. These general critiques of mentalism replace its conception of behaviour with one directed at the social organisation of human activity. Moreover, within these reconceptions specific explicit arguments have been made concerning the programme of Cognitive Science (e.g. Coulter, 1992) and also to the cognitive orientation to HCI (Suchman, 1987).

This questioning of the cognitive approach is more acute than just a concern that activities other then the use of the computer, or different kinds of knowledge, or the participation of others, are not accounted for. The critiques challenge the status and nature of cognitive activities embodied within Cognitive Science, and suggest an alternative conception of mental activities: one which sees these in the ways they are used in practice: as accounts of social accomplishments.

These criticisms have been influential, as they suggest an alternative domain of interest for those concerned with the nature of computer-centred activities: on the in situ accomplishment of social actions. Hence, a growing corpus of studies have emerged drawing on an ethnomethodological orientation, and utilising fieldwork and other naturalistic materials to reveal the organisation of everyday practices. These studies are forming one foundation to research in the developing
field of Computer-Supported Cooperative Work (CSCW), involving the
detailed examination of the use of computer technologies within various
settings. Studies of the accomplishment of artefact-centred activities are
still in their preliminary stages. However, in section 2.7 some of the
issues that are already beginning to emerge from these studies are
discussed. These form the foundation to the analyses in the following
chapters.

2.2 PROBLEMS OF DESIGNING COMPUTERS FOR USE

Amongst the multifarious problems facing developers of computer systems
are those associated with designing a technology that is appropriate to the
tasks, activities and work of the individuals who are going to use it. Over
the past twenty five years or so, the nature of the activity to meet these
requirements has been expressed in a variety of ways; in terms of how
'easy' systems are to use, their acceptability, or their 'user friendliness'.
However, expressing the problem in these ways, offers little to assist
developers of new technologies. At the very least, what is required are
guidelines, procedures and methods for system design, and often
developers appear to be demanding a more rigorous approach, something
more akin to an engineering discipline, or a 'science of design'. The
foundations of any such discipline rest on ways of conceiving the
fundamental relationship between individuals and the activities they
undertake on computer systems.

There have been many proposals for a systematic way of conceiving of
this relationship, including that of a dialogue between 'man' and machine
(Martin, 1967), as a conversation between human and computer
(Nickerson, 1977; Reichman, 1986) or even as 'man computer symbiosis'
(Licklider, 1960). However, more recently the relationship has been
considered more to be an interaction between human and computer (e.g.
Card, et al., 1983; Monk, 1985). This may be because this term appears to
place less emphasis on the linguistic aspects of computer use that
dominated early systems, whilst still allowing for a consideration of the
problems of both the computer and human side of the interface. The
particular advantages and disadvantages of this term can be debated.

2 Anderson et al. (1993) discuss the motivations for a Science of Design for
computer systems in the light of Simon's (1969) suggestion of studying the world
in terms of 'the Sciences of the Artificial'.

26
Nevertheless, a field has emerged called Human-Computer Interaction (or HCI) with dedicated journals, conferences, and other fora and academic courses and professional positions being devoted to its study and practical application.

Perhaps naturally, the initial concerns of the study centred on assessing current technologies. Using a range of techniques based on ergonomics and human factors, it was possible to set up experiments and trials, or simply assess a system against some well-known findings (e.g. Cakir, et al., 1980; Rubenstein and Hersch, 1984). These could identify problems occurring with the use of particular systems, some of which were quite basic. For example, research could identify potential problems with the use of colours on computer displays and the nature of colour blindness. Experiments and trials could suggest where users appeared to be having difficulty accessing particular functionalities, or where there appeared to be confusing features of the interface. However, despite such evaluations becoming considered an essential part of the testing of the system, several concerns began to emerge with the 'human factors' approach to the study of human-computer interaction. First, the findings were often too specific to the system under scrutiny; identifying particular problems with a particular interface. Second, although some recommendations for changing the system could be made from the evaluation, it could be difficult to provide a strong warrant for these. A correction might be made, but this could interfere with other previously unnoticed factors. Third, and related, it was difficult for the evaluation exercises to have any significant influence on the development of the system. Once the system has been developed to the extent of being possible to test, the nature of the changes that could be made in response to the evaluation were unduly constrained.

In the light of these clearly identifiable problems, several proposals for alternative ways of designing systems emerged. Not least amongst these are changes in the design process itself. With a more iterative approach, rather than the step-wise progression through various stages, there would be more opportunities for the assessment of prototype systems and thus more possibilities for influence the emerging design. However, changing the design process would not be sufficient. For there to be a genuine interest in usability from the outset of the development of a computer system, a more generic approach to the analysis of the activities associated with the use of computers is required.
One way of addressing this is to provide general prescriptions and guidelines. Some of these are very general: for computer systems to ‘fit the user’ or for designers of systems to ‘know the user’ (Hansen, 1971). Effectively, such prescriptions are conceived in individualistic terms. Thus, attempts at analysing computer use have involved the examination of such features as the readability of text on a display, the motor skills required to enter information using various devices and the syntactic structure of computer commands (e.g. Card, et al., 1978; Hammond and Barnard, 1985; Marcus, 1982). Although from such studies it has appeared possible to derive precise recommendations for designers, their general applicability to other tasks, systems and devices is unclear, and the nature of principles, guidelines and heuristics to be given to designers being the subject of regular debate (Barnard, 1991a; Gould and Lewis, 1985; Norman, 1986). A methodological requirement has emerged; it is not only necessary to evaluate the various guidelines and procedures, but also to provide some warrant for the claims made on behalf of these prescriptions for design. This requires a general framework with which to consider the relationship between individuals and computer systems. From this analytic framework it should be possible to warrant methods for studying the use of computers and determine relevant criteria with which to evaluate the use of systems. The framework should also provide a foundation for considering the human activities which could be supported by a new technology, thus providing resources for the analysis of requirements and the design of systems. Hence, a science of design is being proposed, against which a system’s appropriateness, relevance and usability is considered throughout the design process. This applied programme has been known as ‘User Centered System Design’ (or UCSD, Norman and Draper, 1986), ‘User Centered Design’ (Landauer, 1995), or even more simply ‘User Science’ (Moran, 1985-1997).

Given mottoes, programmatic and terminologies, extant in the various fields of computer system design, it is perhaps not surprising that an appropriate foundation for an analytic framework should be sought in Cognitive Science. Such a framework would provide for analyses of an individual’s use of a computer system in terms of mental processes. Once these processes were explicated, in terms of findings from experiments, models or theories, then, in some way, this ‘knowledge of the user’ could be utilised for system design. Although the precise nature of the interrelationship between studies of the use of computer and system
design has been the subject of debate (e.g. Barnard, 1991a; Carroll, 1990). Cognitive Science offers the possibility that recommendations, guidelines or designs could be both empirically and theoretically grounded.

One formulation of this relationship is in terms of the development of a design model of a system that is informed by, is consistent with, or reflects a 'user model' (a mental model of the user, Norman, 1986). Given this formulation, the preliminary exercise for students of HCI is to determine a model of computer use utilising the understanding of mental activity and mental processes current in Cognitive Science. The problem of designing usable systems is then mapping tasks performed on the interface to mental processes in the head. Hence, much research in HCI has moved away from the direct concerns of design and towards developing a cognitive framework for analysing computer-based activities; a framework that has a scientific basis (Barnard, 1991b).

As a straightforward explication of the relationship between an individual's actions and the computer system, Norman conceives of the use of a system in terms of two gulfs: the gulf of execution and the gulf of evaluation (Norman and Draper, 1986). These are gulfs between the goals of the user and his or her physical activities, and between the production of actions and their interpretation, respectively. Norman breaks down intermediate mental processes which can be assumed to be undertaken within these gulfs (see figure 2.1). These include generating interpretations of perceptions and specifying productions for action from intentions.

![Diagram](image)

**Figure 2.1** Human-computer interaction (after Norman, 1986, p. 42)
So this is a model of a cycle of activities: users interpret what is on the display with respect to their expectations and goals, and generate specifications of actions in order to perform commands on the interface, thus, changing what is on the display. Through the repeated cycling through of this combination of physical and mental activities an interaction with the computer system emerges. One aim of HCI then is to refine this model, particularly breaking apart what happens when a user is interpreting what is perceived on the system and specifying what actions to carry out next. These problems are seen as cognitive problems requiring an analysis of the knowledge and mental processing of users. If a better understanding of these activities is achieved, particularly of the user's understanding of computer systems then it should be possible to inform aspects of systems design. Norman refers to this relationship in terms of interwoven and embedded relationships between various models of the system; an understanding of the user's model of the system informing the designer's model of the system, which either explicitly or implicitly includes a model of the user.3 The general aim of designers then should be develop systems which reflect or resonate with the users' cognitive processing. Borrowing Rorty's (1980) ocular metaphor, these studies of HCI are attempts at revealing the mind as a 'mirror of the technology'.4

There is no doubt that examining human-computer interaction in this way adopts a particularly constrained view of the activity; focusing on an individual user tied into a continuous activity with a machine. But, it is hoped that this focus will allow for a consideration of the critical aspects of computer use. It moves away from previous concerns with the interface, what and how objects are presented, and towards addressing the

---

3 Norman (1983) describes the project more formally. The study relates the use of a technology (t) in terms of the user's mental model of the technology, M(t), to a researcher's conception of that model, C(M(t)).

4 It has been suggested that the constraints of technological design may also provide a focus for studies of cognition. The relationship between studies in HCI and Cognitive Science is not just conceived as one of applying existing analyses and findings. The use of a computer also offers a complex domain of study for activities considered to be cognitive. Indeed, one of the limitations of work taking a Cognitive Science orientation has been the limited nature and relevance of findings from the current state of research in the field (Carroll, 1991b). Hence, HCI has also been seen as a suitable application domain where Cognitive Science can be extended (Card, et al., 1980).
interaction with a computer. This is conceived not only in temporal terms, as a shift to examining sequences of computer-based activities, but more importantly, in terms of the cognitive resources individuals utilise for operating a computer. Thus, research in HCI that draws on a cognitive analytic orientation seeks to reveal the types of knowledge required by users, the relationships between their goals and actions, the constraints of their capacities to remember and their capabilities to accomplish such activities as searching, problem solving and deciding what to do next. What Cognitive Science provides to the analysis is a common model for such diverse resources, capacities and activities: an information processing model of human conduct. Some of the key features of this model are given in the next section. As Norman’s model suggests these concerns are with the hidden and private world of the user. The cognitive capabilities are only being revealed through a careful categorisation of particular behaviours, most commonly by experimentation and studies of highly constrained activities. This relies on being able to break apart the cognitive domain. It also rests on the general conception of activity in such terms as knowledge, memory, goals and information as being unproblematic; an issue returned to in section 2.6.

2.3 THE COGNITIVE CONCEPTION OF ACTIVITY

The single most influence on psychologists' ideas about cognitive processes at present is the nexus of concepts which has developed around computer programming. (Dodwell, 1971, p. 370)

The goal of cognitive science is to explain how the mind works. Part of the power of the discipline resides in the theory of computability. If an explanation is computable, then prima facie it is coherent, and does not take too much for granted. (Johnson-Laird, 1988 [1993], p. 26)

The tools of cognitive science consist of a set of analysis techniques and a set of theoretical formalisms. The analysis techniques include such things as protocol analysis, discourse analysis, and a variety of experimental techniques developed by cognitive psychologists in recent years. The theoretical formalisms include such notions as means-end analysis, discrimination nets, semantic nets, goal-oriented languages, production systems, ATN grammars, frames etc. (Collins, 1975)

Grudin (1990) gives a broad critique on taking too narrow a focus on the interface between a user and a system. Some of his concerns with wider issues related to interfaces between computers and organisations and groups of individuals will be returned to later in this thesis (particularly chapter 7).
Cognitive Science is a multi-disciplinary research programme, drawing mainly from Cognitive Psychology and Artificial Intelligence, but also involving researchers from other subjects such as linguistics, philosophy, biology, computer science and engineering. The one feature that can be seen to unify this research programme is an information processing model of human action where mental processes are modelled in computational terms. As Marr (1982) states this is not to say that the brain is a computer, but to view the brain in terms of 'information processing devices' (p.361). Although Cognitive Scientists vary with respect to the strength of the relationship they propose between computers and the mind, whether, for example, information processing can be seen as a metaphorical or a scientific explanation, there are some basic and common features in this programme.

First, the interests of Cognitive Scientists centre on aspects of human conduct that can be considered as mental processing. They also suggest that this processing can be broken down into a set of individual modules, tasks or theories. The number, nature and interrelationship between these processes vary between researchers. So in one categorisation there are such 'central processes' as memory, learning, problem solving, decision making, attention, search, scanning and the perception of time.6 In another the 'mind's main tasks' are perception, learning, memory, motor action, deduction, creation, communication, and the creation of feelings, intentions and self-awareness.7 There are doubtless interrelationships between the various processes, but for analysis, particular kinds of mental activity can be a focus of examination.

Second, following Chomsky's (1965) distinctions, Cognitive Scientists do not necessarily seek a performative explanation of human conduct, but rather aim to account for competencies (e.g. Marr, 1982). It may be that models of mental processing could be developed into computer programs, as in the case of artificial intelligence, however, what is required are not programs displaying particular solutions to problems, but more general theories and explanations.

---

6 A list derived from Ergonomic Abstracts (from Schneiderman, 1992, p.24)
7 From Johnson-Laird, 1988 [1993], p. 27.
A cognitive theory should be like a computer program. That is, it should be a precise specification of the behavior, but offered in terms sufficiently abstract to provide a conceptually tractable framework for understanding the phenomenon. (Anderson, 1980 p.11)

Third, and perhaps the most distinctive feature of the work of Cognitive Scientists, is the nature of these explanations. The accounts that they give are in terms of computational models;

A central tenet is that there exists a level of description of intelligent systems at which the organisation and use of knowledge is described functionally in computational or information-processing terms independently of the physical implementation of the system. (Bersen, 1988)

Although there is an emphasis in much of the work in Cognitive Science on the inputs, outputs and symbolic representations required in the processing, there has also been a concern with adopting a deeper model of computation, one which utilises the original conceptions from computer science (e.g. Johnson-Laird, 1988 [1993]; Marr, 1982). Hence, it is worth very briefly reviewing the account of computation developed within mathematics utilised for the development of computer systems, particularly Turing's work on 'Universal Machines'.

### 2.3.1 Turing's Model of Computation as the Foundation to Computational Models of the Mind

Computer systems developed from efforts to define general procedures for solving mathematical problems and generating mathematical truths. In particular, it was the insights of Turing (1937) that led to the conception of an abstract machine to undertake these procedures. The 'Turing Machine' is a notional machine used as a way of outlining a general theory of computation. Its exposition provides a way of conceiving of computable problems by means of a universal computational machine. The machine is a device that 'reads' and 'writes' marks onto an infinitely large external storage medium. Turing suggested that this storage medium could be considered as an infinitely long piece of tape that is moved backwards and forwards through the reading and marking device. General computations could be performed by reading instructions and data from, and writing results out to, the relevant places in the tape. The state of the machine is

---

8 The precise nature of the influence of Turing's theory of computable functions on the development and construction of digital computers has been called into question (Wang, 1957).
defined by the marks and the current positioning of the tape. Because the tape is infinitely long, the Turing Machine can be in an infinite number of states. Furthermore, conceiving the device as a very simple reading and marking machine means that no distinction is necessary between instructions on the one hand and inputs and outputs, data and results on the other: marks written as data can themselves be read later as instructions. This simple, but powerful, conception formed the foundation to a general theory of computation, on which was built the later development of computer systems; with storage, reading and writing implemented as hardware; and input, output and instructions instantiated in various levels of data and software.

That the conception of a general computational machine could also be a way of conceiving of the mind and mental processes was suggested by Turing (1950) himself, and this suggestion forms the basis of Cognitive Science and allied work within Artificial Intelligence. This account of computation provides resources for models of cognition; for new distinctions which could inform the study of mind, for ways of conceiving how mental processing was accomplished and for bypassing long-standing debates in the philosophy of mind.

First, there is a distinction between the device and the symbols, later implemented as the hardware and software, the computer and the computer program. This offers an analogy to that between mind and body, or more accurately the mind and the brain. Software, or at least algorithms, can be considered separable from the physical devices for which they are devised. Thus, a 'new reaction to dualism' could emerge which considered the relationship between mind and body as analogous to computer and program (Johnson-Laird, 1988 [1993]).

Second, the possibility that output from the computer can also be its instructions leads to the description of computation at different levels. Marks, 'symbols' and programs can each be read off as data and processed. A single model of computation can then be utilised to instantiate descriptions of a range of capabilities, for example: aspects of visual perception can be accounted for in terms of processing arrays of marks; recognition of objects as the processing of symbols; and comprehension and understanding as the processing of rules. This has lead to the emphasis

---

9 In Cartesian terms, the hardware appears to be an analogy to the extended mental 'substance' where consciousness, a series of events, takes place (cf. Rorty, 1980).
on the processing of symbols within Cognitive Science. Although there are departures concerning whether explicit symbol processing has to be accounted for (Agre, 1988; Chapman, 1991), work in Cognitive Science pays considerable attention to the ways the information being processed is represented.

Third, the simple mechanical processor of the Turing Machine, that reads marks from the tape, moves the tape and then writes further marks, offers an analogy to the 'Inner Eye' or the 'eye of the mind' postulated by earlier philosophers before which pass thoughts and pains. However, that conception was always prone to arguments of infinite regress; if an inner eye is required, what watches over and controls it. The implementation of devices which were instantiations of the Turing Machine suggested a way of avoiding this problem. For some psychologists, computational models and computational devices offered ways by which mental objects and mental processes could be revealed without prior postulation of epistemological distinctions about the nature of mind. It would be on the findings, models and simulations that their programme would be assessed. The programme of research, Cognitive Science, would involve an empirical study of human conduct utilising the analogy of a computer. This research would generate findings that would address long-standing epistemological concerns.

Let's build and test empirical theories in the best tradition of science. Let's get out of the armchair and into the laboratory or the field. Let's permit the tested science to shape philosophy (as it has done, e.g. in evolutionary theory, cosmology, relativity, quantum theory, and even, in recent years, epistemology) instead of asking the tail of philosophy to wag the dog of cognitive science.

(Vera and Simon, 1993b, p. 132)

Rorty (1980) provides an incisive analytical historical perspective on the emergence of these 'ocular metaphors' for the nature of mind. Rorty notes several reconceptions of prior notions relating to philosophy: Descartes extending the notion of 'thoughts' to cover a range of things including doubts, refusals, imaginings, and also feelings (Rorty, 1980, p.47); Locke using 'idea' 'to cover both sense-data and concepts' (Bennett, 1971, p.25); and "Sensation" being introduced to make it possible to speak of a conscious state without commitment to the nature of external stimuli (Matson, 1966, p.101). These reformulations allowed for the 'mind', 'mental objects' and 'mental processes' to be separated out as resources for a new domain of enquiry. Rorty sees these moves as providing the foundations for philosophy to become analytic. They are also a pre-requisite for the conceptual foundations of Cognitive Science.
2.3.2 The Development of a Computational Model of the Mind

With the emergence of a general theory of computation and with its implementation in actual computer systems, it is not surprising that considerable attention has been paid over the last forty years to following Turing's (1950) suggestions to explore the possibility of devising systems that simulate human intelligence. These attempts at Artificial Intelligence (AI), either as 'strong' AI, where the goal is to specify technologies that have mental processes corresponding directly to human ones, or as 'weak' AI where the aim is just to produce human-like behaviour, have a complex interrelationship with psychological work within Cognitive Science. The motivations behind the psychological and the AI approaches towards Cognitive Science could be considered distinct.\(^\text{11}\) However, it would appear that efforts to build computational devices that simulate intelligence on the one hand, and the utilisation of the computational metaphor to study the mind, on the other are symbiotic. The former provides novel resources for models of cognition whilst being informed of the empirical findings of the latter. Any gaps between the two projects can then be seen in terms of their differing requirements, for model building in AI and for developing theoretical frameworks in cognitive psychology (Johnson-Laird and Wason, 1977).

However, although cognitive psychology has indeed drawn on notions of plans, scripts, schemata, production rules and, latterly, neural networks, it is more difficult to find instances of the reverse relationship. Novel work in AI appears to derive from ingenuity and unguided intuition. Whether the systems and programs work serves as the warrant for their efficacy. Some cognitive psychologists are uneasy with such an approach, pointing out that observable external behaviours could be described by an indefinite number of different computational processes (Anderson, 1980). Furthermore, the warrant for the particular choices of capabilities researchers in AI have identified as intelligent have been questioned. These appear to exclude common-sense, the tacit and the emotional

\(^\text{11}\) Johnson-Laird and Wason (1977) relate the motivations of Cognitive Science to the motivations of cognitive psychology, that is, as a response to Behaviourism. However, they note that practitioners of AI have no such motivations, aiming to develop programs capable of solving complicated problems by any means that happened to work' (p.8). These differing aims and, also different methodological approaches, appear to result in confusions between the two strands of practitioners in Cognitive Science.
(Dreyfus, 1992 [1972]; Weizenbaum, 1976). As noted in many criticisms of AI, following an apparent success in developing systems to cope with small, circumscribed problems, further research became hampered by the limited scope of the enquiry (Dreyfus, 1992 [1972]).

Research from the psychological perspective has sought more to develop explanatory models of human competence. These models are based on ideas concerning computation. So, in Marr's (1982) extensive analysis of the visual process he suggests how from previous studies and experiment, visual processing can be achieved through several levels; from deriving information about light intensity, through the development of primal and 21/2-D sketches and to representations of shapes in three dimensions. At each of these levels appropriate representations are utilised for the processing. Moreover, the processing is broken into modules, so that shapes are dealt with distinctly from surface textures and motion. The warrant for this sub-division comes from experimental studies. Marr emphasises his aim of studying a working visual system, calling his overall information-processing model the 'human visual processor' (p.99).

Marr distinguishes his model both from work in AI such as that of Winograd (1972) and from other information-processing approaches (Norman and Rumelhart, 1974). It is neither too particular to a specific problem nor so general so as to be 'devoid of theoretical substance'. Instead, he emphasises the focus on computational theories of competence, which are distinct from the representations, algorithms, and particularly, the implementation of the model. Rather, the computational theories contain arguments for what is computed and why, and the constraints any operations have to satisfy. It is left to later analysis to outline how these constraints are satisfied.

Marr's analysis of vision has been seen to be a prototype for information processing models: the separation into levels, modules and theories being warranted by psychological research and allowing for different representations, algorithms and implementations to be considered. However, it has been less clear how this approach can be

---

12 For some AI researchers developing a corpus of more general knowledge remains a continuing research programme. Some have attempted to address these issues by increasing the scale of their projects, most notably Lenat and his colleagues' (1986) grand efforts at documenting vast arrays of 'common sense' and encyclopaedic knowledge.
applied to other domains of Cognitive Science. As Marr noted, it may be unclear whether activities like natural language understanding, are modular (p.356) and whether similar constraints could be specified. Instead, more heterogeneous approaches, like problem-solving approaches may be required.

Although the general distinctions discussed by Marr have been maintained, researchers in Cognitive Science interested in HCI have tended to view human-computer interaction in terms of problem solving. The complexity of the activity, the kinds of knowledge that appear to be required and the difficulties of breaking apart particular problems, have led to more generic ways of modelling the activity. Thus, researchers have aimed at revealing 'the structures and processes imputed to a person's mind in order to account for each person's behaviour or experience' (Carroll, 1984) or discovering the 'grammar in the head' (Payne and Green, 1986). These processes are often conceived of in terms of grammars, production rules, propositions and networks, the processing being modelled in terms of symbol manipulation, pattern matching and rules of inference. These maintain the focus on mental activities, explore computational and information processing models, and rely on distinctions being made between representation and process, levels of analysis and different kinds of representations. However, in being concerned with the nature of human-computer interaction they necessarily have to address the performative aspects of the use of computers. Researchers in HCI have a further ambition: to inform the design of computer systems, or as Norman (1983) puts it 'to govern the entire human interface with the system'. This requires attention to be paid to how humans interact with computers in practice. To the levels of analysis outlined by Marr (computational, representation and algorithm, implementation), other modelling activities may thus be required, including considering the designer's model of the user (Norman, 1983; Norman, 1986).

Distinguishing the various models, the user's model of the technology, the researcher's model of the user's model and the designer's model of the user, may lay out the programme of HCI more clearly. However, there still seem to be possibilities for confusion amongst these multiple images, particularly when considering that these models may themselves be informed by computational metaphors. In the following sections, the utilisation of a computational metaphor for modelling an individual's behaviour on a computer system is explored. This is examined in the light
of the foregoing discussion of the programme of Cognitive Science. In particular, an attempt is made to reveal how the mental processes of individuals are conceived in cognitive models of human-computer interaction. This review centres on the influential programme of research initiated by Stuart Card and his colleagues, on what are called the 'GOMS' models. This programme at its various stages, and with its particular concern for being relevant to the design of computer systems, can be seen to be prototypical of much work in cognitive modelling within HCI.\footnote{Amongst the other significant cognitive models for human-computer interaction are Cognitive Complexity Theory (CCT, and the later LICAI, Kieras and Polson, 1985; Kitajima and Polson, 1992; Polson, 1987), Interacting Cognitive Sub-systems (ICS, Barnard, 1987), and Task-Action grammars (TAGS, Payne and Green, 1986). In some areas, these models seek to extend issues not considered by GOMS, for example, involving prediction (CCT) and routine procedures (ICS). In others, the critical distinction appears to surround the representation of the model (e.g. TAGs). Nevertheless, models such as CCT, the later CPM-GOMS (Gray, et al., 1993) and QGOMS (Beard, et al., 1996) have been seen to be a 'family' and applied to several areas of computer use (Olson and Nilsen, 1988). Though other methods differ in their modelling strategies and representation, the analytic approach and conceptual apparatus are similar to the GOMS family of models (see the discussion of 'extensions' to GOMS in Olson and Olson, 1990). Also, see the use of the SOAR cognitive architecture (Laird, et al., 1987), where it is hypothesised that the behaviour modelled, browsing through computer help, is 'GOMS-like' (Peck and John, 1992). Hence, the discussion concerning the general aims of GOMS may also be related to these other cognitive models.}

An examination of the model also provides a background against which the analysis of computer-based activities provided in subsequent chapters can be viewed. GOMS and allied approaches explore the details of computer use, down to a hand's movement across a keyboard, in a precise and systematic fashion. This level of detail is not overlooked in the subsequent analyses, for example in chapter 6. However, the level of detail provided by an analysis is not of issue, rather what will be of concern will the nature of the conceptions which underpin the analysis.

\section*{2.4 A COGNITIVE APPROACH TO HCI}

Card, Moran and Newell are explicit concerning the motivations behind their research: to build a scientific and technical theory of HCI (Newell and Card, 1985). Their work on developing a quantitative model of HCI has been viewed as the 'best established HCI user model' (Brooks, 1991, p.
and the ‘most notable example’ of a theoretical foundation and theoretical concepts within HCI (Barnard, 1991a, p. 103). Hence, it has provoked widespread discussion and influenced much subsequent research within HCI (e.g. Carroll, 1991b; Greif, 1991; John and Vera, 1992; Olson and Olson, 1990; Polson, 1987).

The GOMS (Goals, Operators, Methods and Selection rules) models of Card, Moran and Newell were derived from their earlier work on what were called the ‘model human processor’ and the ‘Keystroke-level model’. As particular key conceptions were developed in this work they are briefly reviewed first (in section 2.4.1). The GOMS models (discussed in section 2.4.2) have recently been extended to account for more complex parallel processes. These extensions are considered in section 2.4.3.

2.4.1 The Foundations to GOMS: the Model Human Processor and the Keystroke Level Model

Compared with Marr’s model of the human visual processor, Card’s model human processor is a rather simple computational model for the user of a computer system (Card, 1984; Card, et al., 1978). Echoing a standard division of computers into input, central processing and output components, ‘the model human processor’ contains three (sub-)processors: perceptual, cognitive and motor. This framework provides the basis for a detailed model of computer-based activities – The Keystroke Level Model (KLM) – (Card, et al., 1980). Focusing on an individual expert carrying out routine and repetitive activities, it examines in depth the operations that surround a user making keystrokes on a computer system. In doing so the KLM reveals in more detail the underlying psychological orientation of the approach and the ways in which mental processes are characterised.14

For the purposes of the KLM, tasks are considered at a variety of grain-sizes, all associated with the individual user. It is assumed that a user breaks a task up into small quasi-independent ‘unit-tasks’, in order to make a task more cognitively manageable. Unit-tasks are divided into two parts: the acquisition of the task and its execution. Thus, the time it

---

14 As with the model information processor, there is a pragmatic aim behind the KLM. It is intended to be easy to use by computer system engineers. Although Card et al’s model is directed at user typing keys on a keyboard, they also utilise the model to account for reading instructions off a sheet, the system’s response time and reacting to changes on the display. As with the earlier model, its design utility is seen essentially in terms of being able to predict the time it takes to accomplish tasks.
takes to accomplish a unit task is considered to be the sum of the times it takes to carry out the two components (i.e. $T_{task} = T_{acquire} + T_{execute}$). Because the focus is on 'routine activity', the execution of tasks is considered the critical component. This itself is taken to be a sum of Keying (K), Pointing (P), Homing (H) and Drawing (D) operations and the response time of the system (R). These can be specified in terms of constants and simple formulae, including a constant for a mental operator (M).

In a 'deliberate simplification' Card et al. assign the mental operator a fixed time of 1.35 seconds. This can be considered to be the time in which a user is 'mentally preparing' to execute an operation. The critical issue then becomes where to place mental operators when modelling sequences of activities (or 'methods'). For example, one method of deleting text from a document on a particular computer system (say, via an Apple Macintosh menu) could be described by the following sequence of 'non-mental' operations, (in KLM form).

- **H[mouse]** 'home' hand to mouse
- **P[word]** 'point' with mouse to word to be deleted
- **2K[button]** 'double click' button on mouse
- **R[selection]** Response time of system
- **P[Edit]** 'point' to Edit menu
- **K[button]** 'click' button
- **R[menu display]** Response time of system
- **P[Cut]** 'point' to 'Cut' on Edit menu
- **K[button]** release mouse button
- **R[menu deletion]** Response time of system
- **H[keyboard]** 'home' hand back to keyboard

Alternatively, deletion could also be performed by the following sequence (without a menu):

- **H[mouse]** 'home' hand to mouse
- **P[word]** 'point' with mouse to word to be deleted
- **2K[button]** 'double click' button on mouse
- **R[selection]** Response time of system
- **H[keyboard]** 'home' hand back to keyboard
- **K[Delete]** press the delete key
- **R[key deletion]** Response time of system

These are just the motor activities of the user and where the system responds. To place the mental operators (Ms), Card et al. give five 'heuristic rules' (p. 400), which can be summarised:
• do not place M's in front of keys that select or type arguments to commands (i.e. they only occur before pointing to commands);
• do not place them when they are 'fully anticipated' (e.g. between pointing to an object and pressing the mouse button);
• do not place them within a 'cognitive unit' (e.g. within a command name);
• do not place them before a 'redundant terminator' (e.g. when two terminators are required, one following an entire command and one following a final argument);
• do not place them within a 'constant string' (e.g. after a key that is always required, for example, in some systems the RETURN key is always needed after a command is typed).

Therefore, the sequences for the operators above would be written something like:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Operator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H[mouse]</td>
<td>'home' hand to mouse</td>
<td></td>
</tr>
<tr>
<td>MP[word]</td>
<td>'point' with mouse to word to be deleted</td>
<td></td>
</tr>
<tr>
<td>2K[button]</td>
<td>'double click'</td>
<td></td>
</tr>
<tr>
<td>R[selection]</td>
<td>Response time</td>
<td></td>
</tr>
<tr>
<td>P[Edit]</td>
<td>'point' to Edit menu</td>
<td></td>
</tr>
<tr>
<td>K[button]</td>
<td>'click' button</td>
<td></td>
</tr>
<tr>
<td>R[menu display]</td>
<td>Response time</td>
<td></td>
</tr>
<tr>
<td>P[Cut]</td>
<td>'point' to 'Cut' on Edit</td>
<td></td>
</tr>
<tr>
<td>K[button]</td>
<td>release mouse button</td>
<td></td>
</tr>
<tr>
<td>R[menu deletion]</td>
<td>Response time</td>
<td></td>
</tr>
<tr>
<td>H[keyboard]</td>
<td>'home' to keyboard</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Operator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H[mouse]</td>
<td>'home' hand to mouse</td>
<td></td>
</tr>
<tr>
<td>MP[word]</td>
<td>'point' with mouse to word to be deleted</td>
<td></td>
</tr>
<tr>
<td>2K[button]</td>
<td>'double click'</td>
<td></td>
</tr>
<tr>
<td>R[selection]</td>
<td>Response time</td>
<td></td>
</tr>
<tr>
<td>H[keyboard]</td>
<td>'home' to keyboard</td>
<td></td>
</tr>
<tr>
<td>K[Delete]</td>
<td>press the delete key</td>
<td></td>
</tr>
<tr>
<td>R[key deletion]</td>
<td>Response time</td>
<td></td>
</tr>
</tbody>
</table>

Thus, in both cases mental operators are only placed where the user is about to point at a word with a mouse. The application of the heuristics is not straightforward and requires some work on behalf of the analyst in order to identify meaningful 'chunks' that are 'fully anticipated' or form part of a 'cognitive unit'. Indeed, it is not clear whether the distinction between 'commands' and 'arguments' holds for some user activities, or when terminators are 'redundant' and strings 'constant'. The heuristics appear to suggest occasions when users are 'thinking', but they also could be considered to mark out places where pauses or breaks in an activity occur, without any recourse to a notion of 'mental preparation'. It is unclear what status these rules have, or what warrant there is for them.\(^\text{15}\)

In order for KLM to go beyond being merely a description of physical and behavioural activities, the model needs to incorporate a mental operator. Card et al. derive a constant for the operator from the

\(^{15}\) Lane et al. (1993) give a brief, but detailed, account of the problems of assigning the mental operator when looking at three closely related activities of users of a spreadsheet program. They note, in particular, difficulties with the terms 'fully anticipated' and 'cognitive unit' (p. 187). Different interpretations give different prediction times.
experiments it aims to model. In effect, mental activity is calculated as the residue left over after the times for observed activities (keying, homing, pointing etc.) have been measured. The model assumes this gap to be 'mental' and then spreads it about the sequences according to the heuristics. Card et al. admit that this procedure is problematic, but suggest that the utility of the constant lies in its use in similar analyses by others.

Although there are admitted simplifications in the model, Card et al. see it as having its foundations in Cognitive Science. The distinctions between acquisition and execution, mental preparation and observable activity invoke the conception of a space where mental activity occurs. However, the nature of this mental processing is unclear. The KLM model suggests cases when mental activity might happen, however, it has difficulties accounting for when mental activity occurs in parallel to, or is interrelated with, physical actions. Even as a preliminary attempt at a cognitive model of computer use, the KLM model appears to draw on the distinction between mental and physical activities too crudely.

The KLM model focuses on routine activities. Its characterisation of simple computer-related activities neglects deviations which may occur in those activities, in particular the model fails to account for users making 'errors', as has been noted by HCI researchers (e.g. Kieras and Polson, 1985; Landauer, 1991). However, the model has more fundamental problems. The account of mental activity as mental preparation prior to the accomplishment of action is a simplification that allows for the model to account for 'gaps' in observable behaviour. Because it does not have an explication of mental activity, the model appears to be akin to a behavioural description.

By taking into account the response times of the system, the KLM model begins to explore the notion of a human-computer 'interaction'. Extending the focus of an individual user's observable behaviour to include, for example, scanning a piece of paper for instructions, appears to require a more sophisticated model of the inner cognitive processes of the user. However, even with this extension to the account of human behaviour it remains unclear how a cognitive component enhances the model. As a pragmatic method there is little need for the KLM model to attain explanatory adequacy, yet such a characteristic may be necessary if the model is to be extended towards other domains of activity.
2.4.2 The Symbolic Processor

The GOMS models are a series of layered models proposed by Card et al. (1983), building on their earlier KLM. Indeed, the 'basic actions' considered by KLM, such as making keystrokes, moving a hand and looking at objects, are modelled on the lowest level (called the M0.5 model). The models layered above this level can be considered as coarser, the operations being modelled being more generic and of longer duration. Card et al. suggest various models at four different levels (called M16, M4, M2, M0.5). On each level each model is made up of four components: Goals, Operators, Methods and Selection rules (hence, the name GOMS). Goals are 'symbolic structures that define a state of affairs to be achieved' (p.38); operators are 'elementary motor or information-processing acts' (p.40); methods describe procedures for achieving a goal; and selection rules allow choices to be made between alternative methods for accomplishing goals.

Goals are considered to be hierarchical, for example, the top-level goal of editing a manuscript can be broken down into individual unit editing tasks. This is the grain of analysis for the highest level, M16, GOMS model. For lower level models, these tasks are themselves decomposed. So, for instance, particular 'edits' could be broken down into getting a task and executing it, the latter consisting of locating a line, modifying the text and then verifying the change. If there are different methods for achieving these tasks then the selection rules come into operation. For instance, on an Apple Macintosh, there are various ways of locating a portion of text, including: scrolling down through the document, moving to a location in the document and then scrolling, or searching for a particular word. In the GOMS models context-free selection rules are defined to chose between these methods. These rules are specified in terms of 'production rules' (or IF THEN rules). For instance, the following selection rules could be defined for the Macintosh example:

16 The numbering scheme of the GOMS family of models reflects their granularity. The number indicates the order of the duration of operators considered at each level. Hence, at level 16, operators have durations of the order of 16 seconds and 0.5 seconds at level 0.5. Letters following the level number (e.g. in M0.5E) arbitrarily distinguish different models at the same level.
Use Scroll Arrows as default

IF destination further than a page away
and less than a third of the document away
THEN use Scroll bar

IF destination further than a third of the document away
THEN use search command

At the lowest level operators are further decomposed: first, in terms of looking, homing to a target, moving a hand, typing and a mental operator (cf. the KLM model); then the mental operator is itself distinguished into searching, comparing and choosing.

In organising the models in this way, close parallels to the work of Marr (1982) can be seen. There is an emphasis on levels, differing kinds of representation and a more modular categorisation of activities. Card et al. are explicit about the nature and purpose of their models, considering the manuscript-editing task they describe as a candidate for a domain of activity that could offer a verification of the theory of 'man' as a 'symbolic information processor' (p.33). Indeed, they suggest that studying an example domain of activity should facilitate the extension of symbolic information-processing theory. Furthermore, the model should be able to make predictions about human behaviour at a computer terminal. These predictions would be in terms of sequences of activities or the time a task should take. In emphasising the symbolic aspects of the model, Card et al. are perhaps distinguishing their objectives from those of Marr. Certainly, apart from utilising a few familiar techniques (e.g. production rules and hierarchies), there is not a prime aim of specifying computational constraints and competencies. Instead, GOMS models appear to concentrate principally on the performative features of computer use.

Card et al. identify some of the difficulties of using the different layers of GOMS models. For example, they note that the accuracy of the models do not necessarily reflect the level of detail by which a particular analysis is carried out; coarser grain analyses neglect particular details whilst fine grain units are more prone to errors in measurement. For the manuscript editing task, which involves examining an annotated sheet of paper and making the appropriate change to the document on the computer system, the coarsest level of analysis is simply in terms of performing a series of editing tasks (or EDIT-UNIT-TASKs). For intermediate models, tasks tend to have broadly appropriate labels such as GET-UNIT-TASK and DO-UNIT-TASK reflecting the $T_{acquire}$ and the $T_{execute}$ of the KLM model. However, even at the lower levels, descriptors are also given vaguely intentional labels such as GET-NEXT-PAGE, MODIFY-TEXT and VERIFY-EDIT. It is
unclear whether this is sufficient to allow these models to be considered cognitive. Presumably, a wide variety of labels could be given for these descriptors which may not have such mentalistic connotations. Furthermore, they are meant to cover a variety of observable behaviour, and it is unclear whether rigorous criteria can be specified for making such classifications. As these labels of generic activities are the only objects handled by the intermediate GOMS level models, it is perhaps worth examining the lower levels when considering its cognitive dimension.

At the lowest level (model M0.5E) the mental operator is broken down into: SEARCH-FOR, COMPARE, CHOOSE COMMAND, CHOOSE ARGUMENT. In their manuscript editing experiment Card et al. note that the 'mental' accounts for 60% of the total time, compared to 20% of the time the subjects are seen to be typing. However, certain difficulties are encountered when trying to categorise activities in terms of these operators. First, as with the KLM, identifying mental activity is problematic: 'pauses' in the observed data of greater than 0.3 seconds are here counted as 'mental'. Second, some observed data do not fit the model, in particular, some mental operators identified in the data could not fit any of the pre-assigned categories (in 71 out of the 581 instances). These had to be assigned a category of UNKNOWN, which always failed to match operators generated by the model. Hence, the accuracy of the model in predicting sequences was less than 50%. Third, when more than one mental operator is predicted in succession, it is difficult to determine the boundary between them. When examining the observed data it did not appear that mental operations combined additively.

These difficulties with the GOMS models appear to reflect Marr's concerns with performative theories of cognition. Without any prior concern with the computational theory behind the activity, derived models are likely to appear ad hoc and unwarranted. However, the nature of visual processing accounted for in Marr's models and that of the activities involved in using a computer are quite distinct. From a cognitive viewpoint, computer use appears to be of quite a different order. Indeed, an account of visual processing appears to be one of the pre-requisites of a perceptual model (or a model on the GOMS 0.5 level). It appears to be much more problematic to break down activities such as searching, comparing and choosing into distinct modules and provide computational constraints and theories for these. It may be that, in order to proceed,
cognitive models have to rely on generic techniques, with different kinds of information processing being considered mainly in terms of the manipulation of different kinds of symbols.

Card et al.'s description of the GOMS models offers an explicit account of a cognitive, and symbolic, approach to HCI. At the coarser levels, it is unclear what status the components of the models have; task descriptions appear to be cover a range of observable behaviour and method and selection rules seem to offer generic procedural accounts of repetitive activities. Although, tasks and methods are often given intentional labels, these appear to be too vague to offer a sufficient description of the user's internal mirror of technology. The lower level models are more specific and appear to involve particular mental activities and processes. Unfortunately, when these models have to fit observable data problems ensue. As with the earlier KLM, observations of the mental appear to coincide with pauses in physical activity. The proponents of these models are not suggesting that such a crude distinction between mental and physical activity pertains, but given their experimental apparatus it is difficult to envisage a more accurate way of matching the models to observed activity.

Card et al. found that when fitting data to predicted operators it appears that the time taken to carry out two mental operations sequentially did not match the sum of the time taken to carry out each individually. This observation may have important consequences for the ways in which mental operations are conceived. If mental operators cannot be combined additively, it is unclear whether mental operators remain the same when occurring alone or when coinciding with other mental activity. However, these models, with their emphasis on routine and repetitive behaviour rely on an assumption that, particularly when associated with physical motor activities, mental operators remain stable. If mental operators can be transformed when in combination what warrant is there that they remain stable at other times? If the definition of mental operators such as 'searching' and 'comparing' is also open to flux, it would be unclear what status these operators would have in the model and, once again, the warrant for such activities being accounted for as mental would be brought into question.
2.4.3 The Parallel Human Information Processor

The problems accounting for parallel activities in GOMS does not merely concern mental operators, but also other operators in the analysis (i.e. the perceptual and motor). With its foundations in the model human processor, GOMS is based on sequential processors for perception, cognition and motor control. Although other layers can be considered to operate in parallel, at the model human processor (or GOMS M0.5) level, the three processors are considered to be sequential, one occurring after the other in a cycle. Hence, the models cannot provide for a user doing two related things at the same time, for example, typing and reading off a screen. To address this problem John (1990) has developed a GOMS technique that allows for modelling more than one processor in parallel and accounting for the dependencies between them. She utilises familiar techniques from project management for this: Critical Path Methods and schedule charts (hence CPM-GOMS). These allow for parallel activities to be mapped out and a critical path (where delays will delay the overall accomplishment of the project) to be determined.

Gray et al. (1993) give a detailed account of an exercise in using CPM-GOMS. This is unusual in GOMS analyses, and cognitive analyses of HCI in general, in that it explores a real-world example: the so-called Project Ernestine. This project was carried out on behalf of a telephone company involved with the introduction of a new computer system to assist their telephone operators responsible for handling routine enquiries relating to connections and billing (or Toll and Assistance Operators). This is a routine activity involving talking to the customer, typing into the

\[\text{17 W warranting such an assumption from previous work in cognitive psychology has been questioned (Olson and Olson, 1990).}\]

\[\text{18 Project Ernestine continues the concerns of the original GOMS methods by utilising the models to derive comparative timings between design options. In this case the authors could compare the time taken to carry out activities between a proposed system and an existing one. The authors state that because of the way the telephone company calculates the cost of the time of the operators, findings from the CPM-GOMS modelling exercise can be seen to save considerably on costs (one estimate being that the 0.63 second difference predicted by the model could save $2 million). This, presumably, is the kind of design decision that fulfils the pragmatic objectives of the analysts. John gives another example of using CPM-GOMS in a model of the activity of typing transcriptions. This model is drawn from previous experiments and debates concerning these kinds of activities, and not a real-world study (John, 1996)}\]
computer system and activating phone connections. The telephone company focuses on the duration of the average phone calls when examining costs, hence predictions concerning estimates of changes in the timing of the call due to the new computer system are viewed as consequential.

The CPM-GOMS analysis is in many ways very similar to the classic GOMS analysis, with motor operators including the movement of each hand and eyes and verbal responses, perceptual operators being such items as 'perceive BEEP' and 'perceive complex-info', and cognitive operators being 'attend-info', 'verify info' and initiate an action (such as an eye movement). At the highest level the unit-task of 'HANDLE CALLS' is broken down into sub-goals RECEIVE-INFORMATION, REQUEST-INFORMATION, ENTER-INFORMATION and RELEASE-CALL. Between the two, other activities are defined, such as LISTEN-TO-CUSTOMER, LISTEN-TO-BEEP and THANK-CUSTOMER. The critical difference to GOMS is that activities can be modelled that occur in parallel, so the model allows for the telephone operator to read off the screen and to type whilst the customer talking. A detailed CPM-GOMS analysis of two systems designed to support the telephone operator, the existing one and a proposed new system, revealed that though the new system was quicker, the savings in time it introduced were not along the critical path. Indeed, the design of the new system introduced activities along the critical path.

This is the kind of finding envisaged by Card and his colleagues for the model human processor and allied models; a cognitive analysis having direct consequences for a development activity: in this case, the general decision whether to deploy a system or not.

In order to generate their predictions, Gray et al. had to make several deliberate simplifications. So, for example, although there were four different cognitive operators they were each assigned the same duration (50 msec). As for GOMS there were problems with the 'placement heuristics of unobservable operators' (p. 262). The heuristics from GOMS were used as a guide for 'best guesses' and then modified to remove 'impossible times'. More interestingly the variation between the length of turns of talk, particularly those of the customer, had to be accounted for by scripting typical utterances.19

19 One problem Gray et al. faced is accounting for learning behaviour, so they discounted the first month of the trial. The differences for the average times over
These simplifications highlight the pragmatic character of the model. Gray et al. have undertaken a detailed analysis of a particular routine activity, exploring the potential changes to the activity introduced by a new system through an analysis of video-tapes and data collected of recorded calls across different types of calls and down to a level of detail that considers eye and hand movements. Indeed, the critical feature of the new system appears to be the positioning of the key to mark the end of call. Previously, this could be accomplished with the user's left hand, which could be moved towards whilst typing with the right hand. The change in key position made this awkward and, hence, shifted an activity onto the critical path (consisting mainly of motor operations). The difference in time between the two computer systems was, after all, only estimated to be 0.63 seconds.

As with GOMS, the extent to which CPM-GOMS models cognitive activities appears to be limited. Once again, the focus is on routine activities that involve continuous use (principally typing) on a computer system. The accuracy of the predicted timings appears to be more due to attending to details of observed activities, rather than an incisive analysis of unobserved mental processes.

In undertaking an analysis of a system used in practice, Gray et al. have extended the domain in which GOMS models have been applied away from experimental or hypothetical systems. As it directly relates to aspects covered in later chapters of this thesis, it is worth re-emphasising some of the assumptions behind the model.

- the only resources users appear to have are a set of perceptual, cognitive and motor competencies. These are governed by goals, and sub-goals to achieve a unit task. The participants have limited resources for developing other courses of action (cf. the use of computer applications in chapters 3 and 4);
- the system activity is modelled in terms of undiscriminated responses and changes of displays. It appears that the functions the system performs or the appearance of items on the display is not relevant to the activity being modelled. Indeed, it is unclear to what extent the system is a computational system (cf. the practices surrounding the use of CAD systems in chapter 4);

the months can then be related to changes in the distribution of calls over the different categories.
• the activity as modelled is routine and involves continuous operation of the system. Even acknowledgement that there may be activities other than those accomplished on the computer system is not provided (cf., the use of a computerised phone system in chapter 5);

• the model only accounts for the behaviour of one individual. This is despite of the involvement of a customer in the activity. Indeed, there are calls which involve the participation of a second operator. These are ignored in the model (cf. the analysis of collaborative activities in chapters 5 and 6);

• despite the focus on a telephone system the talk is neglected. The activities focused on remain on those surrounding the typing on the keyboard (cf. the analysis of the entry of commands within a control room in chapter 6). 20

2.4.4 A General Discussion of GOMS Models

Given that they are explicitly intended to be cognitive models of computer use it is perhaps strange that GOMS models focus on routine cognitive skill where 'no puzzlement or uncertainty of what to do next' is apparent (p. 69). Such a focus appears to preclude an examination of the very conduct that such models should be suited for. One might have expected an exploration of domains where there were more obvious opportunities for activities which could be analysed in terms of cogitation, deliberation, decision making and planning. There are many settings where the use of computer systems are interrupted by periods and gaps which might offer more appropriate occasions for giving an account of mental activity. 21

The work on GOMS has various pragmatic design objectives, but it is also intended to contribute to an understanding of symbolic cognitive

---

20 These constraints also seem to imbue the few design suggestions Gray et al. imply from the modelling exercise. Apart from suggesting that the system response time should be reduced, Gray et al. propose the investigation of automatic greetings and changes to the organisation of the call. The latter could include adding suggestions to the caller on how they should have presented their query. This does not appear to account for the sequential implications that such suggestions might have on the caller's next actions (see Chapter 7).

21 There has been some work within the cognitive science approach to HCI on more exploratory activities. These tend to adopt the same mechanisms, namely production rules, to explain cases of this behaviour (e.g. Kitajima and Polson, 1996).
processes. As such it exhibits many of the features that would be expected; a focus on building models of activities; a concern for the distinction between the process and the information (or symbols) being processed; and an arrangement of the model into levels where higher levels utilise the products of the lower ones. However, what does seem problematic is the general characterisation of activities as information processing. This relies on simple models of input and output with processing in between (the three processors of the model human processor), or of the firing of production rules (e.g. the selection rules of GOMS) or hierarchical relationships between tasks (e.g. the goals and sub-goals of GOMS). It is possible to break apart activities in many ways, the same activity could be defined on many different levels and in terms of differing combinations of symbols and operators.²² Moreover, the nature of processing specified by these mechanisms is unclear. These are either too weak or too general. On the one hand, mechanisms, like hierarchies, merely express a relationship which activates a set of other processes (in sequence), or, like production rules, appear to imply that any kind of processing can occur.²³ On the other, a model in terms of input, processing and output is so general it appears to offer no constraints to the kind of processing that can occur. Indeed, this model though weakly specified would appear to have the power, in the mathematical sense, of any Turing Machine, and hence programming language. They are powerful, or strong, in being able to account for any computable function (Brady, 1977; Hopkin and Moss, 1976). Therefore, they offer few constraints on how processing is modelled. Hence, there are innumerable ways in which the distinction between information and process can be made.

The warrant or criterion for selecting the various tasks, goals, operators and rules used within the models is unclear. In modelling an activity, the general conception of tasks is cognitive, however, this appears

²² Indeed, in one extension to GOMS, QGOMS (for 'quick and dirty' GOMS), levels of analysis can be ignored and times directly substituted for the breakdown of tasks (Beard, et al., 1996). It is stated that this offers the desired accuracy for design.

²³ Indeed, the computational specification of the production rules in GOMS models is unclear. It is possible that the computational power of production rules can be equivalent to the power of a Turing Machine (i.e. specify any function that is computable). In more technical terms, the production rule mechanism lacks a computational semantics. So, for example, it is not clear whether there are any constraints to the kinds of functions specified in the 'action' parts of the rules.
to rest on analysts assigning appropriate intentional labels to activities. More specific notions of mental operations, when examined in detail, face similar problems. They appear to be accounting for gaps in observable activities rather than explicating mental processing, and also rely on analysts’ own competencies at recognising certain behaviour in terms of particular mental operations. Analysts are given little assistance on how to assign particular mental operations from the description of the models. Hence, there are difficulties attributing conduct to cognitive categories and distinguishing between them. Analysts have to use common-sense in order to determine whether an individual is, for example, searching or comparing. Utilising intuitions may not be problematic if the model is considered to be a tool for design, but this is inconsistent with the claims made for the model. Card et al. state that they aim to develop a more scientific approach, moving away from intuitions, anecdotal or subjective accounts of behaviour. However, when examined against these claims for a symbolic processing model of human conduct on a computer, or of Cognitive Science in general (cf. Johnson-Laird, 1988 [1993], p. 26), their aims do not appear to have been fulfilled.

Similar concerns were expressed by Marr relating to approaches towards Cognitive Science developed within Artificial Intelligence (see Section 2.3.2). He considered that the warrant for a particular information processing model should be in terms of computational constraints on the process. In the first instance, it would not be necessary to be concerned with how particular cognitive activities were carried out by the human processor, but rather with the kinds of information that had to be processed. This would offer constraints for the later concerns of considering the representations required and the algorithms needed to process them. For visual processing, it is possible to see how such a programme progresses (cf. Marr, 1982). What arrives on the retina provides the first level of information. Subsequent levels of information can be postulated from experiments undertaken within studies of vision: such as primal, \(2^{1/2}\)-D and 3D sketches. Then what has to be explored is the nature of computation required to process this information to detect shapes, textures and movements, for example.

For analysing computer use, on the other hand, the pre-requisite information appears to be harder to constrain: there are images arriving from the screen, sounds from the machine, feelings from the various devices. There are also words and shapes presented on the screen and
courses of activities to be undertaken. Merely listing the competencies involved appears problematic. Take reading off the screen, for example. Presumably a transformational model of syntax would be one model of the competencies to be outlined for such an activity (cf. Chomsky, 1957; Marr, 1982). But even in a cognitive account, users would have to be considered as bringing with them more than just syntactic competencies, for example, they may have resources from using the system before, concerning the activity to be undertaken on the machine, and common-sense, real-world knowledge. This would require an outlining of the kinds of information, or knowledge, required to use a computer system. Such a consideration of knowledge would indeed seem to be a pre-requisite of a cognitive model. Hence, some researchers have proposed that such a 'knowledge analysis' be undertaken, where different types of cognitive resources are categorised, (Young, et al., 1990). These proposals are examined in more detail in chapter 3.

A wider exploration of the information, or knowledge, required to operate a computer system would invariably raise problems of scope. There are other ways in which cognitive models could be extended. CPM-GOMS has already allowed for the examination of parallel activities in association with the computer. A model could also account for other activities engaged in by users, the 'content' of what they are doing, their communication with others and the possibility of collaborative activities. Rather than gradually expanding the interests of cognitive or psychological models, some researchers have suggested a more radical departure by re-examining the scope of the programme. Of particular concern has been the possibility of incorporating social and communicative concerns within the framework. Whilst some of these developments, like Activity Theory, have quite distinct antecedents to Cognitive Science, others like Distributed Cognition have more obvious relationships to the orientation.24 Of the several proposals drawn from a psychological

---

24 So, for example, several researchers have suggested that Activity Theory, an orientation developed within Russian Psychology, is an ideal resource from which to develop an understanding of the use of tools and technologies (e.g. Bannon and Bodker, 1991; Kuutti, 1991; Nardi, 1996c). This has a broad conception of activities, not focusing on an individualistic and cognitive conception, but exploring how activities can be seen in terms of their socio-cultural development (Nardi, 1996c). Thus, skilled manipulations, cognitive processes, social factors and historical developments may all be relevant to the use of an object or artefact. However, some researchers have emphasised a concern to maintain a focus on
background, this has attracted considerable attention with regard to its applicability for design. Because of this, the more explicit outline of its analytic orientations, and its concerns with collaborative and communicative activities, it is on Distributed Cognition that the next section will focus.

2.5 DISTRIBUTED COGNITION

As with the conventional cognitive approach to HCI, the motivations behind much recent work in Distributed Cognition have been both scientific and applied. On the one hand, there has been an interest in broadening the scope of psychological studies to account for other concerns, particularly to take account of social and communicative activities. On the other, research of the activities of individuals on personal workstations appears to be largely irrelevant to recent developments in computer technologies, particularly those designed to support collaboration and communication. Hence, either methods and techniques for design will need to draw from other research programmes, or the techniques developed within HCI have to be broadened. As noted in the previous sections, approaches to HCI drawn from Cognitive Science appear to be particularly constrained to focusing on the individual user engaging in routine activities on a personal workstation. So, it may be that by seeking to explore a broader range of activities, analysts may also be able to develop their models of what is conventionally considered to be the domain of HCI. To pursue these concerns, both applied and scientific, researchers have begin to focus more on the accomplishment of real-world activities in various work settings.

Although several frameworks have been proposed, drawing on such disciplines as anthropology, management studies and organisational analysis, particular attention has been focused on utilising resources drawn from both psychology and sociology, proposing, for example, to extend Cognitive Science to account for social action (Hutchins, 1990; Olson, 1990). What is distinctive about this programme is that it relies planful, intentional behaviour (Nardi, 1996c, p.84). Hence, a strong 'closeness in spirit' between Activity Theory and Distributed Cognition has been identified (Nardi, 1996c, p.89).
principally on observations of naturally occurring social action whilst also
drawing from the Cognitive Science paradigm.

Despite the very recent development of Distributed Cognition and the
tentative outline of the programme (Hutchins, 1985; Hutchins, 1995;
Hutchins, et al., 1986; Norman, 1993a; Olson, 1990), there have been
several attempts to explore its utility for the development of technology
(e.g. Flor and Hutchins, 1991; Hutchins and Klausen, 1996; Rogers, 1992).
These studies draw on field studies to examine particular settings and
workplace activities, such as design and other offices, aircraft cockpits and
shipboard navigation, with particular attention being paid to the uses put
to artefacts in those settings. The studies, thus, provide many insights
into the use of tools and technology in naturalistic settings, particularly
concerning how objects like navigational instruments and dials are used
as a resource for collaborative work. A range of notions which appear to
be conventionally drawn from Cognitive Science are utilised within these
analyses. So, for example, artefacts can act as 'memories', individuals can
collaboratively achieve 'goals' and knowledge is shared and distributed
amongst individuals. However, the nature of cognitive activities is
significantly transformed in these analyses. Cognition is no longer
conceived just as 'residing in the head' but also as distributed between
individuals. These shifts in perspective are most cogently argued for in

Hutchins outlines the ways in which navigation is achieved both from
accounts of Micronesian navigation and his own study aboard a large US
Navy vessel. He then analyses these in terms of various computational
activities, calculating speeds, identifying landmarks and plotting fixes on
those landmarks, revealing ways in which these are achieved economically, reducing the need for re-calculation or complex arithmetic
activities. Hutchins suggests that through activities such as plotting
against imaginary stars or plotting fixes every three minutes, both
Micronesian and Western navigators manage the complex computations
required. Through these fascinating studies, Hutchins reveals how the
activities are sensitive to the contingencies faced by the navigators, such
as the artefacts they have available and the time in which they need to
accomplish their work.

Whilst taking account of the use of physical artefacts like the chart and
the compass, and the talk between the navigation crew, critical to
Hutchins' analyses is the conception of the activity of navigation as
cognitive. However, this is a distinctive account of cognition; Hutchins hopes

...to show that human cognition is not just not influenced by culture and society, but that it is in a very fundamental sense a cultural and social process. (Hutchins, 1995, p. xiv)

Hutchins reiterates questions concerning a strictly symbolic account of cognitive processes, wishing to move away from the narrow focus of cognition 'residing in the head'. Whilst proposing this transformation he explicitly aims to maintain Marr's architecture for information processing. In particular, despite the shift away from Marr's focus on the individual, Hutchins circumscribes an activity – navigation – applies an analysis at several levels of description and pays attention to first developing a computational model of the activity. Hence, Hutchins outlines the computational constraints underlying navigation; the relationships between distance, time, speed and position. He then examines ways in which Western and Micronesian navigators deal with these constraints differently by developing various forms of representation and processes. For example, Micronesian navigators use reference islands and star paths to determine their position and route, Western navigators use charts and bearings. Finally, Hutchins outlines details of the implementation of the representations and algorithms in the work and interaction of a navigation crew on the bridge of a large naval vessel.

With respect to levels of analysis, Hutchins does appear to follow Marr's programme. However, he makes two critical and explicit diversions, firstly applying cognition to social processes and secondly with regards to computation. Because Hutchins relies on a computational model of cognition, issues surrounding the latter diversion will be considered first.

2.5.1 The Conception of Computational Processes in Distributed Cognition

Citing the work of Turing, Hutchins notes that the original conception of manipulating symbols need not be applied to an internal process, it could involve interacting with the material world or be part of a social system. The problem for Cognitive Scientists is that they confuse symbol manipulation with individual human computational accomplishments. Searle's (1990, 1980) criticism, that an account of symbol manipulation is not sufficient to model human conduct, is convincing because it too constrains the focus on the individual. If the origin of formal systems to
manipulate symbols is considered, and the nature of the activity, then 'what Turing modelled was the computational properties of a socio-cultural system' (Hutchins, 1995, p. 362). For Hutchins, in formal systems, symbols are manipulated in terms of form only, and from these manipulations result other symbolic expressions (p.359). This certainly has resonances with Turing's abstract symbol-processing machine. However, Hutchins is providing an account of human activity, not mathematical problems. He analyses navigation in terms of computational constraints, so that the use of artefacts, the talk between participants and their coordination of activities are all considered with respect to the relationships between distance, time, speed and position.

What is curious is that the computational level of the analysis does not appear to involve symbolic manipulation. There are good reasons for this. It is hard to envisage an analysis of symbols which is both outlining a socio-cultural process and culturally independent. For the examples Hutchins gives, the analysis would have to be implemented in both the use of Micronesian star paths and reference islands and to the 'three minute rule' for taking bearings. This may not matter as the next level provides accounts of representations and algorithms which can be culturally specific. Western models of navigation can then rely on Western mathematical conceptions and symbol manipulation. However, it is unclear without a symbolic account of computation, even at an abstract level, in what way the 'computational level' of analysis is computational. Hutchins gives constraints, which could be expressed as (constraint-satisfaction) rules, but there is no implication of a process.

---

25 Note, the Turing machine provides a foundation for considering a set of mathematical problems, being utilised to prove, for example, whether a problem expressed in a particular way is computable (or non-computable), or whether it is decidable or enumerable. Critical to these proofs is the expression of the problems in terms of symbols and instructions, written and read by a simple notion of a machine. The motivations behind the original formulation of the machine was to provide a means by which general mathematical problems could be considered. This exercise in meta-mathematics needed a constrained conception of computation, provided by a device that processed distinct symbols. Thus, it is Hutchins who characterises Turing's innovations as a 'model' (of a socio-cultural system, and not of a mental process).

26 Hutchins does provide a general account of a constraint satisfaction process (p. 295-310), but this is at the representational level, i.e. with respect to the activities in relation to mediating artefacts and representations by members of the navigation crew.
Hutchins appears to be giving, in his terms, is a formal system for navigation without symbolic representation.

It appears that the computational level is necessary for Hutchins to outline a culturally independent conception of navigation. Here, he can state the constraints which apply to navigational activities. These relate positions to spatial displacements, and distances, rate and time (p.58). As these constraints may, in different ways, be packed or 'crystallized', into procedures, artefacts and representations, they are implicit to the activities and they are oriented to by participants. So, for example, onboard a large US Navy ship, the use of landmarks and the plotting of bearings within the 'three minute rule', can be seen as a procedure that makes use of these constraints, but the crew need not orient to the constraints themselves. Instead, in order to reduce cognitive effort the procedure embodies the constraints. In the Micronesian case, which is non-literate, no explicit representation or process of symbol manipulation is undertaken. Arriving at a destination is achieved by breaking a voyage into segments and in navigating each with respect to the relationship between the objects in the sky and visible landmarks. By some sophisticated analysis, Hutchins shows how this process, particularly the use of phantom islands, could also be seen to operate within the constraints.

What Hutchins gives is an account of a formal system in terms of some constraint satisfaction rules; what could be considered, perhaps grandly, as an explanatory theory of navigation competencies. But in the actions of the participants undertaking this activity, this formal system is only implicit. To accomplish navigation they have to utilise a wide range of common sense and skilled competencies. Hutchins provides a fascinating account of some of these, including the interactional accomplishment of the taking of bearings (p. 238-9), the relationship of the use of artefacts to the orientations of participants (p. 235-6), the use and recognition of local landmarks (p. 134-6) and in the Micronesian case, making sense of the colour of the sea, the swell patterns of the sea and the vicinity of birds.

Thus, what is unclear is the status of the computational model. It appears weak in terms of its computational properties. Although Hutchins emphasises how computational properties, in terms of the processing of symbols, can be considered as socio-cultural, these do not appear to be relevant for the computational theory. Thus, it appears to be more an account of some formal properties; formal properties which are
not oriented to by participants, but are hidden, guiding the historical shape of artefacts, the organisation of procedures and the activities of participants. It is an analytical account that has to be distinct from the practices of participants. Hence, it could be argued why Hutchins’ ‘computational level’, without a clear computational specification, can be warranted with respect to other analytic models of hidden processes, or, why these formal properties have a primary status over the practices through which navigation is accomplished. Although in the explication of his analytical framework the computational and the symbolic are tightly related, the nature of this relationship is unclear in practice.

2.5.2. The Conception of Cognition in Distributed Cognition

As their name suggests, studies in Distributed Cognition adopt a social conception of cognition: cognition as a social process. However, this is not to say that researchers in this area reject cognition as a mental process. In fact, in many ways the conventional notion is maintained. So, although Hutchins calls for a deconstruction of the framework of Cognitive Science (p. xviii) and argues that symbolic manipulation is not the architecture of cognition, he believes that ‘humans actually process internal representations of symbols’ (p.370). Rather than rejecting an individual conception of cognition, Hutchins wishes to embed this notion within a broader social conception. In particular, he is concerned with a ‘division of cognitive labor’ in which ‘the cognitive properties of human groups may depend on the social organisation of individual cognitive capabilities’ (p.176).

As Hutchins emphasises his account of a socially, distributed cognition cognition applied to groups, it is this conception which will be considered here. This is largely accounted for in terms of communicative and interactional accomplishments. So, even the use of artefacts is considered in terms of them mediating communication. Hence, Hutchins’ analysis of distributed cognition pays attention to language and the talk between participants. Hutchins notes that in Cognitive Science:

...language is usually thought of primarily as a human computational capacity that should be understood in terms of the processing that individuals must do in order to produce or interpret it.

(Hutchins, 1995, p.231)

He wishes to understand it in a new light in terms of technologies and the 'larger cognitive system'. Hence, Hutchins is seeking a reconception of cognition that involves the use of artefacts, the distribution of social
activities and interaction between participants. Throughout his analyses of quite heterogeneous activities, however, Hutchins maintains his attention on what he considers cognitive properties. One way in which these analyses are brought together is with a concern with their computational nature and in terms of manipulating symbols and representations. Another concern is with the types of knowledge required by participants and the ways meanings are communicated and negotiated. Hutchins outlines these concerns principally in relation to the activities undertaken by the navigation crew aboard a large US Navy vessel.

For example, in a simple interchange between three individuals - a plotter, a recorder and the watch responsible for taking the bearings - a first utterance by the plotter to the recorder "Tell him to take Point Loma first. It's on his beam" is transformed by the recorder for the watch into "Take Point Loma first, Mark, Beam bearing first, mark it". This, Hutchins sees as an example of language socialisation, of the recorder explicitly invoking a rule concerning the ordering of taking bearings and providing an opportunity for the watch to 'add to his knowledge of the meaning' of the expression 'Beam Bearing' (p. 209-210).

