

The Contextual Evaluation Framework:

A prototype evaluation technique for e-Research



Grace Eden
Linacre College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2010

Acknowledgements

Thanks and appreciation to the participants in each of the case studies which include members of the UK-NEES, the NCeSS, and the BVREH projects. Also the e-Research community in general and the organisers of the All Hands Meetings in particular for providing a channel of communication where social and computer scientists are given the opportunity to exchange ideas. I'd also like to thank members of the OII, and the OeSS project, Paul Luff for the data sessions, and Marina Jirotko for introducing me to a fascinating and important field of research. Finally, love and thanks to family and friends.

Abstract

The contextual evaluation framework (CEF) is a requirements engineering technique that incorporates a particular sociological orientation, Ethnomethodology, in the development of a rigorous and systematic approach for requirements elicitation. This qualitative approach examines how well a system may be aligned with the endogenous organisation of work within a community of practice. Assessing how well a system supports the knowledge, skills and practices that already exist within a community is equally as important as developing solutions that will eventually reconfigure those practices, create new ones and extend modes of collaboration. The aim of this thesis is to address the absence of a systematic approach to quasi-naturalistic prototype evaluation which may be useful to a broader community such as requirements engineers, computer scientists and others not familiar with the details of sociological approaches. Such an aim is in line with the ways in which prototype evaluation approaches, particularly in HCI, have successfully been disseminated throughout the computer science research community - with the provision of guidelines. Likewise, the CEF is conceived of to be implemented in a similar manner.

Its focus is on the analysis of a prototype's relevance as a tool that is in some manner familiar to those who might use it. Specifically, professionals within a discipline share complex skills and knowledge where they learn to use similar tools, instruments and processes necessary for their work. Implicit in these social practices, practitioners gradually acquire the knowledge and skills necessary to become full members of a community of practice. In this way, the processes, objects and artefacts of practice come to possess specific meaning and significance. The CEF examines how this complex architecture of meaning is supported, constrained or transformed when using a prototype and makes possible an assessment of the ways in which participants interpret its usefulness and usability.

Table of Contents

Chapter 1: Introduction	2
1.1 Research focus.....	2
1.2 Outline of thesis chapters	4
Chapter 2: New computing infrastructures for global research collaboration ...	8
2.1 Introduction	8
2.2 Tracing the origins of e-Research	10
2.2.1 Supercomputing and big science.....	10
2.2.2 Metacomputing and the grid	12
2.2.3 A new science and engineering agenda.....	13
2.3 Engaging researchers in e-Research	16
2.3.1 Tensions between visions and practices.....	16
2.3.2 Participation in the software development process.....	18
2.4 Reconceptualising usability.....	21
Chapter 3: Understanding computer-supported work and interaction.....	25
3.1 Introduction	25
3.2 HCI: an information-processing perspective	27
3.3 CSCW: a social computing perspective	29
3.4 Prototype evaluation in HCI & CSCW.....	33
3.4.1 Pluralistic walkthrough	35
3.4.2 Think-aloud protocol	36
3.4.3 Cooperative prototyping.....	38
3.4.4 Situated evaluation.....	39
3.4.5 Quasi-naturalistic experiments.....	41
3.5 Summary.....	43
Chapter 4: Revealing social order in life and in the work setting	46
4.1 Introduction	46
4.2 Ethnomethodology	47
4.3 Conversation analysis.....	53
4.4 EM/CA/IA approaches to the description of observable order	56
4.4.1 Ethnomethodology and language games	57
4.4.2 Conversation analysis and the turn-taking machine	59
4.4.3 Interaction analysis and paralinguistic communication.....	60
4.5 Summary.....	62
Chapter 5: Developing the Contextual Evaluation Framework.....	64
5.1 Introduction	64

5.2	Ethnomethodologically-informed software design	65
5.2.1	Practical guides for ethnomethodologically-informed fieldwork	66
5.2.2	Practical guides for video-based studies of work.....	69
5.2.3	The need for a practical guide to quasi-naturalistic prototype evaluation	72
5.3	Recognizing a prototype's relevance to work practice	74
5.4	The contextual evaluation framework	78
Chapter 6: Seeing through digital images to an analysis of ancient texts.....		81
6.1	Introduction	81
6.2	The study of ancient documents and manuscripts.....	82
6.3	Fieldwork: working with projected digital images	84
6.3.1	Analysing the scribe's hand	85
6.3.2	Comparing letters within the geography of the manuscript.....	91
6.3.3	The role of re-inscription in the interpretation of ancient texts	94
6.4	Prototype evaluation: working with digital images in the VRE	97
6.4.1	Re-inscription activities at the interface	99
6.4.2	Tracing and gesture activities around the prototype	102
6.4.3	Referencing and juxtapositioning other resources.....	105
6.4.4	Transforming practices with the overhead image view	108
6.5	The use of the CEF in the BVREH project.....	110
Chapter 7: Coordinating local and distributed earthquake experiments		117
7.1	Introduction	117
7.2.1	Distributed hybrid testing	122
7.3	Fieldwork: local laboratory experiments	124
7.3.1	Observing an experiment through multiple views.....	129
7.3.2	Interpreting audible behaviours that suggest cracking	133
7.3.3	Ambulatory checks around the test rig	135
7.3.4	Initiating and terminating an experiment.....	136
7.4	Prototype evaluation: distributed earthquake experiments.....	141
7.4.1	Collaborative monitoring using local graphical displays	144
7.4.2	Coordinating the initiation of experiments	150
7.4.3	Terminating distributed experiments and saving data.....	158
7.5	The use of the CEF in the UK-NEES project	160
Chapter 8: Web portal for the social sciences		165
8.1	Introduction	165
8.2	Developing an e-infrastructure	166
8.3	Prototype evaluation: a web portal for collaborative work.....	168
8.3.1	Encountering system limitations	169
8.3.2	Getting to grips with system idiosyncrasies	177
8.3.3	Post-evaluation feedback.....	182

8.4 The use of the CEF in the NCeSS project	185
Chapter 9: A practical guide to the Contextual Evaluation Framework.....	191
9.1 Introduction	191
9.2 Get to know the domain: conducting initial fieldwork.....	194
9.3 The equipment: preparing for an evaluation session.....	195
9.4 The naturalistic setting: identify the participants and place	197
9.5 The evaluation session: identify its purpose and focus	199
9.6 Data analysis: identify fragments of work and interaction.....	202
9.7 Film elicitation: confirm preliminary analysis with participants.....	204
9.8 Assemble potential requirements: informing software design.....	206
9.9 Summary	207
Chapter 10: Discussion and further work.....	208
10.1 Introduction	208
10.2 Successes and Challenges	211
10.3 Conclusion	213
Appendix 1: UK e-Research user studies.....	217
Appendix 2: Transcription notation.....	220
Appendix 3: Video fragments	223
Appendix 4: List of publications	224
References.....	226

Chapter 1:

Introduction

The UK e-Science programme has invested over £250M in the development of systems designed to support large-scale academic research through the availability of shared computing infrastructures and software applications. These provide access to software, equipment and databases across multiple locations with the goal of facilitating distributed, multi-institutional, multidisciplinary information sharing and collaborative activities thus creating the potential for the emergence of new forms of scientific research. However, in order to realise this vision widespread adoption of e-Research systems is necessary. This requires that systems be designed for research communities are both useful and usable and that they do not cause major disruption or negatively impact upon routine work activities. To ensure these goals, I argue that software engineering projects must develop an understanding of the details of researchers' scientific work including the materials and resources they rely upon and the ways in which they collaborate and interact with colleagues. This is a non-trivial matter, primarily due to the complexity associated with identifying requirements in settings where technologies are being designed to support, transform and in some cases create new work practices. Add to this the widely acknowledged view that the high-level vision of an e-Research project may not always map directly to the needs of the people who will actual use the system in real work settings. Therefore, a critical aspect of iterative software engineering practice is to facilitate active engagement with the research communities who may use an e-Research system throughout the design process in order to identify requirements and to modify system prototypes accordingly.