In another example, a gesture by the plotter (or chief), a point towards a document, is seen not only in terms of its uses for the plotter, but also in terms of its 'communicative function' for the recorder: as a clue to what the plotter is looking for (p. 235).

The plotter's action is part of a memory-retrieval event that is internal to the system but directly observable...Because of this, the chief's pointing can be both a part of his private cognitive processing and an element of communication to the recorder about the sort of thing the chief is trying to accomplish.

(Hutchins, 1995, p. 231)

In these examples, and in many more cases, Hutchins reveals the complexity of the activities of the individuals on the bridge of the ship, how tightly coordinated they are and how they achieved in ways which are not made explicit. Hutchins provides an account of these in terms of the meanings of expressions and the 'transportation of knowledge' (p. 218). He accepts that transporting knowledge between individuals, and between individuals and a 'public context', may require some changes in the meanings of words (p. 218) and that meanings cannot be straightforwardly assigned to messages (p. 238). However, his analyses of activities is primarily in terms of assigning meanings to utterances and visual conduct to reveal how knowledge is passed from one individual to another. This
account does not seem to be so far removed from that of others within Cognitive Science who are concerned with communication and the passing of information between individuals (Clark and Schaefer, 1989). It relies on there being stable meanings of utterances (or messages) on which individuals can rely for them to be interpreted. This can be negotiated and subject to disruption, however, the accomplishment of interaction is primarily through a communicative interchange of information.

There is some evidence that Hutchins does conceive of a more radical notion of meaning. He states that ‘The illusion of meaning in the message is a hard-won social and cultural accomplishment’ (p. 238-9). This would seem to apply a more performative conception of meaning and knowledge. However, by maintaining a conception of an individual processing messages and knowledge which has to be passed or transported to others, knowledge and meaning have to be conceived of, to some extent, as stable and propositional.

Hutchins’ focus on naturally occurring activities and interactions does indeed provide a radical opportunity for ‘deconstructing’ Cognitive Science. The insights into the accomplishment of activities in a workplace setting suggest issues which have not been a concern of Cognitive Scientists. However, the programme of Distributed Cognition appears to be hampered by its maintenance of a conventional conception of individual cognitive activities. Even when apparently subsumed by broader social concerns, an account in terms of mental processing constrains the extent to which cognition can be reconceived. What remains is essentially a conservative account of social actions and accomplishments.

One final example illustrates this point. Hutchins provides an account of ‘why we talk to ourselves’ (p. 313-6). Speaking aloud, as when a quartermaster enters numbers into a navigational calculator, provides for transformations so that meanings can be interpreted in an appropriate medium in relation to a task. The vocal mediating structure provides for constraints or procedures in the relevant medium that can be coordinated with the accomplishment of the activity. Despite being accomplished in a setting where other individuals are present, this is primarily an analysis of mental processing. It could be possible to provide an account of such ‘self-talk’ in terms of social and public actions and as practical
accomplishments (cf. Goffman, 1981b). However, although Hutchins’ account of the individual mental process is conceived of as being embedded within a socio-cultural process, it is those conventional conceptions of cognition that seem to have primacy in the analysis.

2.5.3 Distributed Cognition as a Distinctive Analytic Framework

Hutchins does provide a distinctive analysis of the everyday actions and interactions through which an activity like navigation is accomplished in a real-world environment. With its concern on naturalistic materials it would be a significant development within Cognitive Science. But Hutchins’ contribution to the study is more than a substantive broadening of the topic. He aims to explore a range of analytic concerns such as the embodied use of artefacts, the coordination of activities between participants and the achievement of learning in organisations. In setting out such an ambitious programme of work it is not surprising if the analyses of Hutchins and his colleagues are more outlines of potentially novel research issues rather than more extensive studies. However, in addressing each of these concerns the conventional conception of cognition remains to account for individual mental processing. Despite Hutchins’ programmatics for considering cognition as a socio-cultural process, the individual conception appears to provide the resources through which accounts of communication and social interaction are given.

Hutchins also maintains the conventional concern of Cognitive Scientists with accounting for cognition in terms of computational processes. In comparison to other studies of activities less-constrained than vision, Hutchins seeks to take seriously Marr’s criteria for specifying information-processing models at a level which is not tied to representations and can operate across cultural implementations. This appears to lead to a computational theory of cognition in which computational processes are not specified. Hence, it is unclear in what sense descriptions at this level can be considered computational and whether there are any constraints on such theories.

Heath and Luff (1992a) provide a quite different account of the practical social activities surrounding the ‘self-talk’ of participants in a London Underground transport control room. Here, there is also a concern with coordination of activities, but this is with the incipient, or potential, courses of action of others. ‘Self-talk’ can be utilised by colleagues as a resource for participating and contributing to the activities of others.
Hence, by diverting away from key elements of the work within Cognitive Science - the use of a strong model of computation and the individual conception of cognition - Hutchins appears to be in danger of relinquishing its criteria for adequate explanations of human behaviour whilst still remaining shackled to its conventional conceptions. Of course, it may not be necessary to hold onto both of these concerns. For many of Hutchins' observations there neither appears to be a need for an underlying model of computational processes nor for an account of individual cognitive activities. Hutchins can be seen to be outlining some practices and methods by which individuals coordinate their activities in natural settings. In order to achieve these, participants rely on a set of competencies and common-sense resources. It may not be necessary either to account for these in terms of a single cognitive process, or to maintain that a heterogeneous set of concerns surrounding knowledge be considered under one rubric. This is a more radical conception of distributed or social cognition, one which draws from a social science orientation.

In the next section, an alternative analytic framework for considering rational action and human conduct is outlined, one which pays attention to the practices and methods by which participants achieve social actions. Drawing on recent developments in ethnomethodology and allied orientations, researchers have been concerned with how interactions and activities are accomplished in natural settings. Although some commentators, including Hutchins, have included this work within the rubric of Distributed Cognition (p.371), its provenance is from a different discipline with differing concerns and motivations. It is also an area of work which has also been of recent interest to researchers in HCI and CSCW. It is to this orientation to which we will now turn.

2.6 A SOCIAL SCIENTIFIC CONCEPTION OF COGNITION

In developing their programme of Distributed Cognition, Hutchins and his colleagues can be seen to be either broadening the concerns of Cognitive Science or attempting to respecify its key principles. Such developments have been suggested by the limitations of the field, they can also be seen in the light of a sustained criticism of the Cognitive Science programme. In the last ten years a set of critiques have emerged that question the foundational conceptions of Cognitive Science. These have been presented from a variety of disciplinary perspectives including; phenomenology
(Bolton, 1991), ecological (Morris, 1991) and discursive approaches to psychology (Edwards and Potter, 1992; Shotter, 1991) and cultural theory (Lave, 1988). These critiques have drawn on a disparate collection of works including Marx, to outline the practices through which cognitive activities like calculations are accomplished in natural settings (Lave, 1988), and Heidegger to collapse distinctions between mental representations and actions (Winograd and Flores, 1986). In this section we will briefly explore one orientation from which critiques of the conceptual underpinnings of Cognitive Science have emerged. This is an orientation deriving from the original work within the social sciences of Harold Garfinkel (1967), which has been termed ethnomethodology. This has attracted interest not only because of its novel analytic concerns, but also because its focus on the accomplishment of practical actions has been considered as potentially useful in the process of system design. This orientation will be considered by paying attention to one study, Suchman's (1987) 'Plans and Situated Action' - a work that has been significant in three respects. First, it suggests a reformulation of a particular concern of Cognitive Scientists: planning and plan-based models of action. Second, it appears to offer an alternative orientation for the analysis of artefact-centred activities (particularly 'human-machine communication'); an orientation that has been termed a 'Situated Action' approach (e.g. Nardi, 1996c; Vera and Simon, 1993a). Third, it has been utilised as a key resource for an emerging domain of study, that of Computer-Supported Cooperative Work (CSCW).

2.6.1 Situated Action

Suchman draws on ethnomethodological studies in order to re-examine aspects of human-computer interaction. In particular, she utilises work on the analysis of talk as a background against which to consider the nature of the interaction, or communication, with a computer system. The work she draws on, conversation analysis, derives from ethnomethodological concerns. Researchers in conversation analysis have sought to outline the social practices through which, from moment-to-moment, the production of turns of talk is both accomplished and rendered

In a later article, Suchman (1993b) states an uneasiness with the term human-computer-interaction (p.73). In the book, she opts for human-machine communication, but even this formulation is more of a resource for discussion rather than a conceptualisation of the activity (cf. Suchman, 1987, p.118).
intelligible to others. These studies of talk rely on detailed analysis of the ways and methods by which activities are achieved in naturalistic settings. In line with these concerns, Suchman takes as materials for her study recordings of various pairs of individuals trying to use a complex 'intelligent' photocopying machine.

Suchman takes as resources for her analysis what is available and public for the participants, both through their talk and displayed on the machine. From these resources she reveals some of the ways in which the participants achieve an understanding of the operation of the machine, and manage to get the machine to produce copies. Thus, individuals utilise the locally available displays, commands and configuration of the machine to shape their understandings of its behaviour. By taking particular actions, the participants are provided with a further set of resources, which in turn contribute to the accomplishment of the task and their understanding of the machine's operation. These local understandings are contingent, displayed from moment-to-moment, through the participants' talk and actions. They are developed from the actions taken on the machine and also shape future ones.

Suchman compares the accomplishments of the users with the rationale of the designers who developed the machine. The device, in common with many AI systems (and sophisticated prototypes of intelligent interfaces), developed a plan of actions for the user from what was accessible from the users' conduct (through the sensors and buttons on the machine). Thus, the system developed what is commonly known as a 'user model' - an explicit model of the user (cf. Norman, 1986). The various displays and commands on machines were designed with respect to the current state of the model and these pre-specified plans. Unfortunately for the designers, the resources available to the machine were not symmetrical with those of the users. In making sense of the local resources with respect to their current circumstances, participants would utilise common-sense understandings of actions so that, for example, when a further instruction is displayed it can be seen to confirm a prior action, a non-response provides evidence of a trouble, and a repeated instruction could either be request a repeated action or invoke a repair of previous one. It would appear that these sense-making resources would be indefinite, and difficult to pre-specify in a machine. Because of the 'profound and persisting asymmetry in interaction between people and machines' (p. 185), Suchman questioned the utility of developing such
plan-based user models in the design of computer systems. She then reflects back on the origins of such models from Cognitive Science (e.g. Miller, et al., 1960).

From her analysis, Suchman suggests a reformulation of the conception of plans. Rather than pre-specifications of courses of actions or controlling structures, plans can be considered as resources for actions. This turns around the conventional problem of planning for Cognitive Scientists. Rather than aiming to outline plans that later get instantiated, or 'filled out' in courses of actions and interactions at some 'operational level', an analysis could explore how individuals are able to 'bring efficient descriptions (such as plans) and particular circumstances into interaction' (p.188). When considering the relationship of plans to situated actions, primacy is given to situated actions, with plans being one of the many things these produce (Suchman, 1993b, p.72).

In response to this study, some Cognitive Scientists have attempted to re-state their conception of plans, in terms of information and physical symbol processing (e.g. Vera and Simon, 1993a). They seek to give back the primacy to planning, and more importantly, to an account of action underpinned by symbol processing. For them, situated action can simply be viewed in terms of symbolic systems that 'operate adaptively in real time in complex environments' (p.47). In parallel with certain arguments of Hutchins, social processes are important to consider, but these too can be brought into a cognitive simulation (p.43), by dealing with the 'mutual transfer of knowledge or 'simply aggregating the system to treat collective knowledge as shared memory' (p.43).

Such a treatment of situated action leaves the architecture of Cognitive Science in place. Symbols, internal representations and plans can still be considered to determine actions, just they have to be accounted for 'in context', as flexible and dynamic. This account misses the fundamental respecification that Suchman suggests. Her reconceptualisation of plans draws on earlier ethnomethodological considerations concerning the nature of social action. These arguments, themselves, can be seen to have precursors in earlier philosophical debates concerning the nature of cognition and the attribution of conduct to mental processes. In order to see how Suchman's analysis suggest a more fundamental reconsideration of Cognitive Science, and HCI, a brief discussion of the respecifications of cognitive and mental conduct, drawn from an ethnomethodological orientation, is required.
2.6.2 An Ethnomethodological Orientation

In Garfinkel's original formulation the term ethnomethodology does not specify a sociological method or approach, but rather a study of a particular range of phenomena: those procedures and methods which members of a society utilise in order to make sense of, and produce orderly social actions (Garfinkel, 1974). So,

In exactly the ways that a setting is organized, it consists, of members' methods for making evident that settings' ways as clear, coherent, planful, consistent, chosen, knowable, uniform, reproducible connections - i.e., rational connections. In exactly the ways that persons are members to organized affairs, they are engaged in serious and practical work of detecting, demonstrating, persuading through displays in the ordinary occasions of their interactions the appearances of consistent, coherent, clear, chosen, planful arrangements.

(Garfinkel, 1967, p. 34), italics in the original.

From the original programme of study and empirical exercises, a wide range of studies of 'naturally organized ordinary activities' have emerged. These include such apparently diverse interests as:

- analyses of talk-in-interaction, and the practices through which turns at talk are produced and rendered intelligible in conversation (Sacks, 1992; Sacks, et al., 1974);
- an analysis of the tacit practices by which the work of proving mathematical formulae are produced and recognised by professional mathematicians (Livingston, 1986);
- an account of the 'code', or rules of conduct, and its uses for members of a half-way house for narcotics offenders (Wieder, 1974);
- an account of the practices through which coroners assess and categorise potential cases of suicide (Atkinson, 1978);
- a study of the organisation of conduct in learning and producing jazz improvisations on a piano (Sudnow, 1978).

Originally, the term was intended to parallel other related studies emerging at the time, such as ethnobotany and ethnomedicine where members own practices of botany and medicine were considered. Thus, the phenomena of interest to ethnomethodology is members' own methods, their mundane knowledge and common-sense reasoning procedures (Garfinkel, 1974), that is

...not a method of understanding, but immensely various methods of concerted actions and methods of common understanding are the professional sociologist's proper and hitherto unstudied and critical phenomena.

(Garfinkel, 1967, p. 31)
Throughout these and many other studies a range of concerns emerge. These reemphasise Garfinkel's original concerns with the methods by which members make sense of, and produce recognisable, organisational activities, and include:

- that rational and methodological properties of action are treated as contingent accomplishments of socially organised practices (cf. Garfinkel, 1967, p. 33);
- that members make observable the rational character of their actions (cf. Garfinkel, 1967, p.8);
- that in order to accomplish practical activities in a setting members take for granted what a member must "know" (cf. Garfinkel, 1967, p.8);

Thus, ethnomethodological analyses are studies of rational action. However, the conception of rational action, shared agreement, 'what is known in common' and the like, is quite distinctive. These are social methods and accomplishments of a particular kind. They rest on the everyday practices and common-sense reasoning of participants. Moreover, the intelligibility of social actions are revealed and displayed to others in and through their production.30

These foundations have led to a reconsideration of critical concepts within sociology, including the use of formal structures (Garfinkel and Sacks, 1970), measurement (Lynch, 1992), culture (Lee, 1992) and context (Schegloff, 1992a; Schegloff, 1992b) as well as more general concerns in the social sciences (Button, 1991b). Amongst these reconsiderations and respecifications has been a reexamination of cognition, the mind and mental processes — a particular interest in the various books and articles of Jeff Coulter (1979; 1989; 1990; 1992).

---

30 So, for example

Shared agreement refers to various social methods for accomplishing the members' recognition that something was said-according-to-a-rule and not the demonstrable matching of substantive matters. The appropriate image is therefore an operation rather than a common intersection of overlapping sets. (Garfinkel, 1967, p. 30).

Note, that in 'Studies in Ethnomethodology' and earlier writings Garfinkel draws explicitly from Schutz' theoretical work (e.g. Schutz, 1962) with regard to 'taken for granted' understandings, common-sense knowledge and 'seen but unnoticed' background expectancies (Garfinkel, 1967, p.53-7, Heritage, 1984b).
In keeping with the ethnomethodological orientation, Coulter explores the occasions of members' uses of cognitive phenomena, treating these as topics of study. For this analysis he draws on ordinary language philosophy (Austin, 1962), and what might be termed logico-grammatical resources of the kind developed by Ryle (1949) and the later Wittgenstein (e.g. Wittgenstein, 1953; 1958). Hence, he examines such phenomena as 'understanding', 'remembering' and 'intelligence' and looks at how these 'mental predicates' can be combined and used in practice. Shifting the domain of interest in this way involves a consideration of the nature of the uses of, for example, 'understanding'. It is not a verb used to account for some inferential process, but an achievement (cf. Ryle, 1949).

Understanding is a part of knowing how. The knowledge that is required for understanding intelligent performances of a specific kind is some degree of competence in performances of that kind.

(Ryle, 1949 p.53), italics in the original.

Although producing an intelligent performance is distinct from recognising it, the two rely on similar competencies. Similarly, remembering is not an inferential process but a (defeasible) achievement. It makes no sense, for example, to recall or recollect things that have not been witnessed, nor is it legitimate to 'recall unsuccessfully' or 'incorrectly' (Ryle, 1949, p.260-3). Such mental predicates relate to competencies and achievements, not internal, inferential processes operating on 'sources' of knowledge. To assign them to the non-spatial workings of a mind and apply a framework developed in relation to a physical process is a 'category-mistake' (p.20). This is not because there are no such things as mental processes, but because mental processes are not the same kind of thing as physical processes. It is inappropriate to import conceptions, such as mechanical causation from the latter domain into the former:

Minds are things, but different sorts of things from bodies; mental processes are causes and effects but different sorts of causes and effects from bodily movements.

(Ryle, 1949, p.20)

The consideration of these activities as accomplishments, coupled with a concern for the contexts-of-use of cognitive phenomena suggests an alternative empirical programme for the study of knowledge and cognition, one which explores how memories, forgettings, agreements and understandings are achieved as social actions. Indeed, through the analysis of naturally occurring interactions and activities, researchers have begun to explore, for example, how agreements are accomplished in
conversation (Pomerantz, 1984; Sacks, 1987), how forgetfulness can both be displayed for participants and a resource within interaction (Goodwin, 1987), and some practices through which understandings and informings can be displayed (Heritage, 1984a).

Suchman's analysis of plans can be considered in the light of these analytic developments. In her respecification of planned action, Suchman is seeking to turn the analytic attention onto how plans are accomplished *in situ* within interactions. By focusing on plans as achievements, these are being considered as 'knowledge that' rather than 'knowledge how' (cf. Ryle, 1949). Interpreting them in terms of the results of internal processing, no matter when this processing occurs, how dynamic and flexible it is and what external factors contributed it, maintains the conception of mental processing in terms of physical processing, in terms of physical causes, effects, states and processes (cf. Ryle, 1949, p. 20). Thus, it is not the domain of enquiry of Cognitive Science which is of concern. Expanding the field to account for social processes, no matter how benevolently, will not reconcile the differences between it and other frameworks, such as those drawn from an ethnomethodological orientation. Nor will just reconceiving cognition. Instead, the reconceptualisation that is required by recent work in philosophy and the social sciences involves a rethinking of the entire conceptual framework which is utilised to address the practices surrounding cognitive and mental processes and the uses of knowledge.

2.7 DISCUSSION

It is, at first, unclear what the direct relevance of such a respecification of mental processes in general, or planning in particular, would have for developers of new technology. The respecification, though addressing concerns which are fundamental to the underpinnings of much of the work undertaken within HCI, would appear to have little consequence for the design or deployment of technologies. However, the relevance may be not so much in terms of the domain of the study, but more on the orientation adopted for analysing practical, social actions. The potential relationships between the analytic orientation and the development of computer systems will be considered in more detail later in the thesis. In the discussion here, what will be considered will be the potential for developments in the analysis of the use of artefacts, in particular, of
computer systems. The relationship between these developments and the more engineering concerns, such as requirements analysis, and the design and deployment of new forms of computational support for practical activities will be considered in more detail in chapter 7.

An ethnomethodological orientation could provide a novel perspective with which to consider artefact-centred activities, particularly the use of computers. The concerns outlined above, at least, suggest a shift towards naturalistic studies. So, a programme of studies could be initiated which explore the situated use of computer systems. Such a study would explore the resources through which practical activities are achieved through the use of technologies and artefacts, and how these utilised common-sense knowledge, practical reasoning and tacit practices. Thus, an ethnomethodological orientation does not just suggest a reconsideration of an existing framework for analyses of human behaviour, but a novel programme of research activities. In the last few years such a body of work has emerged that has begun to explore the in situ accomplishment of activities in various ‘tool-saturated’ settings (Goodwin, 1993). Because of the interest in the achievements of social actions in interaction, it is perhaps not surprising that these studies have explored, in various ways, how collaborative activities are accomplished in natural settings, drawing from an ethnomethodological orientation and informed by studies in conversation analysis and visual conduct (e.g. Goodwin, 1981; Heath, 1986).

So, for example Goodwin and Goodwin have explored activities in an airport control room: the ways talk, such as instructions, are produced and made sense of within viewings of the local domain including various technologies such as screens, boards and Close Circuit Television monitors (Goodwin, 1992b; Goodwin and Goodwin, 1996; Goodwin, 1996). Whalen (1995b) has analysed the interrelationship between talk and the use of a system for computer aided dispatch, revealing the practices through which the production of particular commands are organised with respect to the interaction between caller and dispatcher, and vice versa. These studies, with others by Heath, Luff and Greatbatch (Greatbatch, et al., 1993; Heath and Luff, 1992a; Luff, et al., 1994), suggest the foundations for an interactional analysis of the use of artefacts.

This work on the use of artefacts in interaction has begun to outline some of the practices which appear to be relevant to the accomplishment of technological work. So, for example, studies have revealed the various
forms of participation individuals have with each other and in relation to the technology (Goodwin, 1992b; Goodwin and Goodwin, 1996; Heath and Luff, 1992a). Hence, individuals may not just achieve tasks in focused collaboration with others, but may rely on monitoring the conduct of others, particularly their activities in relation to the uses of some technology. Many activities also appear to be coordinated with the activities of another, for example, a turn at talk commencing on a boundary after a sequence of typing (Greatbatch, et al., 1991), or the typing on a computer interrelated with the talk in a phone call (Whalen, 1995b). These practices are tacit, 'seen but unnoticed' by the participants.

In this preliminary work, resources from conversation analysis have been utilised to begin to map out fragments of data, suggest some general issues for study and to explore how participants produce and render intelligible not only turns at talk, but also non-vocal activities. Studies of conversation analysis have revealed some of the systematic practices and methods by which participants in a turn of talk both display an understanding of another's prior turn and though its production display a sensitivity to potential recipients. Thus, turns of talk are sequential, both shaped by and shaping the emerging context (Heritage, 1984b). Researchers of naturally occurring activities have begun to explore the possibility of providing sequential analyses of conduct with respect to technologies and artefacts. Needless to say, these initiatives are preliminary and it remains unclear whether and how transformations can be made to provide a sequential analysis. Hence, the early work has outlined some broad early glosses of practices, such as monitoring, coordination and peripheral participation. These issues will be considered in more detail in chapter 6.

Although this work has been undertaken in technology-rich settings, for example control rooms and service centres, it is perhaps strange that the details of the uses of the technology have not been primary to the analysis. This is partly due to the materials available for analysis. In collecting data, particularly video data, it is difficult to focus both on the interactions between participants and the detailed activities with the technology and other artefacts. Hence, the analyses have tended to focus on the public displays, the generally available screens and artefacts that are available (Goodwin and Goodwin, 1996; Heath and Luff, 1992a), and the interrelationships between such activities as turns at talk and the hands movement across the keyboard (Greatbatch, et al., 1993). These
analyses centre on what is publicly available to the participants in a co-present environment, resources that are visible or audible and can be both utilised by the 'computer user' and others in the local setting. In outlining such practices, the flexibility, contingency and variability of collaborative activities has been emphasised. In doing this, researchers in the field of CSCW, have drawn implications for the general requirements for collaborative technologies (Benford, et al., 1996; Hudson and Smith, 1996; Kristoffersen and Rodden, 1995). However, these preliminary studies of the production of technology-centred activities have also laid out analytic concerns and issues that need further exploration. To take just three:

- *monitoring*, or mutual monitoring, itself appears to be a gloss for a range of interrelated activities including: maintaining an awareness of the 'goings on' in a domain; general overlooking or overseeing another's activities, and monitoring for the boundaries or junctures in another's activities (Goodwin and Goodwin, 1996; Heath and Luff, 1992a);

- similarly, *peripheral participation* can cover activities out of the line of direct regard by another, non-focused collaborative activities, and even, by other authors, apprenticeship and learning prior to more full immersion into a professional setting (Brown and Duguid, 1994; Lave and Wenger, 1991);

- *coordination* accounts for the tying together of diverse activities, such as typing or writing, to talk but it is unclear to what extent these non vocal activities can be said to display an understanding or render intelligible a prior activity (Greatbatch, et al., 1993; Heath, et al., 1993).

These early observations are no doubt related: the coordination of an utterance, so that it is tied to a completion of a typing sequence, requires (often peripheral) monitoring of another's activities. The very conception of an activity as peripheral or as monitoring suggests some ambivalence to its nature which may not be available for further analysis. However, because of these concerns, what may be noticed about these early studies of complex domains is that the detailed nature of the activity in relation to the artefact, particularly the computer system, fades from view. Although this may be private, unseen or only partially available, the computer systems may be providing resources which shape the ongoing activities.

In an example given by Whalen (1995b), the organisation of the items on a screen and the orderings through which operations can be carried out
on a system can be seen to shape, to some extent, the talk on the phone of a dispatcher of emergency services. The activities on the system, unavailable to the caller at the other end of the line, appear to be relevant to the public and social accomplishment of the activity.

Paying attention to the details of computer use, the displays available to participants and the activities carried out using them is a concern of this thesis. This is explored not only in cases, like Whalen's, where the activity is distributed between different locales (chapter 5), but also in co-present settings (chapters 4 and 6). In making sense of these activities, participants rely on common-sense reasoning and practical resources in order to accomplish activities on the system.\(^{31}\)

Such a focus brings attention back to the domain commonly of concern to HCI. Where related to the development of computer systems, preliminary studies of the accomplishments of activities in complex domains have been principally considered in relation to collaborative systems. By exploring the details of activities achieved on and through computer systems, it appears that the concerns of HCI and CSCW may not be quite so distinct. The use of a computer system is shaped by social practices, and social actions and interactions are shaped by the resources available on the technologies. The reformulation of the concerns of HCI, led in particular by Suchman's work, may not mean that researchers would be better directed to addressing themselves to the field of CSCW. This may, indeed, avoid some of the critical issues with respect to both the analysis of artefact-centred activities and the design of new technologies. Instead, what is required is an orientation that can underpin both substantive domains, and reveal the practices and methods through which individuals, with others and through the resources they have around them, 'can produce stable accountable practical activities i.e. social structures of everyday activities' (Garfinkel, 1967, p.185).

2.8 SUMMARY

Researchers in HCI have been concerned with the uses of artefacts, particularly computer systems. By looking at the ways particular objects are used it may be possible to replace them with computational alternatives or to suggest reasons why current systems are not being used

\(^{31}\) An initial examination of these skills and practices is considered in chapters 3 and 4.
effectively. Similarly, detailed analyses of the uses of computer systems should be able to provide resources for improving particular designs or, more generally, guidelines for the development of systems. To aid them with this programme, they have turned to Cognitive Science for methods, models and a conceptual apparatus with which to consider computer use.

This chapter has explored the methodological foundations offered by Cognitive Science, not only with regard to its computational foundations, but also in terms of the ways these have been instantiated within models of HCI. Although models such as GOMS do provide a way for analysts to attend to the details of computer use, they have a limited domain of interest. It is also unclear to what extent they, in fact, do provide a cognitive model of the activity under investigation. This chapter has noted not only the rather limited extent to which cognitive activities figure in the models, but also the homogenous ways in which mental processing is treated, merely as gaps and pauses between other physical activities, where mental activities take place. Moreover, when some of the original objectives of the information processing programme are considered, particularly with regard to providing explanatory theories of cognition, it becomes unclear to what extent cognitive models within HCI satisfy these.

Interestingly, a model of cognitive processing that explicitly attends to providing an explanatory and computational account of cognition, Distributed Cognition, also seeks to extend the domain of the study and transform its underlying conception of cognition. Though still a provisional framework, it appears problematic to maintain an individual conception of cognition and a computational model of symbol processing, whilst also seeking to account for the contingencies of social actions and interactions. From one perspective, the perseverance with individual cognition constrains the framework, from another it appears to be unnecessary.

A quite distinctive approach to considering the concepts under scrutiny – cognition, mental processing and the mind – is as topics for analysis, through participants' own linguistic and social practices. Such an orientation in philosophy, linguistic philosophy and the social sciences suggests that the perspective taken by Cognitive Science fundamentally misconstrues its principal domain of concern. Mental activities are not processes occurring in some inner domain but social accomplishments. By assigning these to the individual is to mistakenly categorise these in the same order as physical activities and processes.
From the ethnomethodological orientation it has not only been possible to draw out critiques of cognition and allied conceptions relevant to Cognitive Science, but also to reconsider its domain of study - human behaviour. Hence, it may be that such an orientation may provide a framework for analysing how activities are accomplished on, with and through new technologies. Some such studies have been undertaken in a range of domains. What these have in common is a concern for the in situ accomplishment of activities and the resources participants utilise, through their interactions, to make sense of the conduct of others in relation to the technology.

At present, these studies are at an early stage. However, their distinctive orientation raises a set of issues with regard to various ways in which computer systems, and other artefacts, can be utilised. The resources through which such activities are accomplished remain unclear, but do suggest a range of ways in which computers can be considered to be 'used' or 'interacted with'. As Suchman notes, the formulation 'human-computer interaction' may be inappropriate for such studies. Nevertheless, computer systems, like documents and other artefacts, do provide resources through which participants accomplish activities and make sense of the conduct of others. In the remainder of the thesis, a range of settings with different configurations of technologies are examined. The resources, the common-sense reasoning, the practical knowledge and the tacit practices, that individuals utilise in these settings are the principal focus of these studies. By revealing how technology-centred activities are achieved through the interactions of participants, and how they inform those interactions, potential resources for the development process might be provided. At the conclusion of this thesis, the relationship between this analytic framework and the design process is considered. Such an approach suggests the foundations for an alternative way of conceiving and implementing 'user centred design'.
Chapter 3

The Practicalities of Menu Use:

improvisation in screen-based activity

I began to see and then find use for further work in the observation that note choices could be made anywhere, that there was no need to lunge, that usable notes for any chord lay just at hand, that there was no need to find a path, image one up ahead to get ready in advance for a blurring out. Indeed, to conceive particular terrain places up ahead seriously undermined the singing that I sought to sustain. Good notes were everywhere at hand, right beneath the fingers.

(Sudnow, 1978, p.94)

I began to find, in the undulating nature of my entrance and pacingly tuned interdigitations, that I could undertake new sorts of shaped and rated courses with well-at-hand route segments.

(Sudnow, 1978, p. 130)

3.1 INTRODUCTION

The critiques of work within HCI that derive from psychology and, more particularly, Cognitive Science have led researchers in the field to consider different approaches to the study of computer use. These include not only developments to the existing frameworks, but also the use of resources
from different disciplines. The work of Suchman, outlined in the last chapter, has been critical to these; suggesting other features of computer use to focus on, a different, contingent conception of artefact use and alternative studies, principally from social science, to draw upon. In this chapter, a preliminary investigation of human-computer interaction is undertaken which draws on recent work in the social sciences.

Suchman (1987) utilises resources from ethnomethodology and conversation analysis in her analysis of 'human-machine communication'. Section 3.2 briefly reviews work within HCI that has been influenced by Suchman's consideration of the 'situated actions' surrounding the use of computational artefacts. Some of this work applies the resources, particularly of conversation analysis, more directly. This application of the 'findings' of conversation analysis has recently been subject to criticism, with attention being focused on the status of the objects revealed in talk and their appropriateness to the analysis to activities in human-computer interaction (e.g. Button, 1990; Button and Sharrock, 1995). The focus of this chapter is to explore screen-based activities without the assumption that the use of objects on the screen has any similarities to activities in talk. The domain of this analysis is one familiar to researchers in HCI, the use of menus.

The key elements of this study within the field are outlined in Section 3.3. Most analyses of menu use have been highly constrained and have been accomplished through an experimental paradigm, however, recently there have been some exceptions. By adopting a more naturalistic approach, the findings from these appear to be surprising and have led to a rethinking of the underlying models concerning computer use.

The materials utilised for the analysis in this chapter are similarly naturalistic, being video-recordings of a simple exercise in using a popular computer application, PowerPoint. In Section 3.4 the details of this exercise are given with some background concerning the PowerPoint application. The instances of menu activity are analysed in some detail. In order to support the analysis a transcription system for screen-based activity has been developed. The critical features of this system are also described in Section 3.4.

As menu use, in particular, and screen-based activity, in general, has been a topic of considerable interest within HCI, a terminology has emerged related to these activities which draws on framework from Cognitive Science. Hence, uses of menus are often considered as searches
for particular items or general browses, which in turn are conceived in
terms of the user's goals, plans and knowledge, that is the 'knowledge that'
a user possesses (cf. Ryle, 1949). In order to reconsider this conception, a
different strategy is utilised for this investigation. A provisional typology
of menu uses is outlined in Section 3.5. This typology is not intended as a
classification of user behaviour, but rather as an analytic device for
revealing some of the resources that individuals utilise in order to carry
out activities on a computer system.

Menu activities of increasing length and complexity are considered in
terms of eight basic temporal and sequential types. By examining cases of
these types some of the resources individuals utilise when using the
menus are revealed. In particular, the menus offer the user resources for
a continually changing set of possible course of actions. These can provide
valuable means for accomplishing screen-based activities, and also can
allow a user to improvise and digress. The reading and use of the menus
can both shape an ongoing course of activity and be shaped by it. To
perform activities on the menus, relies on being able to make sense of the
possibilities being presented in terms of a shifting set of resources and an
ongoing activity that can be transformed. This relies on practical
reasoning and common-sense knowledge, that is 'knowledge how' (cf. Ryle,
1949).

Section 3.6 briefly discusses the relevance of such a study for other
approaches towards human-computer interaction. This discussion will
draw particularly on work by Young and his colleagues (1990) which aims
to categorise the knowledge users require to operate a system. It appears
that though these models do seek to explore a set of wide-ranging
concerns, this would have to be by formulating the user's knowledge and
competencies as propositional knowledge. The study undertaken in this
chapter seeks to reveal some of practices through which screen-based
activities are accomplished. The focus of the study, on a particular
circumscribed activity, makes it problematic to utilise analytic resources
drawn from conversation analysis. The status of the analysis is discussed
in the light of these concerns. Nevertheless, through the examination of
the details of naturalistic screen-based activities the analysis suggests
how even such a simple activity as the use of menus may be practically,
skillfully and contingently managed and organised.
3.2 BACKGROUND

The critiques of Suchman (1987), and others (Lave, 1988; Winograd and Flores, 1986) have generated a range of responses from researchers within HCI, ranging from those that reject entirely the critique (e.g. Vera and Simon, 1993a) to those who have sought to transform the study to adopt 'social concerns' (e.g. Brown and Duguid, 1994). In less extreme responses, there are have been suggestions to either extend cognitive models in HCI or to utilise approaches drawn from the social sciences to analyse human-computer interaction. In the former case, researchers like Young et al. (1990) have sought to outline a range of different classes of knowledge utilised when carrying out activities on a computer system. In the latter case, attention has focused, in particular, on drawing upon the resources utilised by Suchman, namely ethnomethodology and conversation analysis.

For Young et al. the disadvantages of models, such as GOMS, are due to their not taking account of the information available in the 'external situation' and how it combines with the user's internal knowledge (Young, et al., 1990, p.115). This requires a better understanding of 'interactivity' within HCI. Thus, they suggest, a model is needed which outlines how internal knowledge of different kinds 'interacts' with external information provided on the device. This proposal is considered through the analysis of particular instances of menu use in Section 3.5.

The work of practitioners who have sought to utilise conversation analysis within HCI have a similar set of concerns with the interaction between a user and a system. So, for example, Norman and Thomas (1990) use 'findings' from conversation analysis, such as rules for turn-taking and the repair of troubles, as a resource for proposing design guidelines, Bowers and Churcher (1989) analyse electronic mail in terms of 'turn-taking' and Frohlich and Luff (1990) design an interface to an expert system which has system and user 'turns', and sequences for the 'repair' of trouble, changing 'topic', 'openings' and 'closings'. In these early studies a direct application of findings from studies of social interaction has been undertaken, with organisations of talk between conversational interactants being used as resources for suggesting an organisation of the contributions between a user and a computer.1

1 There are a range of other studies of human-computer interaction also deriving from conversation analysis which include Cawsey, (1990); McTear, (1985); Payne,
This general approach has not been confined to those principally concerned with the design of interactive systems. For example, Thomas (1990) compares an analysis of the problems of individual users of a computer system with an analysis of troubles in a conversation. Similarly, Payne (1991) draws from an analysis of turn-taking for resources for analysing the contributions to a human-computer interaction (Payne, 1991).2

Payne takes an approach to HCI that draws on a model of conversation proposed by Clark and Schaefer (1989). In particular, Clark and Schaefer embed in their model the notion that a single turn can be doubly contextual (see Heritage, 1984b). That is, a turn is both context-shaped, displaying an interpretation of preceding turns, and context-renewing, contributing to the context in which the next turn will be understood. Thus, Clark and Schaefer propose that contributions to a discourse have an acceptance phrase and a presentation phrase, that is, a single utterance can be the acceptance phrase to the prior contribution and the presentation phrase of the next. These contributions can then be represented on a 'contribution tree'. Payne proposes similar 'interaction trees' for HCI. Again, a single action by a user or the device can be both an acceptance and a presentation.

Although from less direct sources than Thomas and others, Payne adopts a model of conversation for the analysis of HCI. Although, for Payne, a contribution may have a double duty, he still adopts essentially a turn-taking model for human-computer interaction. Such an approach is attractive. As mentioned in the previous chapter, the potential

---

1991). Of principal interest for these researchers has been the work on the social organisation of turn-taking and repair in conversation, particularly that formulation of the analysis given in a few specific works by Sacks and his colleagues (i.e. Sacks, et al., 1974; Schegloff, et al., 1977).

A similar set of resources have been considered by researchers looking at different aspects of computer use, including Computer-Mediated Communication (McCarthy, et al., 1990) and speech understanding and generating systems (Fraser and Wooffitt, 1990; Gilbert, et al., 1990).

Payne does not directly cite work from conversation analysis, but utilises work by Herb Clark and his colleagues on language use and discourse (Clark and Schaefer, 1989). Clark draws extensively from conversation analysis including adjacency pairs (Schegloff and Sacks, 1973) and side sequences (Jefferson, 1972), as well as the organisation of turn-taking (Sacks, et al., 1974). Further resources from conversation analysis are outlined in his later, more extensive, work (Clark, 1996a).
similarities between human interaction and that with a computer has been a common resource for those in HCI seeking to conceptualise the activity of computer use, as well as for designing novel systems. It would, therefore, appear to be useful to draw from detailed studies of social interaction for rules, procedures and methods to analyse human-computer interaction and for the design of interfaces. However, such a direct utilisation of the work of conversation analysts has been called into question.

Button (1990), in remarks echoed by Sharrock and Anderson (1991), is sceptical of the value of using ‘rules’, ‘sequences’ and ‘routines’ revealed by conversation analysts for computational models. He suggests that the differences between both the nature of human interaction and that of human-computer interaction, and the ‘rules’ posited by conversation analysts and the rules required for computer systems make it unlikely that the ‘findings’ of conversation analysis can be usefully applied to the design of systems. In particular, he draws attention to the status of a rule in conversation analysis as a resource utilised by participants; a resource whose application is contingent to the moment-to-moment circumstances. Such rules are not deterministic, and are not followed by participants, but are oriented to. Their application reflects a moral order. Even when the organisation appears to be subverted or diverted from, there still is displayed in the contributions of the participants an orientation to the ‘rules’.3 The rules a system is designed to follow are of a different nature; their design cannot take account of the participants’ practical purposes, their ongoing concerns for undertaking a social interaction.

So, the contributions for system and user suggested by Payne, are of a different kind to the contributions to a conversation. In conversation, the order of a conversation is revealed through the participants’ production of talk: by displaying in a turn of talk an understanding of a prior utterance and by being designed from moment-to-moment with respect to a recipient (Sacks, et al., 1974). Thus, a turn of talk, whilst it is being uttered, can be transformed in the light of a contribution by a co-participant. The production of a turn of talk reveals members’ practices, methods, for making sense of, and displaying an understanding of, the activities of another. Contributions to and by a computer are not of this kind. Not least because the nature of the resources are asymmetrical. When

---

3 ‘Rules are oriented to features of action, they are contextual, situated practices of use’ (Button, 1990, p. 84).
exploring human-computer interaction, it is no longer possible to analyse an 'interaction' in terms of participants using the same resources to produce talk and display understanding of that talk. Nor is it possible to rely on a co-participant's turn to display an understanding of a prior. Therefore, rules, sequences, and routines revealed by such an analysis of conversation are no longer likely to be appropriate for such a different kind of activity. Hence, considering human-computer interaction as turns may be useful only as a metaphor with which to consider the activity, rather than an analytic device, or a design tool. An 'interaction unit', consisting of an acceptance and a presentation, is a category constructed by the analyst, not one to which users nor, indeed devices, 'orient'.

Rather than directly applying findings and conceptions from conversation analysis, there may be other ways of following the programme suggested by Suchman. Although it may not always be possible to utilise as a resource the interactions of users whilst using a system, there may be other naturally-occurring materials that may provide resources for an analysis. In the following chapters, video-recordings taken from different work settings provide for the detailed analysis of the interactions and collaborative activities in which computer use is embedded. The analysis presented in this chapter takes a different approach.

Rather than examining human-computer interaction in terms of system and user turns, initiation of repair, changes to topic and opening and closing sequences, it draws on conversation analysis, ethnomethodology and cognate developments in the social sciences to outline a methodological framework for analysing some of the resources participants utilise in order to accomplish screen-based activities on a computer system. This orientation is further explicated in later chapters, but by taking a domain familiar within the field of HCI, it is hoped that the distinctive nature of the approach can be sketched out. The particular focus for this analysis is on the use of menus. As this is an activity given some considerable attention in HCI, a brief review of the way that this activity is conceived in the field may be useful.

3.3 ANALYSES OF MENU USE

When confronted with the problem of designing interfaces that give a user a wide range of disparate options, HCI practitioners have commonly
turned to the 'menu' as an orderly and convenient way of presenting these options. It is not surprising that a considerable amount of research has been carried out on the usability of various designs for menu systems. Most of this research has tested competing options by designing experiments where users select items from different configurations and then measure their speed and accuracy at these tasks. The menu designs tested in this way have ranged from the mundane to the exotic. For example, Landauer and Nachbar (1985) looked at response times of users when items are alphabetically or numerically ordered, and when the number of menus and the number of items on each menu are varied. Teitelbaum and Granda (1983) have examined the effect on the time users take to select items from a menu when information associated with the menu (e.g. the title and guidance) is displayed in a constant location and when it is varied from screen to screen. Various alternative mechanisms have been explored for manipulating menus, including those that are pulled down, those that have to be held down and those that stay down when selected (Macleod and Tilson, 1990). The use of colour for menus has also been examined, McDonald et al. (1988) examining the speed and accuracy of users of different layouts. Circular menus have been tested (Callahan, et al., 1988) and this has led to other research into menus that scroll, shift and rotate (Mills and Prime, 1990). Although after some of these experiments, users are asked for their preferences (e.g. Macleod and Tilson, 1990; Mills and Prime, 1990), these judgements are ascertained by structured questioning of the users when they have only performed the trials into accuracy and speed. Despite all this work, there has been very little research into the way menus are used as part of the general activity performed on a computer system.4

One exception is the study by Mayes et al. (1988), one part of which was an experiment where users were asked to carry out the whole process of creating a document. This included turning the machine on, formatting and printing it. At various points in the experiment the users were asked to anticipate what would appear on the screen including the contents of menus. Despite having a range of experience with the actual application, all the users performed worse than expected. Even 'frequent users' (ones who used the application more than twice a week and most using it once a

4 An alternative approach has been to explore the menu metaphor more generally, and investigate what else computer menus could be designed to support (Norman and Chim, 1989).
day) had difficulty recalling about 50% of the 'gross details', such as the names of menus and the existence of menus and menu items. A follow-up study by Mayes et al. showed users had few problems performing a task that involved using the menus and menu items that they had had difficulty recalling. There appears to be a paradox, users are skilled at using an interface of which they can recall few details.

Mayes et al. suggest an explanation for this that maintains a cognitive account of the users' behaviour. The users still have knowledge of the system, but this is somehow inaccessible, or 'forgotten' (p. 284). Thus, there are different kinds of knowledge of the system that users can possess. They can have 'lexical' knowledge of the names of particular functions, for example, but they can also have 'semantic' knowledge that certain functions exist. In a move common in Cognitive Science, expertise is associated with a transformation of an explicit form of knowledge into a 'compiled' or 'internalised' version.

The possibility that users may have knowledge of different kinds has led others to propose other categories and classes which may also be relevant. So, Young et al. (1990) have proposed a preliminary analysis of eight different classes of knowledge which appear to be relevant to the user of an interactive computer system.

By proposing different kinds of knowledge, Mayes, Young and their colleagues are seeking to extend cognitive models of computer use. These include knowledge of the affordances of screen objects and general everyday semantics considered in terms of knowledge about the system or of common words. The models are still concerned about laying out what knowledge users have and how this 'leads them to exhibit their individual behaviours' (Young, et al., 1990, p. 117).

It may therefore be useful to explore the kinds of behaviours which are being sought to be modelled and the extent to which these can be considered in terms of the resources that individuals utilise when performing activities on a computer system. In line with the suggestions of Mayes et al. and Young et al. it would be useful to explore a more open activity, rather than the restricted tasks of conventional experiments in HCI. It would also appear to be useful to still provide some focus to the investigation, hence, an analysis of the naturalistic uses of menus.

The approach taken in the next two sections is to lay out the kinds of activities that models such as those of Young and his colleagues are seeking to account for. Rather than try and seek to assign various classes
of knowledge to behaviours, instances of menu use are examined. The nature of these activities and the resources that individuals appear to utilise in these cases are considered. To examine instances in detail, a transcription system has been developed that provides a way of revealing aspects of the users’ screen-based activities, and in order to provide a framework for the observations that are made, instances of menu use are considered in terms of a provisional taxonomy. The resources for this analysis are drawn from a ‘naturalistic experiment’ where individuals were asked to carry out a simple task on a computer system.

3.4 AN INITIAL EXERCISE FOR CONSIDERING HUMAN-COMPUTER INTERACTION

To gather some appropriate materials relating to human-computer interaction, a straightforward exercise was undertaken where people were recorded using an Apple Macintosh\(^5\) to prepare overhead projector slides.\(^6\) Eight users with variable experience with the Macintosh and the application, PowerPoint\(^7\), were given some headings and asked to prepare three slides about each of those headings. Having been told the name of the application that they should use, they were left alone to carry out the exercise. They were interrupted after thirty minutes if they had not finished by then.\(^8\) The exercise was recorded using two cameras and a microphone. One camera recorded the user, the screen, the keyboard and the mouse and the other focused only on the screen. The exercise was fairly open-ended and ‘naturalistic’. Apart from a general specification of

---

5 A trademark of Apple Computer Inc.

6 Using the Macintosh provides materials which are potentially useful when considering the proposals by Mayes et al (1988) and Young et al. (1990). The design of the Macintosh interface provides for easy movement between different menus. It is also the system that Mayes et al. used in their experiment and in terms of which Young et al. consider their model. Moreover the choice of system and application also has the possibility of raising a range of broader issues within HCI, concerning the direct manipulation of text and graphics, the use of screen-displayed tools and the use of a mouse and a cursor.

7 A trademark of the Microsoft Corporation. Details of PowerPoint can be found in the user manual (Microsoft Corporation, 1987). The version used was 2.01.

8 Given that the application is to produce slides, none of the users could be called ‘frequent’ users of PowerPoint, in the terms of Mayes et al., that is, using the package more than twice a week. However, all the users had some experience of using a Macintosh and some had used the PowerPoint application before.
the activity, users were not given a set of tasks to perform, nor did users have to perform any extraneous activities, such as producing a verbal protocol or answering questions about what they were doing. However, the audio recordings did capture anything they did say to themselves. Figure 3.1 gives a sense of the materials recorded from each of the cameras.

![Figure 3.1](image)

**Figure 3.1:** A frame from the 'mixed' recording. In the bottom left hand corner is the image taken from the camera focusing 'over the shoulder' of the user and onto the screen.

Inset in the general shot of the user and the system is the image recorded from the other camera. This was also recorded onto a separate tape so that details of the screen-based activity could be analysed with respect to the objects appearing on the screen. The general shot provided resources for analysing the user’s orientation to the system, for example, whether they were looking at the screen.

Figure 3.2 shows the PowerPoint application once it has been opened.

![Figure 3.2](image)

**Figure 3.2:** The PowerPoint Application
To the left of the screen is a set of tools for graphical operations such as creating lines, circles, ovals, rectangles and labels (the tool labelled 'A'). After selecting one of these tools, the user has to specify the shape and position of the appropriate object in the main window using the mouse and mouse button. This main window, labelled 'Untitled' at first, is where the image of each slide is built up, so all text that should appear on the slide should be typed into this window. Beneath the tools is a 'Slide Changer' for changing between slides. To the right of the screen are two buttons for viewing collections of slides (or presentations) and a scroll bar for moving up and down a particular slide. Along the top of the screen is a menu title bar. When the cursor is over a title, and the mouse button is pressed down, a menu appears (or is 'pulled-down'). The contents of these are shown in Figure 3.3. As well as offering the usual facilities of a Macintosh application (e.g. filing, printing, 'cut and paste'), they allow the user, amongst other things, to change the style of text, create set backgrounds to the slides (master slides), annotate slides, change the style of lines and boxes and colour the slides in complex ways.

Figure 3.3: The PowerPoint menus

PowerPoint's menus, in common with other Macintosh applications are arranged in a particular order, general operations to the left ('File', 'Edit' and 'View'), text formatting in the centre ('Style' and 'Text') and drawing

---

9 When more than one slide is created a bar appears allowing the user to move easily between the slides, the slide number being displayed to the bottom-left of the window.
menus to the right ('Draw' and 'Color'). The 'Window' menu allows the user to select between open presentations and windows. The leftmost menu, the 'Apple' menu, contains applications and accessories that can be used at any time, even when another (i.e. PowerPoint) is being used. PowerPoint's 'Help' option is provided in this menu.

Each menu has items arranged in groups of related operations divided by lines. Each menu item can be preceded by a checkmark (indicating that the menu item has already been applied to the current selection) or a diamond (indicating that the item is the default). It can be followed by a right-pointing arrow (indicating that a further sub-menu is available for more precise choices) and a symbolic representation for the keystroke alternative for the menu item. Menu items can also be grey indicating that the item cannot be selected at that time.

Faced with the complexity of the conduct revealed by looking at the recordings of PowerPoint use, it is necessary to develop a transcription system. It is impossible to provide a literal description of all human conduct in the recordings, so the transcription system has to be selective. The principal concern is with developing a system which provides a shorthand for locating the emergent and contextual properties of screen-based activities. This system is used alongside the actual data, that is the video and audio recordings, and has evolved alongside the analysis of certain actions, in particular the use of menus. It draws in part from the orthography created by Jefferson (1972) for the transcription of talk in interaction and it is described in more detail in Luff and Heath (1990). Fragment 1 gives a flavour of the transcription system.

The rightmost menu, shown in Figure 3.2, whose title resembles an image of a tape is for a general Macintosh feature called 'MacroMaker'. This allows for recording and repeating keystroke and menu selection patterns. It is not part of PowerPoint and was not used (nor opened by) any of the participants in the exercise.

So compare with some notations used within HCI for describing tasks on interfaces for the purposes of design (e.g. Siochi and Hartson, 1988).

It is possible to skip over menu titles and menu items by moving over them very fast. For example, in lines 5-6 of fragment 1 the user skips over the Style menu. There are also occasions when it unclear from the recordings where the cursor is. In keeping with the flavour of Jefferson's transcription system parentheses are used for doubt in transcription.
1 BC 53

The user clicks on the main window, labelled 'Untitled' (line 1), and following a gap of over 3 seconds (line 2) he successively opens the File, Edit, View, Text, Draw and Color menus by moving the mouse over the various titles on the title bar (lines 3-9). Finally, he moves his cursor quickly down the Text menu (line 10) and releases the mouse over the 'Untitled' window (line 13). The titles of menus, for example 'File' or 'Edit', are used to show when they are selected by the user and highlighted on the screen. They are written down the page, whereas when the user highlights particular menu items they are written to the right of the menu title. Colons are used to capture the timing of the activity, each colon representing one tenth of a second. Pauses between activities are given in seconds in parenthesis. Up and down arrows represent when a user presses or depresses the mouse button; a bracketed heading such as '[Untitled]' is for the region of the screen over which the mouse button was clicked. Users' 'self talk' is shown in double quotes, such as "Errn:::", and follows the orthography of Jefferson (1972).

So in fragment (1) the user selects the File menu for 0.9 seconds (line 3), then moves the cursor to the right and holds the Edit menu down for 0.8 seconds (line 4). After holding both View and Text menus down (lines 5-6), he quickly moves further to the right with Draw and Color being very briefly displayed (lines 7-8). He then moves to the left, still keeping the mouse button pressed and holds down the Draw menu (line 9). After moving left again he drags the mouse down the Text menu (line 10) and says "Errn:::" (line 11). He does not select anything, releasing the mouse over the main window half a second after moving off the bottom of the menu (lines 12-13).
The instance of screen-based activity in fragment 1 reveals some of the complexities of naturally occurring menu use. It also points to certain problems with some of the conventional terms regarding menus utilised within the HCI literature. Fragment 1, for example, could be viewed as a 'scan' of most of the available menus at the top of the screen, it could be a 'browse' through the options, or a 'search' for a specific item. Similarly, the simple holding down of a particular menu could be termed a 'scan' of, or a 'browse' at, the various menu items available, or even, as in Mayes, et al., (1988), a 'hesitation'. In order to reduce the potential for confusion between these terms, descriptors such as 'scans', 'searches', 'sweeps' and 'browses' are avoided. Instead, a typology is introduced relating to the uses of menus. Commonly occurring features of these are distinguished in terms of 'packages' of activities, described with respect to their sequential and temporal properties: i.e. the length of time spent at particular menus, the number of menus visited and the order of visits. The purpose of outlining these types of menu use is not to suggest that they specify organisations to the behaviour, let alone ones to which the participants orient. Rather, they are used to draw out some of the resources utilised by participants to accomplish particular activities through the use of the menus.

The typology consists of eight types of gradually increasing complexity: the selection; the hold; the pick; the straddle; the shuffle; the chug; the ramble; and the meander. These in turn are described in terms of four simple, single menu uses: opting; brushing; peeping and holding. Opting for a menu item is where a user releases the mouse button when the cursor is over an item that is highlighted (i.e. that has not been 'greyed out'). Users 'brush' a menu or menu item when they move the cursor over the title or the item very briefly. In the case of a menu title, this is just long enough for the menu items to be displayed, and in the case of a menu item this is just long enough for the menu items to be highlighted (typically this is less than a third of a second). Users 'peep' at a menu or menu item long enough for all the items or sub-items to be seen (typically, this is less than a second). A 'peep' is marked by a user pausing when moving the cursor along the title bar or down a menu. A 'hold' is when a user stops moving over a menu header or item for some considerable time (typically over a second).
These simple menu uses are primarily for distinguishing between different types of movement over a menu. As on a Macintosh application, and on similar interfaces, it is straightforward to move between menus, a single menu use can easily be transformed into a visit to many menus. Such uses of menus tend to be ignored by experiments that only measure the time to opt for an item and whether that was correct. It appears that individuals may utilise the resources available to them in quite complex ways, even when only peeping at a single menu. Similarly, the potential complexity of the uses of multiple menus may be ignored by glosses such terms as ‘searches’ and ‘browses’. These compound into one activity what may be a heterogeneous range of actions. As a starting point a straightforward kind of menu use is considered.

3.5.1 The Selection

The menu selection is the type of menu use that is usually portrayed in the conventional literature on menus. It could be considered as the most ‘direct’, or straightforward, way of using a menu, consisting of three components, pressing the mouse down over the title bar, moving down the menu to a menu item and finally opting for that item. Due to the ways Macintosh menus are manipulated there is always a brush of the menu title as the mouse button is pressed and a brush of the menu item as the mouse button is released. In (2), the user selects a block of text by dragging the cursor over it (line 1) and then selects the item on the Text menu (line 3). This justifies a block of text to a left margin.

(2) BN 79
1 ↓ > [Text] > ↑
2 (1.6)
3 ↓ Text: Left:: ↑
4 (1.3)

The selection of the text prefigures the forthcoming activity on the menu. Perhaps surprisingly, there are few instances of such direct selections in the data. More often, users peep at the menu title before selecting an item:

(3) BN 95b
(1.3)
↓ Text::::: Left::: ↑

or peep at or brush a menu item when moving down the menu;
or hold or peep at a menu item before it is opted for;

In fact, when users opt for an item they usually pause in some way on the menu from which the item is chosen, as in the fragments above. Of course, the physical operation of the system features in this activity. It takes some time for the system to display the menu, the user has to change the direction of the mouse to move down a menu, and there are limits to how far the mouse can be moved in any one direction. The users' 'pauses', however, tend to be noticeable 'breaks' in the ongoing activity. It could be conjectured that these reveal users' lack of experience or knowledge of the system or application. Indeed, as such, these perturbations could provide a useful resource for those endeavouring to model the cognitive capabilities of users of a computer system, as in the GOMS analyses presented in the last chapter, or as evidence of a 'search' taking place, or for a lack of knowledge concerning the contents of a particular menu. However, there appears to be no simple match between the experience which users reported and these breaks in activities. All the users, no matter how often they had used the Macintosh or the particular application had some perturbations on a menu before opting for a particular item. Such perturbations occurred throughout the duration of the exercise, even in cases where users had been examining the same menu moments prior to opting for a particular item.

This may support the observations made by Mayes et al. (1988) that even frequent users of an application have difficulties recalling the details of the interface. From their experiments, Mayes et al. suggested that users may not have to remember mental replicas of menus. Instead, their skills may be in 'remembering only enough to make the right choices accurately' (p.287). So, rather than perturbations in the course of activity being an indication of the lack of expertise of the users, they may point to the opposite, at the users' proficiency with the Macintosh: experienced users may know in which menu to find the 'target' of their search and merely pause to home in on their target.
Alternatively, it may not be the case that users are 'looking for' or 'searching for' a particular item. There may be some equivocality, or ambivalence, to their activities. In each of the above cases the user has, just prior to the instance, opened the same menu. Indeed, in these cases the users appear to be engaged in an activity which could be characterised as 'trying out' various options. In (3), the user has just 'justified' the text, in (5) the user has just held open the Text menu for over 3 seconds and then goes on to select 'Center', and in (4) the user has been engaged in a multitude of style changes.

(4) LD 116 - Transcript 2
1 (5.8)
2 ↓ Style: Plain Bold Italic Underline Outline: ↑
3 (2.4)
4 ↓ Style Plain: Italic Underline Shadowed: ↑
5 (2.6)
6 ↓ [Untitled] ↑
7 (1.4)
8 ↓ Style: Plain Bold Italic: ↑
9 (2.3)
10 ↓ Style: ↑
11 (1.6)
12 ↓ [Untitled] ↑
13 (0.2)
14 ↓ 'Witty' ↑ (. ) ↓ 'Witty' ↑
15 (1.6)
→ 16 ↓ Style: Plain Bold Italic Underline: Italic: ↑
17 (1.1)

Thus, the user in (4) is not unfamiliar with the Style menu, having visited it four previous times in her recent history (lines 2, 4, 8 and 10). It would then seem strange to characterise any delays in terms of searching for an item, or even in looking for a lexical match for a semantic specification, as she has just selected the same item from the menu. The selection is part of a course of activities which appear to be related to the appearance of text on the slide. Prior to the selection of Italic (in line 16), she has selected Outline (line 2), Shadowed (line 4) and Italic before (line 8). Instead, a characterisation of the single menu activity has to be seen in relation to her prior activities or selecting various options on the same menu, of clicking on the main window (line 6 and 12) and on the word 'Witty' (line 14). Hence, she not only is selecting a single Style for the text, but this is part of a course of activities where she is looking and selecting the various options available on the same menu: each selection being viewed in terms of its consequences to the objects on the slide.
The Style menu, as with many on a Apple Macintosh, by its very design, offer a range of related items in one place (e.g. Figure 3.4).  

<table>
<thead>
<tr>
<th>Style</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain</td>
<td>☑P</td>
</tr>
<tr>
<td>Bold</td>
<td>☑B</td>
</tr>
<tr>
<td><em>Italic</em></td>
<td>☑I</td>
</tr>
<tr>
<td><em>Underline</em></td>
<td>☑U</td>
</tr>
<tr>
<td>Outline</td>
<td></td>
</tr>
<tr>
<td>Shadow</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 3.4: The Style Menu in PowerPoint](image)

The possibilities the menus offer are read in the light of a previous activity on the screen. The Style menu offers a range of related actions that can be undertaken. These can be read as offering a way of changing the style of the text, but they can also be read as a more flexible set of resources, for developing a course of activities where styles are tried out, options are attempted and the results on the screen looked at.

Users can read the menu for their purposes at hand, the choices offering potential developments for the particular activities in which they are engaged. In (4), the selection of the Style menu is part of a course of activities, it emerges from a series of other activities on the menus and with the text.

By focusing on a single menu selection it is possible to attribute possible kinds of mental processing to pauses in the activity. So, in the GOMS models the pauses could be counted as cognitive operators for mental preparation. A similar account might be provided by those wishing to explore more complex analyses of the knowledge required to interact with a system. Even a simple instance of activity reveals that such an analysis would have to provide an account of the knowledge of the application (what is on this menu), the general features of the system.

---

13 The apparent consistency between different applications provided on the Macintosh can be attributed to following the Apple Human Interface Guidelines (Apple Human Interface Group, 1987). Among other things these outline ways in which menus should be designed and presented.
(what typically is on this type of menu) and the ongoing task (changing the style of text), and also some account of the prior activities on the system. However, the nature of these tasks and activities have to be considered in particular ways, for these do not remain fixed and can appear ambivalent. The user in (4), for example, could be considered to be italicising a piece of text, changing the style of the text, trying out some styles of text, discovering the operation of the system or selecting a suitable appearance for the slide. It would be problematic to consider the single menu use in isolation or too have too fixed a definition of the task that is underway.

There could be many accounts for the perturbations which occur in menu selections: in terms of the physical properties of the device, as delays due to the operation of the machine; the mental processing of the individual, as time for mental preparation, or just as junctures in the activity. However, it may be misleading to consider a pause as indicating an unseen activity guiding a forthcoming action. Instead, even in the course of an action on a single menu, different menu activities could develop, one of which is opting for a menu item.

In the following sections we will see how, from similar small beginnings, a range of activities on menus can emerge.

3.5.2 The Hold

Holding a menu varies from a menu selection in two ways: the menu is held down; and an item is not opted for. The menu is held open 'for view'. The mouse may be released after a hold over the title bar (6), or over the 'greyed out' (unselectable) items in the block. In these latter two instances, both on the Text menu, the mouse is released over the unselectable items which only apply to particular types of interface objects (i.e. for indenting text, showing text rulers and altering line spacing).

(6) LD 103
(2.5)
↓Text:..........................↑
(1.1)

(7) UN 163
(2.9)
↓Text:.......: Left Center Right (0.5)↑
(0.4)

(8) CB 29 - Transcript 1
(1.6)
↓Text:............: Left Center (0.7)↑
At first, a hold of the menu may appear to be an aborted menu selection where a user is looking for an item that does not, in fact, appear in that particular menu. Thus, dropping down the menu and releasing over a 'greyed-out' item, as in (7) and (8), could be seen as a way of not selecting an item, different to releasing the mouse over the title.

Alternatively, holding down menus could also be seen as occasions when users are drawn to items on menus, but also draw away from opting for one. As with the simple selection, the user may be actively engaged in using the system: for example, a user opening, holding and then closing the menu without moving the mouse may not have 'failed' to find an item. A user could be involved in other, or a range of other activities, whilst holding down a menu.

In (8), prior to the hold on the Text menu, the user selects an object on the screen (containing the text 'How I see Myself' - line 1).

(8) CB 29 - Transcript 2
1 ↓ ['How I see Myself'] ↑
2 (1.6)
3 C: “ thats a word processor object if ever there was one”
4 ↓ Text: Left Center (0.7) ↑
5 (10.2)

On arriving at the Text menu, he holds it down for 1.7 seconds and then moves the cursor quickly down the menu, releasing the mouse button after pausing for 0.7 seconds (over the ‘greyed out’ items Indent >> and Indent <<, line 4). This activity commences whilst he is saying ‘object’ in his utterance ‘thats a word processor object if ever there was one?’ (line 3). Earlier to this fragment, the user has been utilising a range of the application’s resources, not only has he frequently opened menus, but he has also used the on-line help system. He keeps the window of the help system open (which is the one related to ‘indenting and tabbing text’, see Figure 3.5) whilst returning to the menus several times, the fourth occasion of which is (8).14

14 The utterance ‘thats a word processor object if ever there was one’ appears to echo the instruction ‘Select a word processor object, then choose Select Text Ruler on the text menu’ in the help (Figure 3.5).
Indenting and Tabbing

You indent and tab text using the Text Ruler. Select a word processor object, then choose Show Text Ruler on the text menu.

Each indent symbol controls the indenting for a different level. Drag a symbol along the ruler to set its indent distance. To indent text, choose Indent > (right) or Indent < (left) from the Text menu and begin typing. To indent (or "un-indent") existing text, select it, then choose Indent > or Indent <. The shape of the symbol indicates the left margin for its level like this:

```
left text
  ...text
  ...text
left text
```

For more about this topic, refer to page(s) 169-177 in your PowerPoint User Manual.

---

Figure 3.5: The 'help' for indenting and tabbing in PowerPoint (v. 2.01)

Hence, it appears that the user's activities are related to the possible resources offered by the system, particularly with respect to manipulating 'word processor objects'.

As previously mentioned, the menus on an Apple Macintosh are more than just lists of available commands or words. They provide a range of resources for the user. Each menu is divided into groups and each item can be marked in various ways, including being 'checked' and being 'greyed out'. The checkmarks can be used to provide information about the properties of items previously selected and the 'greying' of items can provide information on the availability of the various options that could be applied to selected items. Thus, by 'greying out' particular items, menus can reveal not only which actions are possible to perform at the moment, but also a selection of other actions which are possible at other times. It appears that the user in (8) holds down the Text menu as part of an active exploration of the possibilities afforded by the system. When simple text is selected, the Text menu in PowerPoint appears as follows:

---

Figure 3.6: The Text Menu in PowerPoint when a simple block of text is selected
Options like justifying text, finding or replacing text and spell-checking are applicable to text. Options such as indenting lines of text, altering line spacing and displaying text rulers, on the other hand are only applicable to 'word processor objects'. So, by displaying both the items that are applicable and inapplicable to the object selected at the time, the Text menu can be used to reveal the type of an object. It can also suggest operations which can be accomplished on other types of objects. Thus, the menus can be utilised as a resource for making sense of the workings of the system and can draw users into other courses of activities. So, in (8) the user's hold of the menu (line 4) appears to be in the light of the changing appearance of the menu: what is 'greyed out', what it makes visible and what may or may not be a 'word processor object'. Following the hold on the menu, the user goes on to select another object on the screen (the tool for creating 'labels' - line 6) and then returns to hold down the Text menu again.

(8) CB 29 - Transcript 3

5 (10.2)
6 ↓ [A] ↑
7 (1.4)
8 ↓ Text:..............................

The label tool does not create 'word processor objects'. Again, opening the Text menu reveals that the items applicable to these objects are 'greyed out' and not available.

In (8) the user could be considered to be trying to indent text, to create a 'word processor object' or to discover properties of a system. The user appears to make use of the dynamic and changing nature of menus as a potential resource for subsequent activities. However, identifying one particular 'goal' of the user, even a vague one, appears to be problematic. His activities appear to be related to indenting text as well discovering whether a particular object is one to which you can apply word processor-like actions. Indeed, there may be some ambivalence in the nature of the activity from moment-to-moment. It can be transformed in its course.

Hence, when a user holds down a menu, it may not be just in order to look for an item to select (a 'search' that fails), it may be with respect to a range of practical purposes and embedded within an emerging course of activity. The menus provide a resource for actively making sense of the system. In displaying options which cannot be selected, PowerPoint is typical of those applications. Indeed, this is one of many features that is meant to contribute to the Apple Macintosh's 'look and feel'. These options
can draw users into a course of activity, which itself develops. Skilled users of a Macintosh can make use of the changing nature of the menus to make sense of an unfamiliar application or unfamiliar aspects of an application. It would appear to be inappropriate to attribute such an activity to a particular goal or plan, instead, the menus are read in the course of an emerging course of activity. An alternative characterisation would more be in terms of improvised activity.

3.5.3 The Pick

Picking a menu item can be considered as a composite of the previous two types. The menu is held down whilst the cursor is over the menu title, or over a menu item, and then the user goes on to opt for an item (e.g. 9).

(9) TX 65
(2.3)
↓Edit:........................ Cut Copy:..............↑
(0.9)

Often, picking an item involves traversing up and down the menu, holding on items along the way.

(10) BN 97
(2.3)
↓Text:......... Left Center:...........
Text:................ Right: Center: Right: Center:......
   Right:......................↑
(2.9)

In (10) the user, after going down the menu once, returns to the title bar. Without releasing the mouse, he then holds the menu down for nearly four seconds, goes down to Right, up to Center, back down to Right, back up to Center and, finally, opts for Right, after holding it down for nearly another four seconds. Fragments (9) and (10) are relatively straightforward cases of picking items. A pick may also involve holding down one of the menu items which has a sub-menu attached to it, and then moving up and down that sub-menu (e.g. the Size sub-menu in fragment 11).

(11) XT 69
1  (2.3)
2  ↓ Style:.... Plain Bold Italic Underline Outline Shadow Font:
3  Size:..............: 10: 9: 10 12 14 18 24:..............: ↑
4  (3.3)

In (11), the pick appears to be composed of two parts: the movement down the Style menu (line 2); and the pick of 24 from the Size sub-menu,
with holds on the Size and '24' items (line 3). Here, the shape of the activity on the sub-menu resembles the shape of activity on the main menu revealed in other instances (e.g. in fragment 9). There are other similar instances of complex, single menu activity in the data. In (12), the user has just finished the contents of a slide. He then goes down and up the View menu three times before releasing the mouse button over the title.

(12) BN 82 - Transcript 2
8 (3.2)
9 ↓ View: Full Size: 66% Size 50% Size 33% Size: (1.7) Notes (#1)
10 Notes Master:: Handout Page::: Notes (#1) Slides (#1)
11 Notes (#1): 50% Size: 66% Size:
12 Full Size::: 66% Size::: 50% Size:::
13 View:: ↑

For the user's particular circumstances prior to arriving at the menu there are a wide range of appropriate next activities. He has completed his first slide, created a new one, and added some text which is centred. He has then selected this text and justified it to the left (line 3).

(12) BN 82 - Transcript 1
1 ↓ > [Text] > ↑
2 (1.6)
3 ↓ Text: Left:: ↑
4 (1.3)
5 Text: Left:: Center:: ↑

On opting for 'Left' the text moves almost to the left hand margin of the slide, further to the left than on his first slide. The user immediately returned to the Text menu and centred it again (by the selection in line 5). This is the position of text when he goes to the View menu in (12, line 8).

Figure 3.7: The user holding down the View Menu (line 12, fragment 12)
The user in (12) could be seen to be looking for inspiration from the View menu, the items over which he holds suggesting candidate next actions and candidate next courses of activity, his movements up and down the menu corresponding to him being drawn to various possibilities. For example, the menu provides an option for viewing several slides together (i.e. the Slide Sorter), for manipulating several slides at once (i.e. the Title and the Slide Sorters) and for altering the size of the image of the slide (i.e. Full Size, 66% Size, 50% Size and 33% Size). The user in (12) first moves towards the Title and Slide and Sorter options (lines 9-10) and then back up to the slide sizes (line 10-12) and down them again (line 12). These options not only suggest next actions to take, but potential courses of actions, or trajectories. On creating a second slide, there are various courses of activities that would appear to be appropriate and relevant. These include aligning the text on the second slide with the first (or vice versa), going on to create a third slide, or editing or formatting the text on the second slide. The View menu can provide the resources either to assist with these activities or, more accurately, to project a course of actions to accomplish them. The Slide and Text Sorters, for example, offer capabilities for adding and comparing slides. The various Size options offer the potential for reducing the size of the slide displayed on the screen, and therefore, making the objects on it easier to manipulate.

However, in (12) the user appears to be faced with a particular problem: one in which text is placed too far to the left. His activities on the menu, nevertheless, suggest an ambivalence to the nature of the activity in which he is involved. Even the choice of the View menu appears to be problematic. Of course, it could be accounted for in terms of goals, knowledge of the menus, the capabilities of the system and the interaction between these and the menus. Vaguely formulated goals concerning a general search for a solution to a problem and accounts of lacking knowledge of the menus could be suggested. However, such accounts would lose the continually shifting nature of the activity, where goals, courses of actions and purposes appear to be continually transformed in the light of the menus and possible readings of them. In (12) there appears to be some ambivalence to whether the activity on the menus can be related to the change to a new slide or the position of the text. The View menu could provide ways of dealing with both of these, and could, indeed, assist with the in situ discovery of the nature of the problem at hand. It is unclear whether assigning a vague a priori goal to the user
provides anything to the analysis. Rather, what appears to be relevant to
the user is the in situ uses of the menus; how in their reading they can be
utilised to develop a course of activity, in other words, how they can be a
resource for innovation.

3.5.4 The Straddle

In the data there are several instances of users opening one menu, then
quickly going onto an adjacent menu and opting for an item on that menu.
For example, in fragment (13) the user flicks over the File menu before
picking ‘New Slide’ from the Edit menu.

(13) BC 56

(0.6)
File

down
Edit: Undo Omit Master: Set as Title: New Slide: ↑

In (14), the user holds down the File menu before opting for Help from
the ‘Apple’ menu.

(14) CB 24

(2.3)
File:

-----: Alarm Clock: Help: ↑

Straddles can themselves be viewed as a package of two single menu
activities. For example, (14) could be considered as a hold on the File
menu followed by a menu selection of Help on the ‘Apple’ menu. However,
it is not always clear where the division between the two menu activities
occurs. A user may also quickly go up and down a first menu before going
onto the second, as in fragments (15) and (16).

(15) XT 71

1 (2.5)
→ 2 ↓Text: Left Center Right Right Center Left
3 Style Plain Italic Outline Shadow: Font Size: 9:: 10 12 14::: 18::
4 24::: 36::: ↑

(16) BN 77

1 (2.4)
→ 2 ↓Style: Bold::: Plain:::
3 Text: Center::: Right:: Center Left
4 Text Left::: ↑

In these instances a move towards making a menu selection appears to
be curtailed. Immediately, the user moves to an adjacent menu and moves
down that one. The actions on the two menus seem to be closely tied
together, the first providing an environment for making sense of the second.

The design of the Macintosh makes switching between menus very simple, only requiring a user to move the cursor up to the title bar of an adjacent menu without releasing the mouse button. So a transformation of an activity from one menu to another could be seen to start in the visit to the first menu (marked by an arrow). In fragments (15) and (16) the user moves down the first menu immediately after he selects the title (0.2 and 0.1 seconds respectively), the selection and the descent being part of the same movement. Within a second, the user changes direction and moves onto the next menu. Immediately as he arrives on the second menu, he descends that one. After some perturbation, a descent of the Size menu (lines 3-4, Fragment 15) and a return to the title (line 4, Fragment 16), the user opts for an item in the second menu. These instances appear to suggest that in one movement users go to select an item, find themselves on the wrong menu, try the next menu and then go on, with slightly more circumspection, to opt for the item from there.

It is interesting to note that, although the instances of straddles in the data are spread across most of the users, they are not spread across all the menus. Particular pairs of menus appear to be 'straddled': the Text and the Style menus, the File and the Apple menus, and the File and the Edit menus. If we take one particular pair, the Text and Style menus, it can be seen how these can be read together.

![Figure 3.8: The Text and StyleMenus](image)

In PowerPoint, the Style menu offers a range of options, many of which can be applied to text objects (e.g. Plain, Bold, Italic, Underline, Font and Size). Similarly, many of the items in the Text menu could be considered as 'stylistic', particularly those which relate to line spacing, indenting and
justification. Indeed, in other Macintosh applications similar options are provided, in different configurations, on menus with titles such as 'Format,' 'Font,' 'Type' etc. So, given a particular activity at hand, users read the menu titles for potentially relevant descriptors for classes of appropriate next activities. Both Text and Style could describe a class of operations for transforming text; the File and Apple menus are, within the Macintosh domain, relevant locations to find on-line help; and the Edit and File menus can both act as descriptors of a class of actions of which defining a new slide could be part.

When arriving at a menu where the title appears to be relevant to a developing course of activities, for example the Style menu in (16), the appropriateness of the contents may not be immediately apparent. In the light of this reading of title and contents, another menu title may appear to offer relevant resources. Though the course of activity may appear tied to a prior activity, a selection of text, the developing activities may still not be straightforward - even as straightforward as the characterisation of a search failing initially and then being successful. In (16), on arriving at the second menu, the user drops down the menu beyond the first block of options, peeping at Center and Right on the way. He then goes up the menu pausing over Left before it is selected.

It is a problem for designers to arrive at an appropriate classification of actions for the various menu options. They are faced with either presenting a large number of options on one menu which may be difficult for a user to cope with (and possibly break guidelines for interface design), or distribute options between potentially confusing categories. For a user engaged in a course of activities, the menu titles provide potential classifiers of possible next actions on that course. On opening a menu, the user is faced by a range of names for possible next actions. These contents can be read as items to opt for, and also as a collection of related objects. The contents of a menu may not only be utilised to make sense of particular items in the set, but also read together to make sense of the classification as a whole. Users then have to make use of a set of practices for making sense of menus, these practices relying on common-sense and skilled readings of items and collections of items. A title like Style can suggest a set of options relating to the style of objects on the screen. Only when presented with options like 'Plain', 'Bold' and 'Italic' can a more precise nature of the menu be revealed: it relates to text, but only its appearance (and not its shape). The Text menu then can be seen to be
relevant as an alternative. However, that an alternative is required may only emerge in the light of a first visit. The menus together not only provide a solution, but also help define the problem.

Hence the process for understanding the menus is an emergent one, undertaken in the light of particular problems at hand, and developing through the changing appearance of the menus as they are opened. The division of options between menus and the relationships to their titles in PowerPoint may not be straightforward, but then again there are a large number of options and ways in which they can be divided. The purposes for which menus can be visited is indefinite. The skills of users, whether familiar or unfamiliar with the application, is in making sense of these resources for their own practical purposes.

In the cases so far considered, characterising the use of a menu as 'selecting' or 'searching' for a command glosses the ways in which even that apparently simple activity is achieved. The contents of a menu are read in relation to particular activities at hand and serve as a resource for shaping subsequent actions. Moreover, the possibility that their contents can be transformed in subtle ways allows menus to be utilised so that users can make sense of both prior actions and the general operation of the system. The use, or reading, of a menu is not only embedded within an ongoing course of activity, but also serves to shape that developing course.

3.5.5 The Chug

There are many cases in the corpus where the user proceeds in a single direction across two or more adjacent menus and holds down more than one of these on the way. These are characterised as 'chugs'.

Fragment (17) is a chug through the File and Edit menus, (18) is a chug through four menus and (19) through seven menus.

(17) TX 66 - Transcript 1

(3.7)
↓ File:...............................: Edit:...............................: Undo Cut Copy Clear
Bring to Front Paste as Picture:............................... ↑
The chug might be considered to be a slow, systematic search for a solution to a problem that a user is facing at a particular moment. However, it is hard to think of a single problem to which all the menus from Edit, View, Style, Text and Draw (in fragment 18) or all the menus from Window to Edit (in fragment 19) could offer a potential solution. The instance (19) comes from very early on in the exercise: just prior to the fragment the user has justified the text to the left, the first action he has opted for from a menu. The click in the main (Untitled) window deselects the previously justified text and leaves nothing on the screen highlighted, an activity which does not prefigure a specific next action, rather the combined deselection/selection appears to foreshadow the possibility of leaving options for a next action open. His steady progression through the menu, particularly the long holds over Color, Draw, Text and View, does not seem to be oriented to a single ‘goal’. Instead, it is as if the user ‘ends up’ selecting the New Slide option. This selection is problematic if the entire menu activity is characterised as a ‘browse’, however (as in Young and his colleagues’ hypothesised examples).

Like searches, browses are broad classifications of menu activities. In (18), the user has just typed some text and she goes to the Edit menu (line 1), peeps at the View menu (line 2), then she holds over the Style, Text and Draw menus (line 3-5).
Figure 3.9: The Text, Style and Draw menus

As with straddles, the options presented when opening one menu can lead the user onto another. The Edit menu is a potentially relevant place to go after typing text, however, the characterisations of Cut, Copy, Paste and the like, may not necessarily be appropriate. The Style and Text menus might suggest the availability of alternative resources. In (18) the user goes on from holding these to the Draw menu. This visit appears to be more problematic to consider in the light of the foregoing activity with the text. Nevertheless, an item, Shadow, is selected from it which can be applied to text.

This selection is not unequivocal. The user holds on the menu title for over three seconds. It is as if the user is ‘drawn’ into this selection.

Chugs through the menus, as in (18) and (19), do not seem to be best characterised in terms of either browses or searches. The user may have just undertaken a particular activity where a menu option would provide an appropriate next action, however, the user can appear to be ambivalent towards the options being made available to them. The user may be looking at the broad range of items on the menus, but ends up selecting one. The items on the menus can be seen, at one and the same time, as relevant or irrelevant to the activity at hand.

It may be that a browse or a search is too crude a categorisation for particular menu activities. The nature of the uses of menus is transformed in its course, broad classifications for a range of activities being inadequate to capture the details, and the work involved, in practically utilising these resources.

3.5.6 The Shuffle

In a ‘shuffle’ a user visits two or three different, adjacent menus, but in the course of this activity, changes direction once or twice. So, for
example, in (20) the user returns to the Text menu (line 3) and in (21) the user changes direction and returns to the File and Edit menus (line 5)\textsuperscript{15}.

\textbf{(20) TX 59}
\begin{itemize}
  \item 1 \downarrow Text: \ldots \\
  \item 2 Draw: \ldots \\
  \item 3 Text: \ldots \\
  \item 4 Style: Font: : Size: 9:: (18):: \uparrow
\end{itemize}

\textbf{(21) BC 58}
\begin{itemize}
  \item 1 \downarrow [(Slide Sorter)] \uparrow \\
  \item 2 (2.7) \\
  \item 3 \downarrow File: \ldots \\
  \item 4 Edit: \ldots \\
  \item 5 File: \ldots \\
  \item 6 Edit: \ldots \\
  \item 7 View: \ldots Full Size: 33\% Size: \ldots Slides #1: \ldots Notes #1: \\
  \item 8 Slide Master: \ldots Slide Sorter: \uparrow
\end{itemize}

Because of the wide range of options available in the application, characterising an activity that involves the entire range of menus as 'a search' for a particular item, or a solution, would appear to be problematic. However, a user shuffling through two or three menus might be characterised in such a way. After the event, the activity appears to be more tightly focused.

Prior to (21) the user completed his first slide, created a new one, typed in a title and some contents into the second slide and then highlighted a block of text. He now moves back to the first slide by selecting it within the 'slide sorter' (line 1)\textsuperscript{16}. By highlighting some text, the user appears to foreshadow a forthcoming activity with respect to that text. It would be now be possible to change its style, 'cut' it or copy it, for example. Moving to another slide suggests the last of these possibilities. Indeed, the File and Edit menus (lines 3 to 6) could suggest ways of moving text from one slide to another\textsuperscript{17}. After moving to the View menu, he holds over a few items, and selects 'Slide Sorter' (line 8). The slide

\textsuperscript{15}In the following fragments, the menu visit after a change in direction is marked by an arrow.

\textsuperscript{16}In PowerPoint, the 'slide sorter' is a way of viewing and organising a set of slides. The slides appear as small images. Selecting one of these images with the cursor, as in line 1 of (22), will result in 'moving to' that slide (i.e. that slide being presented so that it can be edited. To select the 'slide sorter' itself, there is either an item at the top-right of the main window, or an item in the View menu.

\textsuperscript{17}Either through the standard 'cut and paste' (on the Edit menu) or through such items as the 'Paste from...' on the File menu.
sorter is another way of beginning an activity of moving text from one slide to another. The user goes on to copy his first slide twice and delete the one he just started. This is one way of copying the same text onto each of the slides. There are other ways, such as selecting all the objects on a slide and copying them onto new ones, which may be quicker.

In (21) the user does appear to face a problem and goes to the menus to provide a solution. However, neither the problem or the solution need be precisely specified. Once again, the menus can draw the user into opting for an item, providing a posteriori, a possible definition of the problem (and solution). By making a choice the user can get drawn into a particular course of actions.

In (23), the user appears to be trying to alter the style of some text on his slide.

(23) CB 2 Transcript 1 (simplified)

1 Text:
2 Style:
3 Text:
4 Style: Underline Outline Shadow Font:
5 Text:
6 Style:
7 Text:
8 Style: Underline Outline Shadow Font:
9 Text:
10 Style: Underline Outline Shadow Font:

An item of text is selected and he only visits two menus, each of which provide options for changing the nature of text. Hence, fragment (23) could be considered to be a ‘search’ where, though vague, an appropriate problem could be specified. The following more detailed transcript, includes the user’s ‘self-talk’, and begins to reveal the ways in which this activity may change whilst a user is ‘interacting with the menus’.

(23) CB 2 Transcript 2

1 C: “err: (0.2) alright; um:: (0.2)”
2 Text:
3 Style: “Centre”
4 Style: “Bol:(d) (0.6) som”
5 Text:
6 Style: Underline Outline Shadow Font:
7 Style: Underline Outline Shadow Font:
8 C: “Here we are”
9 Style: “font”
10 Style: “font”
11 C: “that's too big (1.1) what about 24 ”
12 (1.5)
At various points in (23) the activity could be characterised as reading off the menus, checking what is on them, searching, finding or correcting. So, in the first two visits to Text and Style the user echoes items appearing on the menus ‘Centre, Style, Bold’ (lines 3-5). On returning to the Text menu, he utters ‘somewhere in here’s going to be a font’, then ‘Here we are font’ when moving down to the Font item (lines 5-8). This item has a sub-menu. When this is opened he utters ‘font Helvetica’, and when he opens the Size sub-menu he says ‘size you’ve decided thirty six’. Whilst selecting ‘24’ he utters ‘I think that’s too big (1.1) what about 24’.

In many methods for human-computer interaction, such talk would be considered a useful resource for analysts, a ‘verbal protocol’ as discussed in approaches developed within Cognitive Science and Psychology (Ericsson and Simon, 1980). Needless to say, the status of such reports can be questioned if they are considered to be reports of mental processing. However, in this case, it suggests at least how a ‘self account’ of an activity can be transformed in its course, as menus are opened and their contents are inspected. After the event, it could be said that a goal of changing the style of the text had been satisfied by reducing the size of the text, but this would miss how the activity emerges. In this instance, in some sense, the user is interacting with an interface. The objects on the screen and the contents of the menus are made sense of with respect to one another. The former make relevant readings of the menus, the latter are read with regard to what is on the screen (e.g. the text and its font). The use of the menus is an emerging activity, shaping and shaped by the ongoing activities of the user.

Even if they are crude glosses, scanning and browsing can be useful as ways of considering activities on menus. Unfortunately, what is usually associated with these terms are cognitive accounts of the activities; matching targets to a priori goals, the mental processing involved in choosing between different methods and the knowledge of the screen objects, the application and the system - in other words, knowledge that. An alternative conceptualisation is in terms of knowledge how, how individuals manage to make use of the system for the in situ accomplishment of activities. A browse or a scan, would then be considered as an accomplishment, from the active utilisation, from moment-to-moment, of the resources available.
3.5.7 The Ramble

A 'ramble' is an extensive jaunt through the menus, encompassing a large number of different menus.\(^{18}\) Most of the ramble proceeds in one direction, but after three or four menu visits there is usually some change in direction. After this change, either the ramble proceeds along in the new direction (as in line 7 of fragment 24) or returns to the old one, the diversion being only to revisit a single menu (as in lines 4-7 of fragment 25). Often, towards the end of the ramble there is a review of the last few menus visited (as in lines 13-14 of fragment 25).

(24) LD 122
1 ↓ [Untitled] ↑
2 (7.3)
3 ↓ View:..............................
4 Edit:..............................
5 File:..............................
→ 6 Edit:..............................
7 View:
8 Style:..............................
9 Text:..............................
10 Draw:......................: Opaque Framed Filled Shadowed Sized To Text
11 Ignore Grid Show Guides Show Edges: Show Guides:↑

(25) LD 106
1 (3.7)
2 ↓ File:..............................
3 Edit:..............................
→ 4 View:..............................
5 Style:..............................
→ 6 View:..............................
7 Style:..............................
8 Text:..............................
9 Draw:..............................
10 Color:..............................
11 Window Untitled:................ Window:
→ 13 Color:..............................
14 Draw:↑
15 (8.5)

Just prior to fragment (25), the user has been clicking and dragging the mouse over objects and pieces of text in the main window. From the nature of these actions, she is apparently trying to move a portion of the text she has typed. The user starts at the left hand title (excluding the

---

\(^{18}\) For the purposes of this investigation, a ramble extends over at least six different menus.
general Apple menu) and proceeds across to the rightmost title (excluding the general MacroMaker menu). When the user reaches the Window menu, she changes direction (line 11), reviews the Color and Draw menu, releasing the mouse over the title of the Draw menu. This ramble, as with others in the corpus, is bounded by the span of the PowerPoint menus (from Edit to Window).

It appears that in (25), and in other instances, the user is actively engaged in some activity in the application and then goes to the menu in relation to that activity: the movement of the text. In (24), after changing in various ways the style of the text, the user goes to the View menu, and in (26), after (re-centring) some text, the user goes to the Edit menu, and then rambles and sub-menus, including some just visited (e.g. the Styles sub-menu, lines 5-6).19

(26) BN 86
1 (5.5)
2 ↓ Edit:......................................................
3 View:.............................................
4 Style:..................................................: Plain Bold Italic Underline Shadow:
5 Styles:: Helvetica 18:..............................: Styles:: Define Styles:::
6 Styles Shadow Italic Bold Plain
7 Style
8 Text:..........................................................
9 Draw:.............................................
10 Color:.............................................: Fill:..........................................
11 Window:..........................................
→ 12 Color:..........................................
13 Draw:↑

As with other instances of menu use, it is difficult to characterise the activity as a whole in terms of some pre-specified goal. Even though prior to the use of the menus the users were engaged in some activity, through the use of the menus they appear to be drawn away from this. On departure from the menu, either nothing is selected (as in 25 and 26), or an unrelated item is picked (such as 'Show Guides' in 24). In other cases looked at so far, the menus appear to be a resource for innovation and improvisation, suggesting courses of activities. In these rambles, the menus appear to offering little with respect to the ongoing activity. The

19 In (25) and (26), the user on reaching the Window menu changes direction, holding down the Color menu again, then peeps back at the Draw menu. There are other rambles where this occurs (cf. the return to Draw and Text menus in fragment 1). It is as if the user is attracted to the recent and more exotic sights of the ramble.
ease with which they can be opened and perused allow for menus to provide for novel courses of activity to emerge. It also can provide for distraction, the user being drawn into viewing options which it is increasingly difficult to make relevant to the prior activity on the system.

3.5.8 The Meander

The meander can be considered to be a composite of the shuffle, the chug and the ramble. In a meander, in the course of making several visits, a user repeatedly returns to one of them. In (27) the user visits six different menus, changing direction three times and opening the Text menu three times.

```
(27) XT 68
1 (1.6)
2 JText:::::::
3 Draw::::::::::::=
4 Color:::::::
5 Window:::::::
→ 6 Color
7 Text
8 View
→ 9 Style
10 Text
→ 11 Style::::::::: Plain Bold Italic Underline Outline Shadow Font
12 Size:.............................. ↑
13 (2.3)
```

The meander, therefore, appears to characterise the constantly changing nature of menu activity. Fragment 27 begins with a chug through the Text, Draw, Color and Window menus (lines 2-5). On reaching the end of the menus, the user brushes across the Color, Text and View menus (lines 6-8), changing direction once more, before shuffling between Text and Style (lines 9-11). The user then holds down the Style menu, goes quickly down to the Size item opening the sub-menu and after 3 seconds releases the mouse button without opting for an item. Thus, there are two distinct components to this fragment: the chug from left to right (lines 2-5); and the shuffle between Text and Style menus (lines 9-12). The transformation from one to another is marked by the brush across the menus (lines 6-8).

---

20 For the purposes of this investigation, a meander extends over at least five different menus, where at least one menu is opened at three times.
In fragment 28, the instance can be considered a composite of several parts: a hold of the Text menu (line 2); a chug through the menus (lines 7-18) and then a shuffle between Text and Style (lines 20-25). Again there are brushes which mark the transitions between the components (lines 4-6 and line 20).

(28) BC 52

1 (1.1)
2 ↓ Text: Left:
3 Style:
→ 4 View
5 Edit:
6 File:
→ 7 Edit:
8 View: Full Size:
9 Style:
→ 10 Text:
11 Draw:
12 Color: 
"ern: ↓"
13 B: Window:
14 "let's see"
→ 16 Color: 
"how does this thing work"
17 B: Draw: 
"no: ↓"
18 B: Text
19 Style: "tex:"
20 B: Text: "um"
→ 25 Style: Styles Define Styles
26 (0.2)
27 ↑ [Untitled]
28 (0.6)
29 B: "err: I can't even remember how to get text in here"
30 (0.2)

In previous instances it has been noted that menus appear to draw users into different activities. An activity in which a user seems to be engaged appears to become transformed in the course of using the menus. In these instances of meanders, the changes in activity are more dramatic, the user brushing over menus and between two different shapes of menu use. In both instances the users appear to be engaged with some problem
associated with the text they have just typed. In (27) the user has been attempting to alter the style of the text (including trying to centre it) and in (28) the user appears be trying to create a box in which to type text. In each instance, the user appears to be drawn into a chug through the menus, only when they arrive at the last PowerPoint menu do they change the course of action. In (27) this is marked by a brush back towards the Style and Text menus and in (28) it is marked by the user's 'let's see how does this thing work' (line 15). These appear to close off the present activity and reorient to a new one.

In meanders, as in the other instances, there are difficulties attributing the menu use to a single activity. The users' prior activities do suggest some resources for making sense of their uses of menus: for example a manipulation of a text box foreshadows a move to the Text menu; or the completion of a slide is prior to moving to the Edit menu; or the small size of the slide on the screen precedes the visit to the View menus. However, from these beginnings it is harder to provide an account of the menu activity as a whole. Instead, this appears to be transformed in mid-course, so from a visit to a neighbouring menu a slow progression along other menus emerges. This, in turn, can then be characterised as a general browse of the available functionality, but then from this, a shuffling between menus emerges, which itself could be seen as a search for something in particular.

Meanders involve broad and sustained activities with respect to the menus. At times visits can be seen to be relevant to a prior activity, at others the appropriateness of a visit to the menu is less clear. The user appears to be drawn into visits to other menus. The reading of the menus can suggest other options, other course of activities open to the user which may be of relevance. These readings can be more of general interest. Along the way, the user might be drawn back towards, once again to consider menus previously visited. However, these visits are now in the light of other alternatives which now have been seen. In (27) and (28) the users return to a place on the menu near to which they started, but in neither is anything selected.

In the course of using the menus the ongoing activity can be transformed. Menus appear to draw users into novel courses of activities, sometimes they can be a resource for innovation, at others, for distraction. Investigating the details of this rather mundane screen-based activity does suggest its emergent nature. The uses of menus, the ways in which
they can be read and applied, are contingent with respect to the prior activity, potential future courses of activities, and the practical purposes of the users.

3.6 DISCUSSION

The foregoing investigation of menu use has revealed how rich even a straightforward activity like menu use can be, when examined in detail. Even a use of a single menu is complex when the range of potential purposes that the users could be orienting to, the various ways menus can be read and the possible relevance of other objects and prior activities are considered. Even if a model of the activity is preserved in terms of goals and the selections of different methods to attain them, a great deal of flexibility in the model is required. So, for example, goals either need to able to be transformed at any time, or to be so vague that almost any action could satisfy them. Selection rules would also have to be specified that choose between an indefinite number of methods, or the criteria for selection open enough so that novel options could be considered. It is perhaps not surprising that models such as GOMS are considered to be too restrictive to account for human conduct on a computer system.

The provisional model of Young et al. (1990) does suggest a sophisticated way in which models of computer use could be extended. If the foregoing analysis of menu use is considered, it does appear that the resources utilised by the participants have many similarities to those suggested by Young and his colleagues. They distinguish between eight classes of knowledge including: knowledge of the screen; the ordinary meanings of common words; the specialised meanings of words associated with the computer system; knowledge of how to translate everyday meanings into these specialised meanings and knowledge of how to divide a task into sub-tasks. Such common-sense resources appear to be made use of when making sense of a particular menu item, sets of menu items or an object on the screen. Young et al. suggest that it is the different patterns of knowledge that users have across these categories, in interaction with 'external' knowledge arising from the situation, that can account for different kinds of behaviours.

The foregoing investigation suggests a few difficulties with such an approach. The first would be to catalogue such knowledge. It would be hard to envisage how to delimit a category such as the common-sense
meanings of words, or the 'knowledge of' tasks. Second, it would be hard
to transform such knowledge into the kind appropriate for a model, i.e.
propositional knowledge. The use of menus appears to make use of
common-sense competencies, embodied, tacit and skilled practices, the
nature of which would be lost when formulating it into a form to be
utilised in a model. Third, the characterisation of the interaction between
what appears as external and what is considered internal is problematic.
Presumably, reading a menu item would be a simple case of such an
interaction. However, in the instances above, making sense of even a
single menu item can be seen to be accomplished with respect to what else
is available and the ongoing course of activity. Models such as Young et
al. would require a highly volatile, unstable and indefinite catalogue of
resources. Users would have to be ambivalent about their goals, the ways
they are to achieve them and the relevant knowledge to be applied to a
particular case at hand. In other words, knowledge and their rules
of application would have to be defeasible.21

What models like those of Young and his colleagues and GOMS share is
an account of human conduct in terms of mental processing, particularly
with respect to an activity of mental preparation. In this, either methods
are selected or internal knowledge is matched with external information:
'knowledge of' the interface, or of the world is brought to bear, applied or
utilised.

It is this conception of knowledge that was questioned in the previous
chapter. If this is reconsidered, then a distinctive account of human-
computer interaction might be offered. Drawing from Ryle (1949), the
skills of using an interface could be considered as 'knowledge how' - how
an activity is accomplished through an interface - rather than the
knowledge of, or the 'knowledge that' a user possesses. This has been an
aim of the foregoing investigation. Through the use of a preliminary
typology, some of the ways in ways in which menus figure in the
production of naturalistic activities have been uncovered.

The analysis of the uses of menus with respect to a series of types
should be distinguished from similar classifications of activities
undertaken in the social sciences and psychologically-informed work. The
aim here has not been to construct a set of types that exclude, ignore or

21 A term borrowed from jurisprudence and the philosophy of language (e.g. Hart,
1961). See, Heath, (1982) and Heath and Luff (1996b) for a discussion of
defeasibility in relation to the uses of rules in practice.
abstract away from the details of particular accomplishments. Rather, the classification is intended as a resource through which the details of naturalistic activities can be revealed. It should be clear that the types outlined cannot be considered to be member's categorisations of their own activities (cf. the membership categorisation devices of Sacks, 1992) and nor can they be warranted through the sequential resources drawn upon in the analysis of talk by conversation analysts. The materials explored are not amenable to such analyses. As such, the analysis of menus presented in this chapter is more akin to a detailed ethnographic study, where the materials available are subjected to close scrutiny and the analysis is sensitive to the ongoing concerns of the participants in the setting. The foregoing analysis has sought to reveal the nature of the activities with respect to the use of menus. For example, menu activities can perform multiple duties: the reading of a menu item being an opportunity to see if it is relevant as a next action, but also for making sense of a prior action, or to discover properties of some screen object. Menus can not only used to accomplish particular activities at hand, but also those activities can be transformed moment-by-moment through the use of menus. Users can appear to be ambivalent about the tasks they are performing with menus, they can be drawn into certain courses of actions or away from apparent trajectories.

The uses of menus, even in instances where users visit a single menu, can reveal how individuals utilise the resources around them to make sense of what is underway and what is possible. The uses of menus can shape that understanding, but also that understanding can shape the ways menus are used. The menus are read with respect to the ways the contents are collected together, the relationship between their title and their contents and the individual appearance of each item. The understanding of one item made with respect to what is around it, the sense of the collection by the appearance and readings of the individual entries.

The activities on the menus can have an improvisational and innovatory character. It can be problematic even to provide a suitable post hoc characterisation of these. Instead, menu activities develop in their course, there is an 'at handedness' quality to the conduct. The analysis, then, has certain resonances with an earlier study of jazz improvisation on

---

22 In this sense it has similarities to the study of the use of signs in a hospital by (Sharrock and Anderson, 1979).
the piano. In his insightful analysis of his own development as a pianist of improvisational jazz, Sudnow (1978) how through practice his note choices became ‘just at hand’. Rather than trying to ‘find a path’ or imagine notes up ahead, ‘good notes were everywhere at hand’ (p.94). Menus on a computer system are, of course, of quite a different kind to keys and chords on a piano. They have to be read and made sense of. However, the openness of these simple options to various readings in the light of the particular purposes of the users, means, to some extent, relevant courses are often ‘at hand, right beneath the fingers’ (cf. Sudnow, 1978, p. 94). From these it is possible to shape new courses of activity along the menu bar.

This understanding of the menus then need not be seen as a mental process, but an achievement, an accomplishment made possible through the, in situ, and ongoing activity of using an artefact - made possible by the continually shifting resources at hand.

3.7 SUMMARY

The investigation of menu use has been undertaken within a framework that is, in many ways, similar to conventional studies within HCI. An ‘experiment’ was recorded where users were given a circumscribed activity to perform and an analysis undertaken which focused on a particular phenomenon. This focus, the use of menus, itself being a topic of research and enquiry within HCI.

The nature of the analysis has been distinctive, however. Instead of particular pre-specified criteria, such as the accuracy or speed of the activity being of issue, the activities have been investigated afresh. Through the use of a transcription system, developed for the purpose, instances of naturalistic menu use have been explored. In order not to lay too much emphasis on the goals and purposes that users may have when using the menus, a simple taxonomy of menu activity has been utilised. This reveals the richness and complexity of what appears to be a straightforward activity.

Not only do there appear to be a wide variety of uses for menus, but also the very nature of a particular use can be transformed in the course of its accomplishment. In scanning a menu bar a user could be involved in a wide range of activities including those that could be characterised as selecting an item, searching for an item and browsing through a sequence
of menus. These activities could also be characterised as discovering properties of the system, and objects on the screen, or as solving, defining and redefining the general activity in which the user is engaged. The particular nature of the activity is often ambiguous. Also, even when traversing a single menu or when holding open a menu, the user's activity can be transformed in the light of the particular circumstances at hand. There is often an ambivalence to the user's activity and a user can appear to use menus as a resource for improvisation. These are not only those users who could be characterised as 'novices' seeking inspiration for a next possible step. The menus are such a resource for all users, once a certain skill of moving through the menus is acquired any user can be drawn from one current activity into another.

It has been suggested that it may be problematic to consider these practices in terms of an individual's knowledge, whether internal or interacting with external information. Rather these resources might be considered in terms of the ongoing activity when using an interface, the practicalities of menu use, the in situ demands and possibilities that menus offer and provide. Menus figure in how individuals produce actions on an interface and render it intelligible. In this way, menu use can be considered an emergent, improvisational and innovatory activity.

The nature of the foregoing investigation places constraints on this exploration of contingent activities. In organising their actions and activities, the participants in the exercise no doubt orient to the constrained nature of the experiment. For practical purposes they are accomplishing an experiment, and this is revealed in how they carry out the activity, what they type into the application and how they order their activities. They are isolated from their work activities, other individuals and are performing a contrived activity.

An extension of the study would be to explore the nature of using a slide-making application in the course of everyday activities. This is indeed the focus of an ethnographic study by Nardi (1996b). However, the nature of screen-based activities can also be explored more broadly in a range of domains. In the next chapter, the use of a different graphical application is considered - a Computer Aided Design package - with respect to the practices and activities of the participants in the settings, that is, the members of an architecture practice.
Human-Computer ‘Interaction’ in an Architecture Practice:

observations on computer use and collaborative activities in a workplace

Let him be educated, skilful with the pencil, instructed in geometry, know much history, have followed the philosophers with attention, understand music, have some knowledge of medicine, know the opinions of the jurists, and be acquainted with astronomy and the theory of the heavens.

(Vitruvius, 1960, p. 5-6 cited in Cuff, 1992, p.84)

The ‘bricoleur’ is adept at performing a large number of diverse tasks; but, unlike the engineer, he does not subordinate each of them to the availability of raw materials and tools conceived for the purpose of the project. His universe of instruments is closed and the rules of his game are always to make do with ‘whatever is at hand’; that is to say with a set of tools and materials which is always finite and is also heterogeneous because what it contains bears no relation to the current project, or indeed to any particular project, but is the contingent result of all the occasions there have been to renew or enrich the stock or to maintain it with the remains of previous constructions or destructions.

(Lévi-Strauss, 1962, p.17)
4.1 INTRODUCTION

By examining individuals engaged in a circumscribed activity on a computer system in an isolated environment, the domain presented in Chapter 3 has many similarities to those experiments undertaken in much of the research undertaken within HCI. In this chapter the focus shifts to the production of screen-based activities in a work setting, an architecture practice. Through the analysis of naturally-occurring instances of screen-based activities, the ways in which computers are used to achieve the practical, everyday accomplishments of drawing and architectural work are examined. This approach to human-computer interaction differs from most of the work carried out in HCI, particularly the conventional approaches reviewed in Chapter 2. When examining naturally occurring screen-based activities not only is it problematic to categorise these in terms of a priori specifications, as suggested in the previous chapter, but these activities may be shaped by the contributions of others. Even screen-based activities then, the typical domain of HCI, can be considered as collaborative. The ways in which these activities are collaboratively achieved is the concern of this chapter.

By exploring the collaborative nature of screen-based activities the concerns of this chapter resonate with issues raised by researchers in CSCW, not only because of the analysis of the use of artefacts in workplaces, but also because of the focus on activities associated with drawing. 'Shared drawing', both as an activity and as a domain for technological support has been a long-standing concern of researchers within CSCW. However, the approach taken here explores the interactional production of screen-based activities. This orientation suggests what a more finer sense of 'sharing' and 'collaboration' than that current being considered in CSCW.

As a foundation for the analysis of the collaborative production of human-computer interaction, the nature of individual screen-based activities in an architecture practice is explored. After briefly giving some background to the motivations, analytic assumptions and methods behind this and the remaining studies in the thesis (Section 4.2), the setting and the technologies utilised within it are outlined (Section 4.3). This provides for a consideration of the issues raised in the previous chapter, in particular with respect to how architects organise their conduct on the screen, for example when they use menus (Section 4.4). Two further features of screen-based activity are explored in detail, the ways in which
they manipulate objects (Section 4.5) and how they navigate around the drawings (section 4.6).

In section 4.7 the domain of concern is broadened. Changing focus from the details of just the screen-based activities on the computer system, such activities are considered with respect to the ongoing conduct of the architects. Through the detailed analysis of an instance of activity, the use of the computer is explored in relation to other activities on paper plans, and the talk and visual conduct of the participants. The interactions of the participants are shaped by the shifting resources available to them on the paper and the screen. Moreover, and what may be of more concern with respect to HCI, the screen-based activities are organised, produced and even designed for another co-participant. This chapter concludes with a discussion of the relevance of the analysis to current conceptions of context within HCI and CSCW, and potential resources that may be useful when considering the analysis of workplace activities (Section 4.8).

4.2 BACKGROUND

One consequence of the recent critiques of plan-based, goal-oriented models of system use, by Suchman and others (Suchman, 1987; Winograd and Flores, 1986; Gilbert, 1987; Frohlich and Luff, 1990; Whalen, 1995b) is a turn away from the focus on experiments on individuals performing restricted tasks and towards more naturalistic studies of the uses of technology. Hence, there have been a variety of studies of technology and artefact use in a range of settings, including control systems for air traffic, airports and urban transportation (Goodwin and Goodwin, 1996; Harper and Hughes, 1993; Heath and Luff, 1992a), record-keeping systems in general practice (Greatbatch, et al., 1993), navigation on a large naval vessel (Hutchins, 1990), and computer-aided dispatch systems in emergency services (Whalen, 1995b). This chapter reports one such workplace study, of architects using Computer-Aided Design systems in a medium-sized architectural practice.

Although addressing particular sociological concerns, the workplace studies mentioned above have principally been reported in terms of their relevance for technology design and deployment. The reports of the studies have been primarily addressed to an audience concerned with supporting collaborative activities with new technologies: those involved in
the emerging field of Computer-Supported Cooperative Work (CSCW). Hence, it is perhaps not surprising that the focus of the studies has been towards forms of collaboration and participation. The technologies currently being developed are principally concerned with distributed capabilities, either within a single setting, like a meeting, or geographically dispersed, through computer networks or through video. As the primary interests of researchers in CSCW are with respect to these capabilities, the design of the particular interfaces to the various devices can remain within the domain HCI.

Similarly, the analytic focus of HCI and CSCW can appear disjoint. Workplace and other studies of collaborative work explore the nature of social interaction and cooperative activities between groups of individuals, whilst those concerned with human-computer interaction focus on the individual using the computer. As mentioned in Chapter 1, methods and conceptual frameworks drawn from the social sciences are considered appropriate to the former study, whilst experiments and psychological conceptions can be utilised for the latter (e.g. Dix, et al., 1993). However, this leaves the primarily cognitive conception of artefact use largely in place. The potential consequences of the critiques of Suchman and others are dissipated. Whereas in CSCW the focus has been on the interactional and collaborative nature of artefact use, the concerns and approaches of those within HCI can remain largely unaltered. Moreover, studies of collaborative work have, save for a few exceptions (e.g. Whalen, 1995b), ignored the domain of activities which have been of interest to researchers in HCI: the ongoing production of activities on computer systems. By examining naturally occurring materials gathered in a workplace setting, such activities can be subjected to scrutiny, and the local and emergent nature of screen-based activities with respect to the skills and practices of architects can be considered. This analysis may suggest some ways in which conventional activities which have been investigated by researchers in HCI and User Centred System Design (UCSD), such as direct manipulation and navigation through systems, can be reconsidered.

An alternative development for the field of HCI is to develop psychological frameworks and conceptions so that these account for social activities and naturalistic materials. Such are the motivations for those interested in Distributed Cognition (e.g. Hutchins, 1995, considered in Chapter 2), and related initiatives (e.g. Nardi, 1996a). These leave particular cognitive conceptions of human conduct in place, whilst seeking
to explore conduct 'in context'. The motivations and concerns of such programmes have much in common with other work within CSCW, indeed when considering the collaborative, communicative and informal practices surrounding the use of artefacts it can be difficult to distinguish the approaches.\(^1\) However, when considering the use of particular devices the cognitive and psychological foundations reappear: an individual's activity still appears to be disjoint from a surrounding social and communicative context.

As an alternative approach resources could be drawn from the large literature within the social sciences concerned with the use of new technologies, particularly analyses of work activities that makes use of qualitative materials gathered in real-world institutions and organisations. For example, within the sociology of work and organisations a wide range of studies drawing on field observations and informal interviews in the workplace provide accounts of the interactions between participants and the interrelationships between the introduction of new technology and the organisations and institutions into which it has been deployed (Taylor, 1911; Hughes, 1956; Becker, et al., 1961; Simon, 1960; Burns and Stalker, 1961; Cyert and March, 1963; Friedman, 1977; Hannan and Frieman, 1977; Aldrich, 1979; Dawson, 1994).

Within this body of work a wide range of orientations have been taken towards the nature of organisations and the tasks and interactions of participants in the workplace (Morgan, 1986). Given this diverse literature, it is not surprising that researchers in CSCW, and to a lesser extent HCI, have sought to explore the utility of this corpus of studies for the field (see reviews by March, 1991 and Jirotka, et al., 1992). Particular attention has focused on the possibilities of drawing from the work of Mumford (1986) and earlier of Emery and Trist (1969) which consider organisations in terms of the informal relationships and as socio-technical systems (Cherns, 1976). This notion has been particularly influential on researchers concerned with the design of new technology encouraging a more participatory approach to the development of systems where managers, individuals and designers collaborate in the process (e.g. Bjerknes, et al., 1987; Ehn, 1988; Greenbaum, 1988). Although these studies frequently involve the researchers undertaking observations and interviews with participants, the focus of these exercises is on involving

\(^1\) For example, see Nardi's (1996c) comments on the similarities between Distributed Cognition and Activity Theory and the possibilities for convergence.
participants in the design activity. Through these exercises it is hoped that the participants concerns with new technology will become part of the design, and that the designs will in some way be consistent with their skills and work practices, whether these are tacit or made explicit in the process. Nevertheless, if some qualitative materials are gathered in the process some orientation is required for their analysis.

In the remaining studies in this thesis a particular orientation is taken to the analysis that focuses on the tacit practices of the participants in the work setting. As mentioned in Chapter 2 this draws on previous work within conversation analysis and the ethnomethodology. It is worth reiterating some of the assumptions underlying the analysis.

First, the analysis focuses on the practices and procedures employed by the members of a particular setting. It has been a long-standing problem within the social sciences to assert the relevance of a particular analysis of social actions and activities (Weber, 1947). By detailed inspection of in situ human conduct the relevant ‘features of context’ are those which are oriented to by the participants themselves. Hence, the ‘problem of relevance’ is also one for participants, whose solution is revealed in the in situ procedures, practices, knowledge and reasoning they use and orient to. An analyses then seeks to specify, and provide evidence for, the relevance of features of context which inform the very accomplishment of the participants' conduct.

Second, the analyses in the remaining chapters focus on the interactional organisation of social actions and activities, where these activities may be accomplished through visual and vocal conduct as well as though tasks achieved on or through a computer system.

Third, the principal organisation of these interactional social actions and activities is sequential: a prior action providing the resources for making sense of a next action; a next action revealing the relevance of a prior.

Fourth, the domain of inquiry of these studies – tasks and activities on and through a computer system – may also be sequentially organised. Such tasks as an individual entering a command into a computer may, at times, serve to ‘respond’ to a colleague's prior action. Particular tasks may also provide the resources for next actions by others to be accomplished. In other words, tasks may also be sequentially organised.

Fifth, following an ethnomethodological orientation, the common-sense organisations revealed through the participants' actions are seen both as a
topic and as a resource for inquiry. They provide a focus for the study as well as access to previously unstudied phenomena and the 'methods of understanding' of the participants (Garfinkel 1967 p.31).

The orientation necessarily requires empirical materials, both of visual and vocal conduct. From these lengthy audio-visual recordings instances of particular phenomena are selected, say of possible coordinated computer activities with visual and vocal conduct. As is typically undertaken by researchers in conversation, these instances are then transcribed and collected together and into candidate sets for analyses. These sets are refined, often requiring a return to the original materials and more detailed transcription. In the studies in this thesis it was also necessary to develop novel transcription systems for detailed analysis and comparison, particularly with respect to providing a resource for revealing aspects of computer use found in the materials. Particular instances were then selected for inclusion in the thesis, this often requiring further developing the transcription systems for presenting the fragments.

Also in the development of the analysis instances which appeared not to fit within the candidate sets were subjected to particular detailed scrutiny. These so-called 'deviant cases' provided for extending and enhancing the analysis by revealing how the available resources were being oriented to. Such cases can often help a search for a more general socio-interactional organisation, which can deal with the specifics of the case at hand whilst preserving the integrity of related cases. They can also be used to demonstrate how a procedure or practice is oriented to even when a feature or action, recurrently provided for by a proposed organisation, is either absent or uncharacteristic. So, in Chapter 5 differences in the timing of the vocal activities of participants suggest further how temporal features may utilised as a resource within an interaction, and in Chapter 6 activities that, at first, did not appear to be tightly coordinated revealed resources which participants were orienting to in their collaboration.

After giving some background to activities undertaken in the architectural practices and continuing some of the issues revealed in the study of menu use in Chapter 3, these analytic and methodological resources underpin the remaining studies in this thesis. They provide a foundation for unpicking how the activities accomplished in and through a computer system are interactionally organised, allowing for instances of computer use to be scrutinised in some detail. Thus, the social context of
computer use need not be considered in terms of some spatial conception, as an 'arena' in which individual activities are accomplished, even where the focus is on the 'periphery' (Brown and Duguid, 1994), or in terms of some pre-specified variables, rather the context can be considered as continually being produced and shaped by the ongoing interactional activities of participants.

4.3 Technologies and Tools for Architectural Work

The domain of architecture and, more generally, design has been utilised both within HCI and CSCW as a resource for the development of novel computer systems and as a site for analysing computer and artefact use. So, for example, the practices of designers have been considered a useful starting point for considering the capabilities of collaborative systems, particularly the designers' use of shared artefacts (Bødker, et al., 1987; Harper and Carter, 1994; Tang, 1990; Tang and Leifer, 1988). Such studies are either centred on constrained 'design' tasks in experimental settings (e.g. Tang, 1990; Tang and Leifer, 1988), or rely on fieldwork, exploring the activities surrounding design, particularly the communication between designers and their 'informal' work practices (Harper and Carter, 1994; Murray, 1993; Rogers, 1992).

The materials upon which the observations in this chapter are drawn consist of video recordings of 'naturally occurring' computer use and collaboration in a provincial architectural practice in England. The data consist of field notes and video recordings which were collected on 8 different days in a period of three months. Two video-cameras were used, one focusing on the screen and the other, either on general activities in the design office, on an architect's work area, or on the architect in relation to the computer. Hence, there are nearly 90 hours of video-based materials available.

These materials principally concern the main project in which the practice was involved. This was to produce the working drawings for a large public building - a courthouse. The original designs had been produced by a public agency and the activities of the architects involved producing drawings that had the necessary details for building contractors to use. The project consisted of two parts: the demolition and redevelopment of portions of an old, existing courthouse and the design of a completely new building linked to the old one. The work of producing the working drawings was originally scheduled to take about ten months.

During the period of data collection, around six architects were working full-time on the building. The project was divided between the architects in terms of

---

2 The data consist of field notes and video recordings which were collected on 8 different days in a period of three months. Two video-cameras were used, one focusing on the screen and the other, either on general activities in the design office, on an architect's work area, or on the architect in relation to the computer. Hence, there are nearly 90 hours of video-based materials available.
practice studied is, for the United Kingdom, a medium sized practice, employing about 15 architects, town planners and graphic designers. It is an unusual British practice in that, although it is quite small, it uses computers extensively throughout the process of designing a building including: showing clients what prospective buildings will look like and how they will fit within the local landscape; producing the working drawings for building contractors, and even drawing sketches of original designs. In order to do this, they extensively use Apple Macintoshes, and principally a CAD package called MiniCad to construct drawings.\textsuperscript{3}

The use of the application and the system makes the setting quite opportune to study. This is the same hardware as used in the experiment reported in the last chapter, so that the \textit{in situ} use of that technology can be explored. The software incorporates a direct manipulation interface and a direct metaphor. There has been a considerable amount of research into these ideas (e.g. Ankrah, et al., 1990; Hutchins, et al., 1986; Norman, 1988), primarily because they are considered to help make an interface easy to use and yet it is unclear, precisely, why this might be. Such properties may account for the apparent popularity of the Apple Macintosh for professionals like architects. However, it has also been claimed that the system's popularity is due to its unique design and marketing, reflecting 'creativity' and even a certain 'individualistic and anti-establishment' outlook (Jones, 1990, p.25). By examining the \textit{in situ} use of the technology it may be possible to cast light on the affordances of the system, and particularly how it might support the actual accomplishment of practical activities by architects.

\footnotesize{the drawings that had to be produced: one worked on the old building, one on the plans for the bottom six levels of the new building; one on the plans for the top five levels; one on the sections, one on the elevations and one on the staircases. Two more architects contributed to the work: one was concerned with managing this and other of the practice's projects and an eighth worked part-time on drawings, including those of the details of the windows. This was the general division of responsibilities. The architects, in the course of their work, would make more contingent arrangements concerning which plans had to be drawn so that these could be produced according to the demands of other organisations, such as the fire and planning authorities and the customer.}

\textsuperscript{3} Macintosh is a trademark of Apple Computer Inc. and MiniCad is a trademark of Diehl Graphsoft Inc. More details concerning the version of the software used by the architects (v. 3.0) are given in Diehl Graphsoft (1990).
At various stages during the course of a particular commission, the architects will be given the responsibility to undertake drawings of a particular section, plan, elevation or detail of a building. In the cases focused on here, these include the staircases. As both the number of architects involved and the number of drawings can vary throughout the course of a project, so the division of responsibilities related to the kinds of drawing each architect worked on changed. Furthermore, the architects had to continually collaborate in order to produce a consistent set of drawings.

All the architects on the project used the MiniCad package which shares many features with other Macintosh applications. Figure 4.1 shows the overall screen layout when MiniCad is opened for the first time.

**Figure 4.1:** The MiniCad Screen when opened. The details of the tools on the top left of the screen are shown, enlarged, on the right.

MiniCad's tools are provided on two palettes, initially on the left hand side of the screen. The top one contains drawing tools allowing users to type text, draw rectangles, polygons, arcs and special, recurring symbols such as double doors and desks. The bottom palette is for a set of tools that allow the user to constrain the placing of certain objects on the screen, for example, constraining one line to be perpendicular to another.

---

4 Plans cut through a building horizontally about 1.5m above floor height, sections cut through vertically, usually at places where there are significant changes to the shape of the building, and elevations show the sides of the building viewed from the outside and capturing all its possible faces.
or ensuring that two objects fit precisely together. Below the main window is a ‘data display bar’ that provides information about the status of the drawing at a particular time, such as the current position of the cursor. Above the main window is a menu bar. Figure 4.2 shows the menu bar and the contents of the individual menus.

Figure 4.2: MiniCad Title Bar and Menus

In addition to the Apple menu available on all Macintosh applications, there are eleven MiniCad menus. The File and Edit menus have similar options to those in other Macintosh applications for opening, printing, saving and closing files and for simple editing operations. These menus also include options for smoothing, mirroring and reshaping objects. The Tool menu collects together several miscellaneous functions including joining objects together, combining several lines into a polygon, breaking up a polygon into several lines and changing surfaces for later three-dimensional manipulation. MiniCad allows the user to create a drawing in layers. These act rather like tracing paper, except that the user not only has control over which layers of a drawing can be seen at any time, but also how each layer can be used, so certain layers can be invisible, greyed out, visible but unchangeable or visible and changeable. Options for manipulating layers are given in the menu labelled ‘==‘. Apart from the Delta menu (labelled ‘a’) which provides users with ways of changing a
drawing into a 3D representation, the rest of the menus have similar capabilities to menus in other Macintosh drawing packages.\(^5\)

### 4.4 The Practice Using Menus

Although designing a building using a CAD package is a quite different activity to the production of presentation slides explored in the previous chapter, the architects' uses of menus appear to be similar in many ways. For example, in the following instance an architect is working on an elevation of the old building.\(^6\)

(1) **4B MV 5124 - Transcript 1**

1. Fill:.....................
2. Color: Fill Foreground: (0.5)
3. Color (Fill Foreground) (Fill Background) (Pen Foreground)
   (Pen Background) Use Layer Colors::::↑
4. (5.7)

He opens the Fill menu, holds it down for over 3 seconds (line 1), then goes down the adjacent Color menu (line 2), back up to the title bar, and then down again, selecting 'Use Layer Colors' (lines 3-4). So, even in a setting where an individual uses the system almost continually every day in the course of his work, there appears to be some ambivalence over the nature of the activity. It could be possible to characterise this use of the menus (lines 1-4) as a 'search' for a particular item (i.e. 'Use Layer Colors') where the long hold on Fill is considered to mark a failed step and the eventual selection a successful match. Or, as in the accounts of Mayes and Young and their colleagues, in terms of some type of matching of semantic

---

\(^5\) To list these: the Text menu offers a range of ways of changing the style, font, size and layout of text; the Page menu allows the user to view the drawing in various ways; the Color menu provides 'pens' for drawing and shading objects on the screen in different colours; the Fill menu offers a range of shadings (or hatchings) for objects; and the Line Weighting menu (labelled '/\') gives a variety of styles of lines in differing thicknesses, differing dashing patterns and with differing arrowheads. Though these options are similar to other graphics packages, for example PowerPoint considered in the last chapter, MiniCad offers a larger number of options, with a greater degree of flexibility.

MiniCad also provides for the user to associate spreadsheets (or 'worksheets') with drawings. The rightmost menu, the command menu, provides options for manipulating these.

\(^6\) The transcription system uses the same orthography as that in the previous chapter. A fuller description of the transcription system is given in Luff and Heath (1990)
and lexical, or internal and external knowledge, taking place (Mayes, et al., 1988; Young, et al., 1990). However, the Fill menu only contains a set of different graphic patterns (see Figure 4.2). It is, of course, possible to provide an account of such an activity in terms of plans and goals, but a goal that is both related to the long look at the set of patterns and the selection of 'Use Layer Colors' would have to rather broadly defined. So, as in the cases of menu use in previous chapter, the activity, even for a routine user of a system, would appear to be more fluid, flexible and dynamic.

Over five seconds later the architect goes on to briefly peep at the File menu (line 6) before typing the command key and the 'S' key together (line 8).

(1) 4B MV 5124 - Transcript 1
5   (5.7)                        
6   ↓ File::: New (Open) ↑      
7   (2.3)                        
8   'cl'                        

Again, this could be viewed as a search: for example as a search for an appropriate key command for saving a file displayed on the File menu. Both the command and the 'accelerator' for the Save command are displayed on the menu. The typing (in line 8) could be related to the activity with the menu, but to account for these in terms of a goal, plan, learning or matching knowledge appears somewhat strange. As Save is available when the menu is opened (in line 6), such an account would require the target of the search or the knowledge on the interface not to be appropriate just prior to when the command was eventually performed (line 8). Some aspect of the knowledge would have to be defeasible, appropriate and applied in one case and not the other. Alternatively, the goal could be transformed or changed in some way. The question then would be what relevance such an account of shifting goals would have to the user's ongoing activity.

---

As with many other applications for the Macintosh, MiniCad offers users, what are called, 'accelerators' as quick alternatives to selecting items from menus. These are performed by typing combinations of keys simultaneously (called 'chording'). In (1), the user types the command key labelled 'cl' and the 'S' key together. This performs the same operation as selecting 'Save' from the File menu. These chords are often displayed on the menu to the right of the appropriate item.
As in the case with the PowerPoint menus considered in the last chapter, the difficulty may be not in specifying an appropriate characterisation of the activity, but in characterising the activity in terms of the knowledge and goals that the participants possess at any particular time.

In the following instance, the architect is working on a particular box on a plan.

(2) 6B MV 4302

1  ↓ //::: ((Line thickness 1))::: ((Line thickness 2)):::  ↓
2  Fill::::::  ↓
3  Color:::::::: Put Down Color::: Use B&W Color: ↑

Once again, this activity does not appear to be that conventionally associated with experts using menus. It is also difficult to view it as a search for a solution to a particular problem, the line menu offering quite different possibilities for action to the Fill and Color menus. Instead, the nature of the activity of the user when engaged with the menus appears to develop in the course of using them. Although appearing to go to select a particular line width, the nature of the user's activity transforms as he is using the menus. Characterising such an activity as an interaction between knowledge in the head of the user and knowledge on the menus (as in Young, et al., 1990), focuses analysis on developing some mechanism or causal connection linking the two together and ignores the innovatory, improvisatory and, almost ambivalent nature of the activity.

Studies in HCI frequently focus on one aspect of the interface, like the use of menus, or on particular pre-designed tasks which users are asked to accomplish, often in experimental settings. By concentrating on individuals carrying out such tasks, these studies operate with a relatively constrained conception of screen-based activity. In the following sections a broader view of the interaction between computer and user will be taken, one which explores naturalistic screen-based activities.

4.5 THE MANIPULATION OF OBJECTS

Much of the architects' day-to-day activity is involved in making detailed changes to the working drawings of the building using the CAD package. It is worth examining in detail an instance of this screen-based activity. In the following fragment an architect is making changes to a staircase in a drainage tunnel that leads away from underneath the building. Figure
4.3 is a sketch of the plan of the staircase. The arrow is used by architects to indicate the downward direction of the staircase. Each stair is also numbered.\(^8\)

Figure 4.3: Sketch of the plan of the Tunnel Staircase (line 1, fragment 3, arrowed in transcript 1)

The user has just been moving around the plan, both panning across it and selecting various layers of the drawing. He then shapes a box (a Marquee\(^9\)) over the region of the staircase (line 1),

(3) 7B MV 4284 - Transcript 1\(^{10}\)

\begin{verbatim}
1  ↓ >Marquee:·········:····:⟩↑
2  (0.8)
3  ’£C’
4  (1.4)
5  ‘(£↓ £↓ £↓ £↓)’
6  (1.5)
7  ‘£↓’
8  (1.9)
9  ‘£V’
\end{verbatim}

This marks the objects contained within the box as ‘selected’. He types the command and ‘C’ keys simultaneously (line 3), moves down five layers (lines 5-7), by holding down the command key and pressing the down arrow key, and then types the command and ‘V’ keys simultaneously (line 9). Together the actions in lines 1 to 9 are commonly characterised in Macintosh and related applications, as a ‘cut and paste’, or more accurately, a ‘copy and paste’. The architect copies a portion of staircase

---

\(^8\) These are just visible in the figures, numbered from 1 to 15. By the operations that follow, the architect accomplishes a deepening of the tunnel.

\(^9\) A Marquee is a dashed box which can be shaped by the user to mark out an area of interest on the screen.

\(^10\) ‘\(\gg\)’ is used when the user ‘drags’ the mouse over some area of the screen, in line 1, for 2.3 seconds in marking off an area of the screen.
from one layer and pastes it onto another (see Figure 4.4). The architect goes on to manipulate the amended plan.

He types 'command-J' (line 11), this 'joins' the new portion of staircase to the old. After marking out an area of the staircase that contains all the numbers of the individual stairs (line 13), the architect clicks in the area above the staircase (line 15). He then selects each end of an arrow and then moves it upwards and rightwards (lines 17-21). After positioning a construction line using the '+' drawing tool and removing an indent in the wall below the staircase (not shown in the transcript), the architect moves first the wall at the end of the staircase (line 29) and then extends the wall alongside it (line 31). The shape of the walls now encloses the extended length of the staircase. Figure 4.5 is a sketch of the plan at this point.

The architect then removes a portion of stairs at the left end of the staircase and makes several alterations to the numbers on the stairs. He then selects the right and left hand sides of the arrow (line 35 and 37). After appearing to group the ends of the arrow together (by the 'command-G' combination, line 39) he moves it into a new position on the stairs (line 41, Figure 4.6).
This fragment is extracted from an extended flow of activity. It is possible, post hoc, to view these actions as steps towards a goal that the architect had before commencing, or in terms of sub-tasks that make up some overall task he had to accomplish. In order to extend the staircase, the user has to add a new portion to the existing flight of stairs, and move the existing walls to surround this new portion. The numbers and arrows on the new staircase have also to be matched to the new drawing. The remainder of the activity could then be characterised as local adjustments necessary in order to accomplish the overall goal.

It might be expected that an activity such as extending a staircase would, for an architect, be a routine, general procedure that, with minor alterations, could be instantiated for particular cases at hand. Yet, in the fragment above, and throughout the corpus of data, it is hard to discover such a routine procedure. Although it is possible to construct a post hoc 'plan' for an architect's manipulation of a drawing, even for other manipulations of staircases, there appears to be no generality to these 'plans'.

When the new portion of staircase is pasted onto the old there are a range of actions to perform and a range of possible ways of performing them including: moving the walls and stairs; renumbering the stairs; and grouping, ungrouping and constraining the movement of the various objects displayed on the screen. As the architect continues to manipulate the plan there remains a range of further and different possibilities for action. As with the use of menus mentioned earlier, the nature of the screen-based activity that the user engages in is continually changing throughout the articulation of the activity.

A characterisation of the activity as a pre-specified plan would have to provide for some flexibility in the way the 'sub tasks' are carried out. It would be difficult to envisage operators or selection rules that could define
an ordering between, say, removing an indent from a wall, moving an
arrow, or renumbering the stairs. There would be an indefinite number of
such orderings. Similarly, apparent diversions from a plan, or
'inappropriate' activities might have to be characterised as 'errors',
'aborted sub-tasks' and the like.

So, in foregoing fragment (3) the architect selects the numbers of the
stairs using a 'Marquee' (line 13) and then deselects them with the
subsequent click of the mouse button (line 15).

(3) 7B MV 4284 - Transcript 2
10   (7.2)
11   '8ßJ'
12   (3.0)
13   ↓ >Marquee:........:↓
14   (0.3)
15   ↓↑
16   (0.2)
17   ↓ ((left point of arrow)) ↑

This selection and deselection does not contribute to a change in the
drawing. However, it is also not arbitrary: the Marquee takes some time
to shape over a portion of the drawing. When the Marquee has been
drawn the items within it are then highlighted. Making the selection with
the Marquee also brings into view a set of objects that can be manipulated.
This provides resources concerning not only what can be manipulated but
the range of relevant next activities for the architect. Actions performed
at the time will be applied to the selection. The click away removes the
highlighting, and also provides for a range of other next actions, like
selecting the left end of the arrow (line 17). The Marquee is one way of
discovering what can be done at the current time on the interface and
what is appropriate. By its use the architect can reveal the nature of the
objects on the screen, whether they are disjoint or connected, for example.

Therefore, a particular activity need not have to be related to some
subsuming goal or plan, instead each action can be seen to be linked to the
prior and making possible a range of next actions. A course of activities
could then be seen to unfold temporally and sequentially. There is no
doubt a trajectory to the actions, but it is managed locally from moment-
to-moment, a selection bringing to light possible next actions, revealing
the resources through which the objects on the screen can be manipulated.
Likewise a movement of an object can make apparent not only the further
actions to be undertaken, but the nature of those actions: adding the block
of stairs reveals the ways the accompanying arrows and numbers need to be altered. Even the starting point in this fragment suggests an improvisatory and unfolding nature to the activity. Rather than creating the additional stairs afresh, they are borrowed from a layer above. Not only does this save on the work the architect has to perform, but also helps ensure that the drawings on the various layers will match. The architect utilises the tools that are at hand, with what remains of previous constructions and destructions (cf. Lévi-Strauss, 1962).

The ease of manipulating screen objects on systems like MiniCad on the Macintosh has been characterised in terms of its 'directness' (Hutchins, et al., 1986). Although the user has to move objects on the screen by operating a mouse and a mouse button which then makes cursors, marquees and the like move around the screen, these act directly on the images on the screen, unlike, say, type-written commands. Hutchins et al. explain the properties of such interfaces in terms of the mappings between an individual's goals and the physical tasks required to operate a system. They suggest that there are two aspects to the notion of directness of an interface: distance and engagement. Distance accounts for the translation necessary from goals into physical tasks. Engagement accounts for the 'qualitative feeling' of manipulating the objects of interest. Direct manipulation interfaces make the distance shorter by, for example, avoiding the semantic translation necessary to construct typewritten commands, and they increase engagement by employing visual metaphors of objects related to the user's task on the screen. Other researchers have attempted to further clarify aspects of direct manipulation by, for example, exploring the interrelationships between direct manipulation and metaphors (Ankrah, et al., 1990) or by characterising the properties of screen objects in terms of their affordances (Gaver, 1991).