1.1 Research focus

The aim of this thesis is to develop a requirements engineering technique that supports this process specifically through the evaluation of novel technologies developed, in the first

instance, as proof-of-concept prototypes. Because the goal of e-Research is to support distributed collaborative work, the aim of this technique is to facilitate access to the routine work practices and collaborative interactions of researchers in order to inform the design of systems that facilitate new forms of work and collaboration. As a result of this, my research is closely related to similar work within the field of Computer Supported Cooperative Work (CSCW) which utilises a particular method from the social sciences proven useful for design and development of collaborative systems. The technique that I have developed is socio-technical in its orientation in that it contributes to the development of new technologies by taking into account both the work setting and the details of work and collaboration from the perspective of the people who would use the system. By focusing on an assessment of the usefulness and usability of a prototype within its local and organisational context, critical factors for the successful uptake and use of new technologies may be identified.

This thesis also aims to address the limitations found within traditional requirements engineering techniques that rely upon gathering retrospective feedback from participants. This is usually achieved through interviews, surveys and observation in laboratory settings as participants conduct pre-assigned tasks. The drawbacks of such techniques are threefold. First, the feedback is retrospective and therefore may not identify the ways in which a system either supports or constrains the moment-to-moment, ongoing production of tacit, seen-but-unnoticed work practices. Second, feedback is taken from a number of individual users who are conceived of as being independent from each other. Because feedback is acquired in this way, it may not recover the ways in which collaborative activities are actually performed. Third, the prototype is evaluated within the confines of how a system operates as defined from a high-level requirements specification. This may constrain end-users to the evaluation of tasks which are pre-assigned by a software development team and aimed at validating functionality that has already been implemented. However, this may not validate its actual usefulness and usability, especially the ways in which the system may transform practices or create entirely new ones.

The contextual evaluation framework (CEF) is a requirements engineering technique designed to address these challenges and can be used in prototype evaluation as a way in

which to iteratively acquire end-user requirements. The CEF was developed through my participation in three e-Research case studies that spanned academic disciplines; from the sciences, the social sciences and the humanities. The technique addresses the limitations of traditional requirements engineering techniques by incorporating a particular sociological approach, ethnomethodology, in the development of a rigorous and systematic technique for requirements elicitation. Ethnomethodology is relevant because it provides concepts that can be used to gain access to social phenomena through detailed empirical studies of social interaction. Specifically, the ways in which members of a community cooperatively organise their work activities and the manner in which understanding is mutually achieved.

1.2 Outline of thesis chapters

Three case studies are presented in this thesis that use the Contextual Evaluation Framework in the assessment of e-Research software prototypes. The framework was developed through my participation in three e-Research case studies that spanned academic disciplines; from the sciences, the social sciences and the humanities. The thesis is organised in the following chapters:

Chapter 2: New computing infrastructures for global research collaboration. The e-Research initiative is comprised of multidisciplinary teams working together to develop and embed software applications and infrastructures into a variety of research settings. In order to successfully deliver computing infrastructures that support global research collaboration the requirements of researchers and their communities of practice need to be understood from their own perspectives to ensure that these systems support and enhance both their work practices and collaborative research activities. This chapter will trace the origins of the e-Research vision including its historical underpinnings. It will also discuss how research communities have been engaged in the design of e-Research applications. The chapter will conclude with a critique of the notion of usability and discuss the need for its reconceptualisation.

Chapter 3: Understanding computer-supported work and interaction. The purpose of this chapter is to develop an understanding of the background and role of both human-computer

interaction (HCI) and Computer Supported Collaborative Work (CSCW) in the design of computing systems. Each has a different theoretical and philosophical orientation that informs the ways in which it attends to the gathering of requirements through the evaluation of prototypes. This chapter will discuss research trends and approaches to end-user participation in the design of software applications through various prototype evaluation activities. What emerges out of the discussion is a critique of these approaches and the identification of the need for a systematic and rigorous approach to prototype evaluation that is sociologically-informed. The availability of such a technique may have consequences for the types of findings and insights that e-Research project members may draw upon to inform the development of applications that are meant to support collaborative research.

Chapter 4: Revealing social order in life and in the work setting. This chapter will build upon the previous two chapters through a discussion of three approaches that provide guidance to CSCW researchers interested in understanding the organisation of social order, human behaviour and practical action; specifically ethnomethodology (EM), conversation analysis (CA) and interaction analysis (IA). Each of these orientations has informed the development of the Contextual Evaluation Framework and so it is essential that the analytic position of each be discussed.

Chapter 5: Developing the Contextual Evaluation Framework. The purpose of this chapter is to tie the previous three chapters together by clarifying the need for an ethnomethodologically-informed approach to e-Research system prototype design and evaluation that is at once both rigorous and systematic. To achieve this, the practical guides currently available which focus on informing software design through ethnomethodologically-informed studies of work will be discussed. The identification of a gap in the literature will then be explained and as a consequence the need for a systematic and rigorous approach to quasi-naturalistic evaluation that may be used, in the first instance, within the e-Research software development community. After which, the notions of *relevance* and *familiarity* will be discussed as concepts that have informed the Contextual Evaluation Framework.

Chapter 6: Seeing through digital images to an analysis of ancient text. This chapter presents the use of CEF within the BVREH project case study and will discuss how it was used to identify the tacit practices of classicists as they worked with digital images of an ancient manuscript. Through the use of video observation, both the *in situ* work practices of researchers and their evaluation of the prototype VRE are framed within the CEF to assess the ways in which the e-Research system may support, constrain or transform the practices and activities associated with the deciphering of ancient manuscripts and texts.

Chapter 7: Coordinating local and distributed earthquake experiments. This chapter examines the use of CEF to evaluate the design of scientific software used to conduct earthquake experiments. In this case, the CEF is used to reveal that making observations in an experimental laboratory involves a variety of activities from reading graphs on a computer display to maintaining awareness of the ambient properties of the setting. The CEF is then used within a distributed prototype evaluation to reveal how new procedures are supported through the e-Research system. Specifically, the mechanisms that enable remote researchers to coordinate multiple activities across geographically distributed locations.