To account for the screen-based activities of architects using the MiniCad application conceptions such as metaphor may not be adequate. They may be useful for interface designers when developing the system, but it is unclear to what extent the invocation of the metaphors of pens, erasers, rulers, set squares and tracing paper utilised within the MiniCad interface design contributes to an analysis of the architects' moment-to-moment use of the system. Similarly, an account of the affordances of the objects on the screen may be able to suggest the various perceptual properties associated with different kinds of lines, boxes and shapes, but it
is hard to see how such an account could be related to the unfolding activity of the users.\footnote{Indeed, the original formulation of the conception of affordances is a radical turn in psychology, focusing on the direct perception of objects without intermediate information or mental processing (Gibson, 1979). In recent uses of this work, particularly for technology design, the conception has become blurred (Gaver, 1991; Norman, 1988). These do preserve, however, the primacy given to the direct and perceivable properties of objects. It is unclear whether the concept of 'an affordance of an object' can be preserved (and still remain stable) within a framework where the sense-making practices of participants are contingent on an unfolding course of activities. See also Coulter and Parson's (1990) discussion of the discursive practices surrounding perception.}

Obviously, the positioning and sizing of objects is a critical aspect of the work of the architects. Not only are they working on drawings on which the measurements to be used by building contractors need to be accurate to about 5\(\text{mm}\), but also the placement of objects must fit precisely across layers in a drawing and across separate drawings. In (3) the architect organises his activity so the location to which objects are to be moved are specified before he moves the object. The join command (line 11) does not just connect together two disconnected objects, it makes one object of two distinct objects. So, in the case of walls, the join command not only connects the walls together, but also removes intermediary lines between the connected walls. So, the architect's 'paste' of the new portion of staircase (line 9) is carefully designed for subsequent actions. The architect still has to make adjustments to the positioning of objects on the drawing, especially to related ones, but his manipulation of those objects could be better characterised as extending a staircase rather than sliding or pushing them around the drawing. The activity is one related to the production of drawings of buildings rather than merely the movements of graphical objects. Throughout the course of the day, the architects engage in making a series of changes to the drawings that they are responsible for. Each can have knock-on consequences for other parts of the drawing and even for the drawings of others. The visual appearance; the size, shape and position of objects on the screen is important for managing this work, but these are utilised within an unfolding course of the activities. An activity and its results on the screen shapes the development and production of the next.

It is not just the tools and the objects on the screen that are essential to the way the architect orients to this activity, but it is also the space on the
screen. So, for example, in (3), the user moves the arrow on the staircase up and to the left (line 21). This position is out of the way of the staircase, which the user goes on to manipulate. It also is vertically above the new location of the staircase, visible whilst he is working on the stairs and in a position where it can be easily moved back down. Its very presence on the screen can be seen to mark out future possible and necessary courses of activities.

In the following instance the architect is redrawing a plan of a toilet. After calculating the distance between two walls in the centre of the drawing (‘250’), the architect then moves the number slightly (line 1). deselects it (line 3) and moves the left wall to the right (line 5).

(4) 6B 5249
1 ↓ >((‘250’))······················↑
2 (.)
3 ‘return’
4 (5.2)
5 ↓ >[Wall]······················↑
6 (2.1)
7 ↓ >[Wall]······················↑
8 I: “ummm”
9 (3.5)

Figure 4.7 is a sketch of the plan before he moves the wall and Figure 4.8 is after it.

Figure 4.7: Sketch of toilet (line 4)  
Figure 4.8: Sketch of toilet (line 6)

The architect then moves another wall, originally in the same position as the wall he just moved, to the right hand side of the drawing (line 7). (Figure 4.9 is a sketch of the plan at this point.)

12 It appears that two walls are drawn in the same place, one on top of the other.
He then goes on to work in the space he has now made available to the right of the urinals, adding a thinner, shorter wall. As in (3) the walls are moved into a space clear but related to the area in which the architect is to work. Though in this case, one, and probably both, of these walls are no longer required for the plan. The walls, whose size and shape are still consistent with the rest of the drawing, are possibly, being preserved as a resource for further work. Such preservation of objects recurs throughout the corpus of data. There are also related instances where a copy of an object is moved into an area of free space and manipulated there, for example, to work on its precise shape or to carry out various operations upon it. The MiniCad package in particular, and the Macintosh system in general, facilitates the use of space as a practical resource for working on drawings. Just as an area of a work table may be utilised to perform various detailed activities, an area of space on the screen can provide for such work. It can also offer a location for temporarily placing objects to be moved back into main drawing or for placing objects which require further work on them. As with the objects on the screen and the tools available, the use of spaces in a particular drawing can be improvised in the light of both the foregoing course of activities and with regard to possible forthcoming trajectories of action. Such practices could be considered in terms of the 'directly manipulation' of objects and space. However, this would gloss a range of heterogeneous practices concerning how objects on the screen are utilised for the practical purposes of the users. Moreover, the characterisation of the use of artefacts in terms of the distance between a goal and a physical task neglects precisely those resources through which the activities are achieved.
Another oft-cited feature of usable systems is the ease with which users can 'navigate' around the information and capabilities offered by them. Systems like MiniCad provide menus and tool palettes to make accessible an extensive range of capabilities. Designers attempt to make these menus and palettes both familiar and recognisable, so that they have a fairly unchanging appearance. The menus and palettes also need to reflect the current possibilities for action, changing in subtle ways like shading out actions that are not currently available, for example. Hence, they provide resources for presenting the capabilities of a system in similar ways to the menus of PowerPoint considered extensively in the last chapter.

The information presented by CAD systems would at first appear to be straightforward, drawings consisting of recognisable shapes and text. But, like other systems, such as Hypertext and other forms of complex databases, the information presented can be on different levels, or layers, and the items on each can be related together in subtle ways (e.g. Monk, 1989). MiniCad provides three ways of moving around a drawing: panning across a drawing; zooming into and out of an area of a drawing; and moving between layers of the drawing.

Panning allows the user to move around a drawing that is too big to fit on the screen. It is possible by pan around the drawing either by using the panning tool on the drawing palette or by pressing the appropriate arrow key on the keyboard. This moves the view on the screen to a contiguous area of the plan.

Zooming in and out of a particular portion of a drawing allows the user to focus in on a detail or to see a wider view of the current area of the plan. This is performed by selecting the Zoom In and Zoom Out tools in the drawing palette (on the second row). So, for example, to zoom in on a portion of the drawing, the 'Zoom In' tool is selected with the mouse, the user then specifies the area to focus on by shaping a Marquee over the appropriate area with the cursor and then releasing the mouse. Zooming out is accomplished by locating the Marquee over the area which will
remain the centre of the wider view, after the 'Zoom Out' tool has been selected.13

Layers offer a way of providing different kinds of information on the same drawing. Each layer can be displayed or hidden, so that, for example, particular details such as the plumbing, can either be available for manipulation or hidden from view. The user can move between the layers of a drawing in three ways: using the Layer menu; selecting the 'Lyr' section of the data display bar at the bottom of the screen, or pressing the Command key and either the up or down arrow at the same time.

In the following instance the architect uses all three ways of moving around the drawing (shown schematically in fragment 5).

(5) 4B 5417

1 ((Pan ↑))
2 (3.3)

3 ((Pan ↑))
4 (3.3)

5 ((Pan ← ))
6 (2.5)

7 ‘(84)’
8 (18.3)

13 Note, that this is the opposite way of operation compared with similar functions on other Macintosh applications, where the selection on the menu applies to a previously selected object on the screen.
After working on the title block in the bottom right hand corner of a set of elevations (sketch a), the architect shuffles the drawing up twice by moving the cursor over the edge of the drawing (b-c, lines 1-3) and then across to the left (d, line 5). He types ‘$64$’ (e, line 7), fitting the entire drawing onto the screen (the accelerator for ‘Fit To Window’). He then zooms out further using the ‘Zoom Out’ tool on the Drawing Tools palette, specifying the area the new drawing will focus on with a Marquee (f, lines 9-11). By selecting the top window bar of the window and dragging it down to the bottom of the screen he effectively moves the set of elevations out of view and reveals another elevation (line 13). He clicks on this elevation (g, line 16), zooms out (h, line 18), goes to the menus to select the layers of the drawings to display (lines 20-24) and then goes on to alter the title block of this drawing.

The architect not only moves from one drawing to another in three different ways, he accomplishes this by moving the mouse, using an accelerator, the menus and the ‘Zoom Out’ Tool. Just as in previous instances it is possible to consider this movement as a path from one title block to another; zooming out of the first plan, jumping across to the other drawing and then zooming in on that. But once again this ignores the development of the activity from moment-to-moment.

At the beginning of the fragment very little of the main drawing is displayed, indeed only a few lines are visible on the bottom left of the
screen. The architect pans up and left (lines 1-5) bringing into view a portion of the building. Once the details of this appear, the architect goes on to display most of the complete drawing (line 7-8). This takes some seconds to draw, as details of each elevation have to be presented. When completed, he then marks out the area which the ‘zoomed out’ picture will centre on: an area in the centre of the screen. Once he has specified this area, a much smaller image is presented with the title block. Again this takes time to complete. Only as the last bottom right component is being drawn does the architect move towards the window bar and onto the next elevation.

The architect appears not merely to be moving from one component on one drawing to a similar component on another, but taking account of the images as they appear: these images providing resources for example, to confirm aspects of the drawings’ completeness and correctness. Each activity then appears to be tied to a previous one: confirming the sense made of the prior displays, its location and contents, and making possible further activities.

In the following instance the user has just finished altering some walls at the right hand end of the service tunnel (shown schematically with fragment 6).
By pressing the appropriate arrow keys the architect pans left across the top of the tunnel five times (lines 1-9), then down three times (lines 11-15). He adds some layers to the drawing (lines 17-20), pans down again (line 23) then pans diagonally down to the right (line 25-26). At the end of the actions the architect has a plan including the female staffroom in the main building displayed on the screen.

The service tunnel is a rough 'T-shape' with a bend in the middle. There are staircases outside the building (at each end of the top of the 'T') for underground access to the main building. The architect pans across and down the plan following the service tunnel into the building. He also paces his activity with the system's redrawing of the plan. It is possible to press keystrokes in quick succession, avoiding redrawing intermediate plans along the way. By slowing down the frequency of hitting the arrow keys as he nears the centre of the service tunnel, the architect is able to monitor the system redrawing the plan as he passes the bend in the middle. Progression down the tunnel slows down further once he passes the crook in the tunnel (line 15). When he arrives at the main building, the architect adds layers to the drawing and then proceeds into the female staffroom.
There are other ways in which the user could move between the staircase at the right end of the tunnel and the female staffroom in the main building; for example, zooming out to a plan of the whole drawing and then zooming into the relevant area. Yet, the organisation of this activity rests on the architect's skill of manipulating this particular plan and the system in general. Zooming out and in requires MiniCad to display the entire drawing before a Marquee can be selected to mark the location to focus on. As MiniCad can take tens of seconds to display a detailed drawing, zooming out to locate the position for a next area of detailed work can take considerable time.

In moving down the tunnel the architect can make use of the geography of the projected building, even possibly taking account of features, like the bend in the tunnel, that are not straightforward. He can also coordinate this activity with the projected speed of the system. For moving around the drawing in this way, it is not necessary to have all the details presented. Layers need only to be drawn once the destination is neared. As MiniCad takes longer to draw the details of a plan if it has to draw several layers each time, the organisation of the activity in (6) can be seen to be shaped by a range of contingencies. The movement down the tunnel and the drawing of the details of the layers are shaped not only with respect to when items appear on the screen and the geography of the building, but also with the projected speed of the system.

Navigation around the plan can also be seen as contingent on aspects of architectural practice in general. The building the architects are working on has certain consistencies and symmetries. The architects can make use of these in their drawings, for instance, copying portions of a plan from one area to another. It also means that work on one localised area of a plan has important implications for changes that need to be made to other areas. For example, changes to an internal staircase on one level will involve similar, but not identical, changes to other staircases on that level. In the materials there are numerous instances of movements back and forth between spatially distant areas of the plan, moving along corridors, up and down stairs and across to other corners of the building.

In some senses, when moving around the drawings the architects are 'interacting' with the computer. They have to make sense of the images as they appear on the screen and carry out actions in the light of what is presented. However, making sense of and achieving this interaction rests on the everyday, architectural skills of producing drawings and buildings,
and common-sense practices for economically utilising the tools available to them.

4.7 THE CONTEXT OF COMPUTER USE: MULTIPLE ACTIVITIES AND COLLABORATIVE VIEWINGS

Though the architects utilise the CAD system to construct their working drawings, the use of the system is immersed within activities with other artefacts, including calculators, product manuals and most significantly, paper. Despite all the architects having a computer system on their desks, they still keep next to them large (over a metre in length) paper versions of the plans, sections and elevations, and frequently use them when working on the details of the drawings. Thus, the use of the system is interleaved with other activities: a calculator may be required to work out a particular dimension, a manual as a resource for the size and shape of a component on the drawing. Architects may also make use of the projected nature of activities in order to interleave these with the use of other artefacts. So, for example, an architect may roll up a set of plans whilst the system may be opening, saving, plotting or closing a file - operations that can take some time with complex drawings. He may coordinate this activity with glances to the screen to monitor its progress, utilising junctures in one domain as a resource for moving onto another. The pacing of his activity on the computer system may thus not merely be related to the changes on the display. Consequently, too narrow a focus on screen-based activities neglects the interrelationships between the actions undertaken on the computer system and the circumstances of its use.

In the following instance, an architect, Pete (P), moves between one section of a drawing to another, both showing different staircase sections. These are one part of a set of three sections contained in a single file. Thus, he can move between the sections via an intermediate 'zoomed out' view of the drawings. From a complex drawing of one section, the architect first pans over to the portion containing the staircase (lines 1-5), and then removes layers from the drawing, thus simplifying what is displayed (lines 6-9).

14 See Luff and Heath (1993) for a detailed case of such interleaving of activities.
The architect then selects a complete view of the drawing (containing small images of all the sections) using the 'Fit to Window' command (through the '⌘4' accelerator, line 11). From this he selects the third image (line 15) - a section of a different staircase. After some 3 seconds Pete can be seen to follow the direction of the lowest staircase’s upward with a movement of the cursor.

These drawings are complex and the system takes some time to display each. However, this rather straightforward instance of computer use is organised with respect to the ongoing activities and interactions of architects in the setting. It is tied to a discussion between Pete and a colleague, Richard, concerning the staircases of the building. In the course of this interaction, the two participants make use of resources both on the

The system takes nearly 2 seconds to display each of the new drawings in this fragment.
computer system and on paper plans. The use of the system is interwoven with this interaction.

In the following sub-sections, this instance is considered in some detail: first, with respect to the activities involved in securing alignment to the screen; second, with regard to the management of the shifts in orientation between computer system and the paper plans; and third, relating to how the screen is utilised and made sense of in the course of the interaction. Through examination of this single instance, a computer-based activity can be seen to be thoroughly immersed in the talk and visual conduct of the participants: it provides resources for making sense of these activities and it is also made sense of through them.

### 4.7.1 Securing an Alignment to the Objects on the Screen

A few minutes prior to the fragment presented in transcript 1, whilst Pete is working on his computer, another architect, Richard (R) asks Pete about a particular feature on the plans.\(^\text{16}\)

\[\text{(7) 8A JVC 3094 - Transcript 2}\]

\begin{verbatim}
1 R: are you sure you're flipping the right one around Pete.
2 (0.2)
3 P: yeah
4 (2.5)
5 P: I'm flipping (0.5) look on level one
6 (1.1)
7 R: you you just told me that you have
9 flip>flipped that one around
10 (0.6)
\end{verbatim}

Richard's question concerns some changes the architects have been making to the plans in relation to the staircases, and draws Pete into a discussion over the paper plans sited on the table to Pete's left.

\(^{16}\) The arrows are intended to give an approximate indication of the location of images on the right with respect to the talk on the left.
In order to accommodate a new room at the top of the building the directions of all the staircases leading up to it have to be changed (or 'flipped'). Any single drawing produced by the architects must not only be internally consistent so that objects, such as staircases between the same floors have an equal number of stairs, but also so that different types of drawings (i.e. plans, sections, elevations and details) are consistent with each other. These drawings may be the responsibility of others in the practice. One of the ways architects both maintain a record of the present state of a design, and make public the changes which they are making to it, is by 'marking-up' the paper versions with annotations and comments. Unfortunately, in this case, there is some confusion about the exact nature of the change, the plans being marked up in two colours each indicating the stairs running in a different direction.

Over three minutes later, the architects are still discussing the staircases, and Pete turns back to his computer and says 'you see (0.3) um: (0.6) there (. ) this is the outside level'.

(7) 8A JVC 3094 - Transcript 3

10 P: you see (0.3) um: (0.6) there (. ) this is the outside level
11 (1.0)
12 R: umm↓
13 (1.3)
14 P: and it can start (0.7) (part of the drawing) we are on now
15 (2.3)
16 righ(t) (here we are)
17 (1.4)
18 you've done a section through
19 (1.2)
20 this guy here (. ) right?
21 (0.6)
22 R: yep
23 (.)
24 P: thats the start of the section I had↓ (0.3) to: give myself( them
25 levels) and he (has to start) back there (. ) and go up↑=
26 R: =oh yeah↓ thats the same one

Both the architects have been making changes to their drawings with regard to the staircases, Pete to the general sections across the building and Richard to detailed drawings of the stairs. Pete's turn to the computer system appears in the light of the apparently confusing marks on the paper. In the ensuing talk in lines 10-26, he points to a drawing of the stairs on the screen, over to an area on the paper plan and then back to another drawing on the computer. In his talk and through his visual conduct Pete is locating features on the screen and on the paper plans, in
order to resolve the difficulties concerning which way the stairs are and should be drawn.

Throughout this talk, Richard's contributions appear to be minimal: an 'umm↓' (line 12) and a curt 'yep' (line 23). Only after Pete has uttered 'and he (has to start) back there (. ) and go up↑' does Richard offer a more substantial confirmation: 'oh yeah↓ that's the same one'. The activities, in relation to the screen and to the paper drawings, in concert with Pete's talk achieve, eventually, a confirmation from Richard that a particular object on the screen is 'the same one'. Pete's work in securing this alignment is particularly complex, involving both the paper documents and the computer system. Moreover, Pete utilises two screen images of different areas of the staircase (the images b and d in transcript 1).

Pete's talk is accompanied by a series of perturbations: his initial utterance in the fragment, for example, includes both pauses and an extended 'um:' (line 10). Whilst he utters this, although he orients towards the screen, his right hand remains on the plans. After he utters 'you see um:', Pete turns to the system and points to a location on the drawing on the screen with his right hand. His right arm remains pointing to this general domain until he utters 'and it can start' (line 14). However, Pete's reorientation and the point itself, though they are designed for Richard, do not immediately secure a realignment. Instead, Richard remains oriented towards the paper. It is only as Pete is uttering 'there (. ) this is the outside level', does Richard glance away from the paper plans.

10  P:  you see (0.3) um: (0.6) there (. ) this is the outside level

Despite the emphasised 'there', the brief pause and the point, Richard's glance towards the screen only commences on the word 'outside'. Pete
utters this word not only with vocal emphasis, but also raises his eyebrows as he says it.\textsuperscript{17}

(7) \textbf{8A JVC 3094 - Transcript 4}  

\begin{tabular}{ll}
P points to screen & \textbf{eyebrow flash} & P points left along floor level on screen \\
\hline
| paper | \(\ldots\) | screen |
\end{tabular}

P: you see---um:---- there-this is the outside level....;....;  
R: \(\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldot
It is not just a pointing at the screen that accomplishes this realignment, Pete's talk and visual conduct together provide resources for Richard. Nor is the point a simple activity. Indeed, it appears to be made up of three components: the point to the right of the bottom-level, a first movement to the left, and then as Pete utters 'and it can start' (line 14) a movement back along the floor to the right (see the sketch in figure 4.10).

10 P: you see (0.3) um:  
   there (.)  
   this is the  
11 outside level  
12 R: umm↓  
13 (1.3)  
14 P: and it can start (0.7)

Figure 4.10: Sketch of staircase section (displayed in lines 10-14, transcript 3, whilst P utters 'there (.) this is the outside level () and it can start')

On completion of the second component of the gesture Richard utters 'um↓' and after that Pete commences the third component, back along the floor. However, as the third hand movement comes to a halt so does Pete’s talk. He pauses for (0.7) seconds. Though having secured the alignment of Richard, it is unclear what is the gist of Pete's utterance 'there (.) this is the outside level () and it can start', even when he completes it with '(part of the drawing) we are on now'. Richard's contribution is less than unequivocal, but this may be in the light of the nature of Pete's own emerging utterance.

Pete's gestures mark out a precise location on the drawing, along the line for the floor of the ground (the outside) level of the building, and to the bottom of the first staircase up, but they also secure an appropriate alignment from Richard. Through their talk, orientation and gestures Pete and Richard arrive at a common viewing of the screen. This viewing is accomplished through the local management of their conduct, one's activity being produced in the light of the changing contributions of the other.
4.7.2 Managing the Alignment Towards Paper and Screen

In his pause after the word 'start', Pete goes on to make a series of actions on the computer, first moving his right hand down to the mouse, then his left to the keyboard, then he strikes two keys together and finally moves the mouse again (lines 10-12, transcript 1).

(7) 8A JVC 3094 Transcript 5

<table>
<thead>
<tr>
<th>points right across screen</th>
<th>right hand d’n to mouse</th>
<th>left hand d’n to kbd</th>
<th>P strikes key chord ‘ß4’</th>
<th>P moves mouse</th>
</tr>
</thead>
</table>

screen

P: ;....;...and it can start------ (part of the drawing) we are on now

screen

The keys Pete presses are a 'key-chord' combination for zooming out from the present drawing (‘ß4’ or ‘Fit To Window’). The result of this starts to appear on the word 'now' of Pete’s utterance 'we are on now'. Zooming out from the single section results in three very small images of sections being displayed on the screen in minute detail (see Figure 4.11).

Figure 4.11: Sketch of the staircase sections (displayed in line 14, transcript 3)

The precision of these diagrams means that the computer requires some time to display them, and this is apparent once the image starts to appear. As they are being drawn Richard turns away from the screen, back towards the paper and there is a pause in the talk for 2.3 seconds (line 15, transcript 3). When Pete starts to speak again with 'righ(t) (here we are)', Richard not only turns towards the screen, but moves toward the computer. It is as if his movement is oriented to the particular display on the screen, the minute images of the sections requiring him to move closer to the screen. Indeed, Richard goes on to move behind Pete to get even nearer to the screen. However, whilst he is moving and Pete is uttering
‘through’ in ‘you’ve done a section through’ (line 18) Pete reorients back to the plans, and moves his left hand to the paper.

(7) 8A JVC 3094 Transcript 6

P moves left hand to paper plan

screen paper

P points to left edge of plan

screen

As Pete is beginning ‘this guy here (. ) right?’, Richard is moving behind him towards the screen, which displays the minute drawings of the sections. Thus, Pete and Richard are looking at distinct domains: Richard towards the computer and Pete at the plans.

The utterance ‘this guy here’ appears to be produced in relation to Pete’s hand movement, to point, for example, at a particular area on the paper plans. However, Richard is neither oriented to nor looking at the plans. Pete briefly pauses and utters ‘right?’ with rising intonation. Together these appear to arrest the movement of Richard towards the screen, engender a reorientation towards the plan and encourage a response from Richard, a brief ‘yep’ (line 22). Yet just as he does this, Pete begins another reorientation. Thus, Pete and Richard are only momentarily oriented towards the same general domain, the plans.
Hence, from being both oriented on the screen, Pete and Richard turn back to the plans. Only momentarily are they oriented to a common domain, before Pete turns back to the computer. Nevertheless, Pete's orientation to the plan, his point and his 'this guy here (0.1) right?' do manage to secure a reorientation by Richard and a 'yep'. Pete's utterance does appear to be organised with respect to the conduct of Richard. As it commences Richard is still moving towards the computer. This movement halts after Pete's point begins and whilst Pete is saying 'this guy'. However, Richard remains oriented to Pete's right. Only after Pete's hand movement is completed, with Pete's arm still outstretched, and after he has briefly paused as he is asking 'right?', does Richard begin to glance back towards the plans. Pete pauses for (0.7) seconds, before Richard answers. Immediately on doing so, Pete starts talking again and Richard begins to turn away.

As when turning towards the screen (in line 10), reorienting to the paper is accomplished through the moment-to-moment activities of the participants. The pointing can only be made sense of with respect to the talk and the artefacts in the domain. However, although Pete appears to organise his visual conduct and talk with respect to securing the orientation of Richard to a particular locale, he only appears to require
some minimal contribution from Richard. Once Richard is glancing at the plans and utters 'yep', Pete turns back to the screen. This utterance displays an alignment to the emerging description of the various locations on the plan, providing Pete with adequate resources to continue.

The activities in this fragment are complex because they require, at various times, the participants to shift orientation between the paper and the screen. Pete uses these two artefacts to locate details of the stairs on the particular drawings on which he is working with respect to details on the marked up plans. More precisely, he utilises the two domains for an account of the different directions in which the stairs are going in the two versions. Furthermore, the system displays sections, cuts through the buildings; the paper shows plans, viewed down from above. The participants have to resolve these two perspectives - part of the common skills required of an architect - and make sense of the various changes being made and proposed alterations to be undertaken on the building. The paper plans provide an indication, though in a rather confusing manner, of these changes, the computer displays the current version, as revised by Pete. Making sense of the work that has been performed on the drawings and also what is required to be done, necessitates seeing features of both the computer and the paper versions in relation to one another. Richard's utterance 'oh yeah that's the same one' explicitly displays an understanding of the work Pete has undertaken in bringing out features on his computer system in relation to features on the paper plan. However, this has been achieved through a series of activities by the two participants, oriented at times with the computer system and at others with the paper plans.

This work does not merely involve a matter of seeing the two domains in relation to one another. It requires Pete to get Richard to see particular features as relevant to him in the light of the current circumstances. Richard has to see this particular staircase, one of many, where the stairs lead from and where they lead to, and make sense of the various 'mark-ups' in relation to these. This activity is further complicated by the nature of the drawings that Pete needs to display on his screen. These concern the changes he has been undertaking and cannot be adequately shown on a single display. He has to first zoom out from one level to a general view, before zooming in to another level (cf. fragment 5). He thus has to coordinate his orientations to the various artefacts with the changing display of the computer system. Retrospectively, the shifting between
computer and paper and back to the computer can be seen as reflecting
juxtapositions between the first section on the screen and the plan, and
the plan and the second section. However, this work is achieved through
the visual conduct and talk of Pete, from moment-to-moment, in the light
of the contributions produced by Richard. It is also produced in the light
of the operation of the computer – its timing. Particular utterances and
even words, like Pete’s ‘right?’ (line 19), are produced with respect to his
orientation towards the paper, for Richard to see ‘this guy here’, now. In
this case, Pete’s utterance is produced, and even transformed, in the light
of Richard’s current orientation, towards another domain and eventually
secures Richard’s reorientation and participation with respect to this
domain. In this way the activities, the reading from the screen and the
paper, are interactionally managed and produced.

The viewing of objects on the screen of the computer can thus be seen
to be immersed within a range of ongoing activities, including those in
relation to other artefacts and those accomplished in collaboration with
others. However, it is not merely that a viewing of the screen is shared
and collaborative, but that to achieve this viewing, to arrive at a common
orientation, is itself a collaborative accomplishment. Similarly, a
reorientation away from a computer is jointly achieved. These
accomplishments are delicately managed through the talk and visual
conduct of both participants. Indeed, in the case explored in detail here
the work of managing a common orientation appears to strongly contrast
with the activities accomplished once the orientation has been achieved.

4.7.3 Readings of the Screen in Interaction

The computer display is made sense of, for Richard, through the talk of
Pete. Pete’s ‘there this is the outside level’ (line 10) makes relevant a
particular component on the screen to be seen in a particular way. The
talk is also made sense of in relation to what is displayed on the screen, for
example, its inherent indexicality (reflected at least in the components
‘there’ and ‘this’). The talk and the items on the display are mutual
resources for participants to render intelligible the emerging activities.
The visual conduct of the participants forms a further piece in this
complex interweaving. The orientation towards a screen provides for a
domain to be seen as relevant and a point can locate a particular area, or
location, within that domain. Thus, the talk, the visual conduct and the
artefact each mutually complement each other as resources for a co-
participant to make sense of an ongoing activity. However, these complementary resources do not stand in isolation from the co-participant, they are themselves designed with respect to the conduct of the other. So, in (7) Pete’s hand movement along the floor level displayed on the screen is only produced once he has secured the orientation of Richard to the computer (lines 10-14 transcript 3). Indeed, Pete's movements, turns-at-talk and even his commands into the computer are designed with respect to the activities of Richard: i.e. Richard's reorientations and his responses. In this way, the production of activities, talk, visual conduct, and those related to the computer, both the typing and the reading of the screen, are interactional. They are designed with respect to the contributions of colleagues and also rendered intelligible by others.

When Pete turns back to the computer system he selects the third image on the screen using the mouse; releasing the mouse button in the pause after ‘I had’ in ‘that's the start of the section I had (0.3)’ (line 24). As the next section is being displayed on the screen, Pete is uttering 'and he (has to start) back there (.) and go up' (line 25). Richard is looking behind Pete, to the back of the room and only turns towards the screen on the word 'there'.

(7) 8A JVC 3094 Transcript 7

head section displayed P nods cursor moves
nods on screen head upward up stairs

P: myself ( them levels) and he (has to start) back there-and go up ↑
R: oh yeah↓

R looking to the back of the office

Pete's talk 'and he (has to start) back there' is produced just after the image of the second section has been displayed on the computer screen. As he utters 'there', Pete nods his head upward. It appears to be on this movement that Richard reorients from the back of the office towards the computer screen. When Richard is looking at the screen, Pete moves the mouse up to the right. This movement causes the cursor to move diagonally up the staircase and coincides with the words 'and go up ↑' (see figure 4.12).
Figure 4.12: Sketch of second staircase section (displayed in lines 24-5, fragment 7) as P moves his cursor up the screen and utters 'and go up↑'.

The timing and organisation of this cursor movement is designed for Richard. Only once he is looking at the screen does Pete produce 'and go up↑' accompanied by the gesture.

Once again, the realignment of Richard appears to be in the light of the talk and visual conduct of Pete; particularly his 'and he (has to start) back there' and his head nod. Only on 'there' - an indexical - does Richard begin to turn back towards the screen. The movement of the cursor and the talk echo each other, not only does the cursor move upwards but 'and go up↑' is uttered with rising intonation. The talk and the activities on the screen are interrelated.

Pete's 'and go up↑' is designed to locate a particular feature on a section, a location that is common to both Richard and Pete's drawings and is a place from which to begin to understand the mark-ups and the various drawings available. This is located by securing a common alignment to particular objects on the drawings on the computer and on the screen and managing the shifts of orientation between them. The reading of the screen is accomplished through a collaboration by the co-participants, in the light of these various artefacts.

On completion of Pete's 'and go up↑', immediately Richard utters 'oh yeah↓ thats the same one'. Only at this point does Richard display more than a minimal understanding of the ongoing actions of Pete. This understanding is an accomplishment by the two participants, produced and displayed through their contributions. The simple movement of a cursor up the screen is one component in this accomplishment. It can only be made sense of in the light of prior interaction and activities of the participants, the work taking place around the screen and the paper.
The activities associated with making sense of an item on the screen appear to require an inordinate amount of work by the participants. However, the activity is itself complex, requiring: objects in the plans to be seen in relation to objects on sections on the screen; mark-ups and old versions of drawings to be seen in relation to modified versions; and one area of a building to be seen in relation to another. It then is perhaps not surprising that some displayed recognition of these differing components is produced as they are brought into play in the interaction. The course of the activity proceeds sequentially; the production of an action such as a re-orientation displaying an understanding of a prior turn of talk or a gesture, and also providing the resources for a next activity. The computer system figures in this emerging course of activities. It provides for the possibility of displaying the current state of the drawings. However, this may not be straightforward, and may require careful management.

It may be that the original removing of layers (line 9, transcript 1) can be seen in this light. Certainly, from where Richard originally stands, on a small screen, presenting too many details would make seeing what is relevant problematic. Pete's turn toward the plans, occurring whilst the small zoomed out images of the sections are being displayed, may be accomplished with respect to what appears on the screen. Displaying these drawings takes some time. It may also be that these sections are not directly relevant to the problem at hand. The time for redisplaying can then provide an opportunity to orient to the plans. Similarly, when Pete selects one of the sections, this also takes time to display. Richard's delayed orientation to the back of the room could be seen in this light.\textsuperscript{18} The screen is a small and changing domain for providing resources for topics of interaction. The work that is apparent in drawing another into look at it appears to require participants not only to display that what is available is relevant, but when it is available and relevant. The use of the system is not only shaped by the interaction between the participants, but also figures in shaping it.

The simple movement from one section to another can thus be seen in the light of the collaborative work of the two participants. Though one of them does all actions on the keyboard and with the mouse, the

\textsuperscript{18} In Chapter 7, some design implications relating to these observations are considered.
contributions of the other are relevant to the production of the activity. In different ways, both could be considered 'users' of the system, with different footings towards the system. Focusing on the individual user would neglect the ways in which the activities are produced and how they are understood by the participants. Nor could these activities or viewings be seen merely as 'shared activities', they are produced from moment-to-moment and are continuously in flux. Hence, a simple activity in relation to the computer system does not appear to fall within either the current conceptions within HCI or CSCW (cf. Heath, et al., 1993). Both neglect how the activity is emergent, situated and produced through the interactions of the participants. Activities on the system are coordinated with respect to a complex of contributions by others, contributions which may be a point or gesture, a glance towards a screen or a turn of talk. This detailed analysis of one instance of naturally-occurring computer use suggests some of the ways in which the use of the artefact is situated within a socio-interactional environment.

4.8 THE SITUATED AND THE COGNITIVE IN HCI: THE INTERACTIONAL CONTEXT OF SCREEN-BASED ACTIVITIES

Conventional approaches to HCI, particularly those utilising a framework from Cognitive Science, tend to be unconcerned with the particular environments in which computer-based activities naturally occur. Either when conducting an experiment or when modelling an activity, the focus in HCI remains largely on the individual undertaking a circumscribed, often pre-defined activity. Even when more naturalistic examples are explored, as in Kitajima and Polson's LICAI model of the background knowledge required to use textual instructions in manuals, the hypothesised activities of users are tightly circumscribed (Kitajima and Polson, 1992; Kitajima and Polson, 1996). Removing an object from the screen is characterised in terms of the knowledge required to comprehend the words in the instructions, to identify graphical objects, and to select appropriate action schemata. These activities are divorced from the particular circumstances in which users may be engaged. Instead, generic tasks such as 'searching for an item in a menu' or 'removing a title' are utilised as a starting point for the modelling process. From there, hierarchies of objects can be constructed, goals can be given in terms of

---

19 LInked model of Comprehension-based Action planning and Instruction taking.
actions like 'move' or 'edit' (operating on these objects and on the devices), and 'if-then' rules defined specifying under what conditions actions can be performed. A complex architecture of objects, rules, goals and schemata can be built up specifying what knowledge is required to understand a set of instructions or to perform some circumscribed activity. All this architecture resting on a foundation that some cognitive processing, comprehension or planning is antecedent to some performance taking place (cf. Ryle, 1949).

Restricting the focus on general, pre-defined, individualistic tasks no doubt reduces the complexity of the models generated. To take account of all the various objects appearing on a screen, and the different ways in which they could be defined, specified and grouped would be a considerable exercise. Therefore, it is perhaps not surprising that the activities which are modelled tend to be tightly constrained; focusing on particular objects, often in one window or screen, and operating on hypothesised examples. The accomplishment of everyday activities, produced from moment-to-moment, even mundane ones, such as those given above, continues to be neglected by most researchers within HCI. The use of objects on the screen to achieve the users' practical purposes is thus neglected. Within the current conceptual cognitive framework of HCI, it is difficult to characterise the improvisatory nature of naturalistic screen-based activities, as this framework requires plans, goals and pre-defined knowledge. It is also problematic for these models to take account of the practical skills of users, their common-sense reasoning. Even though there are proposals to include categories of knowledge, such as those involving the knowledge of particular tasks, heuristics, everyday semantics (Young, et al., 1990) or 'general' knowledge, as in the model of Kitajima and Polson (Kitajima and Polson, 1992), how this is to be modelled remains to be specified. Indeed, 'general knowledge' is often assumed to be a set of propositions of the kind that 'graphs have titles' and 'graphs have horizontal axes'. Similarly, attempts at defining background knowledge in terms of rules and schemata make particular assumptions about the nature of that knowledge. As in the model of Kitajima and Polson, this knowledge is often defined in terms of the relationships

---

20 See also examples from within Artificial Intelligence, where there have been several attempts at defining general knowledge in encyclopaedic fashion in terms of propositions. See, particularly Lenat, et al., (1986) and the relevant critiques by Dreyfus (1992 [1972]).
between particular objects, such as screen objects and their components. The models incorporate a static conception of meaning where comprehension, planning and other intellectual activity is accomplished prior to any action being performed. To model situated aspects of behaviour within such a framework requires a stable and fixed conception of knowledge, tasks and activities, at least whilst they are considered a resource for reasoning and inference and performing the consequent action.

Recent debates have suggested that Cognitive Science can take account of 'situated action' by extending particular models so that they attend to the temporal and dynamic features of behaviour (e.g. Vera and Simon, 1993a). However, the problems with these models may then be not so much their fixed or static nature, but in their commitment to what Ryle termed the 'Intellectualist Legend' where:

...the agent must first go through the internal process of avowing to himself certain propositions about what is to be done ('maxims', 'imperatives' or 'regulative propositions' as they are sometimes called); only then can he execute his performance with those dictates.

(Ryle, 1949, p.30)

In order to maintain such a conception it is necessary to conceive of the 'knowledge how' to perform tasks as 'knowledge that' various propositions hold (cf. Ryle 1949). Thus, emergent practices and situated activities are transformed into actions which are the product of the processing of stable meanings, propositions, rules, and so on. Such transformations lose the very nature of the activity they are attempting to account for. Rather than it being possible to adapt and extend Cognitive Science models to account for situated action, the conceptual framework which underpins them is incommensurate with a conception of situated action where meaning and action are conflated and congruent.

Maintaining a static and fixed conception of meaning allows for researchers in HCI to continue to utilise explicit representations and formulations within accounts of human computer-based activities. However, such an approach will be necessarily restrictive. In providing an account of an activity it will always be possible to provide further descriptions of activities with different elaborations at various levels of details (cf. Garfinkel's discussion of the etcetera clause and etcetera thinking, Garfinkel, 1967, p.74). Therefore, it is not surprising that current approaches to HCI are limited in their accounts of computer-based activities to models of restricted sets of screen-based objects and actions.
Models of human-computer interaction are constrained in other ways; they tend to focus on an individual user only engaged in the activity on the computer system, often isolated from others and unconcerned with other demands. These constraints pervade both theoretical models and the empirical experiments.

The limitations seem also to pervade more novel approaches to HCI. For example, following Suchman's (1987) critique of the conceptions drawn on by Cognitive Scientists to analyse human-computer interaction much debate ensued centring on the nature of situated activity. For some, Suchman's critique has been viewed principally as an attack on the static nature of models that rely on conceptions such as plans, goals and tasks (Vera and Simon, 1993a). The solution then is to develop dynamic models of a human's activity on a computer system. Such an account would then be able to cope with more exploratory, and improvisational, uses of a system. Others have focused on the critique of planning and have engaged in efforts to eliminate that aspect from their models (Agre, 1988; Chapman, 1991). A third consequence of this critique has been a focus on naturalistic activities. Although, this has led to an increased emphasis on workplace studies within system design, particularly within CSCW, with this interest some of the original motivations behind the critique appear to have been lost. In particular, the conception of `situated action' has become identified by some authors with merely being situated in a particular context (Nardi, 1996a, Chapter 4). This has led to the assumption that the consequences of the subsequent programme of research are only relevant to the specific domains of study (Nardi, 1996a).

---

21 Agre and Chapman offer a more radical approach to modelling situated action. Rather than attempting to model dynamic and temporal behaviour in terms of fixed rules and propositions, where state information has to be maintained during the inferential process, they utilise a model with no representation of internal state. These models are instantiated in systems which carry out such actions as playing games with simulated agents. Although the computer systems are sophisticated, the characterisation of action appears to be rather behaviouristic with stimuli from various objects in the games world producing actions in the agents. Agre and Chapman's models do attempt to avoid the problems inherent in separating meaning from action, however, the actions they model appear to be as constrained and restrictive as those in more conventional cognitive models. Their systems may be dynamic, but they avoid any consideration of the practical reasoning underpinning everyday actions, the resources utilised by participants to make sense and produce actions, and the socio-interactional nature of activities.
The conception of context in which collaborative activities are accomplished has been conceived of in a variety of ways by different workplace studies. Context has involved the organisations, institutions, cultures and even the history in which activities are situated. Broadly conceiving of context has thus allowed for a range of frameworks to be utilised for understanding situated action, including Distributed Cognition (e.g. Rogers, 1992) and Activity Theory (Nardi, 1996c). However, Suchman explicitly draws on an interactional framework in her account of situated action; situated activities shape and are shaped by an emerging interactional context.

If the focus is moved away from the individual user carrying out experimental tasks and onto instances of naturally occurring activities then a conception of activity in terms of actions preceded by mental preparation, where static meanings are processed, appears to be problematic. However, even the focus on the single user of a computer merely manipulating objects on the screen and moving amongst the system's capabilities appears to be unduly constraining. In a setting like an architecture practice, such activities rely on a range of practices, including common-sense resources as well as those which constitute the skills of producing architectural drawings. But these activities are not accomplished alone; they are achieved with regard to the contribution of others. The nature of collaborative activities has been recently the focus of a range of workplace studies in the field of CSCW. These suggest an alternative conception of human conduct which may be appropriate for analysing computer-related activities as well as be relevant for system design. The nature of these studies suggests that their consequences are primarily for novel, cooperative or groupware systems. However, some researchers have suggested that workplace studies may be relevant for the analysis of conduct in relation to more conventional computer systems (Greatbatch, et al., 1993; Heath, et al., 1994-5).

In the foregoing analysis of the computer-based activities of architects it appears that activities characterised as browsing, scanning, manipulating objects and navigating around a system are shaped by the skills and practices of the participants. The architects use resources developed in the course of their work, such as those associated with their profession, concerning the manipulation of plans and the organisation of particular buildings. Utilising the available computer-based tools, including the software and hardware, requires particular skills. Some of
these appear to rely on common-sense resources, such as the utilisation of
space within the screen, the preservation of screen objects, the use of the
organisation of a building to facilitate shifting from one domain to
another, and the pacing of activities with the computer system's operation.
However, what is revealed through these brief observations is how these
skills shape and are shaped by the emerging course of activity.

The detailed analysis of fragment 7, reveals how a simple movement
from one drawing to another is developed within an emergent context: a
context that shapes and is shaped by the interactions of the participants.
By taking account of the ongoing activities in the domain, it is possible to
reveal not only how screen-based activities provide resources through
which activities are rendered intelligible but also how they are designed
with respect to the conduct of others. Computer-based activities are
produced, and made sense of, through a complex interweaving of talk,
visual conduct and activities articulated on the computer system.
Navigation around the interface, can thus be shaped by and through the
interactions of the participants.

Human-computer interaction then, is situated in a socio-interactional
context; a context that is emergent from moment-to-moment. This is a
richer conception than those where context is conceived of in terms of the
spatial, as a set of variables or as a static container (Drew and Heritage,
1992a; Schegloff, 1992a; Schegloff, 1992b). In HCI, even studies which are
in some way naturalistic neglect how screen- and other computer- based
actions are coordinated with other activities, what common-sense or
skilful resources are necessary to make sense of the work involved in using
the system, and how talk, visual conduct and human-computer interaction
are interrelated. Observations of the work, activities and interaction of
architects suggest how further examination of these issues may be
relevant to those concerned with human-computer interaction,
particularly those involved in recent attempts at focusing on the
contextual, the situated, or the in situ use of computer systems.

4.9 SUMMARY

In this chapter some observations of human-computer interaction in a real
world, organisational setting have been made. Even a cursory glance of
the ordinary use of such systems within everyday work begins to reveal
characteristics of computer use which cannot be incorporated comfortably
into more conventional descriptions of the interaction between human and computer. Attempts to distinguish internal from external knowledge conceptually and empirically unfortunately neglect the innovative and improvisational character of a system's use, and the ordinary competencies upon which individuals rely in undertaking screen-based activities in 'practical situations of choice'. More importantly, these approaches neglect the socio-interactional environment in which activities emerge. Activities on a computer system can be designed for others, and can be accomplished and rendered intelligible through the emerging talk and visual conduct of the participants.

Part of the difficulty in utilising more conventional cognitive models to characterise real-world computer use, derives from their concern with conceptualising the knowledge a user has 'of' a system, rather than 'how' knowledge is used within the situated accomplishment of a range of social actions and activities. Users, and in this particular case, architects, bring a range of resources to bear in producing and rendering intelligible screen-based actions and activities, and more detailed analysis of the procedures, practices and reasoning involved in the accomplishment of technologically-informed architectural work might well provide a more satisfactory characterisation of certain aspects of human-computer interaction. Moreover, by analysing in detail the production of activities on a computer some of the resources through which these are accomplished can be identified. These activities are achieved through a delicate coordination of talk and visual conduct. Through this complex interweaving, participants manage to produce and understand meaningful activities, an understanding which is displayed by their colleagues through their own contributions. This suggests an alternative resource for understanding human-computer interaction through understandings that are accomplished and displayed by the participants' own activities and interactions; activities that are not mental processes, but social and interactional achievements.
Chapter 5

Distributed Collaborative Work
The use of the radio phone system in the Docklands Light Railway Control Room

That's a funny kind of thing, in which each new object becomes the occasion for seeing again what we see anywhere; seeing people's nastinesses or goodneses and all the rest, when they do this initially technical job of talking over the phone. The technical apparatus is, then, being made at home with the rest of our world. And that's a thing that's routinely being done, and it's the source for the failures of technocratic dreams that if only we introduced some fantastic new communication machine the world will be transformed. Where what happens is that the object is made at home in the world that has whatever organisation it already has.

(Sacks, 1992, p.548)

5.1 INTRODUCTION

The study reported over the next two chapters aims to contribute to the developing body of work concerning the study of in situ collaborative activities. These chapters draw on materials gathered in a control room of
Docklands Light Railway (DLR). This is a highly automated setting, where computer support is provided for activities as diverse as controlling the power supply, monitoring passenger alarms, scheduling trains and handling phone calls.\textsuperscript{1} In Chapter 6 the tightly coordinated activities of the controllers sited in the control room are considered. In this chapter activities of a different kind are examined: those which occur on the radio phone system between the controllers and remote personnel.

The analysis will focus on these distributed activities, particularly how they are accomplished by the participants through a range of social practices. Section 5.2 briefly provides some further background to previous analyses of workplace activities, particularly with respect to the concerns of the analysis of geographically-dispersed activities and the design of novel CSCW systems.

In the DLR Control Room, critical to the control and management of the moment-to-moment running of traffic on the transport system, and the handling of crises as they emerge are the conversations on the radio phone system. The remote parties can inform the controllers of important events, such as suspect packages or train failures, and drivers may need to be told of changes to the running order of the traffic. Therefore, the effective accomplishment of these activities is essential for the safe and smooth running of the service. To meet these ends, a computer system has been introduced to assist with the taking and making of calls. This system queues incoming calls and allows the controller to choose which call to take from this list. Analysing the radio phone talk also happens to provide resources for a distinctive way of analysing the use of this system. An analysis of the radio phone talk is the focus of this chapter.

To provide a foundation for the presentation of the analysis, both here and for the following chapter, a brief overview of the activities of individuals in the DLR Control Room is given (Section 5.3). This reveals the critical nature of the radio phone conversations; how they are utilised to manage the day-to-day running of the service and to handle crises and

\textsuperscript{1} The materials on which the analyses in chapters 5 and 6 are based consist of field observations and video-recordings carried out over three days and several shifts in the control room of DLR. Two cameras were used: one focusing on the controllers' desk in general; the other mainly centred on one controller and the phone system. The corpus of materials consists of nearly 40 hours of video-recordings.
emergencies. As background to the analysis of this radio phone talk a brief review of some work which has been undertaken on the opening of telephone conversations is provided (Section 5.4). This gives a foundation for some general observations on the distinctiveness of the radio phone conversations, on the general organisation of the opening of calls, and on some practical problems that contributors to the calls have to face when initiating the talk (Section 5.5).

The resources utilised by the participants are considered in two ways: first, how the initiation of the conversation is managed through interactional devices, available to the participants, both through talk and through other non-vocal means (Section 5.6); second, how these devices may themselves be produced with respect to other features in the call, particularly the pacing of the initial contributions to the calls (Section 5.7).

It appears that participants are very sensitive to the contributions of their colleagues to the radio phone call: the design of particular turns and the intonational contour of the turns provide resources through which the management of the initiation of the call is accomplished. The nature of the calls also appears to introduce some features which need to be considered by the participants (and for analysts), particularly with regard to the nature of a summons to answer (Section 5.6.1) and for eliciting the reason for a call (Section 5.6.2). However, even when there are apparent problems in the call, participants appear to be orienting to the conventional demands placed on them in ordinary social interactions (Section 5.6.3).

The management of the initiation of the calls provides for a consideration of some observable features of the calls with respect to their pacing. Section 5.7 considers the details of the timing of the talk in two domains within the conversation: with respect to the gap just after the first turn of talk (Section 5.7.1) and to the one just prior to it (Section 5.7.2). It appears that features of the talk may be managed with respect to this pacing, the talk unfolding with regard to such details in the interaction. When examined at this finer level, it appears that the general distribution of responsibilities in the calls is very carefully managed. This is a critical and ongoing concern of the participants. The talk is not organised to some pre-specified pattern, but rather the various contributions are shaped with respect to the prior activities of a co-
participant. Moreover, even the gaps and pauses in the talk may reveal a sensitivity to the conduct of a remote colleague.

The pacing of the talk appears not only to be resource for the accomplishment of activities through the radio, but also for other activities in the control room. Pauses in the talk can provide opportunities for undertaking prospective activities with respect to the call, for example, examining the various computer displays available to the controllers. This can provide resources for making sense of the circumstances of the caller, hence the readings of the screens can be seen in the light of the emerging phone call, even prior to any talk from a co-participant (Section 5.8).

Therefore, in this chapter, various kinds of resources utilised by participants in the course of their activities are considered. These include the design of turns of talk, the pacing of talk and the information available on computer screens. These resources are critical for the accomplishment of the controllers' management of activities through the computer system and, hence to, the transportation network. They are, however, resources that are made use of through interactions, through turns of talk. Hence, they rest on common-sense practices and practical reasoning, practices which are socially ordered and organised. Section 5.9 concludes the chapter with a brief discussion of the nature of the analytic resources utilised here with respect to some other recent analyses of distributed activities in technological domains.

5.2 BACKGROUND

For researchers in CSCW the potential contribution of Workplace Studies is that they may yield a better understanding of social interaction and the coordination of activities between individuals. For designers who aim to develop novel computing and audio-visual environments for individuals who are geographically dispersed, such an understanding could suggest potential capabilities that could be supported by the technology. So, for example, analyses of activities in complex environments could inform the ways cameras, videos and other technologies could be reorganised and reconfigured within 'media spaces'. With this reorganisation, the audio-visual environments could better overcome the problems of existing arrangements and provide more flexible ways in which individuals could accomplish collaborative activities (Heath, et al., 1995b). Similarly,
detailed studies of interactions and the use of artefacts could suggest ways in which to develop collaborative virtual environments (CVEs). These provide an electronically presented setting in which representations of individuals can interact with each other and with certain objects available in the virtual environment. Studies of social interaction and activities could suggest, for example, how objects need to be presented in relation to the 'individual' in the system, how movements between different kinds of collaboration could be presented, and some of the requirements for achieving 'pointings' and 'collaborative viewings' in CVEs (Benford, et al., 1996). Hence, studies of activities of individuals who are co-present could inform the development of systems to support activities between individuals who are dispersed.

Although there is an interest from researchers and developers in resources concerning the nature of collaborative work, there is an inadequate corpus of findings to draw upon that could be relevant for design. However, in the last few years some studies have been undertaken that may begin to address the interests of designers (e.g. Benford, et al., 1996; Bowers and Button, 1995; Filippi and Theureau, 1993; Hughes, et al., 1992; Hutchins, 1990). Of particular interest, have been those that utilise video-recordings to reveal the interrelationships between naturally-occurring tasks and interaction (Goodwin and Goodwin, 1996; Greatbatch, et al., 1993; Heath and Luff, 1992a; Whalen, 1995b). These studies are beginning to outline a distinctive approach to the study of activities, for example, how an individual's activities can be considered as designed for others, and with respect to the conduct of others.

The study reported in this chapter seeks to extend this corpus of studies by exploring the use of a technology to support distributed activities: the computerised radio phone system used on the DLR.

The phone system has several interesting features, including a 'call back' facility. In effect, all calls from remote parties are queued. This has the consequence that, no matter who initiated the call and has a 'reason' to call, the controller almost always utters the first turn in the call. This may appear strange given previous studies of telephone conversations which have explored how delicately the initial turns are managed in a call. These studies reveal how through the management of the opening of the call co-participants initiate the talk, give the reason for the call and, hence, set up a distribution of activities between the participants and a
trajectory for the remainder of the call (Schegloff, 1968). Given the nature of calls on the DLR, how they are initiated would appear to be critical to the way they unfold, specifically concerning the reason for the call.

How participants manage and accomplish such actions in their talk through this novel communication and computer technology would provide a distinctive approach from those considered in Chapter 2 which examine similar activities. In CPM-GOMS, for example, talk on the phone, in the domain of call operators, is considered too varied to be a topic of analysis (Gray, et al., 1993). In the analysis of Hutchins, talk, either in the co-present setting or mediated by technology, is considered in terms of communicative acts and functions.

Through their contributions to the talk on the phone the participants accomplish activities regarding the handling of the traffic on the line. These activities are achieved through an interaction where one activity both displays an understanding of a prior turn, and also is designed for a co-participant. The participants' activities, therefore, provide an analysis of their colleagues' contributions. This analysis can also be a resource for revealing how the calls are organised. As the use of the computer-supported communication system is displayed through talk, these contributions can also be a resource for analysing the participants' activities through the system. The talk provides what is publicly available to the participants in the call. In this way, turns of talk can be a resource for analysing the naturally-occurring and necessarily distributed activities that take place in the setting.

This may be interesting if the technology, both the radio and the computer system, requires an organisation to the talk not found in similar conversations, like telephone talk. Though there may be differences introduced by the technology, a consideration of the in situ organisation of the openings of telephone talk may shed some light on the organisation of both radio talk and the operation of the computer-supported radio phone system. In particular, given from previous studies of telephone calls that the resources utilised in their commencement are particularly critical to the way in which they develop, it might be expected that transforming these critical resources may require a different nature of activities in the initiation of a call. This could have consequences for how the management of what gets done by whom in the call is accomplished.
The detailed analysis of a different domain of social interaction, a radio conversation, may further be useful for reflecting back on the resources utilised by participants in the openings in conversations and telephone calls.

As with the CAD system examined in the previous chapter, the controller's operation of the radio phone system is not isolated from activities with other artefacts or with other colleagues. Hence the organisation of the calls need also be considered with respect to the activities of others in Control Room. This is the focus of Chapter 6, but in the present chapter the operation of the phone system is also considered in the light of the publicly available resources. By considering the interrelationship between the talk on the phone and the use of the system, the resources through which participants make sense of the call, and the ways the call can suggest ways of 'reading' what is presented through the technology, can be explored. Through a complex web of contributions, interactions between individuals in the control room and personnel in different locations, with varying responsibilities and concerns, can be managed.

5.3 THE GENERAL ACTIVITIES OF CONTROLLERS ON THE DLR

The taking and making of phone calls in the DLR is the responsibility of one of the Control Room Supervisors (CRSs), or 'controllers', who sits at a large console which takes up most of the space in the control room. This controller (here called Ci) usually sits on the leftmost seat of the console, nearest to the computer-supported phone system. Next to him sits another controller (Cii), who tends to be responsible for traffic movements in the shunting yard and discussions with the personnel who sit at the far right of the console, the Control Room Assistants (or CRAs). They manage the passenger announcements and the distribution of mobile radios to those personnel who are responsible for each of the trains, called Train Captains (or TCs). There are usually two other personnel in the control room, a Control Room Manager (CRM) with strategic responsibilities for the line and control room; and an Electronic Works Coordinator (EWC) who is concerned with the technical operation of the

*Appendix A gives further background to the DLR, including a glossary of terms used in the radio conversations (Section A.5).*
line. Figure 5.1 shows the general location of these personnel within the control room.

![Diagram of DLR Control Room Staff](image)

Figure 5.1: The DLR Control Room Staff

The principal concern of the controllers is to maintain a smooth and safe service for customers. This requires them to monitor and alter the running of the line and to establish and maintain contact with personnel out on the system. To assist them with these responsibilities the controllers have two computer systems: an automatic train scheduler (or ATS) and a phone system (PS). The various devices that make up the hardware of these systems are located in the middle of the console at which the controllers sit (Figure 5.2 gives the general organisation of the technology available to the controllers). Both controllers have a keyboard which can operate the ATS, with a display to it in front of each of them. Because the phone system is positioned to the left of the leftmost controller (Ci) it is this controller who tends to take and make calls.
At peak times, the calls the controllers have to deal with can nearly be continuous, with Ci dealing with one call after another. These calls may concern the operation of the service, such as in (1) when a train captain has to be informed of a change in the destination of a train.

(1) 43a (PT) 2134B 2174A - transcript 1 (simplified)
Ci: five two?
(0.5)
Ci: five two (receiving)?
(0.7)
TC: yeah five two
(1.3)
Ci: err: Gordon we’re going to run you down to Island Gardens from there you’ll be a Stratford service over.

or concern traffic movements in the shunting yard, or Operations and Maintenance Centre (OMC), as in (2):

(2) 45 (T) 2345B 2380A (xi) - transcript 1 (simplified)
Ci: (ee) five three:?
(6)
E: five three to base? (.) I’ve got vehicle twelve being moved forward then waiting your instructions ( ) over
(0.3)
Ci: yeah that’s roger on that (Re: g) follow (green Ess specs) from err (0.2) Dee You run (.) (and we’ll put you on (machine road) over.

As well as these routine occurrences, controllers also have to deal with a wide range of problems and emergencies that arise during the course of the day. These include not only the well-known difficulties of operating public transport systems in a large city, such as bomb scares and suspect packages...
and requesting police officers to deal with troublesome passengers refusing to pay fares ('revenue disputes')...

but also problems with the lifts,...

lost property...

TC: ( yeah ) four eight to base (0.2) I'm enquiring about a (ladies) set of keys (. ) (you haven't had any one hand them in) over? (0.8)

Ci: yeah (I have a) roger on that I've made enquiries ( with our ) Cee Ar A::s and they've no:: keys (have been) handed in (. ) () (1.3)

TC: (oh very good) over.

(3) 14Co - transcript 1 (simplified)
Ci: three two go ahead↑
(3.9)
TC: (yeah this is ) three two to base (0.2) umm: (. ) (someone has) umm (0.2) (a black) briefcase (on trackside) (1.2) coming out of (. ) All Saints (two) (0.2) (on the way to) Devons Road-about two hundred yards down↓ (between) on trackside over? (2.5)

(4) 8iii 3570B - transcript 1 (simplified)
Ci: seven three?
(2.3)
TC: seven three errrm: (.) (will you be able to get me a Bee Tee Pee at the depot.

(5) 10Co - transcript 1 (simplified)
Ci: six two:oo↑?
(1.4)
TC: ( can you give me ) can you give me errr (0.4) the information (regarding the lifts err) at West India Quay ( ) one on the southbo:und↓ one at>one at Stratford (one on westward) over?
(1.8)
Ci: yeah I can err (0.2) I can tell you that one (0.2) lift at West India Quay is out of order but which one I don't kno::w↓>its lift one (0.8) errrrrm: >and the lift at Stratford has err not been reported as defective so I presume it is worκing over?

(6) 52 (T) 3635B 3662A (xii) - transcript 1 (simplified)
Ci: four:: eight?
TC: (four eight ) I have a ( lady on board) (saying she's lost a) set of keys ( ) over. (0.2)

Ci: any keys handed into lost property? (2.3) ((Cii shakes head))

Ci: four eight? (2.3)

TC: ( yeah ) four eight to base (0.2) I'm enquiring about a (ladies) set of keys (. ) (you haven't had any one hand them in) over? (0.8)

Ci: yeah (I have a) roger on that I've made enquiries ( with our ) Cee Ar A::s and they've no:: keys (have been) handed in (. ) () (1.3)

TC: (oh very good) over.
and a wide variety of other difficulties with the power supply, phones, radios, public announcements, the relief of train captains etc. They also have to keep management informed of significant events...

(7) 22Co - transcript 1 (simplified)

Ci: four two: T
(0.5)
CRM: (Peter Cribb.)
(0.6)
Ci: yes is that Peter over?
(1.2)
CRM: receiving↑
(.)
Ci: yeah errrm↓ (0.2) just to put you in the picture↑ we have just had a report of a suspicious item↓ between All Saints and Devons Road stations.
(0.2)

and maintain a record of less significant failures and delays (F&Ds) with the system. Hence, there are a large number of apparently distinct activities that controllers have to cope with, many of which can arise at any time in the course of the day. The work of informing individuals of changes to the service, of dealing with incidents and even of generally maintaining the state of the service largely happens over the phone. Matters are raised in the calls which are then managed through the interaction with the remote party.

Although controllers have to operate the ATS and deal with the concerns of CRAs, CRMs and other staff in the control room, the critical aspects of their work is accomplished through the talk on the phone. As will be seen in Chapter 6, even the operation of the ATS is also tied to this talk. Through the conversations on the phone controllers do not merely communicate or convey information, they carry out actions: transforming the service and managing crises. The accomplishment of their work is accomplished through these interactions. Therefore, in order to explore the nature of work and interaction in the control room it is important to examine, in more detail, how these phone conversations are organised.

It may be noticed that, after the initiation of the call, the participants go on immediately to give the reason for the call: the reporting of a suspect package or lost property, a change in the schedule, or a request for assistance or for information. It would appear to be important in this domain that the 'reason for the call' is produced promptly so that it can be acknowledged or dealt with there and then, and then both participants
can move on either to related activities or something else. Indeed, in most cases, the reason for the call can be considered to be the sole 'topic' or issue addressed in the conversation. If it is a request then it is replied to in the next turn, if some news is given then this is acknowledged in the next turn, and if it is some statement of a problem by the remote party and a potential solution is given by the controller in the next turn then the remote party will typically repeat this solution as an acknowledgement. The parties will then briskly move into some sort of closing with typically 'that's received', or 'thanks for that' or just a 'thanks' from each participant. Thus, in terms of the transition of turns, radio phone conversations on the DLR are typically short and rest on the initial delivery of the reason for the call. This will be a focus of the ensuing analysis.

As has been noted, in the DLR Control Room the phone calls are mediated by a computer system, a technology that allows calls to be queued and taken in any order. Any analysis of the organisation of the radio phone conversations is also an analysis of the 'use' of this computer system. As conversations, and, in particular telephone conversations, have been a focus of study for many years by conversation analysts, particularly conversational openings and initiations, these studies may be a useful resource for beginning to understand how this technology is being used by participants. The mediation of a computer system in the interaction may have an impact on how the talk is organised in this domain. Analysts of telephone talk (e.g. Schegloff, 1968) have both identified how talk on the telephone is organised with respect to the

---

3 It may be that the consideration of an object like a 'reason for a call' may be itself be a problematic analytic category. This will not be dealt with at any length in the ensuing discussion. Suffice it to say that the participants in the setting manage to analyse such components for what is relevant and what has to be dealt with quite unproblematically, they acknowledge what is provided, sometimes repeat it and in other ways deal with what is given them. There may be difficulties with hearing the 'reason for the call' or even with associating the 'reason for the call' with the responsibilities of the recipient, however, this component does invariably seem to be treated in this way, and not, say, as some preamble to another 'reason'. The production and recognition of this is a skilful and collaborative achievement by the parties involved. See Sacks' (1968 [1992]) discussion of the differences between announced reasons for calls and findable reasons for calls, and the 'interesting' cases for people to make calls 'without reasons for a call'.
particular features of the technology, and also how the organisation still relies on more generic moral and social commitments by the parties. A consideration of talk in this more familiar domain may contribute to the analysis of radio talk. It may also be interesting to consider to what extent the resources utilised by participants in telephone conversations can be relied on in radio phone conversations.4

5.4 THE ORGANISATION OF TELEPHONE CALLS

The organisation of the telephone call has been of particular interest to researchers involved in the analysis of conversation. Telephone calls provide an accessible resource of naturalistic materials which can be recorded cheaply. Moreover, although participants in a call have constrained access to each other (only through audio), this access is not only symmetrically available to the participants but can also be similarly available to analysts. Thus, it can be seen as a resource for studying naturally occurring social actions and also for revealing features of other forms of talk. For example, Schegloff (1968) analysed a substantial corpus of conversational openings on telephone calls, which provided a resource for a more generic analysis of conversational resources.

Schegloff notes one particular consistency in the corpus of data he analysed, that answerers talk first. This is despite caller and called having differential 'informational resources'.5 The caller, for example, knows who is his or her intended interlocutor, the called can only speculate. Indeed, although Schegloff has a corpus of roughly 500 examples, only in one does the caller speak first (reproduced as 8).

(8) #9 (Police makes call) (from Schegloff, 1972, p.356)

((Receiver is lifted))
(1.0)
→ POLICE: Hello.
OTHER: American Red Cross.
POLICE: Hello, this is Police Headquarters...

He utilises this 'deviant case' to reveal a more generic aspect of conversations, that summons demands answers. In a phone conversation

4 For other analyses of radio conversations, see the analysis of intonational contours in radio talk (Goodwin, 1996), and the accomplishment of multi-party conversations through voice loops in a complex interactional setting (Watts, et al., 1996).

5 See also comments by Sacks in his lecture of Fall 1967, (Sacks, 1992, p. 632).
the ringing of the phone serves to summon the interlocutor, much as other 'attention-getting devices' such as terms of address (e.g. 'John'), courtesy phrases (e.g. 'pardon me') and physical devices (e.g. a wave of a hand), serve to summons potential co-participants in other interactional domains.

So, in (8) though the caller, the police dispatcher, answers first, rather than the called, the talk can be seen with respect to prior (non-) activities in the call, the one second delay. Thus, the dispatcher can be seen to be repeating the summons of the phone ringing, following the delay after the called has lifted the phone. Once this second summons 'Hello' is answered, the caller then gives the reason for the call. The 'deviant case' can thus be seen to reproduce the generic case, the repair revealing that participants remain oriented to summons demanding answers. Moreover, by considering the case with respect to phone calls, the 'distributional rule' of 'answerer speaks first' can be considered more generally, in terms of the obligations placed on participants within interactions. The initial summons and its answer can be seen to provide for coordinated entry into conversation and for its continued orderliness. So, for example, Schegloff summarises some observations concerning the first two turns as follows.

A summons item; obligates other to answer under penalty of being found absent. insane, insolent, condescending etc. Moreover, by virtue of orientation to properties of answer items, i.e. their character as questions, provides for users' future obligation to answer, and thereby to have another turn to talk. Thus, preliminary or prefatory character, establishing and ensuring availability of other to interact.

Answer summons, thereby establishing availability to interact further. Ensures there will be further interaction by employing a question i.e., which demands further talk or activity by summoner.

(Schegloff, 1968, p. 376)

Thus, through the coordination of the opening of the conversation the organisation of the subsequent talk can emerge; an answer to a summons allows the summoner the right to talk again. A question, such as 'what?' as an answer to a summons provides further resources for the other, the summoner, to talk again.

The suggestion that the ring of a telephone provides comparable resources for a co-interactant as a summons may shed some light onto the organisation of calls in the DLR Control Room. For, even if the technology may transform the resources which can be utilised, it would be interesting to explore whether the more in-depth analysis suggested by Schegloff of
the obligations demanded by a summons permeates this other interactional domain.

5.5 TAKING AND MAKING CALLS ON THE DOCKLANDS LIGHT RAILWAY

At first hearing the calls in the DLR Control Room appear to be highly formalised and repetitive. They frequently involve the repetition of key components like instructions, the completion of turns with the word 'over' and the identification of participants by numbers. The initiation of the phone call appears to be particularly curious involving numerous rings and beeps, lengthy pauses and the talk commencing with the controller uttering the phone number of the remote party. The particular design and operation of the phone technology gives a partial account of some of these curiosities.

The DLR uses a radio phone system allowing members of staff to contact, or to be contacted by, the control room anywhere on the system. However, as the phone system is not a typical telephone network, it does have a few peculiarities. For example, the control room can broadcast messages to all members of staff with a radio phone. In its more usual operation, when a remote caller wishes to contact the control room, he or she makes a call which is then queued (or more accurately listed) on the computer system. This queue is displayed on the screen to the left of the principal controller (Ci). The controller can then make calls either by selecting calls from the list (in any order) or call any other phone that is in operation. So, in effect, the controller has to make a 'call' with the system even when 'answering' the call. Hence, whether the call concerns a controller telling a train captain of a change or an incident being reported by a member of staff, the first turn is uttered by the controller. As well as

---

6 It was stated by the controllers that the telecommunication system was a version of one used for communication by bus operators, and this, in some way, accounted for its particularities. One example is the problem of drivers having the ability to hold the line at the end of a call. This may be necessary in emergencies on the buses, but presented potential difficulties in the everyday running of DLR, as controllers can only move onto further calls when the remote participant disconnects the call.

7 For each call the phone system screen displays the time of the call, the name of the caller, the radio number and the train number, if appropriate (in different columns, each call being assigned a row in a table).
the changing display of the phone system, the system emits various rings and beeps at various stages in the call making procedure. So, in the following call a train captain is enquiring about which automatic ticket machines are in operation.

(9) 1 (T) 2B / 3B - transcript 1

1  ((long ring))
2  (1.5)
3  ((beep))
4  (3.5)
5  ((Ci picks up phone - short ring))
6  (3.5)  ((Ci selects call))
7  ((long ring))
8  (0.5)
9  ((long ring))
10  (1.2)
11  Ci: (three eight) go ahead?
12  (1.4)
13  TC: ( ) A:: Tee eM operating over.

The initial ring (line 1), a five tone sound, is the train captain ‘calling in’. The following beep (line 3) comes from the computer-supported phone system and is accompanied by changes in the screen display, also notifying the controller that a call is waiting. Unlike a conventional telephone system the phone does not continue to ring, and some three and a half seconds later, the controller picks up the phone (line 5). This itself is also accompanied by a ring, a short two tone sound. The controller then selects the call from the phone system using the mouse. Two further rings are generated by the system, the first signalling the call being made from the control room (line 7), the second being for the train captain (line 9). The controller follows this by uttering the radio number of the caller. This appears to demand a response, which it gets. The train captain gives the reason for the call 1.4 seconds later.

Outgoing calls are a little simpler.

(10) 3 - transcript 1

1  ((Ci picks up phone, selects call))
2  (7.0)
3  ((long ring))
4  (0.4)
5  ((long ring))
6  (1.2)
7  Ci: (train base to train) six one?
8  (1.3)
9  M: six one?
10  (1.3)
11  Ci: yeah er take a minute (ordering that) all clear if
12  you'd like to standby for giving instructions...
Again, the two sets of rings follow the selection of the call on the computer-supported phone system, and the controller commences talk on the phone with the callers' number. After the caller replies, the controller then gives the reason for the call (lines 11 and 12).

From the foregoing fragments, it is apparent that phone calls in the control room can be distinguished from conventional phone calls by several straightforward features. These include:

- One party only, the controller, begins to talk on the phone.
- The callers are identified by and identify themselves through numbers rather than names.
- There are a variety of phone rings and beeps which have different characteristics to telephone rings (i.e. ringing is not continuous, and hence, does not stop when the called is available to answer).
- A computer system mediates the incoming calls and displays information concerning potential recipients.
- Incoming calls are kept by the system as requests to make contact rather than direct 'summonses' to the control room (i.e. calls need not be dealt with on a first-come-first-served basis).
- In all calls one party is always the same, namely the controller. It is not possible for anyone other than the controller to make a call to other members of staff on the radio phone.
- On incoming calls, the controller can identify who is calling through the computer-supported phone system.

At first glance, the common features revealed by Schegloff for telephone conversations do not appear to hold: answerers do not speak first, the caller does. However, as noted, the first speaker is the controller, whether or not the call has been queued or initiated. Through the phone system technology the categories 'caller' and 'called' become problematic: these could relate to the initial request for a call or the particular call underway. In 'outgoing calls' these would be the same; in 'incoming' ones they would be different. Hence, the conception of a summons would appear to be problematic for incoming calls: this could be related to the original request or the particular call. Moreover, the ringing of the phone is not continuous, and the 'picking up' of a receiver cannot be regarded as an answer to a summons.
However, when the instances above are considered there does not seem to be any undue difficulties in moving through the opening of the interaction. The opening sequence may appear to be extended, say in comparison to those in telephone calls, but, given the differences between when a call has been requested by a remote party and when it is initiated by the controller, there do not appear to be obvious ways of explicitly eliciting the reason for the call, or, perhaps more curiously, of explicitly marking an incoming from and outgoing one.

When the particularities of the technology are considered further, the characteristics of the calls in the DLR appear to be even more curious. If the member of staff and the controller can assume or have information concerning the identity of their co-participant, why does it appear that so much work is involved to establish contact? In all but one of the fragments above, where it can be heard, both controller and the remote party repeat the radio number. As there has just been a phone ring, the number of the remote party is displayed on the screen and, on incoming calls, both parties are ‘known’ to each other, this may be just an example of ‘formal’ talk, the organisation of turns being prescribed beforehand. Even if this so, it would still be interesting to investigate how these formal procedures are utilised in practice, given the ‘information resources’ available to the participants in the radio phone call.

When considering the operation of the technology one particular issue emerges. Given the potential for a delay between a request for a call and the call itself and the operation of system, it would appear to be problematic to distinguish between incoming and outgoing calls. The talk in each has to be initiated by the controller. However, there appears to be no explicit routine device for this. In the cases so far considered in two (fragments 3 and 9) the controller utters ‘go ahead’, but this is not routine. Conversely train numbers seem to routinely be uttered with a rising intonation, but this occurs across all the calls, both incoming and outgoing (e.g. fragments 1 and 10). Nevertheless, it would appear to be a practical problem for the participants not only to elicit the reason for the call, but also to manage that elicitation. It is of course possible that controllers may want to contact a train captain on the queue, or that after some time

8 In the other case (7), a CRM, replies with his name. In all other instances in this chapter, where it can be heard, the radio number is uttered at least twice by the two parties in the opening of the call.
a reason for a call may no longer relevant. It might be expected, given the apparent formality of the call, that some formal device would be specified to mark 'call backs' from 'new calls'.

In the remainder of this chapter, the ways in which this work is achieved will be explored.

Before considering how these problems are managed through the initial activities of the participants in the calls, it may be worth providing a general outline of how the openings are organised (see Figure 5.3a). 9

---

9 Given the operation of the system the distinction between incoming and outgoing calls may be misleading. However, to facilitate the ensuing exposition, calls which have been requested through the phone system by a remote party will be considered as incoming. The reason for the call is produced by the remote party. Where the reason for the call is produced by the controller the call will be considered as outgoing. Hence, the distinction between incoming and outgoing calls can be seen to be emergent rather than assigned a priori through the technology. In most cases when the reason for the call is sequentially relevant, the parties, through their actions, display an understanding as to who should produce the turn. Hence, the call at this point becomes incoming or outgoing. In the problematic cases, as we will see, the difficulties are managed through the sequential organisation of turns of talk. The parties display first an understanding that a reason for a call is relevant and appropriate, and, then eventually produce one.
Incoming Calls

TC: <long ring of controller's phone, change on system display> (= 1.5 seconds )
<beep of system>
(after some period, intervening actions etc.)

C: <takes actions to call remote party> <long ring of controller's phone> (= 0.5 seconds )
<long ring of remote party's phone> <pause>
('train base to') radio number < pause >
TC: (<feedback>) radio number + <reason for call >

Outgoing Calls

C: <takes actions to call remote party> <long ring of controller's phone> (= 0.5 seconds )
<long ring of remote party's phone> <pause>
C: ('train base to') radio number < pause >
TC: (<feedback>) radio number <pause>
C: <reason for call >

Figure 5.3a: A general schematic for the organisation of radio phone calls on the DLR

Apart from the initial ring to request a call, the general outline of the calls is similar: the controller selects the other party on the phone system; the phone rings twice; he then pauses, utters a turn which includes the radio number of the other party who, after a little while and perhaps some feedback, then replies. The reply is different in the two cases. For the incoming calls the other party, typically a train captain, provides the reason for the call, after uttering his or her radio number. In outgoing calls, the reply is minimal, including a radio number. The controller then gives the reason for the call.

5.6 THE INITIATION OF RADIO PHONE CALLS ON THE DOCKLANDS LIGHT RAILWAY

Amongst the many ways that the radio phone in the DLR differs from the telephone is the differential access participants and potential participants have to the call. Most critically, the radio can be a broadcast medium where all remote participants can hear the controller's turns of talk. This occurs when the controller makes an announcement to all 'train captains'. Train captains can also 'switch channel' to listen in to the calls broadcast by the control room. In this mode they only hear the controllers' talk to
other colleagues on the system, but it does provide some access to the 'goings on' on the train system. Thus, only the controller can hear the remote recipient through the radio.

For train captains outgoing calls are in some ways like the telephone, only the controllers can hear that a call is being attempted. However, because of the call-back nature of the system it can be unclear to the train captain whether the subsequent rings are related to the call he or she requested. The rings might foreshadow a return call to someone else or herald an outgoing call. Thus, the rings merely inform potential recipients that a call is being attempted, the particular interlocutor is not apparent until the relevant call sign is uttered. Even when this is uttered, because of the delay between the request for a call and the call, it may not always be clear whether and how this particular call relates to the prior request.

Hence, it would seem that it might not be clear who is initiating the call and, thus, who should provide the reason for the call. The analysis of telephone conversations suggests that the initial turns provide for a trajectory of subsequent actions: a summons demanding an answer, the answer to the summons providing for the caller to give the reason for the call. This organisation does not immediately seem apparent in the radio phone conversations on the DLR. In the next sections, particular activities in the radio phone conversations are considered: the summons (Section 5.6.1) and the production of the reason for the call (Section 5.6.2). As with telephone conversations the production of the reason for the call is sequentially relevant given the prior activities in the call. However, these activities are distinctive to this particular domain, they are designed with respect to the technology and the resources available to the participants. In Section 5.6.3 cases are considered which do not fit with the general

---

10 The use of this is problematic when considering that most of the time train captains are in trains occupied by passengers and are meant to be dealing with such matters as checking tickets and dealing with passenger queries. On the Bakerloo line, where drivers are in cabs the 'broadcast mode' of the radios is critical to the running of the line (Heath and Luff, 1992a).

11 It may be that the participants may orient in some way to the different length of delays before the call is returned. Some such resources are considered later in fragment (24), but without a detailed history of the operation of the phone system a more explicit analysis of such possibilities is not feasible. It could be that different delays project different ways of dealing with the call, or even differentiate the sequential demands of the upcoming call.
organisation given in Figure 5.3a. Nevertheless, even in these, whilst the
distribution of responsibilities for activities in the call are locally
managed, the participants appear to be orienting to the sequential
demands of activities in the talk.

5.6.1 Summonsing the Other

In the following instance, the second controller (Cii) has, sometime earlier,
requested Ci to ‘(just quickly tell number five to go to the Island)’ (not
shown in the transcript).

(11) 39Co

1 Ci: one o: fi:ve to the Island↓
2 (0.3)
3 Cii: yeah (...) please↓
4 (3.3) ((Ci selects call))
5 ((long ring))
6 (0.1)
7 ((long ring))
8 (0.2)
9 Ci: three nine go ahead↑
10 (0.3)
11 TC: three ni:ne (err ) a platform announcement
12 saying (we would be going) to Island Gar:dens↓ can
13 you confirm (that ) over.
14 (0.2)
15 Ci: thats affirmative over that was the purpose of my
16 err call to you (. ) your destination is Island
17 Gardens over↓
18 (1.4)
19 TC: ( destination ) Island Gardens↓ ( return) to ( )

Following this request, Ci confirms this change (line 1) and then selects
a call on the phone system. Ci’s turn ‘three nine go ahead↑’ appears to
request the reason for the call and in the reply the train captain gives a
potential reason for calling: a platform announcement has been made
giving a different destination for his train. However, in the next turn,
although Ci confirms the change mentioned by the train captain, he utters
‘that was the purpose of my err call to you’.12 It appears that, at least at
this point of the call, Ci is orienting to the call being outgoing. The train
captain replies with a confirmation of the change.

12 It may be noted that between the initial request by Ci’s colleague to alter the
route of the train, Ci takes a call from another member of staff. He does not
appear to make any attempt to call another train. Hence, Ci appears to be
referring to the call at hand in lines 15 - 17, and not, say, to an earlier failed
attempt at an outgoing call.
Given the nature of the communication system it may seem irrelevant whether calls are considered incoming or outgoing, as in one sense, all calls are outgoing. However, a prior request for a call from a remote party would appear to require the controller to elicit the reason for the request. In conventional telephone conversations the answer to the summons provides the opportunity for the summoner to speak. In DLR radio-phone conversations the rings of the phone are always initiated by the controller and can be heard by others, therefore it would be difficult to suggest that the rings on the radio-phone suffice to summons a particular individual. Instead, it is the initial turn by the controller, which obliges the interactant to answer. So, in (9) and (10) the controller's first utterance provides the occasion on which the train captain can then speak.

(9)  1 (T) 2B / 3B - transcript 2
7  ((long ring))
8  (0.5)
9  ((long ring))
10  (1.2)
11  Ci:  (three eight) go ahead?
12  (1.4)
13  TC:  () A:: Tee eM operating over.

(10) 3 - transcript 2
3  ((long ring))
4  (0.4)
5  ((long ring))
6  (1.2)
7  Ci:  (train base to train) six one?
8  (1.3)
9  M:  six one?
10  (1.3)
11  Ci:  yeah er take a minute (ordering that) all clear if
12  you'd like to standby for giving instructions...

In (9) the 'go ahead' appears to provide the opportunity for the train captain not only to take the turn, but also give a reason for the original request for a phone connection. In (10) the reply by the train captain, minimally a repetition of the call sign with rising intonation, demands further talk from the controller. Indeed, in the following instance (1), when no answer is forthcoming, Ci repeats the call sign. Only when he gets an answer 'yeah five two' does Ci give the reason for the call.
Thus, in each instance various activities appear to contribute to the production of the reason for the call: the rings, the uttering of a call sign, its repetition and devices like rising intonation and 'go aheads'. The phone rings by themselves do not summons the called. In telephone conversations an answer to the summons will then demand the reason for the call, setting 'in line' the subsequent activities in the conversation. Here, more work appears to be required.

Schegloff (1968) distinguishes summons which open conversations from other activities, particularly questions, both in their design and in the responses they demand. They may involve terms of address, courtesy phrases or physical devices and constrain the kinds of responses provided in the answer to the summons. The call sign does appear to do similar work. By being minimal itself it constrains the possibilities for the remote participant until he or she has responded. Hence, in (1) the uttering of the call sign and its repetition appears to parallel the ring and the caller speaking in telephone conversations, like (8). DLR phone conversations then may provide another domain where the obligations of a summons demand particular responses which then can set in line the subsequent conversation. The technology in the control room has transformed the sequence in quite a different way to the telephone. No longer can the ring be utilised as a summons, instead the activity has to be accomplished in other ways, through other devices. A 'summons' as such is accomplished through two activities: the ring and the uttering of the call identifier, both being required by the way the technology operates. The first could be thought of attracting the attention of the audience of remote personnel, the other selecting a particular party (or parties) from that audience. A conventional summons then might have to be considered to be distributed over two distinct activities.
Although the organisation of the initiation of radio phone calls may be provided by the sequential demands of a summons, such an account appears to be only partial, ignoring other activities which have to be accomplished by the participants. In the opening of the call the controller not only has to establish contact with the appropriate individual, he or she also has to elicit the reason for the call. As the controller always makes the call, the selection of who should provide this reason would appear to be potentially problematic. To explore this, it is worth briefly considering different resources the participants have available to them and the circumstances they face.

Although controllers have a computer system informing them of members of staff who wish to 'make calls', they still have to select the interlocutor from an audience of potential recipients. Unlike a telephone call, there is no resource for controllers to ascertain that the intended recipient is available to speak. Even if the train captain has requested to make a call, the delay in making that call may mean that he or she has become occupied in other matters, such as ensuring passengers have safely embarked and disembarked, dealing with a passenger's query or problem, or operating the train. Furthermore, the train captain works in an environment that is potentially noisy and open for interruption, where there are occasions when it is difficult to answer a call, or answer promptly.

The train captain, on the other hand, has few resources with which to ascertain when a request for a call will be answered. Although some of the radio traffic can be monitored, the train captain does not have access to the circumstances facing the controller - the state of the service, the queue of phone calls etc. When a call is requested from the controller, it is not possible to determine when the relevant call will be made or if, when a call is made, whether it will necessarily relate to the train captain's request.

There is another way in which the resources of the interactants are asymmetric. Controllers have a range of technologies that provide access onto the system: CCTVs; displays of the line; calls from other drivers, all allow the controllers to ascertain the state of the service. These offer a

---

13 At the time of recording the DLR would have up to 15 trains out on the system, each with a train captain. Furthermore, members of staff in the shunting yard, engineers, some members of management and a 'mobile controller' also have radio phones.
potential resource, revealing some of the circumstances surrounding a call from a train driver. Train captains have little access to the state of the service, they can only speculate on the reason for the call.

The need to cope with these different problems may account for the distinctive nature of calls on the DLR. In telephone conversations, the called speaking first can be seen as one part of a collaborative solution to the problem of producing a reason for the call. Partly because of the nature of the technology, calls in the DLR Control Room are always commenced by the controllers and when a call is initiated by the remote party, the controller's actions may occur some time after the initial request is made. There would appear to be then not only a problem of eliciting the reason for a call, but also identifying the individual who needs to provide it. Thus, it is unclear whether the ways in which activities are distributed over the initial contributions to conversations and telephone calls can account for the organisation of calls in the DLR. In effect, it would appear to be a practical problem for participants using the technology, particularly the remote party, to distinguish between incoming and outgoing calls.

5.6.2 Producing the 'Reason for a call'

In the opening of phone calls, there appear to be several resources which participants utilise to deal with the practical problems at hand. Following the rings the controllers invariably summons using a call sign, a unique number identifying the radio set of the recipient. These numbers are available on the computer-supported phone system. On hearing the number, the remote party then replies using the same number.14 With the absence of any resource to mark that a party is ready to speak, such as the phone stopping ringing, these turns serve to establish that both parties are available to talk at the current time. They also ensure that the party summoned is the party who is speaking.

There are other recurrent features in these initial turns of talk in the phone conversations. With a few exceptions the number is spoken with a rising intonation, as in the following examples.

---

14 It should be noted that these numbers are not the same as train numbers. Train captains can change trains during a shift, but their radio set number will remain the same. So, in (11) the captain of train '105' has radio set '39'.
controllers may also include other components to their talk, for example. identify themselves, as the 'train base'.

With some slight variations, these are the conventional ways in which controllers begin their talk on the radio phone, completed either with a rising intonation or a 'go ahead'. As Schegloff remarks about summonses on telephones, these are special forms of questions demanding particular types of responses, hence the distinctive 'rising terminal juncture'. If no response is forthcoming, the summons is repeated (as in fragment 1). Speaking the radio number with rising intonation, even if it has to be repeated, is sufficient to identify a potential co-interactant and elicit a response.

The 'go ahead' appears to accomplish another activity, not only summoning the potential co-interactant, but also eliciting the reason for the call.

This then, is one way of solving one of the critical problems for co-participants, how to identify whether this is an 'incoming' or 'outgoing' call and who is to give the reason for the call. Controllers can elicit the reason
for a call that had been requested by the other party, by simply tagging on a 'go ahead' to the summons. This not only makes it relevant for the co-participant to give the reason in the next turn, but also displays the speaker's orientation to this as an effective 'incoming' call.

Though controllers can display an orientation to the distinction between incoming and outgoing calls in their first turn of talk, the initial turns of incoming calls are not always accompanied by a 'go ahead'. In fragments (12), (14) and (18), for example, following the number given by the controller, the remote caller goes on to give the reason for the call (e.g. fragment 14, lines 5-6).

(14) 33 - transcript 2
1 ((long ring))
2 (0.3)
3 Ci: one five?
4 (3.4)
→ 5 TC: yeah one five () I've got a ( light) at ( ) Bow
→ 6 Church ( ) over.
7 (1.3)
8 Ci: Roger on that (0.2) select A: Tee Pee manual please
9 ( for a couple of seconds ) and then go back into A:
10 Tee O: ( see again) ( Bow) over.
11 (1.3)

In outgoing calls, like instances (1) and (10), the reason for the call is produced in the controller's second turn, but does not appear to be foreshadowed by any vocalised device by the controller. Indeed, in these calls the remote party's turn appears to be designed in order to elicit the reason for the call, through the call sign being terminated by a rising intonation or with the addition of a component such as 'receiving'. It is as though, through the response to the call sign, the call emerges as an outgoing call.

The distinctiveness of the radio phone calls from conventional calls is not only in the ways the technology operates (the way the phone rings, the broadcast capability etc.) or because the controller tends to speak first, but it is also because the controllers do not necessarily distinguish the type of call, particularly in terms of who should give the reason for the call. Given the potentially critical nature of the calls, it is perhaps surprising that there is apparently no routine means of either eliciting or foreshadowing the reason for the call at the beginning of the call. Instead, the reason for the call appears to emerge, first by giving the remote participant the chance to provide it and then, if this is not forthcoming, by the controller producing it. In incoming calls the controller occasionally will elicit the
reason for the call with a 'go ahead', however, in most cases controllers
will initiate both incoming and outgoing calls in the same manner.

Nevertheless, in most cases, one of the participants will manage to
produce, quite unproblematically, such a reason for the call. The
controller, on occasions, may appear to elicit the reason from the remote
participant, however, there are a few cases when the controller utters
anything other than the standard repertoire in the first turn.\textsuperscript{15} Given the
resources available to the controller, and the demands of the setting, it
may appear strange that the controller, particularly on what might be
considered outgoing calls, does not appear at least to foreshadow an
upcoming reason for a call. The 'powerful machinery' for distributing
turns at the beginning of phone and other conversations seems to be less
clear in the radio phone talk. Instead, it appears as if there is a rather
looser means for distributing turns when the ringing of the phone, or the
first speaker, cannot be relied on to set out a trajectory for turn allocation.

\textsuperscript{15} One case where the controller does utter talk prior to a response from the co-
interactant is in the following instance.

\begin{verbatim}
  (22)  29Co
  1 Ci: train base to train six one↑
  2   ((loud beeps from emergency system))
  → 3 Ci: yeah can you keep your emergency (mushroom (.) in; over?
  4   (12.5)
  5 Ci: six one↑
  6   (1.8)
  7 TC: (six one) over.
  8   (.).
  9 Ci: yeaas have you stopped your train over↑
 10   (1.2)
 11 TC: yes (.) I have stopped my train (0.2) over.
 12   (.)
 13 Ci: right what I would like you to do: is go to your (leading
 14   eN: Pee Dee Pee;↓) (0.4) and select emergency shunt
\end{verbatim}

Here, when the call its initiated, in the 5 second gap after controller's 'train base
to train six one↑' there are a series of loud beeps (line 2). The talk in line 2 'yeah
can you keep your emergency (mushroom (.) in; over?' then appears to be designed
for the particular recipient, so that the call can continue less problematically.
The 'emergency mushroom' is a button pressed in the train in cases of
emergencies and emits the beeps in line 2. Once the noise stops, the controller
begins again. This may appear strange as it might be expected that the cessation
of the ringing might suggest that the particular recipient is available and
listening. However, as in most cases, the controller appears to delay giving the
reason for the call or commencing any details concerning the call until the remote
participant has replied. Indeed, in this case, the controller appears to delay even
further, until he has confirmed the status of the train, that it has stopped.
controllers provide their co-participant with an opportunity to speak and give a reason for their call. Only in third turns do controllers give the reason for an 'outgoing' call.

The work of initiating conventional telephone calls appears to be rather brief and economic when viewed alongside the calls in the DLR Control Room, a phone ring summons another party, setting up a trajectory for the distribution of turns of talk, at least for the first few turns, and sometimes for the entire call. In the radio phone talk the distribution of the activities in the calls is managed in the first few turns of the call. The technology and nature of radio phone conversation requires participants to perform particular activities which are efficiently accomplished in the first turn of a conventional telephone call. Producing the reason for the call, and more importantly managing who produces it requires distinct activities. These allow the participants also to manage other contingencies which may arise, such as in the delay between the request for a call and the controller initiating the call, other matters needing to be dealt with (as in fragment 1) or the reason for the call no longer being pertinent (like a train that has been delayed beginning to move). Nevertheless, the initiation of the phone conversations in the DLR are routinely accomplished unproblematically. They are managed from moment-to-moment, the participants collaborating in the emergent distribution of turns at talk and the activities they entail.

5.6.3 Problematic Cases

In most cases, this collaborative management of the distribution of the turns is accomplished routinely, the reason for the call being produced unproblematically by one of the parties. However, occasionally, there do appear to be problems at the commencement of a call, as in fragment 23.
The controller has to repeat his utterance of the remote party's call sign 'five one'. Although there is some delay to the reply, this or any other problem with the transmission of the call does not appear to warrant the repetition. There is no evidence, for example, of the controller not hearing the train captain's reply in line 9. The timing of it and its production with the same tonal contours of the original, however, appears to be consistent with there being something problematic not in the hearing of train captain's first reply, but in its character. Indeed, when the train captain goes on to produce a reason for the call (which is unclear and appears to concern the automatic ticket machines, line 13), the controller turns to the back of the room where a machine is available displaying the status of all ticket machines and he reports which are not working (lines 16-17). Hence, when the train captain gives the reason for the call, the talk appears to continue quite unproblematically. It appears that the controller is orienting to the call as an incoming call, one for which the remote caller has to produce the reason. The production of the train captain's response to the call sign in line 9 appears to be problematic in this regard. The repeated call sign (line 11), not just requiring a confirmation of the remote party's identification, but something more, like a reason for the call. Once this is produced (in line 13) the controller can then take the appropriate action, and does so unproblematically. It is also interesting to note that the train captain's initial repetition of the call sign neither has the usual rising intonation of a response to a summons in an outgoing call nor is it accompanied with a component like 'receiving'. Instead, it is delivered with little tonal contour. Hence, the nature of the
train captain's response may then appear to be rather ambivalent. The controller may orient to this, a minimal repetition of his first turn displaying a potential absence in the train captain's turn, without initiating any particular talk on the matter.\textsuperscript{16}

Although there are systematic ways in which reasons for calls are elicited and turns are distributed between participants, this is not to say that turn-taking in the beginning of radio phone conversations is predetermined, or even set to a course by previous actions on the computer system. A request for a call heard by a beep and displayed on the computer system provides resources for a controller when initiating a call. The ringing of the phone to a train captain also provides resources, particularly if this is heard in the light of a previous request. However, the work of initially finding the reason for a call has still to be carried out by participants. The following instance is particularly unusual, as in it the train captain speaks first.

(24) 32ii

\begin{verbatim}
1  (0.6)
2  ((long ring))
3  (6.1)
→ 4 TC: five one (go ahead)
5  (0.6)
6  Ci: yeah five one err I’ve still got no (0.2) news (on)
7  you yet for your relief over.
8  (1.5)
9  TC: okay then well (we’ll just stay here until we get
10  the job done anyway) okay thanks a lot
11  (1.3)
\end{verbatim}

Here, after a long pause the train captain says 'five one (go ahead)' (line 4). It is as if the pause after the ring of over six seconds is noticeable by the train captain. Indeed, the train captain's call echoes the design of initial turns of incoming calls given by controllers, by eliciting the reason for the call through a 'go ahead'. After some perturbations in his turn, 'yeah five one err', the controller does produce the reason for a call.

It emerges through the turn that this call relates to some previous matter (getting relief for the train captain) that the controller has been dealing with, the topic, presumably, of a previous call. Indeed, the design of the talk is produced with respect to this prior matter, as is displayed in the controller's 'I've still no (0.2) news'. The talk is designed as a call-in-a

\textsuperscript{16} An activity that may be characterised in terms of a preference for 'self repair' by the train captain (cf. Schegloff, et al., 1977).
series (cf. Button, 1991a). The controller's turn appears to be designed with respect to hearing the train captain's utterance in the light of a previous call.

Nevertheless, the train captain, uses routine components to initiate the call and identify the remote party. He has not only taken over the turn but the nature of the turn - to identify himself and to elicit the reason for the call, which the controller does. Even in non-routine cases, the production of particular turns places demands on a co-participant which are treated as accomplishing particular activities and requiring subsequent action. A call sign, for example, can serve to summons another, identify oneself and also to elicit the reason for the call in the next turn. The natural ordering of the turns provides resources for participants to make sense of such a component, and the work it is achieving. The components also serve to make up the ordering, being produced in different ways to accomplish various activities in the light of differing foregoing circumstances. Thus, though the beginning of the phone calls may appear to have redundant or irrelevant parts, or to follow some routine, each component has to be made sense of by the co-participant in the interaction, placing demands on their subsequent activities. The beginning of the phone conversation unfolds, establishing contact, identifying the participants and producing the reason for the call, from moment-to-moment it is made sense of and managed by the co-interactants.

5.6.4 Initiating Calls through a Mediating Technology: a Summary

The technology supporting radio phone conversations in the DLR Control Room no doubt provides capabilities which are not (or not yet) available on other forms of comparable telecommunication services, like mobile telephones. It is possible to broadcast messages, for remote parties to listen in to parts of the other conversations of their colleagues, and also for controllers to deal with callers when they wish to and in an order they find appropriate. These capabilities, however, place additional demands on the talk in the early part of the call. Calls to particular parties have to be distinguished from general calls to all potential co-participants and the possibility of others using the open channel requires that remote parties are successfully selected and identified. These require a method of calling individuals by using identifying call signs (radio numbers) and receiving a
response using the same numbers. The design of the technology also separates, for remote parties, the activity of requesting a call to be made from the call itself. This can mean that there is a substantial delay before a call is established, and that controllers, in effect, initiate the talk in all radio phone conversations. As is revealed through the aforementioned instances this requires participants to collaboratively establish who is to give the reason for the call. This work of producing the reason for the call is interleaved with that of establishing contact between participants. Most commonly, it is achieved by participants utilising the first opportunity to provide a reason for the call once the remote party has been identified and successfully established contact. For the remote party, this is immediately after they have given the call sign, for the controller this is in the next turn. Various devices can also be appended to the turns required to identify co-participants, such as with 'go ahead' by the controller on incoming calls or 'receiving' by the remote party on outgoing.

Nevertheless, in the vast majority of cases the establishment of who is to give the reason for the call and the subsequent production of the reason is accomplished through the turn-by-turn management of the first turns of the conversation. The first position is given to the remote party allowing for cases where a prior request for a call has been made. If both have a reason for a call then the controller's will be dealt with after the remote party's. In this way, the production of an incoming call or an outgoing one is emergent, accomplished through the collaborative activities of both participants. When one of the parties does produce a component related to the distinction between the responsibilities of caller and the called, it is produced for the other in the next turn of talk and not to outline a particular course for their own activities.

In DLR radio phone calls, each participant has different resources available to them relating to the putative nature of the call. The controller has a computer system that provides information on who wants to be called and other displays relating to the overall state of the service. These can be used to suggest possible reasons for the calls. However, the remote parties on the trains, platforms or in the shunting yard are engaged in particular local activities unavailable and inaccessible to the controller in the control room. The organisation of the commencement of the call, particularly its flexibility, may reveal ways in which the participants take account of these contingencies. The remote party always has the
opportunity to give the reason for the call first. Even if circumstances have changed and other things have to be dealt with, the organisation of the talk provides for a ‘queued’ request for a call to be considered in preference. Of course, in most cases it will be unproblematic for the co-participants to work out which party has to provide the reason for the call. Remote parties know that controllers have a display of queued requests in front of them, and no doubt recognise the possibility that a ring will foreshadow a ‘call back’. Nevertheless, the organisation of the call in this way at least allows for the circumstances local to personnel out in the world to be dealt with first and for controllers not to have to rely on the displays in front of them to shape the nature of the call. All calls can be commenced in the same way, with the work undertaken in each turn being managed as the call emerges. Moreover, the organisation of the talk requires little talk to be explicitly related to the kind of call being undertaken, not only allowing for the emergence of the reason for the call, and then, presumably, to how that call can be dealt with, but also reducing the need to deal with problems which could emerge relating to the distribution of the work in producing the call rather than the reason for the call itself.

5.7 PAUSES AND DELAYS - SYSTEMATIC RESOURCES FOR THE ACCOMPLISHMENT OF ACTIVITIES IN RADIO PHONE CALLS

An observable feature of calls on the DLR is the curious timing of the calls between controllers and remote parties. Not only are there, on occasions, substantial pauses within the calls, but there also appear to be particular regularities in the timing and pacing between contributions to the phone conversation. In part, these may be due to the nature of the technology and the demands of the activities in which participants are involved. A train captain, for example, may be involved in a range of activities when a call is initiated. Various noises in the local setting, problems in reception and even the use of the phone technology may also delay the production of a response or a turn at talk. Therefore, it is not surprising that the delays and pauses in radio phone conversations, and the tolerance towards them, routinely appear to be longer than those on the telephone or other kinds of
conversation (Jefferson, 1989).\(^\text{17}\) So, even when the call appears to be well established it is not unusual to find lengthy pauses between the turns of talk of co-participants.

(25) 41

Ci: yeah that's roger on that I've got (0.3) one lift out of service at Mudchute. (.) as far as we are aware the other one is working. (over)

→ (1.7)

TC: (alright ) also the lift at South Quay isn't working I've just been informed over.

→ (2.3)

Ci: yeah okay thats received thanks for that I've got n- (0.2) nothing on my board (0.3) that says South Quays out and we'll get someone out to investigate? (0.2) thanks for that?

(0.5)

TC: okay

As such participants may then take account of these contingencies and the delay may not be considered by the participants as problematic. This is not to say that participants will wait indefinitely for a next turn. As revealed in (24) a pause of over six seconds before a controller produces a first turn can be viewed as problematic, and in (1) there is a brief pause of only 0.5 seconds before a controller repeats a summons. The participants in the radio phone calls, therefore, appear to be sensitive to apparent pauses of different lengths by their colleagues in different temporal locations during the call. Exploring these domains, these apparent silences, may reveal some of the sensitivities that participants may be orienting to. An examination of the timing of the calls at various places within the call may reveal features of the organisation of the call, and also how participants analyse the ongoing happenings and items within a call.

Given the foregoing analysis in Section 5.6 two domains appear to be of particular interest: those surrounding the controller's initial turn. The gap following the turn will be considered first (Section 5.7.1). The first turn by the controller demands a response from the remote party. Examining the pause may reveal the type of resources the participants rely on to secure participation in the talk by a colleague. The pause between the ringing on the phone, considered in Section 5.7.2, may help to further explore the nature of the summons discussed in the previous

\(^{17}\) Jefferson (1989) considers the possibility of a candidate 'standard maximum silence' of about 1 second. She does, however, raise concerns about its status within sequential and interactional analysis.
section. The activity of summoning the other appears to be distributed over two actions; considering the gap between them may shed some light on the ways this activity is distributed.

5.7.1 The Second Gap: a Resource for Securing Engagement in the Call and for Prospective Activities in Relation to the Call

After the initial call sign, long delays before the remote party replies are not uncommon. So, for example in (3), a pause of nearly 4 seconds after the first turn by a controller on an incoming call may not be considered a noticeable delay in answering.

(3) 14Co - transcript 1 (simplified)

Ci: three two go ahead†

→

(3.9) TC: (yeah this is) three two to has:e (0.2) umm: (.) someone has umm (0.2) a black briefcase (on tracksi:de) (1.2) coming out of (.) All Saints...

Even though the remote party may have requested the call, the time before the controller acts on it may be substantial enough for the train captain to be engaged in other matters, or at least take time to answer his phone. Thus, delays in answering after the phone has rung and an identifier uttered can be quite substantial. However, this is not to say that an utterance or a reply can be left indefinitely. In the following case, after over 8 seconds the controller repeats, in abbreviated form, the identifier.

(16) XbCo - simplified transcript

Ci: train base to train three two:o†

→

(8.2) Ci: three two:o?

(0.3) TC: (this is train three two) to train base o:ver†

Here, the lack of reply for 8 seconds by the train captain could be seen as an 'official absence' (cf. Schegloff, 1968, p. 363-4). When the identifier is repeated the reason for the call quickly follows. It appears that the controller orients to the delay as an absence. In (24) above, the opposite case appears to hold, where the absence of a particular turn appears to be by the controller. After over six seconds, the train captain takes the unusual action of commencing a call with 'five one (go ahead)'. The participants by repeating call identifiers and taking turns appear to be orienting to a certain pacing to the calls, where some pauses mark noticeable absences in the activities of another. However, the time
between these activities is not fixed, pauses or delays are treated differently by the participants according to the activity at hand and the occasions of the call. Nevertheless, participants do appear to be utilising particular resources in order to make sense of the silence; as marking missing activities at the beginning of the radio phone calls, for example. It is a domain of particular relevance to participants as it is through these first turns of talk that not only the identity of the co-participants is established, but also it is through these they can establish contact on a communication system where reaching a particular party cannot be guaranteed.

One way of considering these pauses may be to examine whether there are any systematic features in the timing between the initial turns of talk in the more routine calls. If the fragments already considered are examined, there do appear to be some crude regularities. For example, the delay following the initial turn by the controller ranges up to 4 seconds in the more straightforward incoming calls. In the straightforward outgoing calls, the delay can range up to around 8 seconds. This kind of difference might be expected between recipients who are 'waiting for a call' and those who are not. Controllers may thus be tolerant to waiting for a reply in different ways to those remote parties who have requested a call and those who have not.

In (16), the gap in the talk in the phone is not completely silent. After about 3.5 seconds a high pitch feedback tone emerges.

(16) XbCo - transcript 2

| Ci: two: o↑; . . . ; . . . ; . . . ; . . . ; . . . ; . . . ; . . . ; three two: o . . . 8.2 seconds ———— this ———— |
| TC: this |
| Ci moves hand over coffee cup mouse |

---

18 These are fragments 1, 2, 3, 4, 5, 6, 9, 10, 12, 14, 18, 19, 20, 21, and 22. In a larger corpus of over 40 instances of incoming calls, there are 3 where the delay is greater than 4 seconds, the longest being 5.2 seconds.

19 These are fragments 6, 10, 12, 13 and 14. In 16 outgoing calls the delay after the controller's turn is over 4 seconds in half the cases, and ranges up to 11.3 seconds. In all the cases where there is a delay of over 6 seconds, the call sign is repeated or the phone re-rung.
This first tone briefly stops before re-emerging at an even higher pitch for another 0.5 seconds. It is after this noise has stopped that the controller utters the call sign again. The delay in summoning the train captain again appears to be in the light of this noise. Indeed, the noise may be related to the activities of the train captain. The train captain may be attempting a reply. It is only when a significant pause emerges after the feedback that the controller utters 'three two'. Similarly, in (26) a feedback noise occurs about 2.5 seconds after the call sign.

(26) 30Co - transcript 1

It is only after the noise has faded does the controller begin a series of activities associated with the phone system: moving his hand to the mouse, glancing at the phone system screen and then pressing the mouse button. This causes the phone to ring again. In (16) and (26) the controllers appear to hold off re-ringing or re-summoning the remote party until the feedback has faded. In (27), the controller again holds off any phone related activity, until the feedback noise has faded.

(27) 26Co - transcript 1

In this case soon after the feedback has faded, the train captain does reply. There are numerous instances where noises appear on the phone line. The controllers have to make sense of these, as they may foreshadow an activity from a remote party. The feedback may indeed be part of the attempt by a train captain to establish contact. So, though there may be a significant delay after a controller’s first utterance, before he goes on to repeat his efforts to make contact he has to make sense of other occurrences on the phone line.
Various components of the initiation of radio phone calls have already been explored in detail which can be utilised by co-participants. These include: the way the rings of the phone and the initial turn by the controller serves to summons the remote party; the work involved in establishing the reason for the call and who is to give it; and the ways silences and noises in the phone call. Each of these components provide resources for both participants to make sense of the unfolding activities which emerge in the call.

So, although a gap between uttering a call sign and its repetition, or the repetition of the phone ringing, may be over ten seconds, this may be filled with activities that can be heard as potentially related to the call. There does appear to be a slight difference in the ways in which controllers handle incoming and outgoing calls, the delays before repair in the former being slightly shorter than in the latter. However, there is no simple metric when a gap becomes problematic. Instead the controllers appear to be sensitive to the circumstances of the remote parties.

In (16), (26) and (27), the controllers appear to engage in a series of unrelated activities whilst a potential reply is being awaited: moving a coffee cup, scanning CCTV displays, or moving a pen on the desk. However, even these may be utilised within the unfolding course of activities. In (16), the movement of the coffee cup leaves free a space to the right of the mouse making it possible to hold and move the mouse. As he utters the call sign for the second time the controller’s hand is over the mouse as if in preparation to re-ring the phone (as in 26). The movement of the object may then be retrospectively useful for an emergent course of action.

The glances at the CCTVs, the ATS displays and the phone system may provide other resources with which to make sense of the call, showing if the train the remote party is on is moving, docked, or ready to depart, for example. A resource which controllers can use in their talk, as in (16).

(16) XbCo - transcript 3

Ci: train base to train three two:o↑
   (8.2)
Ci: three two:o?
   (0.3)
TC: (this is train three two) to train base o:ver↑
   (0.2)
→ Ci: Yeah I do believe that you have just pulled into
    Poplar. (.) Can you see a Bee Tee Pee Officer (0.2) at
all? (I believe he may be attempting to get on your
train over↑
(0.9)

The characteristic features of the talk on the radio phone, like the
structure and pacing of the talk, may in part be due to the nature of the
technology. The requirement to identify a co-participant in a broadcast
medium and the problems of communicating through a radio each demand
work and activities to be designed in particular ways. There are also
procedural requirements: to repeat instructions and to establish contact
through radio number identifiers, for example. However, these necessities
and demands are both utilised for, and transformed by, the practical
purposes at hand. The organisation of the phone call allows for the reason
for the call to be produced with little 'negotiation' required concerning who
should produce it. Noises on the phone can be heard as preparatory to
further contributions and it may therefore not be necessary to explicitly
repeat a summons or a call for identification. Although the calls may
apparently involve unnecessary turns and repetitions, and pauses may be
lengthy, talking in this way may indeed secure the engagement of a co-
participant, and get to the matters at hand without engaging in talk about
talk.

5.7.2 The First Gap: a Resource for Distinguishing Calls

From the foregoing discussion, Figure 5.3a can be refined slightly. The
domain just considered is arrowed ‘1’.

 Incoming Calls
 <long ring of controller’s phone>
 (= 0.5 seconds )
 <long ring of remote party’s
phone>
 2→ < less than 1 second >
 C: ('train base to') radio number
 or
 radio number + 'go ahead'
 1→ < up to 4 seconds >
 TC: (<feedback>)
 radio number + <reason for call >

 Outgoing Calls
 <long ring of controller’s phone>
 (= 0.5 seconds )
 <long ring of remote party’s
phone>
 < over 1 second >
 C: ('train base to') radio number
 or
 radio number + 'go ahead'
 < up to 6 seconds >
 TC: (<feedback>)
 radio number
 <up to 1.5 seconds >
 C: <reason for call >

Figure 5.3b: A general schematic for the organisation of calls
The second domain to consider is arrowed ‘2’. This is the pause between the second ring of the phone and when the controller utters the call sign; the activities which parallel a summons in a conventional phone conversation.

In this domain there appear to be some curious temporal regularities. In outgoing calls the delay before the controller's turn tends to be over a second long and in incoming calls it tends to be less than a second.\(^{20}\) This may be another case of controllers being sensitive to the circumstances of their recipients, taking account of the potential difficulties of, not only replying to a call but also, discriminating that a call is for them. The talk may then be designed with respect to the varying circumstances at hand, so that the potential recipient 'has more time' before the identification is given. This would be a particular difficulty for recipients who had not previously requested that they wanted to call.

There are cases where for an incoming call the pause after the ring is longer than a second.

(8) 1 - simplified

\((\text{long ring})\)

\((1.2)\)

Ci: (three eight)
go ahead?

\((1.3)\)

(19) 9 - simplified

\((\text{long ring})\)

\((2.5)\)

Ci: four eight
go ahead?

\((2.6)\)

(20) 12 - transcript

\((\text{long ring})\)

\((2.4)\)

Ci: five two
go ahead

\((2.2)\)

In each of these the controller adds a 'go ahead' to the call sign. As in other incoming calls the 'go ahead' appears to be designed to elicit the reason for the call from the remote party, which it does. Indeed, in other cases when a 'go ahead' is uttered, it is produced after some anomaly or problem in the call, as when the train captain talk first in (24), or when there is no hearable gap between the phone ringing and the identifier (as in 21).

(21) 11ii - transcript 3

\(\rightarrow\)

\(\rightarrow\) Ci: =three two go ahead

\((3.2)\)

TC: yeah three two ( ) err: ( ) (got a trip)...

It may be that a 'go ahead' not only acts prospectively to elicit a particular type of talk from the co-participant, but also may serve to retrospectively take account of some problem or anomaly in the call. A

\(^{20}\) For outgoing instances see fragments 10, 13, 15 and 14, for incoming ones, see 12, 14 and 21.
simple device like a 'go ahead' may follow these apparently long pauses (or notably short ones), ensuring the call is heard as incoming for the relevant participant.

There are potentially numerous accounts for why there may be a long delay between a summons and an answer. Even after a phone ring, the called party can be occupied in some other matter or located in a position not amenable to replying, for example. The controllers appear to take account of this by delaying any attempt to repair and hearing particular noises on the phone line as potentially foreshadowing a response. They may even hold off uttering the call sign for a period so that co-participants are able to discriminate the subsequent talk from other activities in their local environment. It may also be that these delays are discriminated between forthcoming incoming and outgoing calls: a recipient of an outgoing call, a genuine 'called', being given a substantial period in which to reply and a longer period in which to hear the summons. The overall activity of securing a recipient for a call being not only designed in order to select that hearer from an audience of overhearers, but also to secure a particular kind of participation in the forthcoming call.

5.7.3 Designing Radio Talk through a Mediating Technology: a Summary

To distinguish whether a call is incoming or outgoing would appear to be an activity for the participants that has to be managed through the talk on the radio phone line. The technology in the DLR Control Room rather subtly masks many of the features used in telephone calls to distinguish caller from called: the ringing of the phone, the cessation of the ringing and the answerer always speaking first. In all of these, the technology in the control room operates differently, the radio phone only rings for a fixed period, call requests from remote parties are queued and a remote party has to be called back. This technology is also different from many other radio telecommunications systems where a completely open channel requires caller to both speak first and identify a recipient in the first turn. Though in these cases the work of the summons may be the reverse of that in a telephone conversation, it may not have the unusual results found with the DLR system. The mediating technology, essentially a computer interface for making calls and queuing incoming calls, requires the controller to always speak first. The subsequent work of both obtaining the reason for the call and managing who has to give it has to be
accomplished over the subsequent turns of talk. This is a critical problem for the parties as the reason for the call sets up a trajectory for the forthcoming call. Following a query for information, the announcement of a problem or a request to take a particular action – the typical ‘reasons’ for a call – the controller and remote party may engage in a fairly brief interaction. This usually involves either a confirmation of the news, or the controller suggesting actions to be taken and the remote party confirming them. In a domain where a phone call may be critical it would appear to be essential that the production of the reason for a call is timely.

As outlined in the previous section the work of producing the reason for a call is a collaborative activity, making use of the resources available to the parties, and which are understood to be available. In this section, it has been suggested that the actual timing of the contributions is a potential resource for participants in accomplishing the initiation of the call. In particular, it may be that there are resources which can be utilised very early in the call, before the first turn is uttered. Both parties may be sensitive to these resources and to potential problems in this pacing. Although such an account relies on absences from the talk, it is interesting to note the one case where there appears to be problems in eliciting the reason for the call occurs after a long delay of nearly 2 seconds between the ring and the summons by the controller (lines 4-6), and when the controller does not explicitly elicit the reason of the call (23).

After nearly two seconds after the ring the controller utters ‘five one?’ (line 7). There is a further delay and the train captain utters his call sign without any particular intonational contour (line 9). After another two
seconds the controller repeats his utterance (line 11). With each having contributed to the talk, although uttered in the same fashion as previously, the controller's call sign can only be heard as a request for a reason for the call. This it achieves in the next turn (line 13). Though there appears to be problems in distributing the activities through the call, producing it is still, however, accomplished through sequential resources. The call sign eventually being heard as a request for the reason for the call.

5.8 The Use of Technology in the Initiation of Phone Calls

Through the apparently formulaic and plodding nature of the phone calls, particularly in their beginnings, members of staff on the DLR manage to accomplish contact, divide out the responsibilities between the co-interactants and set up a trajectory for the remainder of the call. The computer-supported phone system technology, no doubt, figures in this, it being a broadcast medium with associated difficulties of trying to contact and maintain contact with a co-participant. However, the technology does provide a useful resource for co-participants, particularly, the display in front of the controllers. This is one of many displays in the control room, including CCTV monitors at the right of desk, a display at the back of room which reports the state of the Automatic Ticket Machines (ATMs) on the line and various whiteboards, including one which reports the lifts that are not working. In most calls, controllers are seen to be look between these different domains. Some of these lookings appear to be straightforward, for example, when a train captain asks about the state of the ATMs or the lifts, controllers will turn towards the display at the back of room or the whiteboards. In other cases, the controller will be engaged in making a change to the train operation on the Automatic Train Scheduler (ATS) or making a call on the phone system and so will look at the appropriate display whilst this is being done. However, particularly at the beginning of phone calls, controllers will switch their orientation between the phone system display (PS) and the leftmost Automatic Train Scheduler Display (ATS1), often looking at these displays several times before the reason for the call emerges, as in (28).
When the phone rings the controller is looking down at a document in front of him on the desk. Only after the phone system beeps does he look up to the display. He then starts to make the call, selecting the call with the mouse. Once the system has begun to carry out the operations of making the call the controller looks towards the ATS display. He returns to this display after he has summoned the train captain. Thus, the controller looks at each monitor twice before the reason for the call emerges, and he is looking at the ATS1 screen whilst the reason for the call emerges. He remains in this orientation until he starts typing into the ATS at the end of the call.

It is not uncommon for the controller to orient towards the ATS system when taking a call, particularly for incoming calls. Indeed, in all the instances of incoming calls in the corpus, the controller is either oriented to (and looking towards) the ATS screen, either whilst he is summoning the remote party or by the time the reason for the call emerges. This screen provides diagrams showing the positions of the trains on the line,
so it is not surprising that it offers resources for dealing with the issues raised in the calls. It appears that the ATS screen can also provide resources for making sense of a call, as in the incoming call in (29).

(29) 6Ph - transcript 1 (simplified)

((long ring))
(1.7)
 ((phone system beep))
 ((Ci moves to phone, looking at screen))
 (1.3)
 ((Ci picks up phone, phone system beeps))
 (3.2)
 ((Ci selects call with mouse))
 (1.2)
 ((Ci turns to ATS))
 (0.8)

Ci: four five?
 (4.7)

TC: four five ( ) over.
 (0.2)

Ci: Yeah err (0.3) there's a bit of a holdup () as soon...

From the two displays the controller can make use of various pieces of information. With respect to train captains, the phone system display presents the radio number of the caller with the number of the train the remote party is assigned to. The ATS identifies the location of trains by this number. So, by scanning the two displays it is possible not only to identify the caller, but also their location. Moreover, the controller can see particular relevant features of the train, such as their mode (if the train is being manually or automatically operated) and whether the train is at or between stations (as in fragment 16). They also provide a resource for practical reasoning concerning the circumstances of the call. From the phone system display, controllers can identify whether the call is from an engineer, a worker in the shunting yard or special staff, such as trainers or mobile supervisors, each of these may have quite different reasons for a call.

In (29) the train captain is on a train on the same line and running in the same direction as a train where there is a 'revenue dispute' with a passenger. The location of the train on the ATS display can provide resources for making sense of an incoming call, even prior to any talk. A train captain being on a train that is 'backed up' in this way will frequently request information concerning the reason for the delay. Hence, by utilising the two displays controllers can prospectively orient to the nature of the call. The displays also provide other resources which are
potentially relevant for the call. So, returning to (28), after the train captain has been asked to repeat the reason for the call, the controller provides what is a common solution to many of the problems on the track - to operate the train manually, in 'emergency shunt'.

(28) 1OPh - transcript 2
1 (0.2)
2 Ci: yeah that's right Steve err once again select shunt ()
3 once you have done so (do continue) forward in shunt ()
4 from Canary to Herons (Park) ( ) (back into A: Tee 0: )
5 ...

By looking at the ATS1 screen, for a while, it is possible to see whether a train is stopped. The looks towards the ATS1 screen prior to the call commencing and the controllers focus on it (presented in transcript 1), may be utilised to make sense of emerging call. In (28), this can provide a resource to be utilised with the talk of the train captain to deal with the problem at hand. Here, the controller also makes use the name of the train captain 'Steve', another item which can picked up from the phone system display (line 2). In another case, the controller can make use of the location of the train captain to assist him with 'seeing for himself' part of the problem.

(30) 28iPh - transcript 1 (simplified)

((Ci looking at PS))

Ci: (five) two?

((Ci turns to ATS1))

TC: (five two)

(1.6)

Ci: yeah apologies for the delay and (keeping you at that intermediate routeboard) (0.2) Norman (0.6) as you can see there's one stuck on the delta I can't get hold of the train captain err soon as I can (deal with) it you can depart.

(1.5)

The phone system then offers a range of resources for assisting with the management of the call, it provides the identifying number of a caller, their name and the number of the vehicle which they are operating. This itself can suggest features concerning the nature of the upcoming call. So, in a call coming in from a member of staff at the shunting yard, a controller looks at the ATS2 display showing a diagram of that location (31, line 5). This glance follows a look to the phone system and precedes any talk on the phone. When the talk emerges it concerns a request for a change of route in the depot.
The phone system technology can thus provide a resource for suggesting where to look for putative reasons for an emerging call. The identity of the caller, their responsibilities, their location and circumstances all have to be made sense of in relation to one another.

In concert with the ATS displays it is possible also to have some access to the circumstances for the call. The reason for a call then, can be made sense of in the light of the resources available to the controller. So, after a phone rings in the DLR Control Room, controllers will glance at the displays in front of them either before they take the actions necessary to make the call, or, more commonly, whilst they are making the call. The information on the two displays have to be made sense of in relation to one another, the identity of the caller providing for a particular way of looking at the ATS, the very occurrence of the call itself suggesting possible happenings on the line. Moreover, the general happenings on the line, disputes with passengers, the reporting of suspect packages and the general operation of the service can all be brought together to reveal possible reasons for a call. The use of the displays is contingent on the

---

21 These observations, and those relating to the initiation of the call, could be considered with respect to the distinctive discussion of the resources of a
circumstances occurring on the line and in the call. Even though the phone system may provide a rather minimal set of information concerning the call and only once a request for a call has been made by a remote party, this can be utilised in concert with other technologies available in the control room. The rather lengthy procedure for initiating calls in the DLR Control Room then can be made use of by the controllers to utilise the rich resources available to them, both displayed on the screen and through their practical skills in making sense of the local scene.

Although the resources offered by the displays are useful in making sense of emerging incoming calls, controllers also orient to the displays in front of them when making outgoing calls. As in incoming calls, controllers tend to shift from the phone system to the ATS1 screen as the call is ringing or whilst they are summoning the remote recipient. There are potential happenings visible on the ATS1 screen which would be relevant for an outgoing call, a train holding up others may start to depart, for example, or the train of the remote party may be seen to dock at a station, making it more straightforward for that party to take the call. The delays and pauses noted in the previous section after the ringing of the phone in outgoing calls may be further tied to the ongoing circumstances of the remote party.

The gaps between talk and the activities on the phone system can thus be utilised as resources for the controllers, as potential locations in the talk where activities in relation to other artefacts can be accomplished. It may be that a glance to the phone or ATS screen may figure in the production of the activities in the call: the gaps may not only shape what is possible to be undertaken by the controller, but they may also be shaped in the light of these other activities. It is hard, however, to conceive of a way of developing such an analysis of 'noticeable absences'. Whatever the case, the activities which are undertaken between the controller and the remote party, and those unavailable to the remote party occurring in the control room are thoroughly interrelated. The organisation of the distributed activities is interwoven with the organisation of co-present ones. The nature of this relationship will be considered further in the next chapter.

5.9 DISCUSSION: INTERACTIONAL RESOURCES FOR THE PRODUCTION OF DISTRIBUTED COLLABORATIVE ACTIVITIES

One resource that appears essential for the production of collaborative activities within the control room are the phone calls between the principal controller and the various remote parties. A hearing of these is relied on for an understanding of the state of the service. The foregoing analysis has been concerned with how participants themselves analyse these activities. Through the organisation of their talk it is possible to reveal what is relevant (and problematic) in the calls. The participants are sensitive to the fine details of the available conduct of their colleagues on the phone line, including: the pacing of the talk, pauses, sounds on the line and the intonational contour of the talk. However, these are not just treated as signals or markers, they are made sense of in the light of the ongoing activities of the participants, with respect to, for example, a foregoing summons or question. These project relevant next activities, the interactional domain in which the contributions to the call are understood. Despite the apparent formality of the talk, it is this interactional and social organisation to the talk to which participants orient. This is displayed in their ongoing contributions to the talk through what they reveal to be unproblematic and problematic.

The participants to the radio conversation are, in perhaps a different sense to the way usually characterised within CSCW, monitoring the activities of their co-participants. They are closely attending to the details of the talk with regard to the particular demands of the moment. In other words they are, through their own talk and activities, displaying, from moment-to-moment an analysis of the conduct of their colleagues. Even in what may appear to be a peculiar interactional domain, they can manage to accomplish activities and display an understanding of them. Although they may be some requirement to repeat particular utterances, like instructions, or end some turns with 'over', it is through the handling of the minutiae of the talk that the conversation is managed.

The participants' analysis of the ongoing talk on the phone line is an analysis of their distributed activities. Through their contributions it is possible to reveal how critical interactional (and line management) activities are accomplished. So, for example, it appears through the design of their talk that producing a reason for the call is a relevant and critical matter for the participants in a call: invariably this is presented as
the first substantive matter in the call. Arriving at this requires an
initiation of a call, an identification of and selection of a participant (as the
technology relies on broadcasts). These activities are accomplished
through a sequential organisation of rings, call signs and the repetition of
call signs. Each contribution may be brief, but the activities which need to
be accomplished through them produces a rather convoluted opening to
the conversation. The production of a participant’s contribution is
analysed by the other party, and an understanding of it displayed in the
next turn, which is then available to the participant.

Thus, participants are displaying an orientation to the socio-
interactional resources and demands available in their turns of talk. The
organisation of the talk is not a pre-assigned protocol of signals, nor is it
merely a set of communication or information exchanges. The interaction
is managed through the production of the contributions: a summons and a
question places demands on the other, a silence can be accountable, and as
revealed in studies of social interaction, there can be preference for
speakers to repair their own contributions. These interactional resources
have been revealed in other interactional domains, such as telephone,
ordinary, face-to-face conversations and institutional talk.

As it is available to the participants, this sequential analysis is also
available as a resource for analysts. As noted by other analysts of
computer-supported phone conversations, the talk that is produced can be
indefinitely variable (Gray, et al., 1993). However, this does not mean
that there is no organisation to the calls. Indeed, in the materials of the
Docklands Light Railway, the participants appear to be sensitive to the
social organisation of the call at every point. Participants organise their
activities on the phone calls with respect on a set of practices for
accomplishing social interactions. Indeed, even when apparent problems
emerge, participants appear to be orienting to preferences, practices and
commitments engendered by social actions A particular silence or
contribution may be problematic, but dealing with it is still accomplished
through routine interactional practices for repair.

The analysis of the talk and activities through the phone system on
DLR is also an analysis of the use of a computer-supported phone system.
Such an approach is distinctive not only with respect to the use of
naturally-occurring materials, but also because of the utilisation of an
analytic framework that these make possible. Unlike studies which have
considered human-computer interactions (e.g. Thomas, 1990) or human interactions mediated by technologies like electronic mail (e.g. McCarthy, et al., 1990), in terms of conversational resources, an analytic framework drawn from conversation analysis would appear to be relevant for the consideration of the talk through a telecommunications system. Both participants have symmetric access to the talk on the line and the resources with which to continually monitor and analyse the moment-by-moment accomplishment of the other's activities. However, the system which has been deployed would appear to transform critical resources relied on by participants in other interactional domains, like ordinary and telephone conversations; as examples, the cessation of ringing does not mark that the other party is available to speak, the same party always speaks first, and the operation of the technology can introduce spurious noises and breaks into the conversation. Nevertheless, the interactional resources relied on the participants appear resilient. Individuals in the DLR phone conversations still orient to the demands of summons and questions, and a preference for self-repair. These may be implemented by different means, but still carry with them social obligations and commitments.

It is unlikely that the social organisation of phone calls was a consideration of the designers of the technology. Indeed, as already noted, the technology was actually designed for a different domain with differing communication requirements. It may be that an alternative design of the technology could provide for ways in which a summons could be produced through one device, or where incoming calls are distinguished from outgoing ones. It may also be that other 'pseudo-formal' procedures could provide resources for the participants to make sense of the nature of the call. Furthermore, it may be that alternative designs and deployments could lessen the 'work' involved in initiating a call and producing the reason for a call.22

Interactional practices may also be a concern in other domains where call taking is critical or routine. In these domains it would appear to be

22 It may be possible, for example, to devise technologies or telecommunications services which provide the parties with more symmetrical resources. Even a capability that provides similar interactional resources to that of a ringing phone ceasing ringing may be useful. Such possibilities are considered in more detail in Chapter 7.
relevant to consider the sequential organisation of talk. In the DLR, the participants manage to accomplish their work and activities utilising conventional and robust interactional resources. This is neither to suggest that individuals will always find a way of working through the technology, nor that such resources can always be relied on no matter through which technology they are achieved. The analysis does, however, reveal the kind of commonly-occurring resources that individuals can make use of in order to accomplish activities on and through technologies.

5.10 SUMMARY

The work performed by controllers in the DLR Control Room is, in the large part, accomplished through the phone system. They have to inform the train captains, shunting yard staff and other personnel out on the system of significant changes and happenings they are making to the service. They also have to be informed of incidents such as those involving problems with passengers, suspect packages and lost property, and have to provide information about lifts, ticket machines, relief breaks, delays, the running of the service etc. Thus, the reason for any call could come from either party, and, in certain critical situations both the controller and a particular member of staff on the line may wish to contact each other.

The computer-supported radio phone system adopted by the DLR has certain advantages, it allows controllers to chose in which order they wish to make and take calls and to delay responding to particular calls. This requires controllers to talk first in the radio conversations, and hence involves participants in accomplishing activities which are not usually required in more conventional telephone and radio conversations. The talk through the system appears curious in several other ways. For example, on incoming calls, where a remote party has requested a connection, the opening of the call appears convoluted, especially as both parties can expect to know who is who. This does not appear to be distinguished from cases where a request has not been made by a remote party. Nevertheless, the participants in the call have to establish who is in effect calling who and, thus, is responsible for giving the reason for the call. Although there are cases where the controller provides some explicit indication of the kind of call, with a 'go ahead', for example, these only appear to be when some problem or anomaly with the call may have been
heard to have occurred. It is far more common for there to be little to
distinguish incoming from outgoing calls. The distinction between an
incoming and an outgoing call only appears to emerge when there is an
absence of a reason for the call in the remote party's reply. Remote parties
themselves may include, in their first turns, devices like 'receiving' or may
terminate the repeated call sign with a rising intonation to make a call
outgoing. However, they too may just give a minimal call identifier as
their first turn. The participants then collaboratively produce the reason
for the call, relying on the preceding circumstances and the ongoing
development of the talk.

The organisation of the radio conversation appears to give preference to
the remote party to give the reason for the call. This may reflect the
different resources available to the participants. A request for an
incoming call is heard in the control room as a ring on the phone and
beeps on the phone system. An item on the phone system screen also
records pending phone conversations, and controllers orient to this screen
prior to calling personnel. However, although the remote parties can only
be contacted by the control room and are aware that controllers have this
technology, they do not have symmetric resources. When a ring is heard
on the open channel the remote party does not have access to the queue of
calls or the particular number being called, so after requesting a call they
have few resources with which to hear the ring as related to their request.
This asymmetry may be reflected in the organisation of the talk, the
controller providing initial resources through which the remote party can
make sense of the call. In most cases, despite glitches in the
telecommunications systems and noise in the respective environments, the
call proceeds unproblematically. The division of activities relating to who
does what is accomplished through the talk and the activities on the phone
system. The talk may appear to be rather formulaic, repetitive and
unwieldy, but the production of a minimal turn of talk like a two digit
radio phone number, its intonation and even the time taken before it is
uttered, may all be relevant to the way it is heard by a remote party, and
the activities for which it is being utilised. The initial turns of talk in the
radio phone conversation are more than signs or signals uttered to
establish a communication channel, they perform particular activities
concerning the distribution of responsibilities to be taken later in the call.
Even though they may appear rather minimal, the initial components of
the radio phone conversation are an essential resource for the distribution
and management of collaborative activities between personnel in the DLR
Control Room and remote staff.

More detailed inspection reveals a range of interactional problems to be
resolved by the parties, including having to select the remote party from
an audience of possible ‘overhearers’ in the broadcast medium, dealing
with the potential delay of responding to a request for a call and managing
who is to give the reason for the call. As call taking and call making is a
principal activity of the controllers, one which may occupy one of the
controllers for most of a shift, the effective management of calls would
appear to be crucial. It is also the medium through which most
communication between the control room and staff on the line happens in
the DLR. Indeed, it is principally through the talk on the phone that
crises are identified and managed and the smooth running of the service
maintained. This is a collaborative activity between distributed
personnel. Calls typically are short, often dealing with a single
substantive matter. Therefore, the establishment of the reason for a call
would appear to be a critical matter for the co-participants. In the
foregoing analysis, the activity of producing this reason has been revealed
to be collaborative, relying on the participation of both parties.

The technology of the radio phone system transforms the resources
relied on in conventional telephone conversations, the ringing telephone
by itself cannot be seen to be a summons to a conversation, the first turn
as both a response to the summons and an object that elicits the reason for
the call. In DLR radio phone conversations both of these resources are
different. Nevertheless, through the organisation of the initiation of radio
conversations the responsibilities between controller and the remote party
on the phone are managed so that reasons for a call are usually produced
unproblematically and the matters at hand dealt with.

Further examination of the organisation of the activities in the
commencement of calls reveals some subtle differences between incoming
and outgoing calls. Controllers may design the pacing of their talk so that
they are sensitive to the circumstances of the remote party. They are not
just sensitive to long pauses and gaps which might indicate problems on
the phone line, but to the differing delays at different points in the
development of the call. These sensitivities reflect the particular
interactional demands and even the demands in different kinds of calls.
So, for example, a long delay prior to a recipient answering an outgoing call may be treated differently from a delay in an incoming one. Moreover, controllers may also design their own pacing with respect to prospective activities in the call, even prior to any talk being uttered.

By exploring the visual conduct of the controllers, it appears that they have resources available to them with which they can access certain features of the remote party's circumstances. Not only do they have information concerning the identity of the remote party and their general responsibilities, in concert with other displays, it is possible to also locate the remote party and to some extent the circumstances they are facing. Controllers can utilise these to not only make sense of an incoming call, but also so that they can give prompt and timely responses to the issues raised. So, for example, the ATS display is read in terms of the local circumstances at hand and helps to make sense of them. The initiation of a call provides resources with which to make sense of the happenings on the display, the possible relevance, say, of a stationery train on the line. The display also reflexively can help make sense of the call, revealing possible reasons why, for example, the remote party may be ringing. The 'use' of each relies on the resources provided by the other. It also relies on seeing the objects on the screen in relation to the foregoing happenings on the line. These practices rely both on common-sense reasoning and skills though which artefacts, like computer displays, are both understood and used, and through which participants can make sense of the activities of colleagues. These practices are also interleaved, and rely on, the systematics in the pacing of the activities in the opening of the calls.