Chapter 8: Web portal for the social sciences. This chapter explores the use of CEF as a way to evaluate an application's interface for usability with no prior observations of *in situ* work practices. In this way, the evaluation is concerned with the use of CEF to observe the use of specific features of the interface to determine the prototype's 'usability', thus, issues of 'usefulness' were not addressed. The NCeSS case study was conducted in this manner partly because of the rather broad remit of the e-Infrastructure project, which was to deliver a 'generic' web portal for a broad spectrum of social scientists and partly to evaluate the usefulness of the CEF as a stand-alone evaluation framework. This is important to consider because many e-Research projects share similar characteristics to other academic and commercial software development projects which are characterised by a lack of time and/or resources necessary to conduct an ethnographic study of current work practices *in situ*. Even in this limited way, the CEF provided access to the sense-making procedures of participants as they arrive at common-sense understandings of features available within the web portal.

Chapter 9: A practical guide to the Contextual Evaluation Framework. This chapter systematizes the components of the CEF based upon the experience and knowledge acquired through conducting each of the case studies. The aim of the chapter is to provide a practical guide with which interested members of the e-Research community may incorporate the CEF prototype evaluation technique within their own projects.

Chapter 10: Discussion and further work. The final chapter of this thesis will discuss the successes and challenges of using the CEF within e-Research projects. It will appraise the ways in which CEF may offer a robust technique for the evaluation of software prototypes and how it may be integrated into the software development process. The chapter ends with a discussion of further work planned as part of an EPSRC funded fellowship under the Phd Plus programme.

Chapter 2:

New computing infrastructures for global research collaboration

2.1 Introduction

Building distributed computing infrastructures within the e-Research programme has required the efforts of large numbers of researchers and considerable amounts of funding. The programme has been referred to within the community under a variety of names with each referencing the domain of research that the systems are designed to support such as e-Science, e-Social Science and e-Humanities. Whatever the domain, the following frequently cited quote encapsulates the purpose of the initiative:

e-Science is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it. – John Taylor¹

The umbrella term ‘e-Research’ has been used by many in the community to refer to all domains that might use these systems and will be adopted here to reference all such efforts. e-Research may be defined as a variety of computing infrastructures and applications that have been or are being designed for the purposes of embedding them into academic research settings for the purposes of supporting distributed, multidisciplinary research collaboration (c.f. Borgman, 2007). More specifically, the e-Research programme has promoted the notion that research efforts across the physical sciences, social sciences and the humanities may be better supported through the introduction of a particular type of technology known simply as ‘the Grid’ (Hey & Trefethen, 2003). Grid technology and its protocols enable the sharing of computational power and data storage capacities by connecting the resources of many high

¹ Dr. John Taylor was the Director General of Research Councils Office of Science and Technology from 1998-2003 and participated in the commencement of the UK e-Science programme.

performance distributed computers. These linked resources make it possible for collections of individuals and institutions to collaborate through Grid infrastructures (Foster *et al*, 2001). The e-Research programme has framed Grid technologies as the solution to distributed resource sharing and long distance collaboration through the provision of computational infrastructures making accessible remote instruments, computational power, data storage and expertise from across the globe.

In addition to providing access to computational resources through Grid infrastructures, the e-Research programme promotes a reorientation across research communities towards a model of collaboration more akin to large-scale research or ‘big science’ projects (Finholt, 2003; Welsh *et al*, 2006). These projects are made up of large numbers of researchers and require the expertise of multidisciplinary teams which include engineers and domain scientists working together to address emerging scientific challenges. It is thought that by emulating the big science model, e-Research systems could facilitate and increase the collective information sharing and collaborative potential of globally distributed research communities. However, in order to realise this vision widespread adoption within research communities is necessary. The condition for *widespread adoption* implicitly requires that e-Research systems be designed in a manner that does not cause disruption, negatively impact upon the everyday, routine practices of researchers or impinge upon their communities of practice.

Challenges to embedding e-Research systems have been discussed widely and will be presented in this chapter. With large budgets and tight time constraints e-Research project members have recognised that techniques for identifying the requirements of researchers is essential for the successful uptake and use of the systems that are developed. Hence, the motivation for this thesis; to design a requirements elicitation approach that can be utilised within the iterative software development process of e-Research projects. This chapter will begin by providing background to the e-Research programme first by introducing, in section 2.2, its historical underpinnings and its progression through various incarnations. Next, section 2.3 will explore the tensions between the e-Research vision and researchers current ways of working. The section will also discuss the ways in which end-users have participated in the

development of e-Research systems. Finally, section 2.4, will conclude with an analysis and critique of the notion of *usability*, which is commonly referred to within the software development community, in some cases, as an ambiguous and undefinable quality where 'box checking' for 'it' through 'usability testing' is often the case rather than understanding its complexity. In this last section an argument is made for its reconceptualisation.

2.2 Tracing the origins of e-Research

The e-Research programme has advocated the embedding of Grid technology as a support for a mode of conducting research which is collaborative, distributed and multidisciplinary. In this section, the evolving partnership between scientific research and high performance computing communities is discussed in order to understand how this alliance has contributed to the emergence of the e-Research vision. First, in section 2.2.1, the initiation of a partnership between scientific research and the supercomputing communities will be discussed. Next, section 2.2.2 will present the emergence of metacomputing and grid technologies as solutions to the increasing computational requirements of big science projects. Lastly, section 2.2.3 will introduce the current e-Research programme as a derivative of this partnership.

2.2.1 Supercomputing and big science

The term 'big science' (Weinberg, 1961; de Solla Price 1963) appeared in the early 1960s and was coined to describe research activities that require large facilities, large budgets and large research teams. These were high-priority government sponsored projects in the US that typically supported research in areas such as nuclear facilities, military projects, particle physics, astronomy and climate research (Galison, 1992).

For many the age of Big Science was ushered in by the Manhattan Project during World War II, when the building of the atomic bomb involved the mobilization of much of the U.S. community of physical scientists in an engineering project of unprecedented magnitude. By the early 1960s, with the advent of NASA and the national space program, the term 'Big Science' was firmly affixed as a label for projects that required large-scale organization, massive commitments of funds, and complex technological systems (Caphew & Rader, 1992: 3-4).

In order to support such computationally complex and data intensive research, national laboratories in the US acquired commercially developed supercomputing machines

(Woodward, 1996). These provided big science projects with high performance computing and mass data storage capacities required to investigate phenomena which are difficult to investigate experimentally. For example, investigating the circulation of the oceans, the atmosphere, tectonic plate motion and nuclear weapons testing (Buzbee & Sharp, 1985) require the use of complex numerical models and data processing techniques to test predictions and to analyse the vast amounts of data that these models produce. Known as *in silico* experiments, they may take hours or days to execute depending upon the efficiency of a model's algorithms and the computing power available with which to process it. As the complexity of models multiplied, along with the availability of larger datasets, the computational requirements of scientists increased.

As a response, in 1985 the US National Science Foundation (NSF) established five supercomputing centres (Freeman *et al*, 2005) each providing researchers the most technically advanced supercomputing available and they have been utilised successfully for over 20 years². However, consistent increased demand for high-performance computing and greater storage capacity prompted several design innovations in scientific computing research from vector computing in 1967 to multiple processors arranged in parallel architectures more recently (Buzbee & Sharp, 1985: 595). The most recent design innovation has been the development of distributed parallel architectures which has led not only to a new way of maximising the technical capabilities of supercomputers, but also to a new way of thinking about the future of networked computing more generally.