The production of activities by the participants in the calls rely on interactional resources. Through the contributions to the calls, the parties to the call not only have to achieve the practical tasks related to managing the line, but also have to manage the actual accomplishment of the call. This chapter has focused on the interactional achievement of producing the reason for a call. This is done through the contributions of both parties, interactionally, locally managed from moment-to-moment. Hence the management of the distribution of activities in the call and the management of the call is achieved through the self-same resources: through the sequential organisation of talk - a common-sense, practical and social accomplishment.
In producing an activity, particularly through talk, participants display an ongoing understanding of a prior activity, and provide for the resources by which that activity can be made sense of. Through their actions participants provide an analysis of another's actions. So, by giving a reason for a call, a train captain provides an analysis of a number uttered by the controller with a rising intonation as providing an opportunity to provide that reason. In the next turn, the controller through his or her actions provides an analysis of the reason for the call. These activities are situated; they are interactional and sequential.

The analysis of the participants provides an analysis available to others. In this case, through their talk and other activities the participants are displaying and revealing their competencies of using a computer-supported phone system. Hence, it appears that a sequential analysis of talk can also provide an analysis of the use of a particular computer system. Though this is a particular kind of system, not commonly considered in HCI, it no doubt provides for a distinctive analysis of its use. The talk is also a collaborative activity, through it the participants, who may be many miles apart, manage, with limited access to each other, the traffic and happenings on an urban transportation system. Thoroughly interwoven with this is a management of their collaboration, for most part seen, or heard, but not noticed. It may not be readily apparent that a phone system is a technology for supporting cooperative work, but by providing for the utilisation of practical, interactional resources it facilitates, usually quite unproblematically, the management and accomplishment of collaborative activities.
Chapter 6

The Collaborative Production of Computer Commands

the interrelationship between activities and talk in a natural setting

It is for others to do the worrying, to take the action. In the world of the great organization, problems are not solved but passed on.


6.1 INTRODUCTION

This chapter continues the examination of activities accomplished on the Docklands Light Railway begun in the previous chapter, but focuses on the details of the collaboration between the controllers within the DLR Control Room; in other words on co-present collaborative work rather than on geographically dispersed activities. These activities will be considered in the light of their use of one particular technology: the Automatic Scheduler System (or ATS).\(^1\)

\(^1\) Background to the activities within the Control Room and the technologies available is given in Chapter 5.
Each controller has a screen in front of them to the ATS which can be switched between a set of common views of the line and which can vary in detail. Therefore, the controllers could view a similar area of the line or send commands to the same area of track. However, in observations of their activities it does not appear that they do this. Instead, they divide the responsibilities between them and coordinate their activities with one another. This division of responsibilities is a practical problem for the controllers. However, managing who does what, and when, is a problem to which they readily find a solution, from moment-to-moment, throughout the course of their work. How the activities of the controllers is collaborative managed is the principal concern of this chapter.

In previous analyses, the management of collaborative activities has been considered in terms of the practices of 'monitoring'. In Section 6.2 this research is briefly outlined. In particular, the section discusses why such practices have been a concern of researchers in CSCW, both to analysts of collaborative activities and to developers of novel technologies, and some of the potential problems with the characterisation of these practices.

As the analysis pays particular close attention to the operation of the computer system, some background is given to the technology in Section 6.3, for example what is typically displayed on it and how it is typically used. The controllers are, in effect, not only viewing a common resource, but using a 'collaborative' system. Although the controllers are not geographically dispersed and, therefore, can see, not only what the other can see, but also their colleague's activities in relation to the system, they can each switch 'perspectives' and perform a symmetrical range of operations. These are capabilities considered by many proponents of groupware, and presumably would have a similar range of problems in their operation, such as the management of the activities through them. It may, thus, be that the collaborative accomplishment of the commands on the ATS system may reveal some of the resources through which participants in a natural setting manage collaborative activities on a computer system. The analysis in this chapter then, could be considered an analysis of a collaborative technology. Whether the technology is considered in this way or not, the analysis of activities on the ATS is a case of a naturalistic examination of the details of the use of a computer system.
As mentioned in the last chapter, a critical feature of the work in the DLR Control Room is the handling of radio phone calls. It appears that these too provide a resource against which the responsibilities for activities within the control room can be managed. In particular, certain calls require next actions to be undertaken on the ATS. In Section 6.4, the resources which the controllers utilise for performing the relevant and appropriate commands are considered. The phone call may not only be consequential on the controller taking the call (Ci), but also on the controller's colleague (Cii). It appears that certain features, like the confirmation of a change by Ci can be utilised as a resource for Cii to take particular actions.

However, Cis may not just be attending to features in the call such as a confirmation of a call, but may be attending to the calls in close detail. In Section 6.5 a case is considered which reveals some of the particularities of a call through which it may be heard to be relevant by another who is not a direct party to the call. It may even be that the controller taking the call may be producing talk to make the details of a call hearably relevant.

Hearing a call between a colleague and a remote party as relevant so that an appropriate action can be taken on a computer system, would appear to be a suitable candidate for consideration as monitoring. In Section 6.6, by detailed consideration of two instances, the resources through which such activities are accomplished are considered. It appears that not only may Cii be overseeing Cis, but that Cis may be attending to the interactions and activities of Cii. Even whilst engaged in an apparently distinct activity, Cis are making sense of particular commands entered by Cii. In Section 6.6.1, the resources they attend to in order to accomplish this understanding are considered. Controllers also may be 'monitoring another's monitoring' (Section 6.6.2).

This analysis reveals a close coordination of activities by the controllers where commands entered into the system are publicly available as operations that have been, or are being, undertaken to the railway line. In Section 6.6.3, an instance is considered where the activities of Ci on the phone and those by Cii appear to be particularly tightly coordinated. Indeed, the talk to the remote party appears to be shaped through its course, in the light of a command being entered into the system. The computer command, in a distinctive way, is collaboratively accomplished.

Such activities requires a continual analysis of the activities of others: whether they are relevant and appropriate, whether they demand a next
action or not, and when. However, the activities are accomplished through little explicit talk about these activities. Moreover, participants appear to rely on their own activities being attended to by their colleagues. In Section 6.7 a case is considered where these methods for common understandings appear to break down. A colleague explicitly ‘volunteers’ to take an action. Detailed examination of this instance reveals that the participants are sensitive to the conduct of their colleagues, and more, that they may transform their activities with respect to others not in the control room.

Through the analysis in these sections, the use of the ATS system is seen to be thoroughly embedded within the interactions between the controllers, and with other personnel, both in the control room and outside: what is on the screen is made sense of with respect to the talk on the phone: what is being talked about is made sense of in the light of the reading of the screen, and what activities another individual performs on the computer system is seen, monitored and understood in relation to the other artefacts available, the talk between individuals and the visual conduct of co-participants in interaction. This reveals quite a distinctive conception to computer use to that usually considered as human-computer interaction. These practices also appear to be unsuitably glossed as ‘collaboration’, ‘monitoring’ or ‘peripheral participation’. Section 6.8 discusses these issues in relation to recent debates within HCI and CSCW.

6.2 BACKGROUND

When observing the conduct of controllers in the light of their phone conversations considered in the previous chapter it is not always apparent why they glance at the computer systems in front of them at particular times. As suggested in the Chapter 5, these glances can be with regard to particular happenings in the call or may prospectively serve to provide resources for making sense of an upcoming call. However, there are cases when the glances do not appear to be tied to an ongoing activity in this way. For example, controllers are seen to orient to the ATS screens in more general calls, in broadcasts, where there is no particular recipient for a call, where for instance, all train captains are being informed of a problem on the line. These lookings appear to be a resource for keeping track of the train movements along the line. From these, with an understanding of foregoing problems on the line, it is possible to see where
there are problems in the service, trains stopping too long at stations, or unexpectedly between stations, or where there are long gaps between trains. The controller can then change the schedule in the system if these problems emerge. Therefore, continuing attention to the ongoing happenings on the line is an important part of maintaining a regular service and also dealing with crises and potential crises. The ATS diagrams, particularly the one usually displayed on the left screen which presents the whole line, are essential resources for keeping track of the state of the service.

Therefore, on many occasions, including during calls to other individuals, or when no call is being made, controllers will be oriented to the ATS1 screen. There appears to be a different character to such activities. These appear to be distinctive to the lookings at the screen at the beginning of a call or when particular problems are being dealt with on the service. Moreover, other participants, particularly the second controller (Cii) will also look towards the same screen. In these cases, they appear to be 'monitoring', not only the activities on the line, but the activities of the principal controller. Indeed, monitoring appears to be pervasive and a resource for forthcoming activities, particularly collaboration: glances to a common screen being an opportunity to begin an interaction between colleagues concerning a related problem, or for one to take some related action like change the schedule of the trains or make a call. Similarly, participants in the domain not only appear to 'oversee' the activities of their colleagues, but also overhear them. The phone calls as well as being available on the controller's phone are relayed through a loudspeaker above the controller to the rest of the control room. Thus, colleagues can get a sense of the state of the service and the actions the Ci is taking. Again, on hearing features of a call, participants in the control room, particularly Cii, will frequently be seen to perform some related activity.

It appears that these practices are critical for the achievement of activities within the control room. Not only do controllers appear to be continually scanning the local environment, overhearing another's phone calls and overseeing the activities of colleagues, but these can only be accomplished in the light of the activity at hand. Controllers have to be able to determine the relevance of their colleagues' activities in relation to the current state of the service and with respect to the particular trajectories that these activities project. Overseeing and overhearing, or
monitoring, then, in this domain as in others (e.g. Goodwin and Goodwin, 1996; Greatbatch, et al., 1993; Heath and Luff, 1992a), is an essential resource for collaboration and for getting the work done. The technology, the talk on the phone system and the displays available appear to be essential for allowing others to make sense of an individual's activity. There do appear to be distinctions which can be made between these activities and, say, 'focused collaboration' where one or more participants actively engage in some activity together, however, a stronger characterisation is harder to realise.

There are many activities which could be considered as monitoring: individuals may monitor a display, or the activities of their colleagues, and they may do this whilst oriented to that domain or elsewhere. There are a range of similar conceptions which can be utilised: participants may 'oversee', 'overhear' or 'listen in to' the activities of another; they may 'peripherally participate' in this conduct. These activities appear to make possible the performance of another appropriate activity, if something relevant emerges. Individuals can then accomplish tasks and activities with little explicit instruction or comment, and their colleagues can rely on them working in this way. It may thus, be apparent why 'monitoring', 'peripheral participation' and the like has attracted the interest of researchers developing systems to support cooperative work (Brown and Duguid, 1994; Robinson, 1993). It may be a desirable feature to provide for similar capabilities in these systems where the users may be geographically dispersed. Hence, video systems could provide a more general sense of 'awareness' for participants (e.g. Dourish and Bellotti, 1992), or Collaborative Virtual Environments could offer multiple ways for colleagues to 'interact' with one another (e.g. Benford, et al., 1996).

There may be problems, however, with attempting directly to apply observations made in a variety of domains (e.g. Goodwin and Goodwin, 1996; Greatbatch, et al., 1993; Heath and Luff, 1992a) to particular designs of technology. These observations have only begun to uncover how such practices are accomplished. Monitoring, and similar conceptions such as peripheral participation, serve as glosses for a wide range of implicit and tacit work practices that underpin collaborative work. Because they are peripheral or implicit, little more than a tentative analysis of these glosses can be offered, often in terms of later activities where it can be seen that a colleague must have overheard or overseen a prior activity of a colleague. Monitoring can, thus, appear to be both
indiscriminate and ubiquitous. It also appears to be characterised in these developments as a rather passive activity: an individual watching or listening to another to provide a 'general awareness' of another's activities (e.g. Dourish and Bly, 1992).

A stronger conception of monitoring could be considered. In the analysis of turns-of-talk in conversation, analysts have revealed how all parties to a conversation, at all times, have to monitor, or analyse, the emerging talk of the speaker for opportunities for self-selection or places for turn transition (e.g. Sacks, et al., 1974). Through these practices conversational contributions are managed from moment-to-moment and the activities of the participants are coordinated. Moreover, in one turn a participant displays an analysis of a co-interact's prior contribution, and, hence, their analysis or monitoring of the ongoing activity. Of course, problems can emerge, but these are managed, and repaired, through practices which rely on a similar sequential organisation: displaying an analysis of a prior in the production of a next turn of talk.

Several researchers have sought to explore the appropriateness of such an orientation to the analysis of activities with regard to computer systems. Greatbatch et al. (1993) examine the interrelationship between the talk between a doctor and a patient and the doctor's use of a computer system. Here, it appears that not only is the talk between doctor and patient shaped by the operation of the system, but that patients orient to the doctor's typing in order to coordinate particular features of their own vocal conduct. Whalen (1992) reveals how the use of a computer system for dispatching emergency services is tightly interrelated with the talk on the phone between the dispatcher and the person reporting the incident; the ordering and timing of the talk being shaped by the demands of the system. Heath and Luff (1996a) have analysed activities in a London Underground Control Room for the Bakerloo Line between the various members of staff who work there. In this study, particular attention has been paid to the ways technologies and displays feature in the collaborative activities of participants who have different responsibilities and resources available to them.

In each of these studies the researchers have been concerned with developing a sequential analysis of the activities of the participants with respect to the use of the technology. Greatbatch at al. consider how a patient pays particular attention to the details of the typing of the doctor, even hand movements surrounding specific keys. The occurrences on the
screen of the computer may not be visible to the patients, however the patients appear to tie their talk to the publicly available conduct of the doctor; the typing and movements of the hand across the keyboard. Whalen considers how the dispatcher's typing is coordinated to the audible talk with the remote party, and also how the dispatcher's talk is organised with respect to filling in particular items on the screen of the computer. However, in both these cases there is an asymmetry to the resources available to the participants, the computer system and the activities on it are invisible, or only partially visible to one of the participants. It is thus problematic to develop an analysis of the activities on the computer as sequentially relevant to the conduct of the other party: the patient has few resources with which to make sense of the typing of the doctor, the remote caller can only hear, on occasions, the typing of the keys. Hence, for the patient and the remote caller, the activity on the computer can only be considered as typing, hitting keys on the keyboard. They may discriminate between particular keys, like the 'return key' on the doctor's keyboard, but they have minimal access to the nature of the activity these actions accomplish. Similarly, for analysts the resources available to develop a sequential account of the activity are limited.

In the Bakerloo Line Control Room which Heath and Luff consider the participants have greater access to the conduct of their colleagues with respect to the technology. A glance at a large public display can be seen to be related to an ongoing activity, not only for analysis, but for the participants in the setting: a look to a display appearing to implicate others in activities with respect to particular happenings on that display. The participants may at one time appear to be engaged in distinct activities, and yet be attending to the conduct of their colleagues. An activity being later undertaken that requires an understanding, a foregoing 'monitoring', of a colleague's prior activities. Moreover, one participant may take a particular activity by a colleague, such as a turn to a phone system, as being related to their own, in so doing accomplishing a collaborative activity implicitly through interactional resources. Hence, activities by an information assistant such as pressing a button on a phone system, can be seen to be sequentially relevant to a controller's activity, an attempt to contact a train driver; the controller going on to provide the information assistant with what needs to be conveyed. The talk of one participant displays an understanding of an activity of another.
This chapter seeks to extend this form of analysis of artefact-based activities. The domain in question, and the activities explored, differ in several significant ways to those in the Bakerloo Line Control Room. Although the DLR and the Bakerloo Line Control Rooms are both settings from which an urban transportation system is managed, the activities explored here raise some distinctive concerns regarding the sequential analysis of computer-related activities. In the study of the Bakerloo Line Control Room, the attention focused on the collaborative activities between controllers and information assistants. Each have distinct responsibilities, a distinction similar to that between Control Room Supervisors and Control Room Assistants in the DLR Control Room; controllers responsible for managing the traffic on the line and communicating with driver, information assistants for making announcements to passengers and communication with other staff. Hence, controllers and information assistants have different technologies available to them: controllers have touch-screen radio phones, on-screen timetables etc.; information assistants have CCTV monitors and passenger announcement systems. So, in their collaborative activities the controllers and information assistants do not necessarily have the same commonly available resources. Hence, they can be seen to engage in practices to make their activities visible, or hearable, to a colleague. For example, controllers may design their talk so it can be overheard, or read out train numbers when they are changing the running order of a train. The particular activities being performed on the technology may not be available to the information assistant. On occasions, the controller and information assistant may look at a common domain: a CCTV monitor or a large fixed display of the line, however the activities they can perform on these systems are distinctive. For the other systems, the configuration of the control room and the division of responsibilities make those details of the activities of co-participants are unavailable, the technology inaccessible to operate, or particular systems inappropriate to use. The collaboration between controllers and information assistants relies on interactional practices through which their activities are made available to others. It is, thus, apparent why such details of the operation of the systems, such as the commands being entered and the changes on the displays, are not of concern within the analysis. These are not necessarily available to the co-present participant. The resources accessible to the controller and information assistant are distinct and so are their
opportunities and responsibilities for operating the technologies that surround them.

6.3 THE DIVISION OF RESPONSIBILITIES: SCHEDULING TRAINS AND THE PROBLEMS ON THE LINE²

At the beginning of each day the Automatic Train Scheduler (ATS) is reset and from then on the Docklands Light Railway (DLR) is meant to run to a preset schedule. It is this schedule which, by entering commands into the computer systems in front of them, the controllers alter repeatedly throughout their shifts.

The Docklands Light Railway system is an adapted version of similar urban transportation systems found elsewhere in the world. It is meant to be an automated system, with driverless trains, and appears to be based on similar ones originally designed for Canadian cities. Apparently, however, these cities do not have the same features as the Docklands area of London. In particular, the DLR winds around both new developments and old areas of the city, combining overland, overhead and underground sections. The twists, turns and contours in the track are alleged to result in the computer system in each train frequently miscalculating the distance required to get the train from one station to the next. This results in trains missing the docking unit at each station. At the time of recording (May 1991), when the railway had been operating for nearly four years, the problems with docking and their 'knock on' effects, were the principal reasons for controllers having to alter the schedule in the ATS.

After a problem in getting the train to dock correctly, Control Room Supervisors (or 'controllers') will typically turn off Automatic Train Operation (ATO), allowing the train driver to shunt backwards or forwards to the correct location (this mode being called 'Emergency Shunt', or EMS) and then reset ATO. This operation causes delays not only to the train in question, but to others on the line behind it, requiring further changes to the schedule. These problems may also account, in part, for why controllers comment that the routes set by the ATS are less than optimal and in frequent need of alteration. To maintain a steady flow of traffic along the lines of the DLR, controllers not only have to set

² Appendix A gives further background to the use of the systems considered in this Chapter and to the analyses presented.
particular routes manually for trains, but they also have to alter their eventual destinations.

Apart from problems with the scheduling, train captains and other personnel can request changes to the service. Train captains may require a 'nature break', or ask for an early 'turn around' towards the end of a shift, engineers can ask for trains to be restricted from operating on portions of the track ('blocks') and managers may require that specific trains, or a particular number of trains, are put out on the track for strategic reasons. These varied reasons and requests for alteration place considerable demands on the two controllers, involving them in dealing with the resulting knock-on effects and further communication with other staff. Furthermore, alongside this work of altering the schedule they still have to manage the traffic on the line and deal with the various crises that emerge.

To cope with these demands, the range of activities and responsibilities are divided between the controllers. At first glance, this appears to be quite straightforward and related to the location of the technology surrounding them. The controller on the left, for the purposes of this chapter identified as the 'principal controller' (or Ci), operates the radio-phone system, as seen in the last chapter, and also usually schedules traffic on the main lines. The controller on the right, the 'secondary controller (or Cii), usually operates the traffic in the depot (the OMC)\(^3\). In the control room the individuals who sit in the Ci position are referred to, by themselves and their colleagues, as the 'Gods of the Line'. Despite tending to maintain their relative locations, this arrangement does not appear to be fixed and rigid, Cii occasionally swap seats with the Ci and take calls. Moreover, there is some ambivalence to which controller performs which activities: the Cii also schedule trains on the main line, and the Cis trains in the depot. The management of these responsibilities and activities are considered at length in the following sections of this chapter.

Given the general distribution of activities between the controllers, it is normal for the two controllers to have different views of the line presented

---

\(^3\) This is given the name, 'Operations and Maintenance Centre' (hence the label OMC on the screens) and is located between the Poplar and All Saints stations. However, in their interactions the members of staff also call it the 'depot' or identify it through the names of particular lines around the Centre.
on the screens in front of them. Thus ATS1, the screen in front of Ci, often displays a schematic map of the line (Figure 6.1)

![Figure 6.1: The ATS1 display and a schematic diagram of what it presents. The ATS diagram transforms the geography of the line. The bottom most line is south of the two upper ones. The middle one goes to the Northeast and the upper into Central London (North West). The arrows to the left of the schematic diagram are meant to show the direction along which trains move along the relevant portion of track.](image)

ATS2, the screen in front of Cii, often displays the shunting yard (Figure 6.2).

![Figure 6.2: ATS2 display of the depot and, on the right, a sketch of what it typically presents.](image)

In each display, above the diagram there is an area where general comments and warnings are presented. There is space available for 5 lines of text, which either scrolls or is cleared by the controllers using a command key. Typical messages relate to Automatic Train Protection.
(ATP) and manual blocking. When ATP is set on a train, the portion of track behind it is automatically blocked for use by another train. Controllers can also set blocks on portions of track so, for example, engineers can work there. Messages note when a train’s route is scheduled to use one of these blocked portions of track.

Below the map the controllers’ typing is presented. This is entered against prompts such as ‘TRAIN AT’ or ‘PLATFORM’. These are displayed when the appropriate text is required, prompts and text appearing on a single line (see Figure 6.3).

```
TRAIN AT ...
TRAIN AT S..
TRAIN AT ST.
TRAIN AT STR
TRAIN AT STR PLATFORM
EXIT LOCATION BOC SET MODE _
EXIT LOCATION BOC SET MODE 1_
EXIT LOCATION BOC SET MODE 12_
```

Figure 6.3: Example Prompts and Commands. The controller's typing is shown in bold. In the first example, the abbreviation for Stratford (‘STR’) is typed. When it is completed the prompt ‘PLATFORM’ appears. In the second example, after a location is specified the stretch of track is required, entered as a number. Station abbreviations are always three characters and platforms one. A portion of track is identified by a packet of three numbers, but these packets can be joined together to apply to longer portions of track.

Although it is common for the ATS1 and ATS2 screens to present the main line and the OMC respectively, they can be used to give other views, for example, the points and signals in a particular area such as a station, a track where trains can cross over, or a complex intersection. The displays, therefore, present a configurable view of the operation of the service, which can be altered to fit the particular circumstances that controllers face.

Although, the ATS system is complex and offers a great range of functionality, the most common activities performed on it by the controllers are altering the direction, mode and destination of trains. The commands which are entered into the ATS are passed to the docking unit at the appropriate platform. This means that in order to alter the route of a train the controller has to enter the command either whilst the train is docked at the appropriate station or before it arrives. The timing of the
commands is thus very important. It is therefore, usual to direct alterations to a station some way ahead of the relevant train, although of course, only once the previous train on the same line has departed from that station. Altering a route of a train in one of the four terminals (Tower Gateway, Stratford, Crossharbour and Island Gardens) or turning a train around tends to be done whilst the train is docked.

The commands typed are mainly made up of track and station identifiers, each consisting of three character combinations. Station identifiers are the common abbreviations for each station found on most of the computer displays in the control room (see for example, Figure 6.1, above) and portions of track are identified by three number combinations. As both track section and station identifiers have three number or letter abbreviations, commands have quite a regular format.

Commands for changing a route commence with a special function key to the right of keyboard, followed by several 'packets' of three characters and finally completed by an 'enter' key. The first 'packet' consists of the relevant station name, and the remainder are made up of one or more packets of numbers, identifying portions of track that specify the new route. The beginning of the second component (and thus, the completion of the first) is marked by a function key on the right hand side of the keyboard, and the entire command is completed with the 'enter' key. So, for example, a controller may type `STRA 422 J' to change the route of the train at Stratford to use the section of track '422', or may type `CAW 412 453 J' to alter the route of the next train entering Canary Wharf to

---

4 See Appendix A (Section A.2) for a general overview of the DLR and a list of the abbreviations of the stations.

5 To enter commands each controller has a conventional extended computer keyboard with a normal typewriter layout, a numeric keypad, one set of function keys along the top and another down the right hand side. These function keys allow the controllers to carry out such functions as changing which map is displayed, clearing the incoming messages that have been accumulating at the top of the screen (the keys on the top row) and commencing various command sequences (the keys on the right).

6 In the transcripts the typing of function keys are indicated by 'N', number keys by '00', letter keys by 'Ä' etc., 'enter' by 'J', and other (and unidentified) keys by 'U'. Keys that are struck with more force are underlined. Pauses between character presses are indicated by '.', each of these marking approximately 0.1 seconds.
now use sections ‘412’ followed by ‘453’, as in the following examples taken from the data.\footnote{Because of the lack of clarity of the data, some characters typed into the system, particularly the specific numbers, are not transcribed. Appendix A (Section A.3.1) outlines the approach taken for transcribing the typing of the controllers.}

(1) 42ii (10A)

```
command initiator function completer
Cii: D E R . . . . . . . . . . . .
```

<table>
<thead>
<tr>
<th>train identifier</th>
<th>track identifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Devons Road)</td>
<td></td>
</tr>
</tbody>
</table>

(2) 13ii (1A)

```
command initiator function completer
Cii: C A W . . . . . . . . . . . .
```

<table>
<thead>
<tr>
<th>train identifier</th>
<th>track identifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Canary Wharf)</td>
<td></td>
</tr>
</tbody>
</table>

A slight variation of these commands applies to the OMC. There locations, both starting and route parts, can be identified by numbers or by two letter abbreviations, commencing with a ‘D’ or an ‘S’ (for “Depot” or ‘Shed’).

Changing the mode of the train, to change to and from ATO, for example, is very similar. The appropriate command key is pressed, then the station where the change is to commence and finally the track to which it applies.

It is noticeable that the typing by controllers has an apparent consistency, both in the articulation of the activity, and to the weight with which keys are hit. Most commands consist of rapidly typed combinations of character or number keys, commenced, interspersed and completed with single key presses often hit with more force. There are other ways in which a ‘standard pattern’ of typing is produced. The three character station and track codes can be heard as typed in ‘packets’ not only because they are quickly typed together, but also because they are hit with equivalent weight. The function and completion keys ordinarily are typed with more force and sound louder.

As with most computer systems the structure of the commands entered by the controllers is fixed by the design. However, there are a great many
ways in which these commands could be articulated. Certainly, the layout of the keyboard contributes to the articulation produced. If only one hand is used to type, then it has to travel some distance from the character keys and the top function keys to the numbers, the side function keys and the 'enter' key. The structure of the commands also contributes to the articulation of the typing. For example, routes consist only of sets of three numbers, so can be quickly typed by the right hand on the separate number keypad to the right of the keyboard. Station identifiers need to be typed on the main keyboard, often using both hands. However, with only a few stations, the abbreviations can be short and entered fairly rapidly.

Both controllers, at times, enter commands into the ATS, either to the main line or to the depot. This appears to be achieved despite little explicit communication concerning the commands between the controllers. They rarely appear to ask one another to type in a command or correct a command made by their colleagues. Hence, the management of the work of scheduling, the division of labour between the participants, on the whole appears to be unproblematic. In the following sections, consideration is given to the ways in which the production of the ATS commands are organised, as part of the activities within the control room, and how these contribute to the management of responsibilities between the controllers.

6.4 A PRACTICAL DIVISION OF LABOUR: RESOURCES FOR THE MANAGEMENT OF ACTIVITIES WITHIN THE CONTROL ROOM

As mentioned in the last section, pairs of controllers tend to work with the two screens in front of them presenting different displays, typically ATS1, in front of the principal controller (Ci), displaying the schematic view of the line, and ATS2, in front of the secondary controller (Cii), the depot. This reflects a practical division of labour between the two colleagues: Ci handling calls on the phone and controlling the main line, and Cii controlling traffic in the depot and communicating with the Control Room Assistants (CRAs) further along the desk. Therefore, when phone calls are taken by Ci concerning the depot, it is not unusual for Cii to then enter some command into the ATS.

In the following instance, Ci is takes a call that concerns the depot. The controllers have the 'standard' arrangement of displays in front of them.
Although the actual reason for the call is unclear, it relates to a train in the depot: not only does the member of staff mention the 'depot tester' (line 4) but also 'DP' (line 3) is one of the locations in the depot where trains are 'stabled' (see Figure 6.2). Ci hears this utterance as a request for an action and confirms it.

Once Ci has confirmed the train movement from the location 'Dee Pee' to the 'depot tester', on the word 'yep', Cii starts to type a command.

Calls to the control room frequently require more than a simple confirmation from the controllers through talk. They may also require, for example, the controllers to change the mode of the train, to revise its route or to alter its destination. To accomplish all of these requires commands to be entered into the ATS. Either of the controllers can do this. However, the resources for taking an action are available to each in marginally different ways. The principal controller can hear the call through the handset as well as the loudspeaker, but, at the beginning and end of calls he has to, whilst holding the phone receiver with his left hand, manipulate the computerised phone system with his right. So, it appears that although the call is more accessible to Ci, on occasions it may be more appropriate for Cii to enter the command. One such occasion may be when the second controller is not otherwise engaged and has the appropriate display presented.

The repetition of a request, or part of a request, by a controller can, therefore, serve a double duty (cf. Turner, 1976). It can confirm that the request has been heard and will be acted on to the remote party. It can
also provide resources for a colleague to take the appropriate next action. In (3), Cii's typing begins just after the confirmation 'Dee Pee to depot tester'. This confirmation being an abbreviated version of the remote party's request, it displays that the relevant details concerning the starting and finishing point, and the direction, of a train movement have been heard. This is also just the information needed for the Cii to enter a command.

In the following instance, another request comes in for a train movement within the depot.

(4) 27iiCo - transcript 1

1  Ci: three four::?
2       (2.3)
3   E: yeah three four: to (line one)↓ (. ) can you give me
4   a route please from Dee Arr↓ down to Dee Ess (0.2)
5       o:ver?
6       (0.2)
7  Ci: thats from Dee Arr: to Dee Ess: (. ) thats-er (well)
8       received↓ (Julie ) over?
9       (0.2)

The nature of the request makes it hearable as concerning the depot, again because the particular locations given identify places within the OMC (i.e. DR and DS, in line 4). Following this request the primary controller (Ci) confirms the traffic movement. After Ci's confirmation (lines 7-8), Cii commences typing a command into the ATS.

(4) 27iiCo - transcript 2

ATS1

Ci: Ess:-thats-er (well) received↓ (Julie ) over?
E: ( ) over
Cii: .. ý.. Q QQ.. ý
ATS1 . . . . . . . . . . . .

As this command is entered only using keys on the right hand side of the keyboard (the function keys for initiating commands and the number keys), it appears to be one directed at the depot. As in (3) the typing of Cii appears to be related to the talk between Ci and a remote party on the radio. Ci repeats particular details of the call providing resources for both the remote party and Cii. These details are relevant to the production of the command - the starting point (DR) and the end point (DS) of the movement. Although Cii starts typing only after Ci moves into closing the call, his activity before this appears to be related to the call. Just after Ci gives the confirmation, Cii turns away from looking at the display in front

248
of Ci and towards the one in front of his own position. As his glance arrives at the keyboard he starts typing.

It appears then that the confirmation may provide resources for both remote and local colleagues. As mentioned in the previous chapter, there are various formal requirements for the talk on the phone: one of which is to confirm actions that are to be taken. These, however, can be adapted to the local circumstances, for the purposes of the controller and the possible circumstances of the remote party. These confirmations may also be relevant to others in the control room, particularly the second controller not on the phone. This would appear to be a conventional case of monitoring, Ci’s listening in to the calls to hear what is relevant for themselves. This monitoring is not arbitrary with respect to the talk on the phone. In (3) and (4) the Ciis do not merely respond to a call being about the depot, but only commence a movement towards typing once it has been confirmed by the controller.

In both cases the principal controller goes onto another activity. It appears that, although they confirm the changes on the phone, the principal controllers treat the second controllers’ typing as adequate to deal with the matters on which they have been talking about on the phone. Hence, they are also sensitive to the activities of their colleagues. At the least, they see the typing as relating to their own talk. It may be that the timing of the conduct of the second controllers is relevant; that a turn towards the system and then the typing can be seen to be relevant because of its positioning vis-a-vis the talk. It may be that Cis also can utilise as resources the general features of the typing revealed in the last section. Though they may not be oriented to the screen of their colleague, they may monitor the gross features of the typing activity, not only the general area where they are typing, but what is hearable in the typing. Whatever the case, they treat their colleagues’ activities as appropriate to the call: they can also be considered to be monitoring their colleagues’ activities. In this sense, monitoring in the control room can thus be seen to be ‘mutual’ (cf. Heath and Luff, 1992a).

In both (3) and (4), a call comes in over the phone, and related activities are produced by both controllers: a confirmation by the principal controller and a typed command by the second controller. Calls to the control room frequently implicate such a ‘confirmation-command’ package. In (3) and (4), each is produced by a different individual. This distribution of responsibilities is accomplished unproblematically. These activities are
accomplished by both hearing what is publicly available on the phone and talking appropriate next actions. They also are managed through both participants making use of the activities of the other: Ci relying on Cii’s typing to be appropriate and even producing turns of talk to a remote party that, at least, can be heard to be relevant by a co-present colleague.

In the following instance, Cii starts typing sometime after the confirmation by the controller. Indeed, his typing follows the remote party’s final turn in the conversation (line 9).

(5) 5Co - transcript 1 (simplified)
1  Ci: two ni:ne
2     (.)
3  E: (   )
4     (1.2)
5  Ci: thats (a roger::) (fully at greens when I get them
6     from you) over
7     (1.5)
8  E: (   )
9     (0.2)
10    ((Cii starts typing))
11    ((Ci puts phone down))

As in (3) and (4), the call appears to concern the depot, but the precise nature of the reason for the call coming through the loudspeaker is unclear. Ci appears to orient to this, and perhaps also Cii’s inactivity, and, as the remote party begins their reply (line 8), points to two places on the screen. Following this, Cii types a depot command (using the right hand function key and the right hand numeric keypad).

---

9 The confirmation of a train route involving following ‘green specs’ or ‘greens’ tends to be directed to trains in the depot.
Ci’s gesture consists of two components and is quite extensive, the first being to the top of ATS2 (image 5.1), the second to the bottom right (image 5.2). The length of the command, which continues beyond this fragment for another nine characters, is also quite extensive. The pointing and typing appear to be tightly related, the former specifying the nature of the command to be produced on the keyboard.

As the gesture is produced, Cii glances down at the keyboard. He begins to type once Ci’s second pointing has been completed. It may be that the gesture alone is sufficient to produce an appropriate keyboard command from Cii. However, the point does not touch the screen, and even though close, would not specify a complete and precise route. Indeed, whilst Ci is pointing to the second location Cii is looking down to the keyboard. It appears that the point is made with regard to the call, to make sense of the foregoing talk and to outline relevant details of the foregoing call for Cii to take further action. This Cii does, typing a command into the system immediately after Ci’s gesture. Thus, in this case, Ci appears to orient both retrospectively to the hearing of the call for
Cii, and the potential problems in it, and prospectively for Cii to take an appropriate next action. Ci not only orients to the possibility for Cii to take next actions consequent on a call, but is also sensitive to the resources he utilises in order to produce these.

In (3), (4) and (5) a phone call from a member of staff to the principal controller requires actions to be undertaken by control room personnel. In each case, the activity relates to the depot and requires both a confirmation and some change to be made to the system, and in each case Cii types into the ATS a command to move traffic in the depot. In (3) and (4), Cii appears to utilise resources in the talk between Ci and a member of staff. In (5) there are problems with the reception of the call, and Ci's pointing appears to be designed with respect to the practical problems faced by a co-present colleague. Hence, it appears that both participants are orienting to a practical division of labour, where it is relevant. Once an activity which relates to the domain of Cii is confirmed, Cii types in an appropriate command into the system. Ci's appear to orient to this division of responsibility: when a command is typed he (or she) goes onto another activity; when it is not they provide resources for the Cii to type an appropriate command. For a Cii's activity to be seen as relevant it has to be timed with regard to the Ci's activity. This may be by commencing to type or by moving into a position to begin typing.

In each of these cases, the call requires a confirmation and a train alteration and these activities are distributed between the Ci and Cii respectively. However, these activities are not performed in isolation. For Cii to produce a relevant next action requires the skill of recovering the relevant details from the call in order to take that next action. It requires an active listening to the call, for what may be relevant. Moreover, it appears that Ci can contribute to this practice, designing not only gestures to help accomplish this listening, but also talk with a remote colleague to make the relevant details hearable. Ci's are also sensitive to the activities of Cii, taking account of when an apparent 'next activity' is not forthcoming. Thus, in different ways the two controllers appear to be orienting to a distribution of activities and responsibilities between them. This division of labour between the two controllers may appear quite straightforward, and reflected by their selection of views on the screens in front of them. However, to accomplish appropriate actions requires work from both controllers, involving not only a sensitivity to the responsibilities of a colleague, but also to the particular circumstances
they are facing. Although the activities of both controllers can be considered in terms of monitoring their co-participant, this characterisation appears to neglect their engagement and active participation in the conduct of their colleagues.

6.5 PRACTICES THROUGH WHICH CALLS ARE HEARD AS RELEVANT

Despite the controllers' orientation to a division of responsibilities between them, it is noticeable from the materials that Cis do not always work on the main line and Cis on the depot, and that often they do have other views of the system displayed in front of them. For example, in the next fragment, the controllers have their views the other way around: ATS1 has a view of the depot, and ATS2 the main line. As a call comes in from a remote party, Cii starts typing.

(6) 10A 1:00:26 (43) - transcript 1
5  Ci: transit base receiving from transit one five go
6   ahead over
7   (0.5)
8  RP: one five ( try) and pick(ing) up at (route) one
9   one (. ) six at (0.2) Poplar over
10  (0.7)
11  Ci: much appreciated thanks Chas (0.6) (nice to have
12  you:)
13  (1.3)
14  (Ci puts phone down)

Cii's typing commences as the remote party completes the 'reason for the call'.

(6) 10A 1:00:26 (43) - transcript 2

6.1 Cii/Ci 6.2 Cii Ci 6.3 Cii/Ci

body movement to PS, hand to mouse
body movement to ATS2

Ci: ATS2 ATS1 PS ATS2
RP: pick(ing) up at (route) one one (. ) six at--Poplar over
Cii:

PS    ATS2
hands down to keyboard
Indeed, she reorients from the phone system display, to the left of Ci, to the ATS2 screen in front of her just after the caller is giving the route number 'one one (. ) six'. Ci also reorients to ATS2 at this point, and appears to be looking at Cii's screen as she types. As the reason for the call emerges he has been looking first towards ATS2, then ATS1 and then the phone system display. So, whilst the caller is giving the route number, Ci is turning towards the phone system, glancing at ATS1 on the way. As part of this movement, his hand reaches towards the mouse which is used to operate the phone system, arriving as the caller utters 'at'. This movement appears to prefigure a closing of the call.10 However, Ci then suddenly moves back to the right, towards ATS2, and Cii moves her hands down to the keyboard to type. When the caller gives the location 'Poplar' both controllers are looking at ATS2 and Cii is beginning to type. Some feature of the call appears to be consequential to this coordinated movement of both controllers: Ci arrests his movement towards the phone system and Cii begins to type. It is as if the route given in the call, 'one one (. ) six' demands some action from the controllers.

The command the Cii types is concerned with Westferry, the station next to Poplar if going East to Stratford (see Figure 6.4).

As mentioned in Chapter 5, the phone system queues the calls coming in and the mouse is used to select further calls on the queue. As it both identifies the name of the callers and their trains, the phone system display can be used at the beginning of the call to locate the caller (in conjunction with the ATS display). It can also be used prospectively, to provide resources for making sense of calls about to be taken. The reorientations of Ci in (6) between the screens have the character of those considered in Chapter 5, of being utilised to make sense of an emerging incoming call. In this case ATS2 presents the display of the main line and can be used as a resource to locate trains related to the call.

10

---

Figure 6.4 A sketch of the train movements considered in fragment 6. In line 5 it is stationary at Westferry (WES). It is on a route heading towards Stratford (STR) in the East. The caller is at Poplar (POP).
Some time before the incoming call a train captain has reported a problem with his train and Ci commences a course of action to 'stable' the vehicle in the depot (the OMC). As well as the failure of the train there also appear to be problems with the communication between the train captain of the failed train and the controller. When he reported the failure, his confirmation of the action he should take, itself a repetition of a statement by Ci, has to be repeated again. At one point in that call, Ci throws his phone down on the desk. The controller then commences a course of action to 'stable' the vehicle in the depot.

(6) 10A 57:32 (43) - transcript 3

Ci: on arrival Wes(t) Ferry report one three seven set A:
    Tee Manual ( clear to) Poplar (0.2) once at Poplar ( ) O eM Cee. over.

About forty seconds before the call given in transcript 1 a train can be seen to arrive at Westferry. This remains there throughout the course of the call. It appears to be the target of Ci's command. Indeed, just following Ci putting down the handset of the phone (line 14), the train identifier moves along the track eastward (to the left on this part of the screen) towards Poplar. The timing of this movement, following its long stay at Westferry, the typing of a command directed at that station, and the subsequent departure of the train are consistent with its Eastward movement. It also appears to be consistent with the earlier instructions of Ci to the train captain to move the train towards Westferry then through to Poplar to the OMC.

It also appears that the call may be related to this train. Although the remote party's reason for the call is not completely clear, the initial turns of the call are rather distinctive:

---

11 When a train docks at a station it is displayed on the ATS by the block representing the station changing colour to yellow. This can be visible from the video-recordings. Similarly, particular train movements can be seen. These are displayed by a white identifier for a train (its number and destination) proceeding along the track. Track sections which are blocked by the ATP follow the train's location. These can also be occasionally seen on the collected materials.

12 Further resources for analysing the train movements on the line are given in Section A.3.2.
First, in the call sign 'transit base receiving from transit one five go ahead over' the number of the recipient is given (line 5). Second, the remote party reports an action being undertaken that he is undertaking '(try) and pick(ing) up'. He also gives a location, not named but numbered (lines 8-9). This is treated unproblematically by Ci, who goes on immediately to move towards closing with 'much appreciated thanks Chas' (line 11). This closing turn is a third, distinctive feature of the fragment: the caller is named by Ci and '(nice to have you:)’ is also unusual. These features reveal the call to be from a ‘mobile’ CRS or a ‘transit’. These controllers travel around the system assisting where appropriate. They can listen in to the radio traffic and take actions accordingly. It may be that the ‘mobile’ in (6) has heard the previous talk on the radio, particularly with respect to the radio communications and is intervening in the light of these.

Hence, the remote party's talk on the radio can be heard as directly consequential to the participants in the control room. His utterance appears to provide the resources for the controllers to revise the system and set the train at Westferry moving. In this case, Cii is attending to the details of the call. It is not just that talk is monitored for matters of particular concern to a controller, say the depot or the main line, and once these are confirmed Cii can then enter the appropriate command into the system. The Cii has to attend to the details of the call: who is making it, what actions they are requesting or making. In (6) the call can be heard to be distinctive through the particular organisation of the talk and the nature of the turns that are produced. It may be that the call sign and

---

13 Without clear radio communications it is unclear whether the train captain has heard the changes to the operation of the train. The controllers may then be holding the train until more robust contact can be made. The radios have batteries and these can run out. One solution is to get a new radio to the train captain. As the mobile controller appears to propose to pick the train up at Poplar, Cii's command may be to allow the train to proceed. It is unclear what access the train captain on the train at Westferry has to these actions.
even the voice of the remote party can be recognised by the participants in the control room, but his contribution has to be heard as relevant to the circumstances at hand: the particular happenings on the line. In (6) an activity at Poplar is heard to relate to a train at Westferry. This requires the talk on the radio to be heard in relation to the current state of the service. Once the mobile states where he will be picked up, Cii can enter the command into the system to start to move the train at Westferry. The typing of the command is tied to the talk by a remote participant to the call.

However, characterising this accomplishment merely in terms of Cii overhearing the call between the other two controllers neglects how the activity is produced. Ci also marks the call as distinctive. From his first turn, he produces distinctive features that display an orientation to the nature of the call, not only for his co-interactant, but also his co-present colleague. These distinctive features may serve prospectively to mark how the upcoming call can be heard by a colleague. In this case the call is heard by Cii as relevant to her and implicative of some action to be taken. This hearing, then, is accomplished through the collaborative and interactional activities of other participants. Ci then goes on to another activity, Cii's typing of a command being seen to be appropriate and relevant.

It may also be noted that Ci's typing could be considered to be coordinated with the talk on the phone. In this instance, and in (3) and (4), it appears to be tied to particular components of the talk; either a confirmation or specific route number. But it is not simply a coordination between typing and talk, but of activities concerned with the scheduling of traffic on the line. It is the particular command that appears to be appropriate and relevant to particularities in the talk. The talk has a sequential relevance, requiring particular actions to be taken, not only through talk (as in a confirmation or a denial to a request), but also through actions on the computer system. Through the actions of the participants the commands are produced and displayed to be appropriate and relevant, changes to the schedule are accomplished collaboratively.
6.6 THE INTERRELATIONSHIP BETWEEN TALK AND THE PRODUCTION OF COMMANDS

From observations in the cases considered so far, it may appear that although they attend to each other, both controllers are primarily concerned with distinct activities: Ci with the phone and Cii with the ATS. It may also appear that the activities between the two controllers are rather loosely related, one participant listening in or watching the other. In this section, the activities of the controllers are considered in more detail; although they appear distinct they are tightly interweaved. In Section 6.6.1, the ways in which the principal controller engages in the activity of the second controller are examined. In Section 6.6.2, the ways in which the second controller attends to the activities of the principal controller are considered. These two domains of interest can both be considered in terms of one controller monitoring the other. However, in order to make sense of another's activity this has to be seen in relation to a broad set of resources available in the control room and emerging in the setting.

In Section 6.6.3, the entry of commands are explored in relation to the ongoing activities of the participants. It appears that the talk between participants on the phone and the typing into the system can be tightly coordinated. The production of one relies on the production of the other. The accomplishment of a command even though entered by an individual is achieved through the collaborative activities of the participants.

6.6.1 Engagement in Multiple Activities: Cis Monitoring Cii's

In the preceding fragments, it may appear that second controllers' sole responsibility is the system ahead of them, and in order to take any action they have to wait for appropriate phone calls to arrive. However, they do have other responsibilities, for example, along with principal controllers they have to maintain a record of significant changes to the system and of any failures and delays. They also tend to talk with CRAs about problems of common interest, particularly concerning matters which may have a relevance to the crewing of trains. Similarly, Cis have other responsibilities. There is a wide a range of other staff who may need information or assistance from the controller concerning the running of the line. As the following fragment commences, Ci and Cii are engaged in apparently distinct activities, Ci is just completing a conversation with an engineer whilst Cii is answering a request from a CRA. On the CRA's
repeat of the request to Cii, Ci then goes on to take a conventional call from a train captain.

(7) 10A 03:26 (42) - transcript 1 (simplified)

8 CRA: how long are you keeping that A: Tee Pee train there?

9 Ci: base (to) train seven one: (0.1) go ahead over?

10 TC: (Stratford West India Quay over)

12 Ci: Select emergency shunt<proceed forward in that

15 mode<dock at Poplar routeboard two zero seven↑ (0.2)

Whilst Ci takes the call, Cii answers the CRA and then types some commands into the ATS. On completion of the call, Ci turns to her and says "ehhm (0.2) I was going to say: <let the one (one one) go↑" (lines 30-31).

(7) 10A 03:26 (42) - transcript 2 (simplified)

25 (that's a rog) thank you↑ (0.2) ( over)

27 (1.7)

28 ((Ci puts phone down))

30 Ci: "ehhm (0.2) I was going to say: <let the one (one one) go↑ (0.2) ( run through )"

Cii then goes on to type in another command. Ci's utterance appears to be related to some aspect of Cii's prior activity. In the course of the phone call Cii types three commands. In order to investigate what Ci is attending to in the conduct of Ci, it is important to examine in more detail these particular commands and the state of the service at the time of the call. The first of these is typed when the train captain is giving a reason for the call. It is directed to Heron Quays (HEQ).
The second and third commands are typed during the train captain's confirmation and concern Devons Road.

When the call commences there are only three trains on the Red Line between Canary Wharf and Stratford: the train that is just about to leave Stratford going south (the subject of CRA's talk), a train stuck at West India Quay going north (the subject of the call to Ci), and another train going north on the 'delta' towards Poplar. Figure 6.5 is a sketch of the state of the main line diagram when the call commences (line 9).

Figure 6.5  The trains on the North of the Red Line at the beginning of the call in (7).

Ci'i's first command to Heron Quays appears to be addressed to the train that will follow the one currently at West India Quay, and the commands to Devons Road to the train just leaving Stratford. This is the train which the CRA asks Ci'i about (line 8). The CRA's question appears to be in the light of a conversation he is having on the phone. He then

---

14 This is the complex area of tracks between Poplar, West India Quay and Westferry. It appears on the left of the ATS main line screen, in the rough shape of a 'D'.

15 At the time of recording Canary Wharf station was not fully operational. This meant that frequent ATS commands were entered with respect to it, either addressed to CAW or the next stations along the line (WIQ and HEQ). Canary Wharf is also unusual on the DLR in having more than two tracks running through it. This means that there are alternative routes through the station. Other resources utilised to make sense of the happenings on the line are given in Appendix A (Section A.3.2).
turns to Cii and asks ‘(how long are you going to keep that) A:: Tee Pee train (out)’.

(7) 10A 03:26 (42) - transcript 5 (simplified)

CRA ((on phone)): hang on a minute ma:

(1.3)

CRA ((to Cii)): (how long are you going to keep that)

A:: Tee Pee train (out)

(0.3)

Cii: What?

(0.1)

CRA: how long are you keeping that A: Tee Pee train

there↓

(0.2)

Cii: there is a train (leaving ) (0.2) all the ones↑

(0.2)

→ CRA: (oh there is one↓ )

(0.5)

CRA ((on phone)): ( ) leaving Stratford now↑

After 0.3 seconds Cii replies with ‘what?’, the CRA repeats, with a slight transformation, his request. Cii replies, pointing to the screen between them (ATS3),\(^{16}\) as she utters ‘all the ones↑’. The CRA moves closer to the screen, also pointing at it. He utters a statement like ‘(oh there is one↓ )’ before returning to his phone call. The CRA’s initial utterance appears to orient Cii to the issue being discussed in his phone call, she gives a response apparently adequate for the CRA to return to the call and reply to his interlocutor. Their pointing gestures and their talk both appear to be directed at a train in the Stratford area of the screen. Indeed, shortly after this utterance the train leaves Stratford (displayed on ATS2). In order to stable the train, say after some problem, it would be necessary to set the instructions at one of the stations before Poplar. As the train has left Stratford, Devons Road is a suitable candidate station.

It also appears to be this train (‘all the ones’) which the Ci is referring to when he completes the call with “ehhm (0.2) I was going to say: <let the one (one one) go↑” (lines 30-31).

---

\(^{16}\) This screen, which appears to be only used as a display onto the system for the CRAs, is positioned at the end of the controllers’ part of the console, to the right of Cii (and, thus, to the left of CRAi). It usually displays a view of the main line.
As he says, 'I'd put it back into auto:' Cii begins to type, again this is to Devons Road.

Although he has been engaged with the engineer and then the call, Ci's suggestion appears to be related to the train in the discussion between Cii and the CRA and the one she has just altered. Therefore, it appears that whilst engaged in these other activities, Ci is also sensitive to the activities of Cii. Although Ci's utterance 'ehhm (0.2) I was going to say:<let the one (one one) go' may be solely in the light of the conversation between the CRA and Cii, it appears to be designed with respect to the typing of the command, hence the turn being prefigured by 'ehhm (0.2) I was going to say'. Therefore, Ci appears to have been attending to the details of Cii's activities; not only that she types commands, or even that she makes a change with respect to an enquiry by a colleague, but that she makes a particular change in the light of the present state of the service. Ci picks up from the stream of Cii's typing that she has made a change at Devons Road, and this relates to the train coming from Stratford.

In (7), from early in the call Ci adopts an orientation where he can see both ATS displays and the consequences of Cii's typing. Hence, whilst he is engaged in one activity, the call, he can observe what is visible and public in another, the Cii's typing. However, making sense of the typing depends on it being recognised as part of a situated activity, particularly with regard to a particular train and a particular course of action. With respect to the case at hand, making sense of the course of the activity which Cii is engaged in relies on seeing her typing in relation to the happenings on the screen, to the state of the service and in the light of prior activities with others.
6.6.2 Exploring the Scene for Practical Problems: Ciis Monitoring

During the call in (7) the typing of Ciis does not appear to be directly related to the talk of Ci. It is related to Heron Quays and Devons Road stations whilst the remote party in the call is at West India Quay.

From Ci's reply, the train appears to have docking problems at West India Quay. The usual procedure would be to get the train captain to shunt the train to the next station (Poplar if he is going north on the Red Line, lines 14-15) and there set the train back into Automatic Train Operation (i.e. back to the schedule set by the ATS, lines 17-18). Ci also takes the opportunity to inform the driver that they will be turning the train around later at Bow Church, a station north of Poplar, just south of Stratford (lines 18-19). The train captain appears to repeat and confirm this information (lines 21-24).

Whilst this call is underway, Ciis is typing three commands. The first, during the train captain's initial report, being directed at Heron Quays. Following this, Ciis successively glances between ATS2 and the phone system display, her hands resting on the edge of the keyboard.
After a look towards the phone system display, Cii returns to ATS2. Ci then begins to speak, and as he begins to talk Cii quickly glances between the two displays of the phone system and the ATS in front of her, looking at the phone system on another two occasions.

It is as if, after being engaged with the CRA and the command to the train coming into Herons Quay, Cii is looking for resources with which to understand the call - the identity of the speaker and the location of the train. There are two trains to which the call could be related, the train at West India Quay and the one coming in to Poplar. If it is the second, then an action may be required fairly promptly.

It is only once the gist of Ci’s utterance is well underway that does Cii becomes less animated, and her hands move up from the keyboard, on which she then rests her chin.

The two screens would provide resources for locating the caller and the train. The phone system screen presents the number of the caller and the
set (i.e. train) he is on. The ATS displays where that set (or train) is. Moreover, as the turn emerges, the nature of the problem, its solution and its location becomes apparent; selecting emergency shunt at Poplar for a train at West India Quay does not demand an immediate action from the controllers. To make sense of a call requires hearing it relation to the happenings on the line (i.e. on the ATS screen). However, as mentioned in the last chapter, merely locating the train requires reading from both the phone system and the ATS screens. In (7), the train captain's reason for the call is unclear (line 12). It may be in the light of this, that Cii has to recover details from the screens to make sense of the call.

Indeed, it may be in the light of Cii's agitations that Ci continues his instruction.

(7) 10A 03:26 (42) - transcript 10
14 Ci: Select emergency shunt<proceed forward in that
15 mode<dock at Poplar routeboard two zero sevenT (0.2)

There are some slight perturbations in the turn 'select emergency shunt<proceed forward in that mode<dock at Poplar routeboard two zero sevenT'. The talk is produced so that each component runs quickly into the next one. The components are each actions in a sequence for dealing with emergency shunts. Hence, Ci's reply appears to be read out as a list of standard components, the first three of these being the activities usually required in a shunting operation: selecting the mode ('Select emergency shunt'), moving in a given direction ('proceed forward in that mode'), to a given destination ('dock at Poplar'). As this instruction emerges, Cii's movements become less agitated, so that by the end of the second component she remains fixed on the phone system screen, and by the beginning of the third she moves her hands from the immediate vicinity of the keyboard whilst turning back to ATS2. The instruction provides resources with which to make sense of the train captain's call. It identifies the train as the stationary one at West India Quay which requires no immediate action. Indeed, commands to it can only be made once the train on the delta has passed through Poplar. As the location of the problem and the precise nature of the solution is produced, Cii's movements become less energetic. When the next turn of talk, the confirmation from the train captain emerges, she goes on to type her commands to Devons Road.

In this instance it appears as if Cii is overseeing the local environment to make sense of the call, glancing at the various monitors around her that
would offer resources for identifying the nature of the call, the location of the problem, and thus, be potentially relevant for a next action consequent on the call. It also appears that Ci is sensitive to this, he appears to be 'monitoring her monitoring'. Throughout the call he maintains an orientation where he can see both the happenings on the screen in front of him and on ATS2. Moreover, he can also view the ongoing activities of his colleague. It may be that Ci's utterance 'You have a confirmed route (0.3) (points blocked in favour.)' may be sensitive to concerns of Cii. Only once this is produced does Cii go on to type in the command to the train that has left Stratford (the 'one one one').

Hence, whilst Ci is attending to Cii's activities (as revealed in this and the last section), Cii is also attending to Ci's activities. Attention is continually being paid to a colleague's activities even whilst being engaged in a distinct activity. The activities of Ci and Cii can easily be characterised as 'mutual monitoring', 'overhearing' and 'overseeing'. Each carries out activities that are related to the ongoing talk either on the phone or in the control room in which they are not 'direct' participants or recipients. Each controller also appears to be sensitive to the publicly and visibly accessible features of their colleague's work, they can see the screens the others can see and view the details of the identity of the phone caller, the movements of the trains and the results of their partners' typing. However, these activities are more than passive 'overseeing', rather, a colleague's actions are made sense of with respect to an emerging context. The controllers have to read the screens to locate emerging problems, and these screens have to be read in the light of the recent history of the service and the ongoing occurrences in the control room. These monitoring practices are achievements, and critical to these are accomplished understandings of the happenings in the setting, both in the control room and outside.

It is also unclear whether the 'monitoring' activities of each participant can be considered equivalent. In the instances considered so far, the call demands, or requires, an action to be undertaken. The participants both orient to the possibility that Cii undertake this. So, the particular purposes with which Ciis attend to the calls may be distinct from those for Ci. Ciis attend to the activities of Ci in order to identify hearably consequential items; Cis attend to Ciis to provide them with resources with which to hear items as consequential.
6.6.3 The Collaborative Accomplishment of a Command

In the following instance, in an incoming call the remote party requests a change to Automatic Train Protection at Crossharbour station.

(8) 13Co - transcript 1 (simplified)
1 Ci: train base to train at five two:
2 (3.3)
3 RP: five two (check) five two: request permission to (.)
4 change into A: Tee Pee (0.2) driver (0.3) (Ed) at
5 Crossharbour: (0.4) over.
6 (2.9)

As the caller completes his turn, Cii begins to type in a command to Herons Quay. After typing in the first packet, he pauses.

(8) 13Co - transcript 2
TC: (Ed) at Crossharbour: ---- over.
Cii: \[...\]

At the time of the call there are several trains running on the area of track in question (the branch between the Delta and Island Gardens). These are shown schematically in Figure 6.6.

![Figure 6.6: the location of trains as the call is being taken](image)

On the southbound (going right on the ATS screen) one train is between the Delta and Heron Quays, another is between Heron Quays and South Quay and a third is coming into Crossharbour. On the northbound (going left) one train is between Crossharbour and South Quay and another docked at the terminus, Island Gardens. Thus, there are five trains on the area of track related to the train in the call and the commands being typed. Through analysing the movements of these trains in some detail, it is possible to outline a likely interrelationship between the trains displayed on the screen and the activities undertaken by the
controllers. It appears that the caller is in the train moving north (between Crossharbour and South Quay at the beginning of the call).17

Ci does not go on to confirm the change, but rather asks the caller to 'standby'.

(8) 13Co - transcript 3 (simplified)

<table>
<thead>
<tr>
<th>(2.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ci:</td>
</tr>
<tr>
<td>(1.5)</td>
</tr>
</tbody>
</table>

This request holds off any immediate reply, and also projects some forthcoming activity - it is not a denial, for example. There are a range of circumstances for which it would be relevant for call to be put on 'standby'. But, it appears that this turn is related to Cii's ongoing typing. As Ci commences the utterance, Cii starts typing again.

(8) 13Co - transcript 4

ATSI

Ci: (I can see) if you could standby and let me get back to you in a mo::--

Cii: [keyboard input]

Throughout this utterance, Ci continues to look at the screen in front of him, only turning away whilst he puts his hand over the receiver. He then continues to look towards ATSI, only glancing away once he begin his efforts to return to the call. It appears that he is closely attending to Cii's activity of making a change to the main line. In the pause after the first packet of typing by Cii 'H E Q.', Ci begins his reply '(I can see) if you could standby and let me get back to you in a mo:'. This turn appears to be sensitive to the ongoing typing, the pause within it (of 0.7 seconds), and the demands of the call. Ci holds off replying to the train captain for nearly three seconds, until after the first packet is completed and a gap has emerged.

When Ci begins his reply, this emerges neither as a straightforward confirmation nor an alternative action to be taken. Instead, the train captain is asked to wait. His utterance commences with what is hearable as '(I can see)', as if he is reporting to the train captain some ambivalence to the request. As he is uttering this, Cii restarts his typing, beginning to enter the route to be revised. Ci then continues with 'if you could standby and let me get back to you in a mo:'. This maintains the ambivalent

---

17 See Appendix A to give a flavour of this analysis (Section A.4).
nature of his response to the initial request, it is neither confirming the request nor suggesting whether the request will be confirmed or rejected. By projecting a further unspecified delay through 'in a mo::', it does not even specify when such a reply would be likely. Interestingly, this turn of talk is coextensive with the typing of Cii, the route being completed on the word 'mo::'.

These features appear to reveal ways in which Ci's talk is being shaped by the ongoing activities of Cii: not only the delay and then the continuation of the typing, but also the particular activity that the typing is accomplishing - a change to Herons Quay over a specified route. This is relevant to the train captain's request. The talk may also display to Cii this regard, the holding off of a reply until he has completed his typing. Even the '(I can see)' could project such a holding off, allowing Cii to continue. Cii then may also be designing his typing in the light of the activities of Ci, holding off the actual route component until the nature of Ci's talk is apparent. After all, '(I can see) if you', is hearably not the straightforward commencement of a confirmation to the request.

It may be that Ci's initial pause is in the light of Cii's commencement of the command to Heron Quays and only when Cii pauses in the production of the command does he ask the caller to standby. Cii's pause may itself be a delay until a relevant confirmation is given. When this is not forthcoming and Ci's reply emerges, Cii then continues to type the route, Ci attending to the typing as it emerges.

Once the typing is completed Ci turns back to the phone system and renews the connection to the caller.

(8) **13Co - transcript 5**

9  (1.5)
10  (radio ring, phone system beep))
11  (2.3)
12  Ci: fi:ve two, are you still receiving?
13  (4.7)
14  TC: () receiving over?
15  (0.2)
16  Ci: (that change is fine) thats affirmative when you
17  arrive at Crossharbour (0.3) (routeboard) one four three
18  (0.4) you may set A: Tee Pee manual (0.2) and proceed
19  (on your section and clear) (0.2) until further notice↓

The change is confirmed in the talk (lines 16-19). It appears that Cii's typing has accomplished the relevant change. Cii's command appears to be sequentially relevant to the talk. It is produced in the light of talk of the participants on the phone, and that talk continues once it is
accomplished. Ci through his return to the call displays the relevance of
the Ci's command.

Through these closely interrelated activities the two participants
appear to be designing their activities with respect to both the call and
their colleague's actions in the light of the call. The talk and the typing
emerge from moment-to-moment, built in the light of their colleague's
ongoing activities, even as a command is being produced and talk being
uttered. As with previous instances the call requires both a confirmation
and an action, or series of actions, from the controllers: the confirmation
given through talk by Ci in the call, the next actions, in this case, through
amendments to the system's running. As in those other instances, the
nature of the action required appears to make it relevant for the second
controller to undertake it. However, the accomplishment of this activity is
shaped by the contributions of both controllers, emerging from moment-to-
moment.

The division of responsibilities between the controllers is not always
clear and distinct, neither governed by formally or informally specified
responsibilities nor even organised with regard to the technology most
accessible to them. In numerous cases throughout the collected materials,
Ci's enter commands related to activities not directly related to the views
available to them. Nor is it the case that Ci's always make the commands
when Ci's are engaged on the phone. Ci's can be frequently be seen to enter
commands whilst they are taking calls from train captains, engineers and
other members of staff. Rather, it appears that the activities demanded by
the call are accomplished through a situated division of labour emerging
from the circumstances at hand. This division of activities is
interactionally managed from moment-to-moment. So, for example, in (8)
the call requires a confirmation, and yet this confirmation is delayed until
an appropriate activity has been carried out on the computer system. The
call requires a command to be entered, which itself appears to be shaped
in the light of the ongoing talk.

Confirmations and other replies to calls can, as in the previous
instances, inform the nature and timing of the subsequent commands to
be entered by a colleague. It is also possible, as in (8), for the pacing of the
command to shape the timing and nature of the confirmation. Therefore,
the ordering of activities in the light of the call has to be managed from
moment-to-moment through the resources at hand. These may be through
the design and production of the talk, even though this may be addressed
to a remote party. They also may be through the timing and coordination of other activities, so that typing or bodily conduct, for example, can be seen by a colleague as related to their ongoing activities.

In order to achieve such activities requires an active attention to the work of a colleague. Making sense of this work relies on an accomplished understanding of the current circumstances, including a situated viewing of the various resources available to the participants. In the control room, this rests on being able to see the relevance of particular occurrences happening both in the local domain and outside, through the technology available. Making sense of another's activities rests, in particular, on being able to read off the ATS screens what is relevant to an ongoing call. This is a practice required by both controllers, in order to locate a particular caller on the line, the nature of the circumstances they are facing and to shape a possible next action. This reading is facilitated by an awareness of the potential problems which may arise, the current state of the service, and through the activities of colleagues, the changes they are making to the system and requested to make. Therefore, the typing of an appropriate command at the relevant time may be tied into a complex web of activities, involving the continually changing displays of technology and the contributions of colleagues. These contributions may themselves be shaped prospectively with regard to the activity underway, as in Ci's utterance 'I can see) if you could standby and let me get back to you in a mo:'. Indeed, the talk on the phone, both from the remote parties and from the principal controller, is a critical resource on which the coordination and distribution of activities relies.

6.7 MAINTAINING THE DIVISION OF LABOUR: AN EXPLICIT CASE OF COLLABORATION

In the instances considered so far, it is perhaps noticeable that there is little explicit talk between the controllers about the calls. In (5), following a call, Ci points to two places on the screen, on which Cii then enters a command. This gesture appears to be made sense of in the light of the prior phone conversation and Cii requires no explicit talk to accomplish the command. In (7), a suggestion by Ci appears to display how closely he has been attending to the activities of his colleague. In other fragments, the participants' activities appear to be tightly interrelated, their smooth accomplished production appearing not to rely on explicit comments or instructions. The controllers then go straightforwardly on to another
activity, taking another call or issuing another command. Apparently any
actions consequent on the call have been accomplished. However, this is
not always the case.

In this section, a more explicit case of collaboration is considered. In
the following instance, Ci has just turned down an engineer’s request for a
train movement within the depot. He provides a reason for this denial
(lines 10-13), and after the engineer’s reply, suggests roughly when the
train movement could be made (‘if you would like to get back to me in half
an hour’, lines 18-19). As he is producing this, Cii utters ‘I’ll do it now’
(line 201) and starts typing.

(9) 37Co - transcript 1 (simplified)

10 Ci: impossible to say ma:te. (. ) As I say we are
11 investigating a suspect package at the moment (0.4)
12 and also got a train captain who’s lost his keys
13 over.
14 (1.6)
15 E: (okay mate
16 )
17 Ci: (that sort thing) mate (. ) as soon as we ca:: n.(.)
18 (and actually) if you would like to get back to me
19 in half an hour. just in ca:se r( )

20 Cii: I’ll do it now:.  
21 (0.4)

Unlike previous fragments, the Cii explicitly notes to Ci, during a call,
that he is to make a change consequent to the call. Indeed, it appears to
transform the way Ci handles the call. After re-establishing contact (lines
21-28), Ci asks the engineer to repeat his original reason for the call (the
request to move a train, lines 30-31) and then confirms it (lines 36-37).