[P]ast efforts in supercomputing and high-performance networking are being subsumed into a broader, integrated vision of a more capable, ubiquitous, and accessible cyberinfrastructure (Freeman *et al*, 2005: 682).

Motivated by the introduction of parallel architectures and their ability to harness the computational resources of geographical distributed computers, technological visions emerged about the ways in which this innovation could be facilitated more widely. Initially referred to as 'metacomputing' it soon evolved to be known simply as 'the Grid'.

2 The centres now organise their resources through the US TeraGrid, <http://www.teragrid.org/about>.

2.2.2 Metacomputing and the grid

Metacomputing (Smarr & Catlett, 1992) was a key technological innovation in high-performance computing originally implemented as a series of projects that would link the five US supercomputing centres together. The goal was to harness geographical distributed supercomputers as a single resource forming “a *network of virtual supercomputers or metacomputers*” (Foster & Kesselman, 1997, italics in original). This linking of computational resources, which came to be known later as ‘the Grid’ (De Roure *et al*, 2003), introduced a new set of protocols for networked systems that made it possible for distributed computers to share both processing power and storage capacity thus increasing the resources available amongst them as if they were a single virtual machine. De Roure *et al* (2003: 90) explain the ways in which these protocols would transform access to networked information and tools using an analogy; up until now, the Web has enabled the delivery of information using a ‘publishing paradigm’ by providing a platform with which data may be shared in a distributed manner as images, documents and software tools. Grid protocols, by comparison, would harness the computational resources of networked machines creating an ‘interactional paradigm’ so that both instruments and facilities at remote locations could be accessed directly.

Originally implemented to support big science research through the linking of the five US supercomputing centres, visions to extend its implementation quickly emerged with a goal of eventually creating a Grid-enabled network on a global scale (Smarr & Catlett, 1992; De Roure *et al*, 2003), initially to support academic research. Visions for its implementation in both the near and longer term have been elaborated in a series of books (cf. Berman *et al*, 2003; Borgman, 2007; Dutton & Jeffreys, 2010), articles (cf. Foster *et al*, 2001; Finholt, 2003; Foster, 2008) and government reports (cf. Atkins *et al*, 2003; Berman & Brady, 2005; Atkins *et al*, 2010). These visions culminate to where “ultimately, one would hope that the Grid will be the operating system of the Internet” (Berman *et al*, 2003: 40). The vision for implementing Grid infrastructures first across academic research networks and eventually to subsume the Internet may be framed as a *mobilizing ideology* (cf. Kling & Iacono, 1994) employed to

establish support and to provide a rationale for the development of government programmes that would initiate technology projects committed to the design and embedding of Grid infrastructures into research settings, in the first instance.

2.2.3 A new science and engineering agenda

The relationship between big science and supercomputing described in the previous two sections, specifically innovations in distributed resource sharing, led to a consensus amongst the scientific computing research community that Grid technology should be implemented on a massive scale so that its reach may extend beyond big science projects to include all academic disciplines and research projects. Efforts to mobilise this vision were initially communicated in *Revolutionizing Science and Engineering Through Cyberinfrastructure* (Atkins *et al*, 2003) in the US and *The UK e-Science core programme and the Grid* (Hey & Trefethen, 2002) and *e-Science and its Implications* (Hey & Trefethen, 2003a) in the UK. Prior to their publication, the UK e-Science programme in parallel with the US Cyberinfrastructure (CI) initiative established a multidisciplinary science and engineering research programme that was tasked with funding the development of e-Science and Cyberinfrastructure projects (Hey & Trefethen, 2005; NSF report, 2007). In the UK, the programme's vision included success criteria for the 'core programme' tasked with developing solutions to complex technical challenges.

The Core Programme can assist the e-science projects in building their discipline-centric grids and learn how to interconnect and federate multiple grids in a controlled way. If we can build such grids to production quality, operating 24 hours a day, seven days a week, 52 weeks a year, and that do not require computing experts to use them, then e-science really does have the potential to change the way we do scientific research in all our universities and research institutes (Hey & Trefethen, 2003a: 1813).

This description of the Core Programme's aims places both ease of technology use and its relevance to the research communities for which they are designed as preconditions to the success of e-Research. More recently, government programmes have extended the vision to engage social scientists in the use of Grid technologies. In the UK, the National Centre for e-Social Science (NCeSS) was established in April 2004 and in the US, the social, behavioral and economic (SBE) sciences have presented their vision for embedding Grid technology into social science research settings. Implementing the vision of a Grid-enabled social sciences and

humanities is being coordinated in a similar fashion to the original e-Science programme; through the development of government sponsored programmes that coordinate both the funding and promotion of technology projects that develop Grid-based infrastructures with the aim of supporting research in those disciplines.

The next two sections explore how the e-Research vision has been communicated to academic communities, including its key argument that it will transform research outcomes in the physical sciences, social sciences and the humanities. First, section 2.2.3.1 will discuss the notion of the ‘data deluge’ which is a significant challenge for the physical sciences and second, section 2.2.3.2, will present the notion of ‘digital scholarship’ as a means for linking the outputs of the social sciences and humanities researchers.

2.2.3.1 Managing the data deluge

In the physical sciences, one argument used to frame the need for Grid technologies is “the imminent data deluge” (Hey & Trefethen, 2003b: §36.2) which holds that “data generated from sensors, satellites, high-performance computer simulations, high-throughput devices, scientific images and so on will soon dwarf all of the scientific data collected in the whole history of scientific exploration” (ibid: 812). Such vast amounts of data requires mass storage capacity, and Grid-based infrastructures could provide support for this deluge through linking any number of distributed machines that share both computing power and storage capacity.

Another argument for the potential benefit of deploying Grid technologies is its ability to link data stored in distributed heterogeneous datasets which, at the moment, are not being utilised to their full potential.

It has been argued that alongside the two traditional methodologies of science - theory and experiment - computational science has now emerged as a third methodology. With the advent of scientific data warehouses such as virtual observatories we may be seeing the emergence of a fourth methodology, that of collection-based science (Hey & Trefethen, 2003a: 1823).

Connecting heterogeneous datasets through Grid technology could provide researchers access to large amounts of varied data. Such access may facilitate new ways in which they might be analysed allowing for connections between datasets where none had been made before. These

arguments, along with the ability to increase the performance of simulations and models through linking high-performance computers, discussed in the previous section, frame Grid technologies as a solution for supporting data capture, manipulation, storage and sharing across geographically distributed locations.

2.2.3.2 Facilitating digital scholarship

In the social sciences and humanities the design and embedding of Grid-based technologies is relatively new. The argument and tagline for its implementation is: “enabling digital scholarship” (ACLS report, 2006). Digital scholarship has been described as the use of Grid-enabled infrastructure and software applications to facilitate both geographically distributed and co-located research (*ibid*: 7). More specifically, Berman & Brady (2005) envisage:

[T]he development of more realistic models of complex social phenomena, the production and analysis of larger datasets (such as surveys, censuses, textual corpora, videotapes, cognitive neuroimaging records, and administrative data) that more completely record human behavior, the integration and coordination of disparate datasets to enable deeper investigation, and the collection of better data through experiments and simulations on the Internet (p. 5).

Its main funding bodies in the UK state that it would like to see the e-Research vision implemented to enable the use of advanced technologies to support both social science and humanities research in diverse ways³. Grid infrastructure for the social sciences might include the design of applications that enable the use of data mining tools across a diverse range of large datasets or the development of collaboration tools which provide mechanisms for the sharing, annotation and analysis of both quantitative and qualitative data⁴. In the humanities it may be used to provide scholars with access to a variety of digital images of physical artefacts from across the globe thus making it possible to gather a variety of sources quickly reducing the need for extensive travel. The argument presented to the social science and humanities communities is that their research could benefit from the use of Grid technologies through the sharing of interview data, video, and digital images of artefacts. Even with these benefits, social science researchers have examined the challenges of mapping the computational

3 ESRC: <http://www.esrc.ac.uk/ESRCInfoCentre/research/resources/NCeSS.aspx> [accessed 03 October, 2010] and AHRC: <http://www.ahrcict.rdg.ac.uk/activities/e-science> [accessed 03 October, 2010]

4 See <http://www.ncess.ac.uk> for a list of past and <http://www.eresearchsouth.ac.uk/uk-e-social-science> for present software development projects.

requirements of the physical sciences (e.g. the need for high-performance computing and mass storage capacity) directly on to the requirements of the social sciences and humanities (Meyer *et al*, 2008; Carusi & Jirotko, 2009). Some of the challenges to understanding researchers' requirements will be discussed in the following section.

2.3 Engaging researchers in e-Research

This section will discuss the ways in which some research communities have responded to the e-Research vision, including, in section 2.3.1, tensions between its vision and researchers practices. Section 2.3.2 will follow on with a discussion of how researchers have been encouraged to engage in the software development process, by providing them with an opportunity to recalibrate a project's high-level vision so that it might reflect realities on the ground.

2.3.1 Tensions between visions and practices

Previously, in section, 2.2, the origins of the e-Research vision and some of the ongoing strategies to promote its embedding within academic research settings was discussed. However, tensions may develop between an initiative's visions and the realities of researchers' practices. The successful implementation of the e-Research vision may encounter challenges precisely because it includes scenarios of the future that may require changes in practice on the part of research communities. These proposed changes may have immediate implications for both research processes and cooperative work arrangements (Welsh *et al*, 2006). Even so, incompatibilities between high-level visions and practices as they exist on the ground may be reconciled by involving researchers directly in the design of e-Research applications.

As these incompatibilities have been reported, techniques for understanding the scholarly activities and disciplinary requirements of potential users of Grid infrastructure has been recognised as a major research challenge (Proctor *et al*, 2006; Jirotko *et al*, 2006a). Additionally, Meyer *et al* (2008) summarize challenges that include:

- institutional and legal challenges to data sharing

- issues with reward structures and assigning credit for contributing data to public archives
- taking into consideration the policies and practices of individual e-Research projects
- achieving standardization across heterogeneous data archives
- ensuring privacy when researchers are collecting data on human subjects
- conveying trust in both the expertise of people and the content of artefacts

This list of challenges provides an indication of how notions of the Grid's support for data storage and sharing may not be as straightforward as it might appear. Carusi & Jirotko (2009) also discuss a range of ethical challenges and requirements that emerge when considering the sharing of qualitative data in digital environments. Interestingly, while data sharing is central to the visions of Grid technology, evidence suggests that most scientists do not use other researchers' data, whether obtained from a data archive or from a colleague directly. For instance, in Dutton & Meyer's survey of attitudes towards e-Research and research practices (Dutton & Meyer 2008; Meyer & Dutton, 2008) they found that relatively few researchers, qualitative or quantitative, had either used archived data (22%) or data from colleagues (17%). Their sample also suggests a need to support bottom-up research innovation if Grid infrastructures and applications are to reflect the ways in which scientists conduct their research activities. Dutton & Meyer (2008) note their findings suggest that a gap remains between the bottom-up practices of ordinary social scientists and the top-down visions of e-Research advocates.

As a further example, in the US, Bos *et al* (2007) have identified three key barriers to sustaining participation in cyberinfrastructure projects:

First, scientific knowledge is difficult to aggregate. While information has become very easy to transmit and store over great distances, knowledge is still difficult to transfer (Szulanski, 1992) ... knowledge often requires specialized expertise, is difficult to represent, may be tacit, and changes rapidly.

Second, scientists work independently. Scientists generally enjoy a high degree of independence, both in their day-to-day work practices as well as in the larger directions of their work ... The culture of independence benefits science in many ways, but it also makes it more difficult to aggregate scientists' labors. Scientific collaborations must work harder than other organizations to maintain open communication channels, adopt common toolsets, and keep groups focused on common goals.

The third barrier is the difficulty of cross-institutional work. ... Universities often guard their intellectual property and funding in ways that hinder multi-site collaboration ... Very few universities operate across state lines, much less national boundaries.

Other challenges to sharing include lack of data standardisation and data provenance mechanisms (ibid). These barriers require further consideration in order for the embedding of e-Research systems to be successful. From these reports, it appears that the reconfiguration of researchers practices and organisational activities will have an impact on the successful uptake and use of e-Research applications. Additionally, the cost of not involving users in the design of systems more generally has been noted in the much cited Standish Report and Ewusi-Mensah (2003) which state that the key barrier to technology adoption may be mainly attributed to the lack of end-user participation in the design process. The next section will review some of the approaches employed by e-Research projects to involve researchers in the software development process.

2.3.2 Participation in the software development process

This section presents some of the approaches e-Research projects have taken to involve researchers in the software design process. First, in the Integrative Biology VRE project (IBVRE)⁵ project members carried out interviews, focus groups and ethnographic fieldwork studies in order to understand the ways in which both heart and cancer modellers collaborate with workbench experimentalists, mathematicians and each other to design simulations of biological systems. The aim of these exercises was to understand the scientific workflow and research processes of *in silico* experimentalists. They asked researchers to provide accounts of the ways in which the modellers communicate and collaborate with the experimentalists and mathematicians. They also observed ‘a day in the life’ in order to understand the ways in which routine activities are conducted when both designing models and running simulations. These may be considered ‘traditional’ techniques for requirements elicitation.

A more novel technique relies upon the use of analogy as an approach to understand work practices within a research community. Making Tea⁶ (schraefel & Dix, 2009) is a requirements

⁵ <http://www.vre.ox.ac.uk/ibvre/>

⁶ <http://mytea.org.uk/>

elicitation approach developed to identify the experimental practices of synthetic chemists. The approach relies on having the chemists conduct faux experiments in order for the requirement engineer to understand when, why and how lab notebooks are used. In this instance, requirements engineers asked chemists to “make tea” as if it were a chemical experiment. Both software engineers and requirements engineers observed and videotaped the Making Tea exercise to understand how paper laboratory books were used as the experiments were conducted. Although a novel and observational approach, asking chemists to perform a task such as ‘making tea’ as if it were an experiment may not provide an ecologically valid approach to understanding the types of activities that actually occur within this domain, thus skewing the requirements engineers’ understanding of the use of lab notebooks in real-world practice.