(9) 37Co - transcript 2 (simplified)

21 (0.4)
22 Ci: Paul are you still receiving?
23 (.)
24 E: ( )
25 (.)
26 Ci: Paul are you still receiving ma:te?
27 (1.2)
28 E: still receiving↓
29 (0.2)
30 Ci: what was your request again? (0.2) four two seven to
31 whe:re? 
32 (1.3)
33 E: ( four two nine four two eight) back to Dee
34 El1↑
35 (0.3)
36 Ci: yeas its being done now ma:te (0.3) we had a li(t)11
37 (0.2) lull in whats happening↑ (now you will
38 probably move)
39 (0.2)
40 E: right cheers then
So, not only does Cii explicitly display to Ci that he is going to undertake the activity on the computer system, Ci has to return to the caller to confirm information concerning the call. It appears that either (or both) the Cii or the Ci is orienting to the Cii not having undertaken an appropriate action: Cii not having entered the appropriate command when it was relevant; or Cii not having overheard the original request and, despite his typing, not having entered the appropriate command.

In (9) the typing by Cii appears to be timed in relation to his talk 'I'll do it now' and is completed prior to Ci's request to the engineer. Yet, despite the 'now' and the typing that immediately follows it, Ci does not appear to display that the Cii's typing is related. He goes on to ask for the engineer to repeat his specification of the change. On the other hand, Cii does not go on to continue typing during the call. It appears that his typing is adequate to accomplish the relevant next action on the system, and Ci some time later, in the talk to the engineer, he confirms '(now you will probably move)' (lines 37-8).

Throughout the call Cii has the depot displayed on ATS2. However, during Ci's initial denial of the request from the engineer, Cii has been issuing commands to trains at Canary Wharf and Heron Quays which he completes just as Ci reports one of the problems on the line: 'investigating a suspect package'. Cii then clears the screen of warning messages and on 'half' of 'half an hour' brings up a report line displaying route information.

(9) 37Co - transcript 3

Cii accompanies his talk with a gesture pointing towards his own screen and keyboard (the keyboard of ATS2). The command he then types appears to be one directed at the depot (i.e. it is not a standard main line command), and given his talk, and subsequent talk, appears to address the problem in the call. Cii's intervention appears to be appropriately timed, occurring after Ci has stated to the caller that a rather long delay is
required. Cii appears to type unproblematically the required command, completing it even before Ci has re-established contact with the caller.

Given Cii's completion of the command, it is curious that Ci then goes on to request the details of the required train movement. Following his typing Cii continues to look at the display in front of him. He does not type anything else into the system for the duration of the call, or for some time following it.

Although Cii does type during Ci's initial denial to the caller these can be hearably, and possibly visibly, recognised as being commands to the main line and unrelated to the call. The initial typing then provides few resources from which to assess Cii's availability for action relating to the call at hand. Only once he has completed his command does Cii engage in activities that may be related to the call or the depot, clearing the warning messages (during the callers 'okay mate'), line 15) and bringing up a line of route numbers. Indeed, these commands appear to mark a juncture from the prior activities relating to changes that have to be made on the main line. Cii's utterance 'I'll do it now', and his accompanying gesture, occur on the completion of 'just in case' by Ci. It is timely, coming just as Ci has informed the caller of a long potential delay before he can address the caller's problem again, an utterance that may even be preparatory to moving into closing the call. Just as Cii gestures to his keyboard, whilst uttering 'I'll do it now', Ci pauses in the course of his talk until Cii's command is underway. His next turn is shaped in the light of this delay with 'Paul are you still receiving mate?' (line 26).

Thus, Cii's gesture and command entry appear to mark a fairly noticeable juncture in the activities of the two colleagues, both occurring following a series of unrelated activities by Cii, and in the midst of call in which some effort has been expended to state why the command cannot be entered. Hence, it may be in the light of such a change of course of activity that Ci goes on to ask for a repeat of the details of the call.

The subsequent talk, where Ci asks for and then gets confirmation for a request that has already been made, may then not be so much a failure on the part of Ci to recognise the current state of his colleague's activities, but more be designed for the circumstances of the caller. It helps provide an account for the dramatic change of circumstances, from not being even

---

18 They have the familiar 1-3-1-3-3-1 key organisation and they utilise the alphabetic keyboard.
able to consider moving a train for half an hour to doing it almost immediately. Indeed Ci provides an account for this change of course to the caller: 'yeas, its being done mate, we've got a little lull in whats happening mate' (lines 36-7). An utterance that, itself, displays that Ci has been following Cii's actions and their consequences. The remote caller is provided, at least in part, with a display of a coherent sequence of activities going on beyond the phone call to deal with his request. Also, by remaining oriented on ATS2 Ci can not only see the production of the command but the possible consequences on the train at the depot.

In accomplishing the work on the phone, Cis have to attend to the requirements of both the remote recipients and their colleagues in the local domain. When their talk has consequences for both, this may be doubly problematic. In a domain where actions by colleagues can also transform the current circumstances for the caller, it requires particular sensitivity to maintain that all participants, whether recipients or overhearers, can make one's own activities intelligible.

In the talk on the phone, Ci provides a detailed account of why a request from a remote party cannot be met. Such a denial might be considered as dispreferred, in conversation as well as in institutional talk (Maynard, 1992; Pomerantz, 1984; Sacks, 1987). Hence, there may be sequential and interactional demands in the call for producing appropriate next turns of talk (cf. those of questions and summonses outlined in Chapter 5). In transforming a denial to an acceptance also requires interactional work; it too has to be accounted for. The intervention by Cii in (7) may allow for the traffic movement to take place, but it demands from Ci particular activities to be undertaken in the call. The use of the computer system by another can be seen to have an interactional consequence, an explicit intervention by Cii, requiring explicit talk about the activity by Ci. The intervention requires work from Ci to be undertaken in the call with the remote party, but the management of these activities can be accomplished by conventional conversational resources.

At first glance, Cii's intervention could be characterised as 'volunteering' to act. Certainly, 'I'll do it now' displays an orientation to carry out some activity immediately. This turn, however, does not arise from 'out of the blue'. It relies on Cii attending to the activities of Ci, and through the accomplishment of the command it is apparent that Cii does indeed have the resources for issuing the command. A characterisation of
Cii's utterance and actions as being volunteered neglects how it is immersed within the ongoing activity. It is produced through a subtle management of the resources available to the two controllers, including the phone call, the displays and the activities of their colleagues. Activities, including talk from remote participants, but also visual conduct and even computer commands can a sequential import, implicating particular next activities to be undertaken.

In the DLR Control Room, both participants have access to a technology on which actions related to calls from remote parties can be accomplished. It would be hard to imagine a 'formal' procedure that could fully specify ways in which the controllers operate the technology. Simple prescriptions 'to be responsible for' a particular domain could be managed in a variety of ways. Complex and detailed specifications of the distribution of activities would ignore how the work could be flexibly managed. It is also unclear whether the work is managed 'informally'. It would be rather strange to characterise the instances considered in this chapter as 'negotiating' who does what. It is even difficult to identify simple actions as 'volunteering'. Rather, the activities are interactionally managed through the resources that are publicly available. The activities, the typing, the visual conduct and, particularly the talk, are interrelated to each other, and seen by the participants to be tied to one another. From moment-to-moment the responsibilities and actions of each are shaped by and made sense of by their mutual actions. The solution to who does what in this complex environment is a situated accomplishment.

6.8 DISCUSSION

The constant flow of calls into the control room, the recurrent problems with the technology and the various contingencies that arise in the management of an urban transport system, makes considerable demands on the Control Room Supervisors on the Docklands Light Railway. In particular, it presents them with one continual problem: who does what. There are various rather gross solutions to this, for example, stipulating formally, and generally, the responsibilities for each controller, or informally dividing the work prior to each relevant action. However, it is unclear how any general division of responsibilities would assist with specific cases, particularly when these occur in quick succession involving related matters. The timing of activities, and the emergent nature of
activities which have to be undertaken, also make it difficult for controllers to inform each other explicitly, on each occasion, what they are, or what their colleague should be doing.

Nevertheless, there does appear to be some orientation to different responsibilities. The controller on the left operates the phone system and is responsible for taking and making calls to the operators on the line, a critical resource for managing the service. The controller on the right will tend to talk to other staff in the control room, particularly Control Room Assistants. The controllers also tend to display different views of the line on the screens in front of them, reflecting different domains of interest, particularly the main line and the depot. However, controllers do alter what is displayed in front of them, do adopt different configurations of the displays between them and do type commands appropriate to what is displayed on the other's screen. Therefore, it cannot be relied upon that the organisation of the technology itself can be seen to manage, or display, the way work is distributed at any particular moment. Instead, the division of tasks and responsibilities between the controllers appears to be achieved from moment-to-moment in the light of the ongoing contributions of the two colleagues and tied to the emerging talk on the phone.

The work undertaken by the two controllers is accomplished through a complex set of resources being considered in relation to one another. It is worth summarising some of the interrelationships between the activities of controllers (Ci and Cii), colleagues in the control room and the remote parties, and the artefacts (the screens and the computer systems) which have been revealed in the foregoing analysis. These could be glossed as follows:

- in a broadcast phone call between Ci and a remote party a confirmation of a change to the service by Ci can be seen to require a command to the ATS to be undertaken by Cii, and that command is accomplished (fragments 3 and 4);
- in a phone call between Ci and a remote party a confirmation by Ci may require an ATS command but the Cii's contribution appears to be a 'noticeable absence', and is 'repaired' by Ci (fragment 5);
- in a phone call between Ci and a remote party, the talk of the remote party appears to implicate a relevant action to be undertaken on the ATS, a relevant command is entered by Cii (fragment 6);
• whilst engaged in another activity on the phone, Ci 'monitors' a command being entered by Cii, makes sense of it in the light of the current service and prior talk between Cii and a CRA; Ci later 'repairs' that command (fragment 7);

• whilst Ci is engaged in talk on the phone, Cii 'monitors' the screens available, apparently to make sense of the call; Ci appears to transform his talk in the light of Cii's activities, he appears to be 'monitoring her monitoring' (fragment 7);

• in a phone call between a remote party and Ci, Cii starts to type a related command; Ci appears to transform his talk on the phone in the light of the entry of the command (fragment 8);

• in a phone call between a remote party and Ci, Cii starts to type a related command; Cii appears to transform his typing in the light of the talk on the phone (fragment 8);

• in a phone call between a remote party and Ci, Cii explicitly states that he will undertake a related activity; this intervention is delicately managed and placed at a juncture in the conduct on the phone (fragment 9);

• in a phone call between a remote party and Ci, Ci orients towards a domain of a possible next activity for Cii; Cii then states that he will carry out a related next activity (fragment 9);

• in a phone call between a remote party and Ci, Cii states that he will carry out a related next activity; Ci appears to produce talk that is sensitive to both the intervention of Cii and the circumstances of the remote party (fragment 9).

Through these glosses it is possible to see a complex web of interactions and activities, as examples: the talk of the remote party places requirements for the conduct of Ci (a next turn of talk) and for Cii (a command); the talk and visual conduct of Ci prefigures activities on the computer system by Cii; Cii's commands on the computer system are taken account of by Ci; and Ci transforms his call on the phone to the remote party in the light of the commands entered by Cii. The activities of the two controllers, the remote parties on the phone and, also others in the local setting, like the CRAs, are tightly interrelated. However, these interconnections not just coordinations between actions, the activities considered place differing demands on the potential next activities of a co-interactant or colleague. In some cases, a next may be unproblematically
accomplished, in others, it may be ‘noticeably absent’ and then ‘repaired’, and others the production of a next appears to require more delicate or sensitive management of the ongoing activity. The particular circumstances of the moment may contribute to this; what, for example, the consequence of the next activity has on the trajectory of action, whether a prior activity with another has to be repaired, refuted etc. Hence, these activities are interactionally managed. A turn of talk or a command may have implications for others, and these may require a next action, the accomplishment of which is produced and recognised unproblematically. However, the demands, implications and consequences are social commitments, they can be accomplished, repaired and made use of. An absence may also be accountable.

In previous chapters a range of resources have been considered to account for the accomplishment of activities: including the common-sense understandings of actions and readings of interfaces, ways of conducting and improvising actions in order to achieve a practical purpose, the readings of a screen in relation to other artefacts and the talk and activities of another; and the production of activities and the displays of understanding accomplished through turns of talk. These are methods, common-sense, ordinary and practical ways of accomplishing activities on and through computer systems. In this chapter, by examining the talk, visual conduct and activities of a computer system in a setting where the participants in the activity are both co-present in the domain and remote, such methods appear to provide a foundation for accomplishing the work. The talk has to be managed from moment-to-moment on the phone, the talk has to made sense of with respect to the current state of the service and the circumstances of the remote party, made available partially on computer screens. The talk on the phone also can make sense of the changing happenings of the screens: a dynamic environment. The screens can be understood in the light of interactional accomplishments. Similarly, another's typing can be made sense of in the light of the talk in the phone calls. It can be seen as consequential or demanding further actions. A command on the computer system both appears to be produced in relation to a prior activity and recognised in the next as appropriate, by a co-participant.

It would appear to be unsuitable to characterise the interrelationship between the commands entered into the computer and the ongoing phone conversation as a coordination between typing and talk. The controllers
do not just tie their activities into the movements of hand across the keyboard and the pressing of keys, but particular activities made to particular trains, given the state of the service, as in the 'inappropriate' train command to stable a train coming from Stratford in (7). They also attend to the hearing of particulars of the talk as in (6), where the consequences of an utterance are made sense of in terms of who is making that utterance and what it implies for their own activities. Hence, the activities in the control room rely on more than a passive monitoring of another's activities or a mechanical coordination of actions. They are achieved through a sequential, sense making practices.

Controllers have to actively make sense of the call in terms of the technologies and resources they have available to them, for example, what appears on the screens, the prior talk and activities of their colleagues and the routine ways of getting work done. A caller, for instance, has also to be identified and located, the reported problem assessed in the light of the state of the service and previous occurrences and next actions shaped in terms of this assessment. Throughout a phone call resources may emerge relevant for this work, for both controllers. The reading of the technologies then has to be accomplished from moment-to-moment in relation to the talk. The characterisation of activities in the DLR Control Room as 'monitoring' may thus be problematic, suggesting a rather disengaged activity by the 'monitor'. It may also be a rather individualistic conception of the work of the two controllers. As revealed in the instances above, Cis rely on, display an understanding of, and even appear to collaborate in the sense making practices of their colleagues. The activities of the two controllers are thoroughly and actively interrelated, even in the production of a single command or turn at talk.

Examining the potential interrelationships between computer-based activities and the accomplishments of others being achieved in the local environment suggests a distinctive form of analysis to those developed within HCI. Although preliminary examination of the materials reveals a certain consistency to the organisation of typing, the 'breaks' in keyboard activity do not appear to be merely related to the 'individual' activities of the typist, or the time needed to 'process' information on the display. Rather, the activity is often initiated, shaped and completed in the light of the ongoing activities of others. The production of a computer command, therefore, is not organised merely with respect to the travelling time of the hands, the motor skills, or the 'cognitive processing' of the controller but to
other activities occurring in the local environment, namely the ongoing phone call between the principal controller and the remote party revealed through dynamic and continually changing happenings on the computer displays and the talk and visual conduct of a colleague. In this way, with regard to its interactional management, the production of computer-based activities is situated.

Whereas, characterising the controllers' use of the ATS as an 'interaction' with a computer unnecessarily delimits the conception of the activity, considering the work of the controllers as 'collaboration' offers insufficient constraints. The analysis of the activities of the controllers reveals a complex web of practices, resources and contributions through which these are achieved. The commonly held distinction between the individual and the collaborative, paralleled by the domains of HCI and CSCW, would appear to be problematic. Particular activities are accomplished by an individual typing on the keyboard, but these are achieved in relation to the ongoing activities of others. Individuals appear to monitor their colleagues, but this itself can be public and available to analysts.

The analysis of the use of the ATS computer system in the DLR Control Room may be of some relevance to researchers in CSCW. On the whole, the controllers appear to manage the accomplishment of activities into a common computer system quite unproblematically. Indeed, they rarely have to utilise explicit resources and devices to manage the activities between them. Other activities can be accomplished in parallel, like the phone call, without being interrupted by explicit devices concerning who is to do what and when. Part of what constitutes the skill of being a controller in the DLR is assessing the current activities and producing a relevant next activity, or course of activities. It also relies on seeing others' activities in relation to one's own, whether they contribute to it, or can be utilised so that an activity can be jointly produced. The management of the activities is accomplished through the talk, activities and visual conduct on the system. These, themselves, display an understanding of the ongoing circumstances. It may, thus, be problematic to introduce devices and resources for the explicit management of collaborative activities in computer systems. These may unduly split the accomplishment of an activity from its management. They may interfere

---

19 See also the studies of other control rooms, news agencies and dealing rooms (Heath, et al., 1994-5; Heath and Luff, 1996a; Heath, et al., 1995a).
with the production of the activities they are intended to support. Instead, it may be worth considering how to give access to the resources that individuals have available to them with respect to anothers' activities. In the DLR Control Room this is not only the displays in front of a colleague, but also access to the colleague's activities in relation to those displays, and seeing those activities in relation to activities of remote participants. This would require not only access to the resources others have access to, their activities and their activities in relation to those resources, but also ways of shaping one's access with regard to the ongoing activities of colleagues (cf. Heath, et al., 1995b).

Within this complex web of resources and practices in the DLR Control Room, the talk on the phone between controller and remote party appears critical. It is a resource with which both controllers make sense of the operation of the service and the activities of their colleagues. It also provides a foundation for the management of the distribution of activities between the controllers. The talk is not merely a commonly heard information resource but places requirements and contributions on the participants, and implicates them in courses of unfolding activities. A turn of talk is a resource, in one sense, for the participants, but it is also a social action, placing moral demands on a hearer, demands that constitute, and are constituted by, its institutional and conventional setting.

6.9 SUMMARY

An incoming call to the control room can set up the relevance for a variety of actions from the controller, for example: a response; a confirmation of a change to the system; an actual alteration to the schedule, and operation of the service; and, potentially, a recording of any failures and delays consequent on the call. In the instances above, a variety of strategies have been adopted by the controllers to distribute these activities amongst them, but in each this has been achieved interactionally, with respect to the ongoing talk and visual conduct of the participants. The division of labour is interactionally managed. It is also tacitly accomplished.

The foregoing analysis has also revealed how the production of commands into the ATS in the DLR Control Room is tied to the ongoing talk. It is through this talk, principally between the remote caller and the principal controller, that most requests are made, actions requested and
information passed between personnel on the system and the control room. It is also through this talk that Ciis make sense of the service and then produce relevant actions. Maintaining the transport system and managing the traffic relies on timely alterations being made. So, controllers not only have to identify who the caller is, but where they are, what station they will next arrive at, what station a change will then have to be made at and what consequences this will have for following traffic. It is then not surprising that Ciis are seen looking over to the phone system screen and at both screens of the ATS. Ciis look at the screens not only to make sense of the call, but also prospectively to take next actions. Close examination of the materials reveals that Ciis go on to type in commands that are related to the problem in the call. The typing is not only temporally coordinated with the talk, it is tied to the activities in which their colleagues are engaged.

Furthermore, Ciis also reveal through their activities that they orient to their local colleagues making relevant changes to the system related to their calls. It is not just that Ciis type during the call, or at appropriate places in the call, but that they type particular and relevant commands. Each participant is then reading off the screen and listening to the call and making sense of this retrospectively in the light of foregoing occurrences, and prospectively with regard to potential next actions, by themselves and their colleague. The division of labour then is interactionally managed and attended to. It is a problem for which the solution is improvised from moment-to-moment, for this call, with respect to the current resources at hand and the state of the service and in the light of the contributions of colleagues.

This chapter has sought to explore some of the practices through which a computer system is used in a real-world environment. Through close analysis of the details of these activities it has been possible to reveal how closely interrelated this activity is to the activities of others. Commands entered into the system by another can be seen to be produced in relation to one's own activity. It is even possible for a command to be shaped in the light of an ongoing single turn of talk, and for that talk to be shaped by the ongoing typing. By tying one's own activity with another's in these, and other ways, it is both possible for a colleague retrospectively to see that a command is relevant to his own, and also, prospectively, to utilise the coordination to move into other collaborative activities. Such practices are tacit, and are accomplished by being non-intrusive. However, they reveal
how closely interrelated computer use is to the activities of others, particularly to talk.

The use of the ATS system in the Docklands Control Room provides a rather distinctive focus for an analysis of collaborative activities. The ways in which it is used appears to parallel many capabilities envisaged in current developments of technology. But, the problems which might be imagined, and have indeed emerged in the use of prototype collaborative technologies, are not apparent. The management of the activity in relation to the computer is routinely accomplished implicitly and unproblematically. There do not appear to be conflicts concerning who does what and who is attending to which problem. Rather, the management of the activity is achieved through the accomplishment of the activities of controlling, scheduling and managing the service. This is made possible by the two controllers having access to each others' activities, and the ways in which they are engaged in those activities. The accomplishment of these activities also relies on interactional resources, the talk and visual conduct of the various participants. Hence, the various activities of controlling an urban transportation system, operating a complex computer system and managing with the activities of others rely on quite conventional, interactional methods and procedures.
Chapter 7

Resources for Technological Design
relationships between workplace studies and methods for system development

So Taurus was to be the perfect, all-electronic, paperless system that would not just replace the British settlements system but would connect into other settlement systems around the world for international securities. Of course, as with so many other grand British technical visions, they could not get it to work...[M]any of the personal customers for whom it was supposed to be a great advance did not want it: they rather liked the idea of having share certificates to show what they had bought.

Hamish McRae; The Independent 12/3/93.

'The plain fact is that the system was meant to serve the needs of the brokers and jobbers of the Stock Exchange; and it didn't. It usually takes a professional working member of the exchange at least five years to learn how the Stock Exchange works, and I don't see why the analysts and programmers who make computer systems should expect to pick it up much more quickly. A system for a particular business can only be built by people who are experts in that business. Domain knowledge. I think it's called. That's what matters.'

7.1 INTRODUCTION

Because of their concern with the details of social activities, attention has recently been drawn towards the work of researchers undertaking ethnographic studies, particularly orientations drawing from ethnomethodology and conversation analysis. It has been proposed that work in these areas might provide direct resources for design, or more indirect resources for novel methods and tools to support system development. This chapter explores these proposals in more detail, utilising the cases discussed in the previous chapters.

In part, the interest in such approaches comes from shortcomings in traditional methods for design. Section 7.2 focuses on the recent interest in Requirements Engineering to address problems with system design and several novel proposals for enhancing techniques for eliciting the requirements for computer systems. Amongst these there has been a growing interest in the use of naturalistic studies for system design which have certain parallels to recent innovations in HCI and CSCW. Particular proposals from Brown and Duguid (1994) and by proponents of Activity Theory and Distributed Cognition are briefly considered.

Section 7.3 considers the possibility that an ethnomethodological orientation to naturally-occurring activities utilising audio-visual materials may be relevant to the design process. These possibilities are discussed in the light of detailed examples drawn from the previous chapters. It would be strange if, having looked in detail at the uses of technology in several settings, nothing could be said concerning the design of those technologies. From the analysis of the moment-to-moment conduct of participants in a setting it does appear to be possible to suggest particular design options for software functionalities, hardware components or interfaces. These suggestions are discussed with respect to a range of potential resources for various activities in the design process: including the assessment of particular designs and deployments of technology (Sections 7.3.1 and 7.3.2), the design of particular technologies and generic systems (Sections 7.3.3 and 7.3.4); and support for the design and deployment process (Sections 7.3.5 and 7.3.6). Although the studies in this thesis were not intended to contribute to any particular design activity, they will be used to consider how a socio-interactional orientation to the analysis of naturalistic materials might contribute to the design process. The particular resources suggested by the studies outlined here are discussed in the light of other recent proposals drawing from
ethnographic observations and an ethnomethodological tradition (Section 7.3.7).

However, what may be more significant for the process of developing new technologies is a distinctive orientation towards the activities which the technology is being designed and deployed to support. The significance of the orientation may then not be so much with respect to particular design activities, but in how it can serve to provide resources for a reconceptualisation of some of the foundations to current methods and approaches within system design. Section 7.4 considers the conceptual relevancies of the orientation. It points to conceptions in HCI and CSCW that the studies in this thesis, and other related ones, render problematic.

The analysis of naturally occurring social actions has been a long-standing concern of the social sciences, and recent discussions within ethnomethodology and conversation analysis have raised problems associated with the adequacy, warrant and relevance that can be attributed to the findings and results drawn from such studies (e.g. Lynch, 1993). These concerns parallel those regarding the adequacy, warrant and relevance of observations made for design purposes. The discussion in Section 7.4 concludes by briefly considering the possible interrelationships between the requirements for the analytic foundations to the two distinct concerns of social scientists and system developers. It may be that the debate within the social sciences could contribute to a discussion about system design.

7.2 BACKGROUND: METHODS FOR SYSTEM DESIGN

Management were misguided or naive in believing that computer systems in themselves could bring about [such] changes in human practices. Experience in many different environments proves that computer systems cannot influence change in this way. They can only assist in the process and any attempt to force change through the introduction of a system with the characteristics of an operational "straight jacket" would be potentially doomed to failure.


The computer system design process is typically conceived of in terms of a series of discrete, sequential phases.¹ Amongst the initial phases in these

¹ A typical characterisation is in terms of the 'waterfall model' where a design proceeds through the following activities: requirements specification; architectural design; detailed design; coding and unit testing; integration and
There is typically an activity identified as requirements capture, requirements analysis or requirements specification. These activities, with some other variants, have become the focus of a field within Software Engineering called Requirements Engineering.

Researchers in the field of Requirements Engineering have tied some of the critical problems associated with computer system design to the early phases of design. So, possible reasons why systems which have been built do not meet the needs of users, are under-utilised or simply fail have been traced back to problems associated with the original identification of the requirements for the system (Jackson, 1995; Landauer, 1995; Norman, 1988; Norman, 1993b; Page, et al., 1993). Moreover, if problems are identified later in the design process with the requirements that have been specified they appear to be more costly to address (Boehm, 1976). Hence, there has been a growing interest in developing methods for eliciting the requirements of users.

The techniques that have been proposed for Requirement Engineering draw from a variety of resources including cybernetics (Espejo, 1980), socio-technical systems (Mumford, 1983), critical social theory (Flood and Jackson, 1991), however, the approaches which have been derived from these do not appear to be that distinctive from methods adopted in other fields associated with system design. So, for example, researchers in Requirements Engineering have suggested the use of scenarios (Holbrook, 1990), the analysis of users’ tasks and goals (Bubenko, et al., 1994; Karakostas, 1990; Mittermeier, et al., 1990), open-ended or structured interviews with potential users (Loucopoulos and Karakostas, 1995), and the participation of users in the design process (Macaulay, 1994; Mumford, 1983). These have parallels to other similar suggestions in HCI and testing; and operation and maintenance (from Dix, et al., 1993). There are a range of different characterisations of the various stages, and the model has been widely criticised in favour of more exotic and iterative processes (e.g. Boehm's 1988, 'spiral' model). It is unclear whether any of the stages are 'followed' by practitioners in system design, however it may be that they are 'oriented to' in some way (Button and Sharrock, 1996). In this chapter, the activities are only intended as a resource for considering the different ways in which various approaches can contribute to system development.

See, as examples, the reports of the Taurus system for settlements in the London Stock Exchange and the dispatch system for the London Ambulance Service (e.g. Page at al. 1993, and, particularly, the citations given above).
system design (e.g. Bjerknes, et al., 1987; Carroll, 1990; Ehn, 1988; Kirwan and Ainsworth, 1992).

However, there have also been proposals to draw from other resources from within the social sciences, particularly integrating ethnography (Goguen, 1994; Sommerville, et al., 1993) and naturalistic analysis (or naturalistic inquiry) into Requirements Engineering (Potts, 1997; Potts and Newstetter, 1997). These suggestions focus on the possibility of using observations of settings in the process of design, and hence appear to have much in common with proposals for work and studies undertaken within CSCW, where some fieldwork is performed prior to proposals for novel designs or analysis of workplace activities (Ackerman, 1996; Bellotti and Bly, 1996; Berlin and Jeffries, 1992; Moran, et al., 1996; Murray, 1993; Schwab, et al., 1992; Watts, et al., 1996). However, as Anderson (1994) and Button and King (1992) have suggested ethnography is more than just collecting together a set of observations and accounts, or 'just hanging around'. An analytic orientation needs to be adopted in order provide a framework for an analysis.

There have recently been several proposals for framing naturalistic analyses for design activities. So, for example, Brown and Duguid (1994) have suggested an approach to system design which rather than considering the 'core' activities that individuals are involved in, explores the periphery, or the borderline, the properties of objects which surround the competent use of an artefact. So they note the ways in which the noise of using the keyboard may be critical to the use of a personal computer for other individuals in the setting. Thus, Brown and Duguid, propose a conceptual framework that begins to explore the details of the use of tools and technologies, and, in particular, the 'communities of practice' through which they are achieved (cf. Lave and Wenger, 1991). Brown and Duguid argue that designers need to be sensitive to border and peripheral resources when developing innovative technologies.

There have also been suggestions that Activity Theory, an orientation developed within Russian Psychology (Vygotsky, 1978), and Distributed Cognition, considered in Chapter 2, would provide frameworks for considering design of systems to support collaborative activities (Nardi, 1996c; Rogers, 1992).\(^3\)

\(^3\) See also suggestions made by Kuutti (1991) and Bannon and Bødker (1991).
So, as an example, Nardi (1996c), commenting on the Activity Theory's long tradition within Russian psychology, its strong 'closeness in spirit' to Distributed Cognition and the possibility that the two developments might merge (p. 89), suggests it is an ideal resource from which to develop an understanding of the use of tools and technologies within HCI. However, the generality of Activity Theory's concerns, including skilled manipulations of objects, cognitive processes, the mind and consciousness, and social, cultural and historical factors may provide few constraints for an analysis with respect to design. With all these concerns it is perhaps not surprising that when implemented, despite the programmatics, the concern within some of those using Activity Theory for HCI is to maintain a focus on planful, intentional behaviour (Nardi, 1996c, p.84). Similarly, Brown and Duguid's proposal for designers to take account of 'communities of practice' and 'the social' may be too vague to deploy for the purposes of practical system development. Moreover, the attempt to demarcate context with respect to a spatial metaphor (namely 'centre', 'periphery' and 'border') suggests a rather conservative characterisation of context and the artefact (Heath, et al., 1995a).

Although distinctive, the programmatics of Brown and Duguid, and certain proponents of the utilisation of Distributed Cognition and Activity Theory for design, appears to raise some difficulties. The 'periphery', the 'border', 'distributed cognition' and 'activities' prescribe objects of interest for the analysis. These may be useful heuristics for commencing an analysis, but they could either be seen as either prescribing a set of concerns a priori, and face similar criticisms to more conventional approaches such as task analysis, or could be considered too vague to assist with design activities. Despite the promise of these frameworks providing extensive and alternative resources for shaping fieldwork, it may be that the conceptions on which they rest may either be too restrictive or too similar to those utilised by the orientations they seek to replace.

In the remainder of this chapter some alternative proposals for relating naturalistic studies to design are considered, utilising the analyses outlined in this thesis. Alongside other recent studies of workplace settings, an alternative analytic orientation has been adopted which may offer distinctive resources for system developers. Rather than begin with

---

4 See the discussion of the proposals for an analysis of 'Distributed Cognition' given in Chapter 2.
an elaborate conceptual specification, this research has undertaken extensive and careful field work, deliberately avoiding the development of stipulative characterisations of the domain of action. These workplace studies have addressed the *in situ* and emergent character of technologically mediated actions and activities, and have deliberately broken away from the static and limited conceptions of context and artefact that are found in more conventional forms of social and Cognitive Science (see, for example, Greatbatch, et al., 1993; Harper, et al., 1989; Heath and Luff, 1992a; Suchman, 1993c; Whalen, 1995a). They have generated a substantial body of empirical findings concerning the 'communities of practice' which underlie the production and coordination of technologically mediated social actions and activities. These workplace studies also suggest some novel ways of informing methods and approaches taken towards system design and development.

7.3 **NATURALISTIC STUDIES OF INTERACTIONAL ACTIVITIES: POTENTIAL RESOURCES FOR THE SYSTEM DEVELOPMENT PROCESS**

What many of these workplace studies have in common is the utilisation of an analytic framework drawing upon an ethnomethodological orientation to their domains of study (cf. Garfinkel, 1967). Amongst these studies several have utilised audio-visual materials providing the resources for subjecting naturally-occurring activities to repeated and detailed scrutiny.

So, for example Goodwin and Goodwin (1996) explore the accomplishment of activities in an airport control room: the ways talk, such as instructions, are produced and made sense of within viewings of the local domain. The audio-visual materials make available to analysts social actions which are visible to participants. The materials can also be subjected to detailed scrutiny. For example, Whalen's (1995b) analysis considers the interrelationship between talk and the use of a system for computer aided dispatch, revealing the practices through which the typing

---

of particular commands is organised with respect to the interaction between caller and dispatcher, and *vice versa*.

Though exploring the interactional production of social activities, the analysis of both Goodwin and Goodwin and Whalen are not completely divorced from system development issues. The Goodwins' analysis is one component of the Workplace Project undertaken to explore the nature of collaborative work by Xerox Corporation, and Whalen's studies have provided for a discussion of the deployment of expert systems for dispatchers (Whalen, 1995a). It is perhaps understandable why research into the organisation of interactions in complex settings would be relevant to large corporations developing new technologies, particularly within telecommunications. However, these studies, with others by Luff et al. (1994), and Heath and Hindmarsh (1995) may have a broader relevance to the system design community. They are the foundations of an interactional analysis of the use of artefacts: analysis that is made possible through the use of video-based materials. When the artefact is a computer system, such an analysis could contribute to an understanding of the nature of human-computer interaction and collaborative work. When the artefact is a system or a paper document that is considered archaic, then detailed analyses of the practices could be relevant to the design of alternatives (cf. Luff, et al., 1992).

The potential that audio-visual materials provide is to unpick the ways in which activities are made relevant to others and produced with respect to others (cf. Garfinkel, 1967). From detailed examination of the materials, attention can be paid to the different ways in which individuals participate in these artefact-centred activities or coordinate these activities with those of their colleagues. Though primarily focused on the audio-visual materials, these studies also make use of ethnographic materials, providing resources that are not directly available in the data, with regard to the skills underpinning the activity and organisational demands, for example. The fieldwork can both inform the analysis of the audio-visual materials and also its progress can be informed by it. Therefore, the utilisation of video-based materials and fieldwork, coupled with an interactional and ethnomethodological orientation to the analysis,

---

6 Whalen's research was also sponsored by US West Advanced Technologies, a regional telecommunications company in the United States. It suggests some implications for designers of expert systems for call takers. It also can be considered as a detailed analysis of 'users' of the telecommunications network.
offers the potential for a powerful approach useful for a variety of purposes within system design. As mentioned above, together these resources do not provide a method for system design, requirements analysis or technology deployment, but do suggest some contributions such an orientation could offer for system development.

In the following sections the potential contributions of an analytic orientation towards interactional materials is considered in terms of the resources it offers for six activities, or purposes, related to design. These are the resources for:

- assessing particular interfaces
- assessing deployed technologies
- the design of particular technologies
- the design of generic technologies
- deploying of technologies
- supporting the design process

These resources will be illustrated with examples from the foregoing chapters in the thesis. Although none of these studies were undertaken for any particular purpose related to the development of a technology, the specific examples may, however, serve to provide some grounding to the discussion.

7.3.1 Resources for the Assessment of Particular Interfaces

Although the evaluation of interfaces through the use of audio-visual materials within system design is not novel, these assessments are typically organised with respect to specific analytic orientations, particularly experimental psychology and Cognitive Science. Even with the constraints that these orientations provide, researchers still face difficulties in managing materials without either introducing premature categorisation (e.g. Olson, 1990) or losing the detail of how the recorded activity is accomplished. It may be that an alternative analytic orientation could be utilised for considering how activities are accomplished through screen-based activities, one that does not unduly circumscribe an analysis. By the examination of audio-visual materials coupled with appropriate transcription notations it may be possible to subject materials concerning the use of computer systems to repeated scrutiny.
The analysis of the uses of menus within the PowerPoint application in Chapter 3 revealed how menus on an Apple Macintosh are utilised as a resource for innovation and improvisation. The analysis of data may also suggest what appear to be problems with the design of the application. For example, when considering 'straddles', the shift between two adjacent menus, it appears that these often related to particular troubles users are having in the course of on ongoing activity. So, when users straddled between the Style and the Text menu, the File and the Edit menu and the File and the Apple menu, they appear to be having difficulty finding font sizes and font styles, creating new slides and finding help. The reason for this may be the location of these compared to other Macintosh applications. There is no 'Font' menu, so the Style and Text menus appear to be 'possible candidates' for changing the style of text. Getting help and changing slides are very general operations, so they may be expected to be available on one of the menus towards the left of the title bar, but it is unclear which one. This points to a general problem for designers of applications on the Macintosh; how to keep menus consistent across applications and maintain their consistency within the particular application.7

The analysis of menu use reveals the ease with which many menus can be viewed one after another, facilitating finding solutions to existing problems, general browsing of the system's capabilities, discovering properties of various objects on the screen and also suggesting possible next actions. However, there are disadvantages. The very ease of use, coupled with the necessity for menu items to be general and understandable in many contexts, can draw users into digressions and involve them in diversionary courses of activities.

By the detailed analysis of particular instances of interaction with the interface, troubles that users have with the system became apparent. However, it may not be so straightforward to ascribe to generic behaviours an analysis that the user is having trouble. There are cases where users repetitively carry out an operation, like opening a menu, or pause for a long time without taking an action on the interface. In a more quantitative analysis measures might be given concerning such 'problematic' behaviours. There are also cases, even in an experiment,

---

7 It is interesting to note that in version 3 of the PowerPoint application the distinction between Style and Text has been avoided by the removal of the Style menu, and an additional Slide menu contains the item for creating new slides.
where such actions could not be viewed as troublesome, and where a pause and a repeated activity form one part in a skilled course of activity. Moreover, it would be hard to catalogue all such cases of unreasonable or problematic (or reasonable and trouble free) behaviour. Being able to characterise activities as problematic or unproblematic rests on an analysis of tacit and common-sense resources in understanding the interface and carrying out activities on the system. Quantitative analyses of the speed of selecting items or the number of menu visits where items are not selected is not the only means of assessing an interface. A more naturalistic analysis supported by video-recordings may be a useful resource for presenting findings to designers, making apparent particular occasions of apparent troubles, as well as providing resources for warranting an analysis of these behaviours as troubles.

Studies of screen-based activities in naturalistic settings can also reveal the ways in which particular systems are utilised. So, for example, in Chapters 5 and 6, the uses of the phone system and the Automatic Train Scheduler System (ATS) were considered. Observations drawn from the audio-visual materials reveal that both controllers systematically look between the phone system and the scheduler system displays. These reorientations by the participants in the initial turns appeared to surround the problematic nature of making and taking calls through the radio phone system. In order to make sense of an incoming call, for example, a controller may need to locate the relevant train in the call and the possible circumstances facing the train captain. For this, both the train number and the train position is required: the train number of the caller is displayed on the phone system screen, the train position on the ATS. Controllers need to see one in the light of the other. From these observations possible changes to the design of the interfaces to these systems could be considered (see Section 7.3.3).

From the analyses of naturally occurring activities surrounding the use of computer systems it is possible to suggest potential problems with the design of current interfaces. Even without a stipulative characterisation of behaviour, as goals, plans or errors, it may be possible to provide an analysis that reveals features of the screen-based activities of users. The use of naturally occurring materials could provide an alternative or supplementary way of assessing interfaces to computer systems. Furthermore, the audio-visual materials themselves may be of use to make it possible to verify potential analyses. The materials can be
publicly viewed or passed to colleagues and other researchers, to question or provide support for an analysis. Moreover, the materials can provide a grounding for analyses and accounts of screen-based activities, at least in suggesting where care needs to be taken concerning how these analyses are warranted.

7.3.2 Resources for the Assessment of Deployed Technologies

By utilising video-based materials, implications regarding the more general aspects of using technologies in a setting can be suggested. For example, in Chapter 5 even with a technology that appears to be as constrained as the telephone system in the Docklands Light Railway, the analysis reveals particular transformations to the opening of phone conversations through the computer-supported radio phone system. There is an asymmetry to the resources available to the two participants, both in what they have available to them and the actions they can perform on the system. Practices appear to have emerged, including the subtle pacing of the calls and the use of particular features of talk to manage these resources, but it could be possible to consider ways of supporting the participants in this critical activity on the transportation system.

For example, the participants appear to undertake various activities in the opening of the call to secure a reason for the call. It appears that it is problematic to distinguish at the outset of a call between incoming and outgoing calls, something that is typically set in line by picking up the phone in ordinary telephone conversations. In the radio phone conversations on the DLR a summons is accomplished through the discrete ring of the phone and a call sign. The reason for the call has to be managed through the accomplishment of the first participant's two turns of talk. It is unlikely that the designers of the technology envisaged this to be a consequence of the design of their computer-supported radio phone system.

The resources that are available to the two parties in the call are asymmetric. The remote party has little access to the circumstances of the controller, whilst the controller has various displays which provide resources for prospectively giving a sense of possible reasons for a call, including diagrams of the main line, CCTV screens and the phone system display. However, the access the controllers have to the remote party access is limited. There is a potential for the delay between requesting a call and that call being made which can be substantial. In that period the
local circumstances of the remote party may be transformed. The controller has few resources to assess this. The organisation of the opening of the phone call, particularly its pacing, appears to be in the light of these contingencies.

It is possible to envisage ways of attending to these asymmetries. The remote party, for example, could be provided with other resources concerning the call. These could either supplement, or provide an alternative, to listening into the broadcasts from the control room, an activity that can interfere with dealing with train passengers. It may also be possible to either design or, at least, outline some requirements for ways in which it is more straightforward to determine who is calling who. Even with some minor alterations the controller could be provided with some explicit way of marking whether the call is being initiated or returned. This could be presented either visually or audibly to the remote participant.

A focus on the collaborative and interactional production of activities can identify potential areas where a novel development would be relevant, and where resources are lacking in a current deployment. It can also identify critical features in the organisation of work where particular trade-offs, given the current deployment of technology, have to be considered prior to a novel development being implemented. For example, in the DLR Control Room the loudspeaker appears to be a mundane technology which is critical for the collaboration between the controllers, providing a simple means of making available what is going on in the phone calls. Indeed, when it is inadvertently turned off, there is a noticeable disruption the controllers' work. In particular, the controllers appear to be particularly sensitive to the background noise and the other happenings in the control room. This is not only the case for the second controller, but also for the principal controller who can also hear the talk on the phone.

It may be that the loudspeaker offers additional resources to the phone handset for the principal controller, at least it makes apparent what is available to others in the room, that which may be a foundation for future collaboration with others. As well as reducing the times when the loudspeaker has to be turned off, one possibility would be to provide

---

8 The conversations between controllers and train captains are tape-recorded. For some reason, the actual operation of removing the tapes takes several minutes, whilst they are 'transcribed'. This is a regular disruption to the activities of the
headphones. However, this would both limit who had access to the calls within the control room and constrain what those who had headphones could hear in the local environment. As with many potential developments, one focus of attention within the design has to be the ways in which different forms of participation within the control room can be supported, ranging from allowing controllers to coordinate their activities with one another, through to the possibility of occasional visitors getting some sense of the state of the service. This may be provided through straightforward means, through a loudspeaker, for example, or through more novel technologies such as large public display or by advanced systems to support collaborative work. The analysis of activities in the control room reveals that trade-offs may be required when considering developments in the environment: making information more explicit could mean that it becomes more intrusive to the activities of others, making information private could remove an essential resource for collaboration.

One focus of workplace studies is on the ways artefacts are utilised within the local setting. Audio and visual materials can provide a resource that reveals the tacit and interactional practices on which the use of these artefacts rely. Although such analyses may not directly suggest alternative deployments and designs, they can locate areas of relevance for designers. These relevancies need not arise just from the concerns of the analysts. The audio-visual materials and the ways they are subjected to scrutiny can provide a warrant for the relevance of particular features in the conduct. What is relevant for the conduct of the activities in a setting may at least provide a guide to what is relevant for the design and deployment of technologies.

7.3.3 Particular Resources for the Design of Technologies

The analysis of the settings in this thesis, and the foregoing brief considerations of technologies and deployments in those domains, appear to suggest some alternative ways in which the technologies could be developed. In particular, the studies reveal issues of relevance to systems to support collaborative activities, whether these activities are co-present or geographically dispersed. To illustrate these some issues related to controllers. In the data there is a brief period of confusion when the recording of an earlier operation is played back through the loudspeaker, without warning, by the person responsible for the tape recordings.
design will be considered in relation to the activities within the Docklands Light Railway Control Room and the architectural practice, respectively.

In the DLR Control Room the two controllers utilise the displays in front of them not only as a resource for their own activities, such as typing commands to the scheduler, but also for making sense of the activities of their colleagues, for example regarding the talk on the phone. It could be argued that this support could be made more explicit, that, for example, more information is given on a colleague’s screen with regards to another’s ongoing activity. Given that the relevant information is available electronically, there may be simple ways of providing an explicit link between the call and the location from where it is being made. This could involve, perhaps, displaying the train number on the phone system screen or indicating the train involved in the call on the ATS screen, by some form of highlighting. The latter may facilitate the work of the second controller, who at present has a potentially limited view of the phone system operated by the other controller. A single interface could be envisaged, integrating the capabilities of both devices in one display.

Alternatively, the interface to the system could be redesigned completely; moving to a more direct form of input, a graphical user interface (GUI) for commands, a touch screen or even to a large configurable shared display. These possibilities could be integrated into one system and each offers some advantages.

A graphical user interface, with some form of direct manipulation, would provide users with a means of identifying for the other, in the course of making a command, the object of their work - a particular train or a section of track. Similarly, a touch screen could provide additional resources for a colleague to achieve a sense of an individual’s activities, for example a point could locate the general area of the service to which a controller is attending. A large configurable display available to both controllers would allow even greater access to each other’s activities, not only would it be possible to see the general domain of another’s activity but the actual object of a command. Such a configuration would also make shifts between distinct activities more apparent, say when moving from controlling a section of the main line to the shunting yard. It could also allow for an more flexible division of labour, with it being easier to manipulate objects in a colleague’s domain, to change the view to make such changes or to make the activities apparent to a neighbour. Configurable displays, where windows can be sized, shaped and moved
around the screen could, of course, also be organised to reproduce the ways the two individual ATS screens are used at present. This would be essential, given the delicate way in which the production of activities is managed at present.

The scheduler system could also be better integrated with other technologies. For a phone call, for example, a controller may have to view two or three screens, use a keyboard to one system, a mouse to another and the phone handset and keypad. Also, within the controller’s reach are other screens, other keyboards, emergency displays, switches and other types of phones. Some of these devices have specialised uses, and providing them in separate locales may support not only controllers in their management of routine work, but also others, when critical events, like a loss of power supply, happen. However, for some of the technology, particularly the phone system, the distribution of activities between several devices can make their accomplishment problematic. The phone system has a display, a keyboard (occasionally used), a phone keypad, a phone handset and a mouse. Actions like ‘re-calling’ or ‘hanging up’ can be distributed across these devices. By reconsidering the hardware of the technology, it may be possible to bring some of these functions together on one or two devices. For example, in other control rooms there are touch screen telephones to support such work (Heath and Luff, 1992a). These bring together the activities involved in making and taking calls onto only two devices: the handset and the display.

Any redesign of the technology should at least take account of the ways in which the collaborative activities are currently accomplished. Care would have to be taken if activities which are only partly visible at present are transformed so that they become explicit and intrusive. Moreover, if it is relevant for objects to be available for others on the display, they have to be visible at the critical times. It would be problematic if, for example, a window has unavailable because it was being altered or a hand operating a touch screen blocks an item from view.

The possibility of integrating capabilities, either in hardware or software, could be expressed in terms of what is called in software engineering, user requirements. Indeed, considering the making and taking of calls it may be possible to derive some quite specific technical and functional requirements for a new system to support the work. These could relate, for example, to the ways such activities could be performed using a conventional computer system with the minimum use of devices or
operations on the system. The materials collected could also provide a resource when considering the ways the technology could be integrated. If such a strategy for collecting requirements was followed, care would be required concerning the collaborative practices surrounding the operation of the system. So, for example, in both Chapters 5 and 6 it was suggested that a reorientation by the principal controller towards a particular technology, particularly the phone system and ATS screens, provided resources for the second controller to make sense of their activities. Integrating all capabilities onto one device may make such discrimination more difficult for others. However, there do appear to be ways in which the use of the technology could be made more straightforward, whilst maintaining the resources relied on by others, even by simply linking the information provided on each system.

From the initial observations of the work and activities of architects in Chapter 4, it is also possible to draw some implications for the designers of systems to support collaborative activities, whether these are to be accomplished in through a distributed technology or one sited in a particular location. For example, the analysis appears to have particular relevance to developers of CSCW systems aimed at supporting such activities as 'shared drawing'. At present, there is considerable interest in designing particular distributed capabilities, like shared whiteboards, where objects can be drawn and manipulated by several individuals working on different computer terminals. These individuals may be co-present, i.e. located in the same physical space (like in the early Colab system, Stefk, et al., 1987), or they may be geographically dispersed (e.g. the ROCOCO system, Scrivener, et al., 1993). Such systems require complex synchronisation of activities, so that, for example, the application can manage, or control, two individuals changing what appears to be one particular object at the same time. However, the capabilities these screen-based tools offer are still limited, particularly with respect to the availability and visibility of the actions of colleagues. Hence, several researchers have begun to incorporate video into shared drawing systems, either by merely providing extra video monitors (e.g. Bly and Minneman, 1990; Harper and Carter 1994), or by developing sophisticated mixtures of video and electronic images through, for example, half-silvered mirrors and projectors (Ishii, 1990; Ishii, et al., 1993; Ishii and Kobayashi, 1992; Ishii, et al., 1992; Tang and Minneman, 1991a; Tang and Minneman, 1991b). It is hoped that by seeing the other participant, particularly their
gaze, it will be possible to assess the locale they are concerned with on the
drawing, rather than merely the area where their screen cursor is placed.

The observations of the collaborative work between architects, not only
reveal how it is achieved and managed through talk, but also, as in the
DLR Control Room, how sensitive participants are to the orientation, and
changes in orientation, of a colleague. The emphasis on a word like ‘there’
is made sense of in relation to the orientation of a colleague to a particular
domain, as well as to the features within that domain. Providing all these
resources allows for a coordinated movement into a common domain and
for colleagues to ‘see for themselves’ a particular object and its
consequences for further action. It is unclear whether a video image of the
face of a colleague whilst one is drawing, whether it appears on a separate
display or through the monitor, provides similar resources. What appears
to be required is access to the shifting orientation of a colleague. For this,
access to the other has to be variable, so that, for example, a pointing
gesture can be seen in relation to the other throughout its production.
Being available on a different screen or being projected onto a workspace
separates the ‘gesture’ from the activity of a co-participant. The
separation of domains also renders problematic the coordination of one’s
own activity with another’s. Participants appear to require resources
where they can be sensitive to another’s activity, where a co-participant’s
conduct can be ‘seen but unnoticed’, but where such activities can be
utilised to move into different forms of collaborative activity. A more
complex configuration of technology which is more dynamic and sensitive
to the activity at hand may be required, for example using more images of
the other and greater control of the focus of those images (cf. Gaver, et al.,
1993; Gaver, et al., 1995).

However, the use of the CAD system in relation to the use of paper
plans may suggest some more straightforward matters to consider. For
example, in Chapter 4 detailed analysis of one fragment of activity reveals
potential problems with presenting the plans on the screen for co-present
collaborative activities. There appears to be considerable work involved in
securing another’s gaze on the screen. This may not only be because the
electronic screen is physically smaller than the paper plan and the details
more finely drawn. The orientation of the screen can cause problems in
seeing what is presented on the display, and the speed of the system in
drawing complex images can also make identifying what is being
discussed problematic. It is therefore not surprising that the paper plans
are utilised so often when architects discuss changes to the drawings, even though these may be out-of-date. Their size, orientation and stability provide useful resources for in situ collaboration. Hence, developments in display technologies which utilise large screens and projected images often at orientations aligned with the desk, might not just be utilised for activities when participants are making presentations to large audiences or are geographically dispersed (e.g. Elrod, et al., 1992; Tang and Minneman, 1991a), but when they are co-present and engaged in more mundane, everyday activities in the course of their work.9

In both the architectural practice and the DLR Control Room a range of resources appeared to be required in order to achieve 'collaborative' activities. Even a point to a screen had to be understood with respect to the talk on the phone between a remote party and a colleague, the state of the service and the prior activities of colleagues. Considering how personnel make apparent the details of their work to others reveals complexities in these activities glossed in current technological design. For example, cursors on the screen are utilised within current computer systems to provide a resource for individuals to identify where a particular action will take place. It has been common to consider such devices as 'pointers', and when shared systems are developed it is assumed that such pointers could do the work of 'pointing': making relevant to others a particular detail on the screen, the location where activity is to take place and even who is to perform that activity.

9 It should be noted that just by having screens the size of the paper plans and placed flat or tilted may not provide the resources offered by the paper for co-present collaborative activities. There are some capabilities for which some software or interface solution could be envisaged, such as being able to see parts of two drawings at a time and to be able to vary the amount of each drawing which is visible at any moment (cf. flipping over the corner of one page to see parts of another). However, the physical appearance of marks on a screen and the properties of glass make it problematic for others to get a sense of the details of what is being presented on the screen. The case where one architect 'points out' a precise location on a paper plan to a colleague may serve as an example. Not only does a glass screen add a small distance to the object, so the relationship between the point and its object may be seen differently by a co-participant viewing it at an angle, but also the glass can transform the nature of the object. Projected displays reduce this problem, but at present require special lighting conditions.
The materials in Chapters 4 and 6 reveal how complex it may be to accomplish such work. Getting another to orient towards an object requires more than just a ‘point’ to the screen; the gesture has to be designed with respect to the orientation and conduct of the colleague. It also has to be made sense of with respect to the emerging talk and visual conduct. A pointing, then, has to be coordinated with other activities and can make apparent not only what is relevant, but how it is relevant and the practical consequences for a co-participant. The activity is immersed within a socio-interactional context, even when it is accomplished through the use of device on the computer screen. The shape of the activity, its relationship to the object on the screen, to talk and visual conduct all feature in its design and achievement. In CSCW systems, whether for distributed or for co-present users, rather than merely increasing the number of fixed pointers to match the number of users, the approach utilised in most current shared systems (e.g. Bier, et al., 1992), it may be worth investigating more innovative possibilities, for example, providing more access to the conduct of the other (cf. Gaver et al. 1993) and allowing users to design and shape their pointings dynamically and temporally (cf. the devices for pointing in 3-dimensional space of Ware and Baxter, 1989).

Just from the initial observations in the previous chapters emerges a much richer sense of collaboration than that adopted by current designers of CSCW technologies. Rather than individuals merely requiring common views of each other or of an object, studies of workplaces imply a more varied set of computational resources for collaboration; ones where participants can have variable forms of access to each other, to electronic and paper documents (Heath, et al., 1995b). Observations suggest the range of resources required to produce and render intelligible such mundane activities as ‘pointing out’ an object on the screen. In particular, they reveal the importance of interactional activities to the accomplishment of collaborative work. The importance of talk and visual conduct may imply relevant places for specific technological investigations, both in enhancing the resources of current systems used in the work settings, and for the design of more general innovative systems.

7.3.4 Generic Resources for the Design of Technologies

The preceding review of potential design implications has outlined a range of possibilities for the development of technologies. These suggestions include remarks on the ways in which current systems are used, for
example, the problems that users have with a particular application, the tacit and common-sense practices individuals utilise to perform screen-based activities and the collaborative activities that shape and are shaped by use of the technology. From these remarks it has also been possible to make some tentative suggestions concerning improvements to the technology, relocating information between displays or adding capabilities to existing devices, for example. It is also possible to suggest some more radical alternative technologies, including the use of large displays and novel pointing devices. These particular suggestions remain tentative and their implementation would depend on a more thoroughgoing analysis accomplished for the purpose of design. There are also some generic design concerns arising from the studies.

The analyses have focused on the collaborative accomplishment of activities and on revealing those practices which are seen, but may not be explicitly noticed by the participants. The exploration of the interactional accomplishment of tasks in settings such as architectural practices and control rooms suggest various ways in which participants engage in and coordinate their activities. This in turn points to ways in which technology may need to fit with those interactional practices. These could be summarised as follows:

- the resources individuals require are thoroughly related to their ongoing activities and the activities of colleagues. Thus, the technology should not only provide ways of offering information suitable for a individual's particular circumstances, but also information regarding another's activity.

For example, when considering the architectural practice, a colleague's alterations to a plan are made sense of relation to the colleague's orientation and ongoing activities. Similarly, in the DLR Control Room a pointing at two places on the line is only seen as relevant with respect to the ongoing scheduling activities.

- the information offered by the technology may not be of the same order for different individuals. Some participants may require access to complex informational resources, others may only require broader features to be available. Therefore, when designing collaborative technology, it may not be necessary or desirable to provide a common system for each individual in a domain. Others may require asymmetrical resources which could be provided
through quite different means, by cruder displays or by audio, for example.

The requirements of the parties to the phone call in the DLR are quite different. Although each could be provided with more access to the activities of the other, it may not be possible, desirable or relevant to consider similar resources being supplied to each.

- the information required may change with respect to an ongoing activity. Individuals may need to shift between different activities and different forms of collaboration. Therefore, the technology may either have to be flexible and tailorable enough to support shifts between individual and collaborative uses or offer sufficient resources for these to be available together.

In discussions between architects, particularly with regard to the large paper plans, the co-participants can shift easily between activities where they are oriented towards distinct domains to ones where they focus on the same object. Even in the course of particular discussions about specific problems it may not be necessary or desirable for colleagues to remain focused on a common object. Similarly, the collaborative practices of controllers in the DLR Control Room rely on them shifting easily between phone system, scheduler system and other locales.

- the access to another's domain has to be tied to their ongoing activities. It may not be sufficient to provide an individual merely with access to the information that is available to a co-participant at a particular time. This information may need to be seen in relation to a previous history of actions and a trajectory of prospective activities.

For example, in the DLR Control Room a second controller can make sense of the state of the service displayed on the computer screens with respect to the ongoing phone calls between a colleague and a remote party over the loudspeaker.

- the use of the system relies on common-sense and tacit practices. Users bring with them resources for utilising and making sense of technology, providing for the intelligibility of text on the screen, a course of action on a system and the talk of co-participants. In the development of the technology designers need to take account of the everyday (and professional) skills and activities of participants in the domain.
The reading of a menu and the transition along a menu title relies on common-sense understandings of the options presented and also practical reasoning concerning potential courses of action. A movement around a building on a CAD screen relies on an understanding of the building and tying the activity into the capabilities of the system. Accomplishing activities through the computer-supported radio phone system of the DLR relies on displayed and recognised understandings revealed through the talk of a co-participant. Reading items off the screen in the control room is achieved by continually making sense of the interactional activities of a colleague, not only their talk, but also their visual conduct, their gestures and orientation. These practices rely on a foundation not particular to the domain, but on commonly available and acquired socio-interactional resources.

So, the technology has to be commensurate with a complex of activities which may be transformed in their course, and which may be accomplished by various individuals at different times and locales. The field of CSCW is one area where such technological support has been considered. Early collaborative systems focused on supporting common views to information sources, whether the participants were located in the same domain (Stefik, et al., 1987), or were geographically dispersed (Bly and Minneman, 1990; Olson, et al., 1992; Scrivener, et al., 1992). It soon became apparent that there were a range of problems associated with the design of such systems. If users were only allowed to have a common view (so called WYSIWIS) then support had to be provided to change the view. It was unclear how this capability would be presented. Control of the views could be given to specified individuals, but this was considered too constraining. It could be more generally available, to all users, but this could lead to difficulties, particularly when trying to maintain a common view of an object on the screen, and with similar 'shared facilities' being available, like shared pointers, to so-called 'cursor wars' (Stefik, et al., 1986; Tatar, et al., 1991).

One possible way of addressing this problem is to assign, or allow for the dynamic assignment of, roles and responsibilities where individuals have different participant statuses which, at least for a time, are fixed (Greenberg, 1991). With these statuses would come the possibility of different propensities to use and manipulate the system. However, providing additional mechanisms again reduces the flexibility of the
system and requires individuals to engage in explicit actions to manage the system rather than perform the activities at hand. There may be further ways of supporting real-time collaborative activities by considering, for example, crude glosses for the ways in which individuals can make use of differing forms of participation and also shift between these. These glosses could lead to a different way of considering the technological support that could be made available.

For example, if we consider that collaborative activities are achieved through different forms of participation, both with others and within activities, it may be possible to outline, for a design activity, the different participation statuses of individuals. This initial exercise could have certain parallels to the analysis of 'footings', participation statuses and production formats by Goffman (1981a) and its subsequent extension by Levinson (1988). A series of participation statuses and production formats could be outlined and differing support considered for each. As simple examples, technological support could be provided for the cases of focused collaboration and also for non-ratified participation, such as monitoring (Jirotka, et al., 1991). In the latter case, consideration could be given to the technology making obvious key components in another's activities, such as boundaries in activities, marking the commencement and completion of actions.

The case of scheduling the Docklands Light Railway can suggest how this could be designed in an interface so that general activities and boundaries are apparent, but not intrusive to the ongoing activity of others. It appears that controllers are sensitive to the typing of commands by their colleagues. Such sensitivity could be further supported by marking command completions and boundaries aurally, with minimum changes in tone for a key completion or by patterning the keystroke design so that different commands had different 'sound shapes'. They could also be marked visually, with minimum alterations in the appearance on a colleague's display, perhaps by just marking when an activity is being undertaken by another on a particular object. As these suggestions make apparent, the possibilities need not be technologically sophisticated, however, they do require some consistency to the design of such mundane interface issues as command formats. Although 'consistency' is a common concern of HCI design for 'personal' interfaces, it is rarely considered in terms of how these devices are used in natural settings, particularly with regard to the activities of others (cf. Greatbatch, et al., 1993). So, although
some technological possibilities may not require radical changes in hardware or innovations in software, they may need some reconsideration of current design principles.

As well as exploring participation status, the production format of an activity could be considered, examining for example, the differing requirements for composing an action and presenting that action to others. However, although these considerations deriving from Goffman may be a useful starting point, it would also be important to take account of the ways activities are produced sequentially. So, after commencing with a consideration of rather crude participation statuses and production formats, attention could then be paid to how support for these can be provided simultaneously and to the transitions between them. This may require a more general consideration of the domain in which the activity is accomplished. For example, one resource that is tied to the emergence of more focused forms of collaboration is changes in bodily orientation and the monitoring of these by colleagues. Quite crude technologies support these shifts when participants are co-present, either by providing the individuals with mobility\footnote{For example, the ‘whisper chair’ of Ichigawa, et al., (1995). This prototype technology ties the control of video-audio access between colleagues to the chairs of the users. To provide for different forms of participation within a meeting the ‘whisper chair’ is activated by the motion of the sitter. This was only experimental and no assessment of the exercise was given.} or by making documents which are mobile, accessible electronically (by real-time scanning of paper documents, for example, Wellner, 1992). Mobile technologies, of quite different kinds, therefore, may provide support more flexible transitions between different forms of participation.

It may also be that distributed technologies could also support more flexible shifts in orientation between different kinds co-participation by providing variable access to a remote colleague in relation to their activity (Gaver, et al., 1995; Heath, et al., 1995b). However, participants do not merely rely on the visual conduct of their colleagues to monitor their orientation, the talk also helps establish a common domain of activity. It may be that shifts in orientation are also made apparent through changes in the tonal contour of the talk. Hence, consideration may need to be
given not only into how changes in the visual access can be made apparent to others, but also in how audio support can be made more variable.\footnote{Devices may be provided by the spatial projection of sounds, by using directional microphones and stereo presentation, for example.}

The studies in this thesis not only provide a set of resources for reconsidering the design of specific technologies for particular domains, but can suggest possible generic directions for system design. A growing corpus of studies, of which those in this thesis form a part, have begun to reveal aspects of the use of technology in a variety of domains (Greatbatch, et al., 1993; Heath, et al., 1994-5; Heath, et al., 1995a). These studies reveal aspects of collaborative work that appear generic, including: the interactional production and coordinated articulation of tasks; forms of participation in tasks, the use of visual boundaries in activities, the public and the private nature of activities, and the complex interweaving of talk, visual conduct and the use of artefacts. These rely on ordinary, common-sense and interactional competencies.

Although it is useful to draw out these features, if the details (and specifics) of the analysis are not to be lost, it is important that they are seen as embedded within particular cases. Such practices constitute and are constituted by the (professional) competencies of individuals in the setting: the architects do not merely view a shared domain of objects, but collaborate over drawings of buildings, the controllers on the DLR do not just coordinate typing with talk but are involved in the collaborative scheduling and managing of a transportation service. These cases need to be seen in terms of the purposes for which they were engaged: a study of collaborative work or an assessment of a particular technology etc. before other potential implications, say for changes in technological support are considered. However, with these provisos, it is possible that maintaining a set of cases which preserve, to some extent, each of the settings' particularity, may be a starting point for other applied and practical analyses. They may, for example, allow for flexibility, revealing the organisational relevance of particular design options and provide a warrant (or at least a link) between various assertions, comments or recommendations, and a more systematic and rigorous analysis. The use of audio-visual materials may thus provide a straightforward resource for linking design considerations to particular occasions of conduct.
7.3.5 Resources for the Deployment of Technologies

Although it appears to be of marginal concern of system designers, the deployment of technology is sensitive to the organisational demands of participants in a setting. Accounts regularly report the problems that participants face when technology is deployed in a setting (e.g. Landauer, 1995; Neumann, 1995; Norman, 1988; Norman, 1993b; Page, et al., 1993; Wiener, 1993). It may be that workplace studies could provide resources prior to a technology being deployed, highlighting potential areas of concern, where particular training and skills may be required and also suggesting various features of the work which may be consequential to the deployment, for example. Such issues are addressed by practitioners of participant design (Bjerknes, et al., 1987; Ehn, 1988), but it may be that the analysis of audio-visual materials could supplement these activities.

For example, consideration of the phone system in the DLR Control Room in Chapter 5 not only reveals how the conversations through it are an essential resource for the management of the line and the handling of crises and emergencies, but that the accomplishment of these conversations rely on socio-interactional practices. This appears to be the case in a range of similar domains involving talk with remote participants over various kinds of phone systems (e.g. Goodwin, 1996; Heath and Luff, 1992a; Whalen, 1995b). Computerised and other kinds of telecommunications support are provided in all these domains for the participants to make and take calls. It would appear that analysis of the accomplishment of talk in such domains could be a useful resource for not only the design of new technologies to support conversations with remote parties, but also their deployment. For example, in yet another domain where a computer-supported phone system is being deployed, that considered in Chapter 2, Gray et al. (1993) suggest ways in which the talk between the participants could be transformed to fit with the operation of the system. Studies of social interaction could suggest what are some of the sequential implications of automating greetings or getting a call operator to say to a caller 'In the future, if you provide your name first when making a collect call, it will speed your call' (Gray, et al., 1993, p.290).12

---

12 Gray et al. (1993) times this utterance as taking around 3 seconds, ignoring the time involved in dealing with the likely response of a caller faced by the moral commitment to reply to this suggestion.
When considering such a new technology it would be essential, for example, to investigate the possible relationship between a transformation of the technology and the organisation, timing and pacing of the phone calls. Chapter 5 outlined how these are tightly interwoven. It may be that the organisation of the phone call may suggest places when and where other information could be made available. There may be a potential to exploit long gaps and repetitions, tied to the operation of the phone system, for the provision of information on other displays. But it would also be necessary to consider how changing the resources available to one of the participants may change the work that has to be accomplished in the opening of the phone call. Altering who speaks first, what information they have regarding the other party and, even, the sounds and noises available can transform the nature of interactional activities accomplished in call, turn-by-turn. This requires a consideration of the social accomplishments and demands of the turns of talk in a conversation.

Such an analysis would also have to be undertaken of the demands and requirements of other parties, not just the 'user' of the computer-supported system. In the DLR case, this would require an examination of the activities of other participants, particularly of train captains, to investigate how additional resources could fit with their current activities. It would also require a consideration of the consequences on the activities of others not directly involved in the operation of the phone system, for example, others in the control room. Transforming the technology may also transform the resources available for collaboration.