In a different project, agile methods are being assessed as a new approach for obtaining requirements in the development of scientific software (Pitt-Francis *et al*, 2008). The Cancer, Heart and Soft Tissue Environment (CHASTE) project is developing scientific software for computational biology and has fostered a close collaboration between heart and cancer modellers, who have expertise in the research domain, and software engineers, who have expertise in the optimisation of code. The project team has reported that an agile approach is far more useful than plan-driven software development methods.

The eDiaMoND project developed a system for distributed data sharing and collaboration within clinical medicine in field of breast screening. Project members engaged clinicians in the design process through ethnographic fieldwork which was conducted at each work site, prototype demonstrations which provided an opportunity for clinicians to provide feedback, scenario-based prototype evaluation where application developers would act as evaluators based upon their knowledge of researchers work practices and quasi-naturalistic experiments where clinicians were asked to conduct their routine, everyday work activities using the prototype (Warr *et al*, 2007). This project employed a variety of techniques at different stages of the design process in order to understand the requirements of clinicians who might use the system.

These are examples of just a few of the approaches taken to involve researchers in the design of e-Research applications⁷. In each of these projects researchers were involved in a variety of ways so that requirements for these new systems might be identified. There is no single ‘right way’ to elicit requirements, however it is important to identify and refine requirements throughout the design process including at the stage of prototype evaluation. Such efforts to involve potential users of a system in the software development process early and often may provide them with an opportunity to recalibrate a project’s high-level vision (Bergman *et al*, 2002), which in effect may lead to the system better supporting the realities of work practice as it occurs on the ground.

This thesis is focused specifically on developing an approach to requirements elicitation that involves potential users of e-Research systems in their design, through the evaluation of prototypes within naturalistic settings of work. Providing researchers with an opportunity to use an e-Research system and give feedback on both its usefulness (e.g. does it include functionality that supports their work) and usability (e.g. does it provide an interface that is easy to understand) presents an opportunity in which its features and functionality may be refined so that it more closely matches the work requirements of researchers. The notion of *usability* has been used as a criterion with which to gauge the appropriateness or fit-for-purpose of a system and to validate it against identified requirements. Providing researchers with an opportunity to evaluate e-Research prototypes may bridge some of the tensions and incompatibilities between high-level visions and the work practices of research communities. However, in order for the contributions of potential users to be meaningful within the context of software design practice the very concept of usability needs to be reconsidered. The next section will conclude this chapter with an analysis and critique of the traditional notion of usability and will present an argument for its reconceptualisation.

⁷ There have been also been a number of other studies conducted to understand the requirements of the potential users of e-Research systems. These have been undertaken to identify requirements researchers may have in areas such as security, infrastructure functionality, and the development of use cases in order to describe research processes. For more detail, see Appendix 1.

2.4 Reconceptualising usability

The notion of usability is typically conceptualised as corresponding to the ease-of-use of an application's interface. In the domain of Requirements Engineering, usability evaluation has been viewed as a distinct activity undertaken separately from the design and development process. It is categorised as a non-functional requirement and emergent property of the system which is identified through use. Due to its emergent characteristic, the properties of usability are traditionally thought to be discoverable only after all of a system's individual sub-systems have been integrated. Because of this, usability has been considered as part of the final assessment phase of a software development process, that is, once design has been completed and where it is referred to in some cases as 'user acceptance testing'. However, alternative views to this conception of usability have emerged within the software engineering community itself, most notably through Nielsen (1992; 1993). Even so, this orientation to usability has been inherited from early approaches in Human-Computer Interaction (HCI) that focus on developing guidelines for interface design in an attempt to provide predictive guidance and measurements that might improve the usability of a system (cf. Shneiderman, 1998).

The *International Organization of Standards - ISO standard 9241* defines usability as "the extent to which specified users are able to achieve specified goals with effectiveness, efficiency and satisfaction within a specified context of use". More specifically, Nielsen (1993: 26) identifies five attributes of usability:

Learnability – it should be easy to learn so that the user can quickly start getting work done using the system.

Efficiency – once the user has learned the system a high level of productivity will be possible.

Memorability – it should be easy to remember, so that the casual user is able to return after a period of not having used it without having to relearn how to use it.

Errors – it should have a low error rate so that users make few errors and if they do make errors they can recover easily. Also, catastrophic errors must not occur.

Satisfaction – it should be pleasant to use so that users are subjectively satisfied. When using it, they like it.

Both HCI researchers and practitioners incorporate these attributes as criteria used to evaluate the usability of a system. Each of these attributes focus on the evaluation of a system's interface and are quantitatively measured by counting the number of *errors* made, the *time* it takes to complete tasks, *success* rates to task completion, number of *features* remembered, and *Likert* scaling of satisfaction levels (ibid: 194).

This conceptualisation of usability as conceived within HCI has been challenged by researchers in Computer Supported Cooperative Work (CSCW) and Participatory Design (PD) where the notion has been expanded beyond the evaluation of the ways in which individuals interact with an interface to include social interactions between groups of people who engage in the accomplishment of work through cooperative activities (Grudin, 1990). This reorientation towards the ways in which systems support group work has led to a reconceptualisation of the notion of usability which has prompted the evaluation of systems in their support for social and organisational conduct (cf. Bannon, 1991; Heath & Luff, 2000). This shift is based on the idea that “the location of the ‘user interface’ has been pushed farther and farther out from the computer itself, deeper into the user and the work environment” (Grudin, 1990: 261). In response to ‘the computer reaching out’, CSCW and PD researchers have focused their unit of analysis on understanding how systems support *cooperative work*.

CSCW should be conceived of as an endeavor to understand the nature and requirements of cooperative work with the objective of designing computer-based technologies for cooperative work arrangements. The fact that multiple individuals, situated in different work settings and situations, with different responsibilities, perspectives and propensities, interact and are mutually dependent in the conduct of their work has important implications for the design of computer systems intended to support them in this effort (Schmidt & Bannon, 1992: 11).

This reconceptualisation has led to a move away from laboratory-based settings in which to conduct traditional usability testing and towards conducting fieldwork-based case studies. These case studies have taken place in a number of domains including for example; the management of planes in air traffic control (Hughes *et al*, 1994), the organisation of documents in civil engineering (Trigg *et al*, 1999) and interaction in hybrid spaces (Luff *et al*, 2006). Such studies have examined the ways in which members of a setting organise their work practices, how communication and collaboration is achieved and the ways in which

material and technological artefacts are used to accomplish activities. The results of these studies have identified requirements for software development projects that design systems in support of cooperative work primarily through the resources available via networked computing rather than through stand-alone applications.

This reconceptualisation of the notion of usability, where the criterion is one of assessing a system's ability to support skilled practices and cooperative work reorients the software design process so that "rather than refining a pre-specified task or aiming to transform the way work is accomplished, the technology should provide a resource through which collaborative activities can be seamlessly interwoven within the everyday work of the participants" (Luff *et al*, 2000: 8). In approaching the design of systems from this perspective, the embedding of technologies into work settings might cease to be viewed as a disruption and be seen rather as an opportunity in which to design new tools that support work and interaction in novel ways that nevertheless maintain the support and interest of its potential users.

This chapter has provided the context from which the motivation for this thesis originates; with large budgets and tight time constraints e-Research projects have recognised that identifying researchers' requirements for applications that support distributed, collaborative research is crucial to ensuring the successful uptake and use of these systems. Hence, the motivation to develop a requirements elicitation technique that may be used to evaluate e-Research prototypes so that emerging requirements might be incorporated within an iterative software development process. Providing researchers with an opportunity to evaluate e-Research prototypes may resolve some of the tensions between e-Research visions and researchers' practices.

To recap, section 2.2 of this chapter introduced the e-Research programme with an analysis of its historical underpinnings and the development of its vision. Section 2.3 went on to highlight some of the tensions between the e-Research vision and researchers' practices and the ways in which potential users have been involved in the software development process. This final section has discussed the notion of usability as it has been most commonly understood and the need for its reconceptualisation. The embedding and subsequent adoption

of e-Research systems within academic disciplines is crucial to the success of this initiative. Involving the researchers who are being asked to use these systems in the design process could facilitate their understanding of the technological possibilities and could also provide project members with an opportunity to understand the types of research the system is being designed to support, creating an environment of *mutual learning* (cf. Ehn, 1990; Greenbaum & Kyng, 1991; Bødker *et al*, 2004). The next chapter will continue the discussion with an examination of the primary prototype evaluation techniques available that are informed by either HCI, CSCW or PD and which may be incorporated in a software development process in order to involve potential users in the design of systems.

Chapter 3:

Understanding computer-supported work and interaction

3.1 Introduction

The purpose of this chapter is to develop an understanding of the background and role of both human-computer interaction (HCI) and Computer Supported Collaborative Work (CSCW) in the design of computing systems. Each has a different theoretical and philosophical orientation that informs the ways in which it attends to the gathering of requirements as well as to the evaluation of prototypes. To begin, Grudin (1990) provides a good broad overview of how computers use has progressed from roughly the 1950s to the 1980s. His general premise is that systems have evolved from individual and solitary uses to cooperative and social ones with the introduction of networked computing, groupware systems, and mass exposure to computing. He goes on to describe five phases in the ‘evolution of the interface’.

The first two phases are the ‘user-developer’ periods where software engineering experts worked directly with hardware and programming tasks - *the interface at the hardware & programming task*. At this level, interfaces are assessed within a community of experts. Grudin states that these first two phases were dominant through the 60s, 70s and early 80s⁸. In the third phase, which began in the early 80s, there was a substantial increase in the use of computers by non-programmers which led to the emergence of HCI as an important research programme within computer science and software engineering academic departments. With the introduction of the mouse and graphical user interface it is described as - *the interface at the terminal*. At this time, the focus turned to keyboard and display interaction and included

⁸ However, it can be argued that this phase still exists amongst expert communities in the sciences who build their own software solutions within scientific settings.

assessments of human perception, basic cognitive processes and motor abilities with research informed by ergonomics, cognitive theory and experimental psychology. The fourth phase, turned towards the ‘dialogue’ between the end-user and the system - *the interface at the interaction dialogue*. In this phase mismatches between end-user goals and system representations are identified by applying concepts from cognitive psychology to the evaluation of systems. At this time, the ‘think-aloud’ protocol was incorporated in HCI research to identify higher cognitive processes. Finally, Grudin concludes with a description of the fifth phase, where the focus of study is oriented towards how computers support group communication and coordination within the places in which they are used - *the interface at the work setting*. It is at this time, that researchers became interested in the social context of computer use and began a research programme which focused on acquiring an understanding of the skilled practices and social interactions that participants in a setting rely upon in the achievement of cooperative work and how the introduction of computing systems might affect these processes⁹. At this stage, Computer Supported Cooperative Work (CSCW) emerged in 1984 as a research programme quite distinct from HCI.

This trajectory however is not a linear progression and is used more for explanatory purposes. In fact, each of these ‘interfaces’ are still relevant today and HCI and CSCW research focus on these different aspects. This chapter will present an overview of the analytic orientations of and the prototype evaluation techniques used within both HCI and CSCW. First, section 3.2. will discuss the theoretical underpinnings of HCI with its origins in the cognitive sciences. Next, section 3.3 will present Computer Supported Cooperative Work (CSCW) and its origins in sociological approaches to understanding human interaction, work practice and technology. Then, section 3.4 will discuss prototype evaluation techniques used in both research areas and will discuss the attributes of each and challenges in their use. Finally, section 3.5 will provide a summary of the key differences between the two approaches.

⁹ This again is a broad generalisation as sociotechnical where considered as early as the 1970s by Enid Mumford and the Tavistock Institute.

3.2 HCI: an information-processing perspective

Cognitive theory informs the analytical orientation of HCI (Rogers, 2000; Crabtree, 2003) which focuses on gaining an understanding of human perception within an information processing framework. The theory has produced research programmes such as cognitive psychology and artificial intelligence where mental processes are modelled in computational terms as inputs, outputs and symbolic representations. This notion of human activity has been described most notably by Card *et al* (1983) through the ‘model human processor’ framework which describes human cognition as being made up of “three interacting systems: the perceptual system, the motor system and the cognitive system” (Preece *et al*, 1994: 65). Additionally, the model suggests that people process information in three stages. First, people make contact with objects, this is the stage of *perception* where objects are referred to in both short and long-term memory through visual and auditory channels. Second, cognitive processes *interpret* objects and determine actions through representations accessed within memory. Third, *motor processes* respond based upon cognitive decisions (e.g. thinking ‘push yes’ results in the arm extending and the finger pressing the button).

This framework of human action was used to develop the GOMS (goals, operators, methods and selection rules) model (Card *et al*, 1983). The model has been used in HCI research to predict the actions of end-users as they carry out tasks through a software interface. The model takes into account the goals of users and what they want to accomplish, operators of the system that enable actions, methods which are learned sequences of action that the computer supports and selection rules which are the choices that the end-user makes in order to complete tasks. Using the GOMS model, perception and response to stimuli are measured and later predicted, “enabling calculations to be made of how long a user will take to carry out various tasks” (Preece *et al*, 2002: 96). The model has been used to estimate and predict how quickly users might learn, be trained on, and execute system operations (John & Keiras, 1996). GOMS has been used as the basis of Task Analysis (TA) which focuses the unit of analysis on the sequence of steps or actions that need to be taken in order to achieve a goal when using

computing systems (Shepherd, 2000: 16). This framework focuses on maximising the usability of task sequences as they occur within a computing system.