Analysis of the detailed in situ accomplishment of activities can have implications other than those for possible technological designs. They could suggest some of the consequences of organising talk in various ways, for example, of the particular requirements demanded and produced by 'formal' components. It may also be that studies of social interaction could provide resources for the training of participants. In the DLR case, the resources could inform the training of controllers and train captains with respect to their practices through the phone system.\(^{13}\)

Moreover, the analysis of workplace activities could provide resources for those considering changes to the division of responsibilities between participants, for example those of the CRSs, CRAs and CRMs in the DLR

\(^{13}\) In a similar way to that of the analysis of interactional conduct of general practitioners using a computer system informs the practical training of doctors in their use of technologies within consultations (Greatbatch, et al., 1995).
Control Room. On both DLR and on London Underground, there are numerous proposals for reorganising these responsibilities with respect to who is concerned with strategic management, information dissemination, how many individuals are required to perform some activities, etc. A consideration of the work of the controllers in Chapter 6 reveals how their 'public' activities, their collaborative activities on the phone and the ATS, are resources for other individuals in the control room. Transforming the responsibilities of the controllers, or the number of controllers, would transform a tacit resource which others rely on.

Hence, naturalistic analyses may not just have implications for the design and assessment of technology, but could provide resources for the deployment of systems. This could be by undertaking particular case studies in the domain in question. It may be also that the growing corpus of studies of workplaces, coupled with analyses of the accomplishments of related social activities, may provide additional resources for those concerned with deploying systems. These could at least provide some background to the issues that may be consequential for the introduction of a new technology, for example how activities are collaboratively accomplished and how an individual's activities may relate to those of colleagues. The corpus then could provide ways for framing a more circumscribed study in the relevant domain.

### 7.3.6 Resources for Supporting the Design Process

It has been suggested that one problem that is likely to arise from utilising naturalistic studies for design activities is how to provide better ways for designers and ethnographers to communicate with one other, particularly as each come from different disciplines and have different purposes. For example, one practical approach would be for designers and ethnographers to collaborate over a piece of fieldwork, for example, by ethnographers 'debriefing' designers about a particular setting (Randall, et al., 1994) or by designers setting out the questions they need answering. It may also be possible to identify a variety of different 'uses' of ethnographic materials at various stages in the design process. For example, early in the process fieldwork can inform designers about a particular domain, and later it could be utilised to evaluate various design decisions (Hughes, et al., 1994). The materials gathered from a fieldwork exercise may also be structured so that tools could be developed to assist analysts in the drawing up of a requirements specification. These materials may outline
particular aspects of the work revealed through an ethnographic study, the layout of the domain, the rôles of participants or 'workflows', for example (Hughes, et al., 1995).

By focusing on the problems of communication it is also possible to suggest the potential for tools to support fieldworkers and methods for shaping the analysis of observational materials (Hughes, et al., 1995). For example, sophisticated tools could be used to map out local work settings, annotating these with transcripts and notes and in outlining flows and relationships between activities (Twidale, et al., 1993). These suggestions often parallel the development of practical methods for design drawn from other traditions (cf. Maclean, et al., 1989). There may be problems, however, in utilising such tools given the analytic orientation of the ethnographic studies they are intended to support. 14 The rich insights of these studies have been attributed to their focus on the sense-making practices of members revealed through the observation of members' accomplishments in particular domains. This is in contrast to other orientations which bring to the analysis a priori categories, prescriptions and stipulations which pre-specify what can be considered relevant for an analysis. It is unclear how any method, tool or technique developed through an ethnomethodological orientation could be developed that did not contain some means of prescribing the particular features of a setting to be relevant. Even quite general techniques like asking designers to formulate a set of questions to frame an analysis, providing tools for mapping out the spatial organisation of a domain, and techniques for examining workflows, may constrain what may be revealed by an ethnographic study. Such tools may provide useful resources for social science researchers from an ethnomethodological tradition, however they may be less useful for practitioners aiming to undertake an analysis for the purposes of design. They may be used as stipulations for particular circumscriptions of the analysis. Unfortunately, these tools may draw attention away from the details of how naturalistic activities are achieved in the workplace, neglecting the tacit, coordinated and emergent nature of the production of activities.

14 One problem that is faced by Design Rationale (Maclean et al. 1993), for example, is that it does not distinguish between different activities which can be characterised as 'design' and the range of activities for which general tools and techniques for 'design' may be put.
Although tools which support graphical presentations of a setting and of workflows could make apparent to designers appropriate features of an analysis, their use would have to ensure that they do not highlight particular aspects of workplace activities; presenting activities as flows of individual tasks, or the context of a work setting in terms of a physical space, for example. It may be that the suggestions for tools and techniques to support ethnographic studies may point to more general difficulties with the ways in which studies of workplace activities may be presented, it being hard to conceive of graphical and schematic representations which preserve the details and richness of the original analysis. It may be also hard to envisage how plans of rooms annotated with copious fieldnotes could be made sense of and utilised by designers. As with other design tools, these capabilities may impede rather than facilitate the work of programmers and system designers, or, in order for them to be of use, they may be utilised for quite different practical purposes.  

Audio-visual materials may provide a resource for making materials related to particular cases available, for more than a single purpose and a single audience, whilst maintaining attention on the detailed nature of the conduct. Audio-visual materials make possible presentations, alongside other relevant material to other analysts and can be a resource for others involved in the design process. They can be utilised for insights into a domain, to draw out requirements for a innovative technology, to provide a resource for a design activity, for assessing a system, or to contribute to a deployment strategy. Video provides a particularly rich resource to offer others in the design process with access to a setting. The videos and the supporting analysis can be seen as a resource for envisioning new systems, both particular and general. They could form the basis of an alternative to, or a valuable and animated extension of, such design processes as scenarios, storyboards and walkthroughs (Carroll, 1990; Lewis, 1990;  

---

15 See Button and Sharrock's (1994, 1996) analysis of the use of a CASE methodology and a requirements representation tool for the practical purposes of software engineering. Button and Sharrock's studies of software engineers suggest another way in which an ethnomethodologically-informed study might contribute to the practices of systems developers. Through an analysis of the practices of software engineers and programmers it may be possible to derive direct outcomes such as suggestions for tools that would support their everyday accomplishments, and methods that would be sensitive to the practical contingencies facing designers and implementors.
Yourdon, 1989); the materials providing a warrant, or at least a link back to the requirements, and the settings of possible use.

Analysis of the details of technology use can contribute to the assessment of a technology, revealing tacit features which may not be obvious to designers, or explicated by users. Video-based materials could thus be useful when considering an alteration to a system, an upgrade or improvement to an interface, input or output devices, or a replacement for an entire system. They can be a powerful resource for suggesting some of the difficulties encountered in using a technology, but also can reveal the ways in which current technologies and artefacts fit smoothly into working practices. The audio-visual materials can also be relevant when deploying a technology, from minor contributions such as the orientation and placement of technology through to the implications of likely changes to responsibilities, demands on users and training.

Video is the critical feature of these contributions. It can provide for systematic, rigorous and detailed analysis of naturally occurring activities, warrantable analysis tied to the activities of participants and some access to particular domains for a wide variety of design purposes. Of course, it could provide for more, for example the video materials could provide a resource for analysis from a range of orientations: including simple task analysis, Cognitive Scientific approaches, and no doubt Distributed Cognition, Activity Theory, and the orientation suggested by Brown and Duguid. This is viewing the contribution of video as solely providing a supplementary resource for design activities. More important is the contribution the video makes possible to the nature of analysis that can be undertaken, not just showing behavioural features of conduct, but also revealing the detailed organisation of naturally occurring activities as the practical accomplishments of participants. An interactional framework provides resources with which to warrant an analysis and warrant its relevance through the participants' own contributions. This suggests a way of providing some rigour, and making systematic, the problematic activity of analysing ethnographic materials.

7.3.7 Discussion: Naturalistic Analysis and Design Considerations

By attending to the details of activities revealed in recordings of naturally-occurring activities it is possible to draw out implications for the assessment of particular interfaces, systems and deployments of technologies. From the analyses it is also possible to raise issues that may
of concern for design. Revealing the nature of the practices of participants in a setting suggest some features that, at least, designers should be sensitive to. By considering several of these studies, including the four cases in this thesis, some more general issues emerge that seem to have resonances with the concerns of developers of novel technologies. Indeed, recent technological reports draw on related studies referring to such issues as 'peripheral monitoring', 'shadowing', 'fluid transitions between forms of working' and 'mutual awareness' with respect to technologies as diverse as virtual reality environments, shared groupware systems, and toolkits for system design (Benford, et al., 1996; Bowers, et al., 1996; Dourish, 1996; Gutwin, et al., 1996).

Despite this technical work there have been criticisms of the ways in which practitioners of workplace studies consider the relationship between the studies and design. For example, Plowman et al. (1995) point to the vagueness in their formulation in terms of 'insights', 'suggestions' and 'options'. However, this criticism appears to belie the variety of purposes to which workplace studies have contributed. They have not just aimed to produce some implications for general designs, but, for example, have been part of collaborative exercise between social scientists and designers, either to develop particular prototype systems (e.g. Bentley, et al., 1991; Heath, et al., 1995b) or to contribute to the general design process of large corporations (e.g. Bowers and Button, 1995; Harper, forthcoming; Suchman, forthcoming). It may be that what Plowman et al. require are more explicit links between the analysis and the design exercises in question. However, workplace studies themselves point to the danger of making such assertions. For example, Button and Sharrock (1994, 1996) reveal how the development of technology itself emerges through the management of a range of organisational and local contingencies. Thus, it would be problematic for analysts to assert a causal relationship between an analysis and an ensuing technology. For similar reasons, there is a reluctance to proffer simple generic guidelines, procedures and methods. Such resources would have to be used contingently, given the demands of

---

16 However, what is perhaps strange about their critique is that Plowman et al. do not question how the design implications are derived from workplace studies, but criticise the authors for being drawn into making such statements, and because of this design implications are framed in a tentative way. Despite mentioning cases where specific design recommendations have been drawn from workplace studies, Plowman et al. use the general framings as a resource to summarise the general problems of moving from the workplace studies to 'technical', design implications.
the system development setting. This is not to say that there is no interest from analysts about the possible technological implications of their work. However, there is a great deal of caution in asserting unwarrantable claims from the analysis.

As Plowman et al. do state, workplace studies offer the potential for uncovering the details of in situ work practices, by revealing, for example, the tacit accomplishments and routine achievements of participants. Revealing practices oriented to by participants, that are produced and recognised with respect to the activities of others, provides a warrant for the relevance of an activity not only for the participants themselves, but also for analysts. These routine and everyday accomplishments could then provide a substantive resource for a design exercise, whether this is utilised to suggest particular requirements for a technology, specific features in the design of an interface, or guidance for the deployment of a computer system. So, for example, the tacit practices surrounding the use of paper architectural plans might suggest requirements for a technology aimed at replacing them. Or, the ways in which the use of current technologies can be improvised and shaped from moment-to-moment in relation to a range of contingencies may provide useful guidance when new technologies are about to be deployed.

There are substantive implications for design activities that can be drawn from the analyses outlined in this thesis and from a range of other studies that explore the detailed accomplishment of activities in workplace settings. It is possible also to draw out implications from such studies for the design of a variety of generic technologies, for example, in terms of the ways in which the use of conventional technologies are articulated to facilitate collaboration; to suggest the nature of activities that might be supported by future systems. In addition to those mentioned in the previous sections, researchers have outlined the consequences for the design of systems to support cooperative work (e.g. Bentley, et al., 1992; Heath and Luff, 1992a); and the resources provided by audio-visual connections (Heath, et al., 1995b), HyperText systems (Harper and Sellen, 1995), workflow technologies (Bowers and Button, 1995; Suchman, 1993a) and mobile devices (Luff, 1994). Although these implications tend to be preliminary remarks and suggestions concerning the capabilities which such technologies could offer, several experimental or prototype systems have emerged leading from these studies, and others form part of longer term technological design exercises (such as those carried out within Xerox
and Rank Xerox Research Laboratories). Workplace studies, then, would appear to have a wide range of specific and general consequences for the system design process.

However, it may not be these proposals that have the most significance for system design activities. The consideration of the details of particular cases may provide more substantial resources for considering the development of tools and techniques for the detailed accomplishment of activities, rather than, say, broad aims such as supporting collaboration and cooperation. Hence, the importance of workplace studies may be in providing an accessible corpus of studies of social action, in developing a framework for conceptualising the activities revealed in such materials, and offering analyses of the detailed ways in which activities are accomplished in natural settings.

7.4 CONCEPTUAL RELEVANCIES

In outlining some of the shortcomings of current efforts of using ethnographies for design, Anderson (1994) (and Button and King, 1992) suggests that this rests on a misunderstanding concerning the nature of ethnographic studies. Ethnographies are more than just fieldwork (or 'hanging around'), but are undertaken within a particular analytic orientation. Anderson continues by suggesting that what designers may indeed require is just fieldwork, observing the happenings in particular domains. Indeed, there are now numerous cases within CSCW where fieldwork observations are reported (e.g. Ackerman, 1996; Bellotti and Bly, 1996; Berlin and Jeffries, 1992; Moran, et al., 1996; Murray, 1993; Schwab, et al., 1992; Watts, et al., 1996). However, even within a design exercise some warrant is required for the observations being made. Furthermore, it may be preferable if the observations were not made in accordance with categories and stipulations prescribed before the analysis was undertaken, whether these are broad such as 'gender', 'power' and 'social norms', or narrow, a 'plan', a 'goal' or 'negotiation'. These requirements are actually requirements of social science analyses and as such parallel earlier debates about formalising knowledge, social behaviour, categorisation and relevance (cf. Garfinkel, 1967; Heath and Luff, 1992c; Weber, 1947). Ethnomethodology is a radical response to these problems. A focus on social actors' situated, practical accomplishments through interaction is inherent in this programme.
Nardi (1996c) criticises what she calls 'situated action' approaches as being too specific, only being relevant to the particular domains under scrutiny. Her criticism reveals continuing problems with an understanding of these orientations, not only by those interested in HCI and CSCW, but also within social science. 'Situated action' is either associated with 'in context' or 'unplanned' activities or contingent (and almost irrational) actions. By understanding the programme in this way, cognitive models can be extended and seen to deal with changes in plans, goals and tasks. Activities are merely being considered as more dynamic and variable. However, this understanding misses the emergent, collaborative and, most importantly, the interactional nature of situated action. Social actions are continually shaping and developing the context in which they are understood and produced, and they are also shaped by it; activities and interaction are thoroughly interwoven.

Other approaches, even those radical suggestions by Brown and Duguid and proponents of Distributed Cognition have impoverished conceptions of the social and context: context being conceived spatially or ecologically, social action as either a combination of individual capabilities or a vague amalgam of the cultural with the individual. These distinctions do not just rest on socio-theoretical concerns, or part of the long-time debate between psychology and sociology, but also on different orientations within sociology. Taking ethnomethodological insights seriously would involve a rethinking of many of the conceptual underpinnings of approaches for HCI and CSCW, and to other areas of Computer Science where there is a need to develop systems applicable to users and organisations (i.e. requirements, domain and systems analysis). Proponents of approaches such as Distributed Cognition and Activity Theory recognise that such a respecification of concepts is required, and that the conceptions of 'task', for example, in task analysis and in most cognitive approaches to HCI constrains both the analysis of human-computer interaction and the development of new methods for design (Hutchins and Klausen, 1996; Nardi, 1996a; Norman, 1993b). Moreover, it has been proposed by psychologists that communicative activities should also be considered more in terms of social actions (Clark, 1996a; Clark, 1996b), artefacts should be considered as supporting activities that are 'external' to the individual (Hutchins and Klausen, 1996; Norman, 1991)

17 See also the debates in the Cognitive Science Journal (Vera and Simon, 1993a) concerning Suchman's (1987) 'Plans and Situated Action'.
and there are important tasks that are peripheral to an individuals' main area of concern (Brown and Duguid, 1994). However, these extensions to the social are conservative, maintaining a conception of communication largely in terms of fixed meanings passed in messages and where the context of activities is largely considered in terms of the physical location of the participants.

The analyses in this thesis, with others drawing from an ethnomethodological orientation and conversation analysis, suggest possible shortcomings in the conceptions of human conduct that either explicitly or implicitly infuse current design methodologies. For example, the analysis of Greatbatch et al. reveals the limitations of current conceptions of 'consistency' within interface design (Greatbatch, et al., 1993). An interface needs to be consistent not only with respect to an individual's activities, but also with respect to the activities of others who are not direct 'users' of the system. As another example, 'workflows' may constrain the ways individuals can accomplish next activities, and fixed categories of information and activities may ignore the contingent ways by which participants can utilise descriptions and specifications (Bowers and Button, 1995; Suchman, 1993a).

More generally, these studies suggest a potential for a more radical shift in the conception of tasks and activities in HCI tasks, one where activities are transformed from moment-to-moment through the interactions of various participants, shaping and shaped by those contributions (cf. Luff, et al., 1992). This would suggest a conception of activities that is more dynamic, more flexible and more open to the contributions of others. Such a conception of task also transforms the way the use of space, the use of artefacts and communication is conceived. Movements in a setting, the location of individuals, the use of objects or utterances between a pair of participants might then be considered in terms of an ongoing activity or a web of interactional activities. Thus, a reconsideration of the conception of task can lead to a respecification of a range of allied conceptions, including those which are general, such as 'space', 'artefacts' and 'communication' and those which are more particular to system use and design, such as the nature of 'interaction' within human-computer interaction, the varied uses of 'information' in system specification, and the capabilities of 'objects' in system design. It also, makes problematic the distinction between the 'individual' and the 'collaborative', the 'shared' and the 'group' within groupware and CSCW.
Rethinking these conceptions may suggest ways in which some methodological resources for system design could be developed. Initially, these may be in terms of some crude heuristics for such things as guidelines for requirements analysis, suggesting ways of rethinking commonly held assumptions concerning particular domains. For example, it may be possible to outline, in a convenient manner, an encouragement for designers to reconsider what appear to be 'individual' tasks; to investigate how the accomplishment of an activity may be accessible, viewed or monitored by another and how an activity may be designed in such a way. It may then be worth considering how the 'public' nature is managed within the local domain, to outline the challenges and constraints that designers of a new technology would have, at least, to consider within the design process.

The relevance to system analysts and designers of respecifying tasks and activities might be more apparent if a novel example of an activity not considered in any detail in the previous chapters is taken. So, if, within the Docklands Light Railway, a technology was being considered for supporting, or replacing, the reports of failures and delays, an activity carried out on paper forms. The system analysts would need to take account of whether the apparently individual activity of writing the report is, at the time it is written, relevant to others in the control room. This may require observation of the activity and analysis of audio-visual materials to investigate how the activity was shaped by others. This might involve examining the ways that contributions to phone calls by others are shaped with regard to the future writing of the report by another, or how the individual organises from moment-to-moment the writing of the report in the light of the ongoing activities of others (like a controller's activities on the phone). When considering a novel technology it would be possible to investigate, at least in a preliminary fashion, whether there were any obvious ways in which the technology might impact on these public activities. It may transform, for example, the ways in which the resources for report writing were made available by a controller (i.e. how failures are made apparent by the controller, seen to be relevant by the report writer, and also how that seeing is itself recognised by the controller). The choice of technology may also have an impact on the ways in which the report writing is publicly produced. Such
considerations would rely on an understanding of the sequential production of the activities within the domain.18

This is quite a different analysis of a domain to that currently proposed in task analysis techniques. A consideration is given to the resources through which activities are produced in concert with those of others, with regard to the purposes of other participants and also how they inform, or shape, another's activities. This renders problematic the conception of a task or an individual activity. This problematising itself, however, can be a resource for analysis. It would be strange if through adopting an analytic conception that calls into question the unproblematic use of a priori categorisations and stipulations, an alternative set of categorisations and stipulations were proposed, which were just more precise, open or flexible. Rather, studies of the detailed accomplishments of naturally occurring activities drawing on an ethnomethodological orientation focus analytic attention on the particularities of the methods through which members render intelligible social actions. Hence, it is on the detail of these methods, not the prior conceptualisation of the conduct on which the analytic attention should focus (cf. Lynch, 1993).

Proposing guidelines for system designers, even if these are provisional and couched with various provisos, rests on the employment of particular conceptions of activities and also imply methods which may have to be employed. However, some suggestions on how to provide some practical assistance can be given to designers (Heath and Luff, 1992b; Jirotka, et al., 1995). Although more general and extensive utilisation of such an orientation for design remains to be undertaken, these proposals at least suggest alternative forms of support for design activities. Though not

18 See a related analysis of the activities of financial dealers (Jirotka, et al., 1993), where an analysis of collaborative tasks, making use of audio-visual materials, contributed to a significant reconsideration of the design of a technology for deal recording. The proposed technology, a voice recognition system, rested on particular assumptions concerning the activities involved in dealing and recording deals. Analysis of video-recordings of this domain suggests how these assumptions were misconceived, and shows the considerable problems the technology would have to overcome. As in the case considered here, a respecification of the problem can refocus the nature of the technology proposed. For example, if supporting the real-time recording activity and on co-present collaborative work could suggest a focus of technologies used in the setting. Mobile technologies and their integration with other 'shared' or 'public' devices may be a useful place to consider such recording work (cf. Luff, 1992).
being constrained to using the results of experiments on individual users, or modelling the cognitive behaviour of users, or utilising pre-determined formulations for analyses, naturalistic analyses can nevertheless provide some resources that can be robust. Their rigorous nature is not achieved through being able to fit the analyses to some pre-defined model and enumerating behaviour in terms of some constrained categorisations stipulated by the analysts, but through the analyses being displayed as being relevant to both the analysts and the participants.

This is not to say that an analytic framework for utilising ethnographic and audio-visual materials for the purposes of design can be developed merely by drawing from current research developed from an ethnomethodological or conversation analytic orientation. Despite various recent contributions from several researchers, the analysis of artefact-based activities remains problematic (Goodwin and Goodwin, 1996; Greatbatch, et al., 1993; Heath and Luff, 1992a; Whalen, 1992). It is unclear whether the displayed understandings revealed through the sequential nature of talk can be so strongly asserted when analysing non-verbal activities. It is also unclear whether it is relevant or appropriate for the analysis of workplace activities to aim to uncover a set of generic practices of the same order suggested by early conversation analysts concerning the organisation of talk (Atkinson and Heritage, 1984). In a cogent critique of such work in conversation analysis, Lynch, following Garfinkel, has suggested that this programme misses the just this, the just that, the haecctilies (or earlier the quiddity) of the very activities which are under scrutiny (Garfinkel and Wieder, 1992; Lynch, 1993). These make up the unique quality of activities accomplished in natural settings. By attempting to uncover a broader set of generalities of talk and related activities, conversation analysts, and to some extent interaction analysts, will neglect in their analysis the unique qualities through which naturally occurring activities are produced and rendered intelligible - Garfinkel's 'unique adequacy requirement' (Garfinkel and Wieder, 1992). Lynch further outlines some of misunderstandings concerning the 'application' of ethnomethodological studies of work (and with conversation analysis). Principally these consider that ethnomethodology is an observational science, with findings which are general and applied, and which can be separable from the domains in question. From his critiques he goes on to propose a post-analytic programme for ethnomethodology to return to some of the original insights.
of Garfinkel (particularly with regard to Garfinkel's (1967) breaching experiments. These criticisms may also be of concern for those seeking to provide some generalisable advice, policies or guidelines for designers. Maintaining a strong foundation within an ethnomethodological tradition whilst seeking to also be relevant to designers would, according to Lynch's thesis, be impossible. Thus, as with employing a strict conversation analytic conception of the sequential, some of the ethnomethodological (or certainly post-analytic ethnomethodological) concerns may have to be foregone if these other purposes are to be attended to.

However, Lynch's critique may still have some import for naturalistic analysis for design. It points to the importance of a case-based approach. Materials drawn from a corpus of settings provides for analyses, even if these are principally ethnographic, that have to be grounded within the production of accomplishments within the domain. These may have to be understood with the assistance of additional ethnographic resources. It may be that more extensive fieldwork has to be undertaken, but it is apparent that audio-visual materials need not only be a resource for such ethnographic work, but also can assist with shaping the organisation and focus of fieldwork. As yet it is unclear whether post-analytic ethnomethodological studies will be of interest to social scientists, indeed Lynch and Garfinkel express both scepticism and antagonism over whether such studies can contribute to academic concerns in general. However, they do suggest that such analyses may bring back to the domains findings of relevance and interest. Unpicking the tacit, common-sense practices of scientists and mathematicians, for example, may be appropriate and relevant to scientists and mathematicians. There are various purposes for which such analyses could be put, for education and training, for example. Workplace studies then may then resonate more

---

19 So, Lynch and Bogen state, in comparison to conversation analysis:

Our descriptive language is ordinary, and, as such, it bears countless implications that we cannot help to specify or control. Our descriptions are assailable, defeasible accounts, uncommitted to any analytical model of conversational pragmatics or communicative ethics. Our ethnomethodological approach is therefore postanalytic in the sense that we presume that, and selectively describe how, the sources of intelligible action and defensible judgement are not contained within even the most elaborate system of prescriptions and specifications.

(Lynch and Bogen, 1996, p. 287)
with recent concerns expressed by ethnomethodologists than other analyses which have developed from conversation analysis.20

Indeed, a rigorous examination of naturally occurring social activities would appear to offer a way of beginning to unpick the 'just thisness' of activities in natural settings, suggesting at least where more extensive observation, even participant observation is required. The focus also would seem to be an extension of the concerns of Sacks into exploring natural activities in their domains of occurrence (Sacks, 1963; Sacks, 1992). In a different way workplace studies which draw from a socio-interactional orientation are seeking to break apart and unravel the machinery of social action.

These debates also have a relevance for the more mundane work of those concerned in the various activities involved in the design of new technologies. For in order to reveal the requirements for a technology, to design a tool appropriate for a particular activity, to assess the utility of a built system, to plan the introduction of system and even, more generally, to understand how technology is utilised to accomplish everyday activities, attention has to be paid to the details of the accomplishment of activities within the domain. Current methods, and the conceptions that underpin them, offer limited resources to assist with such design related activities. The activities of interest are social and collaborative, and so it would appear to be relevant that social science may contribute to rethinking methods and approaches within the system development process. In the use and design of computer systems it is particularly apparent that the activities surrounding these artefacts are ongoing, improvised and emergent. By examining their use in a variety of contexts it has been possible to show how these computer-related activities are thoroughly interwoven within the interactions between participants collaborating in different ways within work settings.

Although it has been suggested that focusing on the achievements of participants unduly constrains the novel possibilities of technology, close attention to how these are accomplished reveals how rich and complex workplace activities are. It may not be so straightforward to assert that only conservative technologies will emerge, or be required, from such

20 For example, the programme of work exploring institutional talk drawing from conversation analysis has been recently critiqued as itself relying on stipulative and a priori categorisations, particularly of organisational identities (Francis and Hester, 1996). See also discussion by Watson (1994).
analyses. It may be that in order to support the activities accomplished in and through interaction requires radical developments in technology, for example in their capabilities to provide variable visual access to another, mobility within a setting and public resources. However, exploring innovative design activities and processes requires methodological and conceptual developments. It may be that the contribution of social scientists may be more relevant here. Several such developments have already been proposed within the fields of CSCW and HCI, but in some ways the focus of these fields, on the individual and the collaborative, prescribes their concerns. The analyses presented in this thesis do not fit comfortably within either field, being explicitly addressed to the collaborative and the individual, towards 'shared technologies' and screen-based activities.

It may be there are other audiences for which workplace studies are relevant, for example, software engineering, or more specifically Requirements Analysis. Here, the distinctions between the individual and the collaborative are of less importance, given the practical concerns of the fields, and the critical need is for methods and approaches for uncovering how activities are accomplished in everyday settings. Workplace studies could thus inform the conceptual foundations to the field, as well as informing practical methodologies and approaches towards design.

7.5 SUMMARY

The previous chapters have been concerned with various configurations of technologies in several settings, including software applications on personal workstations, for transparencies and architectural drawing, a system to support the making and taking of phone calls between personnel and a technology for several individuals to view and change the running of trains on an urban transportation system. Attention has focused on differing aspects of these technologies: the use of menus to accomplish a quasi-naturalistic task; the production of working drawings; the interaction on the phone system when mediated by the technology; and the activities of controllers collaboratively using a system. Thus, the studies can be seen to relate to a range of concerns both within HCI and CSCW.

From the observations in the settings it is possible to draw various implications for the design of systems in these domains. The variation in
the settings means that these consequences are of quite different kinds, from notes on the particular advantages and disadvantages of a specific system through to considerations of potential extensions and alternatives. These include offering suggestions for alterations to the interface, for different configurations of functionality and additional areas where technological support may be possible.

Of course, the nature of these studies, the fact that they were not carried out as part of any particular design process, means that the implications can only be preliminary and provisional. In order to develop a system, even a prototype, extensive further analysis, specification and design activities would be required. These would need to relate to the practical purposes of design. Nevertheless, a brief exploration of some specific implications provides a foundation for a consideration of the relationship between naturalistic studies of the workplace and design.

In particular, what appears important, both for the analysis and for the purposes of design is to provide some warrant for the observations being made. Casual observation of the workplace or just 'hanging around' is not satisfactory. Several ways of framing analyses of naturally occurring activities have been suggested. It appears that whilst some of these, like Distributed Cognition and Activity Theory, take seriously the need to examine the social conduct of participants they still rely on critical specifications of activity drawn from psychology and maintain a conception of the mental and cognitive processes of individuals. An alternative programme has emerged that draws from a particular social scientific orientation: ethnomethodology. Studies drawing on this framework focus on the practical accomplishment of social actions, and recently, particular interest has been paid to the possibility of utilising such studies for design. These attempts have raised a set of concerns not only related to the practical purposes of communicating between the individuals involved in such projects, but also regarding the transformations necessary to make studies of social actions appropriate to designers and in warranting analyses for design.

In this chapter, various resources have been outlined which may be relevant to different design activities. These resources emerge from the analysis of naturally occurring activities through the use of audio-visual materials and fieldwork. Such materials make possible an analysis that explores the sequential organisation of social action, and allow for a detailed examination of the accomplishment of particular activities.
through talk, visual conduct and the use of artefacts. The analyses in the previous chapters have drawn on this framework to reveal some of the ways in which technologies are utilised in natural settings. In this chapter, these analyses have been used to illustrate potential implications for design. It appears that certain substantive implications can be drawn, relating to the ways in which particular technologies have been deployed, possible designs for alternatives and more general suggestions for developments in system design. However, the more significant challenges implied by such an approach may be those regarding the conceptual and methodological underpinnings of fields like HCI, CSCW and Requirements Analysis, fields which seek to examine the use of technologies in order to develop better and more appropriate technologies. Studies, such as the ones outlined here, suggest resources with which to reconsider how activities are currently conceived in these fields, and also how activities on and through technologies are accomplished. They suggest a different way of considering human-computer interaction, computer-supported collaborative work and user centred system design.
Chapter 8

The Social Organisation of Human-Computer Interaction

a brief discussion of experiences and possible directions for future work

There is no stability in this world. Who is to say what meaning there is in anything? Who is to tell the flight of a word? It is a balloon that sails over the tree-tops. To speak of knowledge is futile. All is experiment and adventure. We are forever mixing ourselves with unknown quantities. What is to come? I know not.


8.1 INTRODUCTION

This chapter concludes the thesis with a few remarks about how the reported studies were undertaken and some possible areas which may be worthy of future research. In carrying out these studies a variety of materials have been utilised, including both ethnographic observations and audio-visual materials. The use of audio-visual recordings, in particular, raises critical issues not only with respect to how the data are collected, but also how they are analysed and presented. Section 8.2 briefly discusses some of these problems and practical solutions that have
been adopted throughout the course of the research in order to address them. Inevitably, various alternative directions for analysis could have been taken and further work could have been carried out if more time was available. Section 8.3 outlines some of the issues which it may be interesting to consider given the materials that have been collected. It also summarises some other possible areas of future research, in relation to the analysis of artefact-centred activities and the practical consequences of such studies for the design process.

8.2 EXPERIENCES IN THE USE OF AUDIO-VISUAL MATERIALS FOR THE ANALYSIS OF HUMAN-COMPUTER INTERACTION

8.2.1 Data Collection

One problem which arises when collecting audio-visual materials of screen-based activities is how to gather data that not only reveals the details of the activities on the screen, but also makes available for analysis the orientation and participation of the individuals in the setting. In all the settings considered in this thesis two cameras were used, one focused on the computer system and another on the local environment surrounding it. Using two cameras immediately sets up two problems: what camera angles to adopt and whether to mix the recordings.

In each of the settings different solutions to the positioning problem were adopted, taken with respect to the demands of the particular domain. So, for the analysis of the PowerPoint application in Chapter 3, one camera was focused, over-the-shoulder, on the screen, whilst the other was positioned to the side, allowing access to the orientation of the user in relation to what was being carried out on the computer. In the architectural practice, reported in Chapter 4, a camera angle was adopted which was also over-the-shoulder, on the screen-based activities of the architects. The actual position of this camera was changed as different architects in the office became the focus of the data collection. The other camera angle varied throughout the study. Initially it was a wide-angle view of the drawing office, so that the particular architects being focused on by the other camera would always be in view. For the remainder of the data collection the second camera focused on one particular architect, and was positioned at an angle (front-on) where his activities on the computer could be seen in relation to his activities on other artefacts in the local environment. This view centred principally on his drawing board on
which the architect's paper plans were placed. In the DLR Control Room, examined in Chapters 5 and 6, one faced the direction of the principal controller's position and the technologies surrounding it. The other camera was more broadly focused, recording not only the principal and the secondary controller, but also the activities at the other end of the console where the CRAs were seated.¹

The various angles that were adopted in these studies not only reveal the constraints of the different settings such as where cameras could be positioned, where potential obstructions to views were likely to be, and the position of the technologies, but also give an indication of the range of problems faced when trying to collect materials recording screen-based activities. It is difficult to select an angle that facilitates the analysis of activities with respect to the screen whilst not constraining the type of materials collected.

If too broad a focus is adopted then the resolution will not be fine enough to allow for a detailed analysis of the happenings on the screen. When collecting the data, it may not always be apparent what screens should be the focus of attention. In the DLR Control Room, for example, there were four screens surrounding the principal controller, three of which were displays to systems that were continually of use. Because of the orientations of the various screens, no single camera position could provide for all these to be available.

If too narrow a focus is selected, say just on the screen, then the other angle has to provide for a sense of the individual's activity on the computer, such as whether the mouse is being used or where the user is looking. This can preclude using the second camera for a more general shot including other participants in the domain. This is one reason why in the DLR Control Room a close focus on the screen was not adopted.

Therefore, in each of the settings various choices had to made that would provide materials of adequate quality, where the cameras were not obtrusive and which provided a significant set of resources for analysis. This, inevitably involved some compromises.

¹ Note that in none of the settings was video output taken directly from the computer monitors. On the computer systems analysed this was either not technically feasible, was considered to be too intrusive to set up, or would be potentially dangerous.
A possible way of improving the quality of materials would be to use more cameras. However, even with two cameras there were difficulties analysing one tape in relation to another. With two sets of recordings the materials can either be viewed one after each other or mixed into a single image; the mixing being done either the recording takes place or after. For the PowerPoint exercise (and for some of the data collected in the architecture practice) a recording was made which mixed the images at the time of collection. This did not preclude the need for two tapes, however. The detail required in the analysis still meant that a separate tape was necessary for the recordings of the screen. The mixed recording then offered few advantages for detailed analysis over the recordings from individual cameras. Even though there were problems in coordinating the analyses from the different tapes, in the remainder of the data collection activities the recordings were taken from two individual cameras. This also simplified the data collection activity. It may be worth noting that using even more cameras would have further exacerbated these problems.

8.2.2 Data Analysis

The choice of whether to collect mixed or multiple images involves other trade-offs. Not only can mixing recordings at source require a transformation (or a restriction) to the shape of the image being gathered, it involves a pre-specification of the organisation of the images to be collected, which may inhibit later analysis. Using multiple images, on the other hand, can inhibit the continuity of the analysis, requiring frequent shifts in focus between the recordings.²

The recordings were made on analogue tapes and analysed on conventional video players. In the course of the studies undertaken here technical improvements were made in the quality of the recordings.³ It

---

² It can be useful to 'time code' tapes before recording takes place, but as well being a time-consuming activity the coding of tapes does not assist with the difficulties in locating items recorded in a sequential manner, or with the need to switch between viewings of different tapes. (At the time of data collection the only accurate way of time-coding video-tapes was in a separate process before recording. 'Time-coding' is not the same as recording the time on the tape.)

³ The initial recordings were made on VHS tapes using a Panasonic M1 camera, in one case using a (Sony) high resolution camera, and the later materials were collected using the higher quality 8mm format (Sony TR45 and TR55 camcorders). Since the data have been collected, Hi8 devices are more widely available which offer even better resolution.
may be that the recent availability of digital video cameras could not only support the collection of materials of better quality, but could also assist with some of the problems outlined above. For example, digital cameras could provide materials which, though recording at a broader focus, could also allow for the later analysis of details. This could be achieved through the use of software packages which allow for the manipulation of moving images in digital form. Although packages like Avid VideoShop and Adobe Premiere were used within the course of this research, they were not utilised extensively for analysis, the quality, particularly of the audio, not being of suitable quality. However the availability of such tools does suggest some potential ways of analysing materials in the future; these packages, for example, allow a great degree of control of the video image. It may also be possible to devise ways of tying the playback of two video streams together, so that materials gathered from two cameras can be viewed in parallel, on a very large computer display or through multiple video monitors.

For supporting the analysis undertaken here several technical options were considered; for remotely controlling a video player from a computer and for recording the materials on write-only storage media. The former was too slow (the response time being greater than a tenth of a second) and the use of the latter did not fit with the ways video data are typically edited and analysed. For example, the analysis of materials usually proceeds iteratively, an initial collection of fragments either being refined or leading to another sweep of the corpus. It may be thus not worthwhile to make a non-erasable copy of fragments which are no longer relevant for analysis, particularly if use of the format is expensive or time-consuming. Copying all the materials onto the medium may make use of the corpus of recordings more manageable, but the scale of the exercise may make this infeasible.

4 Trademarks of Avid Technology Inc. and Adobe Systems Inc. respectively.
5 It may appear strange that the audio quality is poor, but the focus of the technologies appears to be on presenting and playing back moving images. Hence, there are frequently 'breaks' in the playback of sound whilst the images are being presented on screen. Even the loudspeakers provided on, and available for, conventional computer systems appear to be of poor quality.
6 Since the initial trials of this technology took place, CD-ROMs have become cheaper and CD-ROM recorders more widely available. If the practical and financial issues can be mitigated, independent of the medium of storage, digital collections of audio-visual materials would greatly assist with identifying and
As outlined in the appendix the analysis of the details of dynamic on-screen materials can be problematic. Partly this is due to the nature of the materials collected and the issues that have been discussed regarding data collection. However, by adopting some innovatory techniques relevant materials for an analysis can be revealed. Video can provide a valuable resource for gathering ethnographic background to a setting. For example, video materials can provide resources for considering the natural history of events under scrutiny; tracking particular activities, for example. Moreover, they provide details of activities which would be difficult to gather through other means, for example in the materials concerning the moment-to-moment running of a transportation service video gave simultaneous access to the location of the trains and their movements, the happenings on the phones and the use of artefacts in the control room. In the work reported here, for the analysis of the materials the resources that have been utilised are those which appear to be of relevance to the participants through their production of talk and visual conduct, and by the displays of the analysis of that conduct by others. However, the concern for exploring visual conduct of participants requires that the strict demands for analysing talk-in-interaction have to be relaxed (cf. Schegloff, 1992a). Nevertheless, in the analyses reported a continuing concern has been not to introduce objects into the analysis for which only a weak warrant can be asserted for their relevance. Use of video materials may provide for ways of not only exploring the requirements for a systematic warranting of fieldwork observations (Schegloff, 1992a), but also for investigating the demands of the 'unique adequacy' requirement (see Chapter 7 and Garfinkel and Wieder, 1992; Lynch, 1993).

In the collection of materials for analysis, other kinds of data were utilised. These materials included various documents including the manuals for the computer applications, and the software applications themselves. They also involved informal interviews with relevant participants in the setting and others, including managers and personnel, locating fragments within a corpus, particularly when the analyst needs to return to the original corpus.

For example, pointing and gestures have been seen in terms of their iconic or emblematic qualities (cf. Morris et al. 1979). Providing a stronger warrant for such assertions would require access to how co-participants displayed an understanding of the gesture, for example, in their next actions. In other words, the analysis is warranted with respect to the interactional domain of the activity.
situated in other locations. In one case, within the study outlined in Chapter 3, a verbal protocol was also collected after the exercise from one of the users examining the video-recordings just gathered of his own behaviour. This provided the user with an opportunity to produce a post hoc analysis of the materials available on the screen - his own conduct. As, in the original recordings, the user also 'volunteered' a verbal account of some of his own activities, his post hoc analysis appeared to be more an analysis of this talk. Except for being a particular exotic resource for analysis, it was hard to see how to make use of this material.

8.2.3 Data Presentation

In the foregoing chapters attempts have been made to present details of instances taken from the video-recordings of the settings. These have utilised transcription notations developed for the analysis of talk by Jefferson (1972), for the analysis of visual conduct (Goodwin, 1981; Heath, 1986) and those specifically designed for the materials under investigation (for screen-based activities, in Chapters 3 and 4, and for command entry, in Chapter 6). Furthermore, in the course of the analysis various notations and shorthand were used to provide some sense of the data. These included using diagrams for mapping out the movements across the menus or through the screen displays, for Chapter 3 and 4, various schema for unpicking the organisation of talk on the radio for Chapter 5, and maps of the line and diagrams of keystrokes for Chapter 6. For presenting the analyses, sketches of displays and 'frame-grabs' from the recordings have been utilised to give a flavour of the conduct in the materials. The presentation of these materials has been developed from earlier work reporting the analysis of audio-visual materials (cf. Goodwin and Goodwin, 1992; Goodwin and Goodwin, 1996; Heath and Luff, 1996a). Although the constraints of presenting materials in a static form means that the nature of the visual conduct, and also the talk, is lost, an attempt has been made at giving a sense of the significant and delicate, aspects of the behaviour revealed in the materials. This has required some editing of the materials and the presentation of series of multiple images.

---

8 Clips were 'grabbed' from the video-tape onto an Apple Macintosh using either Avid VideoShop or Adobe Premiere. Then, particular still images were 'copied' into Adobe Photoshop where various transformations were made relating to the resolution, contrast, edges, lighting variations and shape of the image. The final image could then be copied to a conventional word-processing package (in this case Microsoft Word).
Fragments taken from the data have been digitised and then individual images have been transformed so that they are clearer when they are presented on paper.9 No systematic way has yet emerged for presenting such 'frame-grabs' alongside an analysis. For this thesis, a case-by-case consideration was given to whether such presentation was necessary, and how that presentation could be organised. Typically, when considering visual conduct, mapped horizontally along the page, the images proceeded from left to right and when alongside transcripts of talk the images were laid out down the page. A series of images were usually required to give a sense of the movements outlined in the analysis. However, exceptions were necessary. The requirements of particular analyses suggested other layouts, for example, when a further single image was required, or when diagrams of screen displays were desirable. It was also necessary, at times, to provide images from another camera angle, additional close-up details of a particular area of the image, or just the close-up details themselves.

From these initial attempts and earlier work on similar materials (Heath and Luff, 1996a), it seems that a more systematic investigation of ways of presenting such materials is required, particularly how delicate features of conduct can be presented alongside a more general view of the local environment in which that conduct takes place. It is also apparent that there are particular requirements for the skills necessary to use the image manipulation package and limitations to the technological resources available. Further study may be useful for considering the requirements of technologies to support the presentation of video materials. It may also be interesting to investigate the skills of others who have to deal, in their everyday work, with similar materials, for example, individuals involved in video, photographic images and high quality document production. Such an analysis may suggest resources that could be utilised for presenting not only static images, but also more animated materials.

9 Typical transformations involve sharpening the images, changing the contrast, brightness and lightening variations and close-ups of particular areas of the image. Adobe Photoshop (Adobe Systems Inc.) provides the resources for making these transformations.
8.3 Future Work

As mentioned, the concerns relating to the collection, analysis and presentation of materials raise in themselves future topics of study. The studies in this thesis, and others utilising audio-visual materials in concert with ethnographic materials, suggest some other requirements for tools, techniques and methods for video collection, analysis and presentation. Not only could recent developments in multimedia presentation software and hardware devices provide some useful resources, but more specific systems could be developed to support the activities that surround the use of audio-visual materials. For example, there are needs for more systematic ways of tying fieldnotes to the recordings when collecting audio-visual data (cf. Lamming, 1991), for accessing a large corpus of materials, indexing particular fragments, relating the fragments to the original sources, analysing the details of fragments and presenting the materials. Although some of these requirements are specific to the analytic orientation adopted in this thesis, it may be that more generic capabilities may be relevant for qualitative analysis in the social sciences, for requirements analysis (cf. Hughes, et al., 1995) or for the presentation of animated materials, in general.

Such tools may be relevant when considering the ways that workplace studies themselves may be deployed within the design process (cf. Anderson, forthcoming). From such studies it may be possible to derive methods, guidelines of heuristics for the practical design-related activities of developers. There is an understandable reluctance on behalf of social scientists to suggest such resources. Indeed, it may be that more practical experience is required before suggestions can be made. Several ethnographic studies have been utilised for the purposes of developing prototypes (Bentley, et al., 1991; Heath, et al., 1995b; Luff, 1994), but it would be fruitful to investigate, in a more systematic and extensive fashion, the warrants behind the observations made and the links to particular designs. It may be that audio-visual materials would be a useful resource for this. Similarly, with respect to the deployment of new

---

10 See, in particular the discussion in Chapter 7 in relation to Button and Sharrock's (1994, 1996) analysis of software engineering. However, even straightforward guides to using video recording equipment for analysis and background to ways for examining such materials are unavailable for designers and others interested in undertaking such research (cf. Goodwin, 1992a; Heath and Luff, 1992c; Jirotka, et al., 1995).

338
technologies, a particular case study could be examined where use was made of audio-visual materials before and after a technology was deployed (cf. Greatbatch, et al., 1992). Areas to investigate could include: the extent to which the details revealed in the analysis are required for design activities, where generic or specific case materials are useful, and how the findings of the study could be presented for different audiences.

In all the domains considered it is inevitable that the materials provide resources for other analyses to be undertaken, or the analyses to be extended. In the data on screen-based activities other foci on different kinds of screen-based objects and devices could be explored, including graphic objects and the ways particular items are manipulated. In the architectural practice the accomplishment of the activities with respect to other organisational constraints, such as the demands of others outside the practice, and the structural and aesthetic demands of constructing a 'working' design could be considered (cf. Cuff, 1992). The materials of the Docklands Light Railway Control Room could provide for a more detailed analysis of the coordination of activities on the phone system with other activities, an examination of the collaborative activities between Control Room Supervisors and others in the control room, and more detailed consideration of the interrelationships between happenings on the line and the screen-based activities of the participants.

However, it may be more fruitful to gather materials in a domain where the screen-based data provide for a development in the forms of analysis from those carried out here. In the analyses in this thesis, due to the nature of the domains, the on-screen materials have been principally concerned with graphical objects. Not only have the textual items presented on the screen often not been available for analysis, but their relevance has not been apparent through the conduct of the participants. It may be that analysis of a setting where text is a critical resource for the ongoing activity may provide a particularly fruitful domain for the analysis of screen-based activities.11 Environments where participants have to read and manipulate on-screen texts within a collaborative setting may be particularly rich in providing materials where not only the dynamic affordances of the computer system, but also the nature of the objects on the screen, offer resources for analysis. There may be other domains where rich naturally-occurring resources, including fieldnotes,

---

11 See the recent analysis of activities in news rooms (Heath, et al., 1995a). However, even here the collection of on-screen materials has been problematic.
documentary materials and audio-visual recordings relating to the use of computers could be gathered. These could provide for the possibility of extending further an analytic orientation which subjects the details of activities to close scrutiny whilst also aiming to reveal the particular unique quality of activities in the settings under consideration.

8.4 SUMMARY

In undertaking the studies outlined in this thesis various directions of research have had to be neglected. Even with the materials available there are other domains of interest which could be explored. Moreover, the studies also suggest other settings where the analysis of the in situ production of human-computer interactions could be extended.

The foregoing remarks have also outlined some of the difficulties in gathering audio-visual materials in natural settings, subjecting these to analysis and then presenting the outcomes. Although it appears that particular technologies could support such activities, either those which are available or ones which could be specified, it is apparent that such innovations would have to proceed alongside the development of the analytic approach. The studies in this thesis, with some others, are preliminary attempts at examining the nature of artefact use in naturally occurring settings. The exact trajectory of this programme of research is, as yet, unclear. It would be useful if technological and methodological support for such activities is sensitive to the ways in which these are accomplished at present, and to the demands of the users. It would be unfortunate if any technological and methodological initiatives in this domain suffered the fate of many of the technologies and procedural recommendations examined by workplace studies. Hence, the requirements have to be revealed by consideration of the activities of the practitioners of workplace studies. In other words, what may be required is a systematic analysis of the practices of those engaged in video collection, video analysis and presentation, the resources they utilise and how they manipulate the artefacts around them. Although an examination of the conduct of social scientists by social scientists may seem rather circular, particularly as it appears to require analysis of the particular kind undertaken by the individuals to be studied, it may be an essential pre-requisite for enhancing the support for workplace studies. There are other resources that can be utilised, however. In following
Harvey Sacks' and Garfinkel's programme it may be useful to explore domains where members face similar issues as part of their everyday concerns (Garfinkel, 1967; Sacks, 1992). As mentioned in this chapter, a useful exercise may then be to consider those other domains where the collection, analysis and presentation of audio-visual materials is a matter of everyday concern.

In summary, detailed studies of the uses of technology also appear to suggest some requirements for novel ways of supporting such studies. The analytic approach adopted in this research, the ways it has been undertaken, the resources used and how the analyses have been presented have revealed some possibilities for these innovations. These innovations may be technological or methodological. The remarks in this brief chapter not only have suggested further activities which could be undertaken with respect to extending the study of technologies-in-use, but also some ways in which these studies could be supported.
Appendix A

Methodological Background

analytic concerns in the examination
of naturally-occurring screen-based activities

A.1 INTRODUCTION

This appendix is intended as a background resource to Chapters 5 and 6, particularly with respect to the analysis of the uses of the Automatic Train Scheduler System (ATS) in the latter chapter. The complexity of the materials in the analysis of the ATS requires that some simplification is required in the accounts of the activities in the control room. It may be that also some further background is needed for making sense of particular occurrences in the fragments. For example, the analysis of the conduct of the controllers relies on an understanding of the geography of the line, the state of the service at particular times, and the activities they are undertaking on the ATS. From the audio-visual materials some of these resources are not readily available, particularly the details of the appearance of the typing on the screen. The controllers attend to the particular location of trains and the conduct of their colleagues, so it is
important for the analysis to have some understanding of these details. It would be impossible, given the continuous flow of activities within the control room, to capture this wealth of detail in field notes. Although some notes were taken of the general incidents occurring on the lines, such as bomb scares, alarms, passenger disputes etc., further work was necessary with respect to the collected audio-visual materials so that further details in the materials could be recovered.

Section A.3 outlines a way of tracking the keystrokes of the controllers. Although it is impossible to see from the video-tapes the characters of the commands being entered, it is possible to see the hand movements and the rough location of the keys which are pressed. By mapping out the various possibilities of station identifiers (17 in all) on the keyboard, it is possible, for commands to stations, to propose candidate stations to which the commands apply. Details of this analysis are given in Section A.3.1.

It is possible to see some of the happenings on the screen, particularly train movements on some of the displays. By coupling this with a hearing of the phone conversations one can provide a provisional account of relevant train movements at the time of the call. The types of resources relied on in this analysis are outlined in Section A.3.2.

Together these two types of analysis are utilised as background to the study of the collaborative uses of the ATS in Chapter 6. Due to the quality of the materials these analyses tend to be provisional and tentative. However, by exploring alternative accounts of train movements and happenings on the line a clearer picture of the state of the service at particular times can emerge. A flavour of the background analyses to the train movements is given by examining one of the fragments considered in Chapter 6 (Section A.4).

To follow the analyses of conduct it is useful to provide some further broader background to the Docklands Light Railway to that given in Chapters 5 and 6. Section A.5 gives an abbreviated glossary of terms used in the fragments. The section that follows (Section A.2) gives a brief overview of the line, its geography and the way it is presented on the ATS screens, including a summary of the abbreviations to station identifiers used in the commands and on the screens.
A.2 THE DOCKLANDS LIGHT RAILWAY

The Docklands Light Railway was opened in August 1987. It runs in the Docklands area of the East of London connected to the London Underground network at Tower Gateway, and later Bank, stations in the west, and to the British Rail network at Stratford in the North. One particularly distinctive feature of the DLR is the trains. Unlike the London Underground or the national rail network, the trains are not driven or 'operated'. Instead, they are automatically controlled by a computerised system that, when a train is docked at a station, supplies it with the distance (in terms of the number of wheel turns) it needs to travel to the next station. However, although trains do not need continual operation, a Train Captain (TC) rides on every train who is responsible for checking tickets, dealing with 'revenue disputes', assisting passengers, and the like. The Train Captain also drives the train when the system fails. The Light Railway consists of two lines: the 'Red Line' running from Island Gardens in the south to Stratford in the north; and the 'Green Line' running from Crossharbour to Tower Gateway (and later to the Bank) in the west (see figure A.1).
The automatic operation of the trains makes it possible for a computer system, called the Automatic Train Scheduler (ATS), to control the running times and running order of the trains. Furthermore, it also allows for a system to generate automatic passenger information concerning the destination of a particular train and its estimated arrival time. Therefore, the straightforward nature of the line, and the automatic operation of the trains, scheduling and generation of passenger information, would appear to suggest that few control room staff are required. Indeed, it would seem that this was one of the original aims of the managers, systems designers or implementors, who conceived of the line being controlled by just one individual. This controller would be responsible for monitoring the power supply, operating the service, reacting to alarms, making additional announcements and dealing contingencies, crises and emergencies arising from TCs and passengers.

However, at the time of data collection (May 1991), during the day there were six staff permanently located in the control room at Poplar: two Control Room Supervisors (CRSs) responsible for the power supply, taking
and making phone calls to TCs and scheduling the service; two Control Room Assistants (CRAs) responsible for passenger information, the TCs' rosters and the provision of radio telephones; an Electrical Works Coordinator (EWC) monitoring technical problems on the trains; and the Control Room Manager (CRM) responsible for the 'strategic' operation of the service. Chapters 5 and 6 outline some of the ways in which the collaboration between these staff is accomplished. Chapter 5 focuses on the activities in the phone conversation between the principal Control Room Supervisor and the Train Captains. Chapter 6 explores the interrelationships between the activities of the principal and secondary Control Room Supervisors and the remote parties in the phone call. For the latter, an analysis of the activities on the ATS system is undertaken.

The ATS displays the line in a compressed form so that it fits on a computer screen more easily (see, for example, the sketch of the diagram of the main line, figure A.2).

![ATS Display Diagram](image)

**Figure A.2** A sketch of the ATS display of the main line. The arrows indicate direction of travel.

On the ATS screen the stations are identified by abbreviations. These are given in Table A.1.

<table>
<thead>
<tr>
<th>Station</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Saints</td>
<td>ALS</td>
</tr>
<tr>
<td>Bank</td>
<td>BAN</td>
</tr>
<tr>
<td>Bow Church</td>
<td>BOC</td>
</tr>
<tr>
<td>Canary Wharf</td>
<td>CAW</td>
</tr>
<tr>
<td>Crossharbour</td>
<td>CRO</td>
</tr>
<tr>
<td>Devons Road</td>
<td>DER</td>
</tr>
<tr>
<td>Heron Quays</td>
<td>HEQ</td>
</tr>
<tr>
<td>Island Gardens</td>
<td>ISG</td>
</tr>
<tr>
<td>Limehouse</td>
<td>LIM</td>
</tr>
<tr>
<td>Mudchute</td>
<td>MUD</td>
</tr>
<tr>
<td>Poplar</td>
<td>POP</td>
</tr>
<tr>
<td>Shadwell</td>
<td>SHA</td>
</tr>
<tr>
<td>South Quay</td>
<td>SOQ</td>
</tr>
<tr>
<td>Stratford</td>
<td>STR</td>
</tr>
<tr>
<td>Tower Gateway</td>
<td>TOG</td>
</tr>
<tr>
<td>Westferry</td>
<td>WES</td>
</tr>
<tr>
<td>West India Quay</td>
<td>WIQ</td>
</tr>
</tbody>
</table>

**Table A.1** The abbreviations of stations on the DLR
A.3 PROBLEMS AND TENTATIVE SOLUTIONS IN THE ANALYSIS OF COMMAND ENTRIES INTO THE ATS

Even with two cameras it is impossible to gather adequate resources in the DLR Control Room to make sense of all activities associated with the talk and visual conduct of the relevant participants and the details of the activity on the computer systems. With one of the cameras focused on the console as a whole (including the CRAs and the CCTV monitors), the other was required to focus on the principal controllers and the technology which surrounded them, particularly the ATS system. This angle provides a view of the principal controller, the screen and the keyboard, but not the details on the screen.\(^1\) Train movements can be seen on the screen, either as the section of track changing colour, a station being highlighted as the train docks, or as moving train (or set) numbers. However, the commands which are typed can only be seen as small blocks as the typing emerges on the screen. The next section (Section A.3.1) suggests how it may still be possible to recover some of the details of what commands are being typed in by the controllers.

Section A.3.2 addresses a different problem. From viewing the screen at any moment it is difficult to get a sense of train movements. Even when a change is seen to happen it is often as a 'disappearance'. A train leaving a station, for example, is marked by the station changing back to its normal colour. So, to get a sense of the train movements a natural history of relevant events on the line has to be reconstructed. Section A.3.2 outlines some of the resources for this work.

A.3.1 Keystroke Tracking

At the time of recording there were 17 stations operating on the DLR, or more accurately 17 stations that commands could be sent to. The format of the commands means that it is possible to examine the hand movements of the controllers when commands to stations are being entered. Although neither number sequences nor commands to the shunting yard can be easily distinguished, most of the abbreviations for stations have distinctive contours. To assist with a provisional analysis of

---

\(^1\) The positioning of the cameras was also influenced by the demands of the setting. It was critical that the cameras did not impede the participants. A more detailed focus was tried, but this camera position could have blocked the controllers' access to the system controlling the power supply to the line.
these commands the contours are divided into six groups based on the
general direction of hand movement: left-left, right-right, right-left
(inside), right-left (outside) and close patterns. The close patterns group
largely overlaps with the others, but assists with distinguishing quick
typing at close proximity.

The work then consists of distinguishing from which group the
particular command belongs and which member it most resembles. In
most cases particular abbreviations can be identified by either the starting
or the ending point of the command (its row or column on the keyboard).
The distance of hand travel provides another resource for identifying
commands, so does whether the hands end at a point beyond which they
started (outside). Breaking apart the problem in this way leaves only a
few abbreviations which are difficult to distinguish, notably between the
commands to Stratford (STR) and Devons Road (DER) stations.

The following diagrams give a sense of the different hand movements
in each of the groups (Figure A.3).

**Figure A.3** The typing patterns of the commands

![Diagram of typing patterns](image-url)
A.3.2 Train Following

As train movements are displayed by discrete changes to the ATS display, these too can be difficult to distinguish in the materials. However, some sense of the movement of the trains can be arrived at from following the occurrences in the radio phone conversations. The geographical distances on the DLR are quite short so attending to the happenings in the recordings surrounding a particular fragment, at most half an hour either way, can reveal some relevant resources for making sense of the train movements visible on the screen. Typical resources are:

- incoming calls stating the present location or direction of any train, or the location where a future alteration is required;
- outgoing calls stating the location or direction of any train, or a location when an alteration is going to be made;
- special identifiers, such as 'mobile' or 'driver Ed';
- the talk between CRSs, CRAs and other staff in the control room concerning happenings on the track, such as driver training ('driver education');
- commands having to be made to a station prior to a train arriving at it, and after a previous train has departed;
- the 'lack' of any other command being made by the controllers in the time before a train arrives at a station, when an appropriate alteration is required;
- the typical time taken to travel between stations; and
- ...a similar set of resources happening following a call. These can be utilised to make speculations about prior train movements.

Thus, from the audio-visual materials some ethnographic background is recovered concerning the details of the happenings on the line. These happenings, however, only have a provisional nature.

A.4 Example Analysis

To give a flavour of the ways in which these analyses are brought together, the instance considered in fragment 8, Chapter 6 is considered.

Just after the remote party's reason for the call, Cii types a command to Herons Quay (a 'left - left' command):
The call continues and when Ci confirms the change, Cii types in two more commands - one to Canary Wharf (CAW) and one to Mudchute (MUD). These are shown on the line below the talk they co-occur with.

Furthermore, following the call, just after the phone is put down, Cii types another command into the system. This time for Crossharbour. Throughout the call, and for a long duration following it, Ci does not type any commands to the ATS. He goes on to take another call. It appears that one or more of the commands typed by Cii could be related to the incoming call.

The state of the service is, at the time of call, shown schematically in figure A.4 (adapted from Figure 6.6. in Chapter 6).
There are five trains running on the area of track in question (the branch between the Delta and Island Gardens, ISG).

On the southbound (going right on the ATS screen) one train is between the Delta and Heron Quays, another is between Heron Quays and South Quay and a third is coming into Crossharbour. On the northbound (going left) one train is between Crossharbour and South Quay and another docked at the terminus, Island Gardens.

Examination of the ATS1 display reveals that during the course of activities occurring around the initial request from the train captain, three trains dock at stations. These are revealed by the colour of the station changing to yellow on the screen. Two trains dock on the southbound, one at Crossharbour (line 10), and one at South Quay (line 12). One docks on the northbound (line 6). However, none of these appear to be the train which is being asked to 'standby'.

The nature of the subsequent instruction 'when you arrive at Crossharbour' (lines 16-17) appears to exclude the train at Crossharbour going south. As the other trains on the southbound section of the track are also beyond Heron Quays, the command to HEQ can only apply to the train coming into Canary Wharf or the train going northbound and entering South Quay. The train at Island Gardens is excluded as it would follow the northbound train through Herons Quay.

Sometime prior to the call, another caller on 'five two' has announced that her train is to be used for testing trainee train captains. Some nine minutes after, from the same call-sign ('five two'), a report is made that the train has arrived at Island Gardens. Just before this report a train is indeed seen on ATS2 to dock at Island Gardens. This train, now moving northbound appears to be the one in the call, the one for 'driver ed' (or driver education).

The subsequent commands, commencing on lines 17 and 19, appear to be directed to the train coming into Canary Wharf going south (possibly to provide a route through one of the three lines through the station) and the train just arrived at Crossharbour going south (possibly also to prepare for the train to enter Island Gardens). The command to Canary Wharf coming after a Herons Quay command would appear to suggest that the HEQ command is for a different train i.e. the train going north. There is also nothing to suggest in the calls that there are any problems going south between Herons Quay and South Quay. Therefore, the command to
Herons Quay appears to be made to the train moving northbound, the one in the call.

A command to the Heron Quays station could both change the route and mode of the next train going south. If the train in question is part of a driver testing exercise then part of this is likely to require changing the mode of the train (so that the operators can 'drive' the train).

Hence, it appears that the incoming call between the remote party and Ci, and the command by Cii are closely related. The talk relates to a train doing driver testing going north, having left Island Gardens. The command is made to a station in front of it (possibly to change it to the requested mode)

A.5 GLOSSARY

**ATM (Automatic Ticket Machine)** There are at least two automatic ticket machines at each DLR station. A control system for the ticket machines called the Automatic Ticket Machine Controller is located at the back of the control room and usually displays which machines are working at the present moment.

**ATO (Automatic Train Operation)** The normal operation of the train (i.e. not under driver control and with ATP set).

**ATP (Automated Train Protection)** A system that prevents the trains getting too close to each other, it is organised with respect to the movement of the preceding train.

**ATS (Automatic Train Scheduler)** The computer system that automatically schedules and routes the trains.

**block** An action that stops two trains using the same part of the track. Blocks are shown in orange on the ATS display. The signal system used at the time of recording by the DLR sets signals centrally. Blocks can prevent trains running on a section of track, so that, for example, engineering work can be carried out. The ATP system also blocks a section of track following a train.

**BTP (British Transport Police)** The police who patrol on British transport systems.

**CRA (Control Room Assistant)** The control room staff responsible primarily for making announcements to passengers.
CRM (Control Room Manager) The manager of the control room who usually sits on a desk behind the CRSs and CRAs, responsible for strategic management with respect to such issues as the reliability of the service, the number of trains running and the types of trains running.

CRS (Control Room Supervisor) The individuals responsible for the general, moment-to-moment operation of line.

Delta The complex junction between Westferry, West India Quay and Poplar. So called because it is D-shaped.

Emergency Shunt (EMS) A mode of the train set by the train captain that gives the train captain control of the train. When in emergency shunt, ATP is not working and speed is restricted to 20 kph. The train can go in either direction.

EWC (Electrical Works Coordinator) EWCs are responsible for the supervision of technical problems with the trains.

F and D (Failures and Delays) A fault and delay log that has to be maintained by the CRS on a form. Any delay over 3 minutes should be logged.

mobile The 'third' CRS who is responsible for travelling around the line and dealing with problems, such as disputes and bomb scares, where they occur.

mushroom A red, mushroomed-shaped switch that cuts off the power to parts of the line. There is also one in the train for stopping in emergencies.

OMC (Operations and Maintenance Centre) The depot where trains are stored, tested and maintained. It is located between Poplar and All Saints stations.

routeboard A sign by the side of the track giving the number of the section of track. Used to give train captain's precise locations and directions. These numbers also appear on some of the detailed ATS displays.

set number The number of any service given to the train by the ATS.

TC (Train Captain) The individual responsible for the train. TCs can drive the trains when there is a fault. Their normal responsibilities are concerned with dealing with passengers and checking tickets.
Acknowledgments


367
Walk-Up-and-Use Interfaces", in Proceedings of CHI 90, pp. 235-242.

and Technology. 3: 201-240.

Routledge and Kegan Paul.

Maidenhead: McGraw-Hill.

Working Paper, Rank Xerox Cambridge EuroPARC.

Luff, P. (1994). The AA: Preliminary Notes on the Activities, Communication and
Organisation of the Work of the Operations Centre and Mobile Workers. WIT
Report.

Luff, P. and Heath, C. C. (1990). The first steps towards a transcription system for
analysing the social organisation of human-computer interaction. Working Paper,
Rank Xerox Cambridge EuroPARC.

human computer interaction in an architectural practice, in Technology in

screen based documentation in collaborative activity', in Proceedings of CSCW '92, pp. 163-170.

the naturalistic analysis of human conduct and requirements capture, in
Requirements Engineering: Social and Technical Issues, Jirotka, M. and Goguen,

ethnomethodological phenomena, in Ethnomethodology And The Human Sciences,

Social Studies of Science. Cambridge: Cambridge University Press.

Management at the Iran-Contra Hearings. Durham and London: Duke University
Press.

Requirements Engineering: Social and Technical Issues, Jirotka, M. and Goguen,

behind the artifact', in Proceedings of CHI '89, 30th April - 4th May, pp. 247-252.


Mittermeier, R. T., Rousopoulos, N., Yeh, R. T. and Ng, P. A. (1990). An Integrated Approach to Requirements Analysis, in Modern Software Engineering:


Nickerson, R. S. (1977). On Conversational Interaction with Computers, in User-
ACM.

NJ: Lawrence Erlbaum.

W. H. Freeman and Co.

Associates.

Erlbaum Associates.


University Press.

Norman, D. A. (1993a). Distributed Cognition, Chapter 6, in Things That Make Us
Smart, pp. 139-154. Reading, MA: Addison-Wesley.


Behaviour and Information Technology. 8: (2), 125-134.

Conversation Analysis, in Computers and Conversation, Luff, P., Gilbert, G. N.

Manuscript, University of Michigan.


GOMS, Human-Computer Interaction. 5: 221-265.

Changes the Character of a Design Meeting', in Proceedings of CSCW '92, pp. 91-
98.

Ambulance Service. South West Thames Regional Health Authority.