Another theoretical framework predominate in HCI is the notion of ‘conceptual models’, which extends the ‘model human processor’ (the three interacting systems of perceptual, cognitive and motor function) to include *internally constructed frameworks* of the world (Norman, 1988). Conceptual models are the ideas and thoughts constructed within the human mind that relate to how objects work, what events mean, and how people behave. Norman (1988) describes three attributes that the mind grasps in order to construct internal representations of the world. These are *affordances* - perceived and actual properties of a thing; *constraints* - things that prevent or limit action and understanding, and *mappings* - the metaphors, analogies and cultural standards used to reference the world. This theoretical framework has been operationalised within HCI research through the “seven stages of action” (Norman, 1988: 45-53) which is used to reveal mismatches between end-user conceptual models and a system’s representations at the interface.

Both of these frameworks developed by Card *et al* (1983) and Norman (1988) have largely predominated HCI research. However in 1995 Hutchins proposed, in his monograph *Cognition in the Wild*, a new orientation to the understanding of cognition, distributed cognition (DCog), which repositions cognitive processes as distributed across actors (both human and artefact). In doing so, DCog removes the limitations inherent in examining only individual cognitive processes. Instead, DCog positions the notion of communication as the exchange of each actor’s internal, cognitive ‘representational state’. This in turn, leads to the production of a ‘shared representational state’ that is made up of multiple representations that the entire cognitive system (both human and artefact) has the ability to access. In this way, artefacts and people are framed as ‘communicators of representational states’ that are constantly working towards the production of shared representational states. Distributed cognition focuses on the question, “How do they go about knowing what they know?” (Hutchins, 1995: 131). That is, how do all actors (human and artefact) in a system participate in information processing, use and exchange through input/output mechanisms and symbolic representations. DCog has been adopted by HCI researchers as a way in which to take into account social interaction. Even so,

this reconceptualisation of cognition however novel remains faithful to the notion that social action is an internal, cognitive process where information is processed through input/output mechanisms within the brain and then shares them as messages transmitted and received as mental representations.

HCI research, through its alignment with cognitive theory, conceives of interaction as a tight “cognitive coupling of users and computers” (Crabtree, 2003: 10). It attempts, through the modeling of input/output mechanisms and symbolic representations, to reveal the “internal constructions of some aspect of the external world” (Preece *et al*, 2002: 93). However, the information processing model has been considered inadequate by some researchers in the field as the following quote demonstrates:

We are not asking about individual meanings and understandings - individual subjectivity. We are, instead, asking about intersubjectivity, the ways that we organise our social and organisational worlds so that we can find them understandable, meaningful, significant in the standard, patterned, institutionalised and hence shared ways that we do ... We begin not by asking how - as an individual cognitive process - we see the things around us, but how as a collective endeavour we organise the world of meaningful objects. How is the world made recognisable for us as the world which we unproblematically inhabit? What we want to say is that knowing how to look is like knowing how to speak - knowing the language games, the culture, the practices embedded in any environment (Anderson & Sharrock, 1993: 158).

This alternative orientation focuses analytic attention on socially acquired interpretative practices used to make-sense of the world. In contrast to cognitive theory, sociological orientations aim to understand interpretative processes and knowledge acquisition as developed socially through our contact with other people rather than individually through brain-based representational states and input/output processes. In this way, sociological orientations to understanding human action and communication seek to understand how interpretative practices emerge and are maintained amongst groups of people through both language and interaction. The next section will present a discussion of how CSCW is informed by the social sciences and its approaches to understanding phenomena.

3.3 CSCW: a social computing perspective

This section will present the conceptual underpinnings of one strand of CSCW research informed by Ethnomethodology (Garfinkel, 1967). Other strands of CSCW research include

those that are informed by Strauss' Theory of Action (cf. Fitzpatrick, 2003), Actor-network Theory (cf. Hanseth & Lundberg, 2001) and Activity Theory (cf. Nardi & Redmiles, 2002). For the purposes of this section however, the focus will be on Ethnomethodology (EM) because it is the theoretical orientation that has informed the research for this thesis¹⁰. CSCW has traditionally focused on two research areas (Wilson, 1991). The first is concerned with the design of *groupware systems* such as communication tools (e.g. video conference, conference calls, email), shared workspaces and shared data repositories. The second strand focuses on understanding *cooperative work* and investigates individual work patterns, collaborative processes, organisational design, and end-user involvement in application development & prototyping. Both areas frame their investigations within a two-times-two matrix (Carstensen & Schmidt, 2003) used to characterise the time and space dimensions of both co-located and distributed group work (Table 1).

	Same place	Different place
Synchronous	Face-to-face meeting	Video conferencing
Asynchronous	Assembly line (Post-it note)	Email system

Table 1: Johansen's CSCW matrix¹¹

The matrix represents time as being either synchronous or asynchronous and space as either co-present or distributed. It is a useful tool for conceptualising how social interaction and communication might be coordinated under these different space-time conditions (e.g. face-to-face, video conferencing, post-it note, email, etc). In order to do so, the policies and practices of ethnomethodology (EM) have been used by CSCW researchers to understand and analyse social interaction within these varying space-time conditions, which includes, for instance, the ways in which the social organisation of work might be transformed through embedding technologies designed to support both interaction and communication in novel ways. More specifically, EM-informed CSCW research aims to recover, describe and analyse the practices and procedures that participants rely upon in order to accomplish activities relevant to the

¹⁰ For a comprehensive review of the other sociological theoretical orientations mentioned see Randall *et al* (2007).

¹¹ Johansen, J. (1988): Groupware. Computer support for business teams. New York: Free press.

work setting. The outcomes of these studies provide detailed ethnographic accounts of the skilled practices and cooperative work arrangements participants rely upon in the achievement of their routine activities within the workplace. These accounts are then usually presented to software engineers for the purposes of identifying requirements and informing software design. Both EM and EM-informed software design will be discussed in greater detail in chapters 4 and 5, respectively. For the time being however, it is essential to discuss another important orientation to understanding cooperative work for the purposes of software development - Participatory design (PD).

PD emerged in the 1970s as a Scandinavian response to the dislocation and deskilling of workers when computing systems were initially introduced into workplaces (Kensing & Blomberg, 1998). Originating as a workplace democracy movement, within trade unionized workplaces, its aim is historically rooted in the facilitation of worker empowerment within the technology design process through worker participation in software development (Muller, 2002). It is an action-oriented research programme and its goal is towards the “democratization of system development” (Ehn, 1991: 299). This is achieved through the development and deployment of techniques which facilitate direct end-user involvement in the design process.

Traditional notions of the software engineering lifecycle have focused primarily on developing software in a linear manner usually from requirements specification documents written as part of a Waterfall model (Boehm, 2006). This type of linear software development process is seen as inadequate within the PD community because it may facilitate a tendency to operationalise corporate visions of how efficiency in the workplace might be improved and may be skewed to favour management’s perspective rather than to investigate the consequences of envisaged designs on work practice and job satisfaction, hence its focus on ‘workplace democracy’ (cf. Ehn, 1990; Greenbaum & Kyng, 1991). In contrast to the Waterfall model, PD reorients the software development lifecycle from software product to work process-oriented design. As Bannon (1991) suggests, traditional software development projects aim to develop a technology (a product) to meet an organisational business objective (e.g. increasing productivity, cutting costs, optimising hardware/software, changing work processes)

with minimal consideration for the wider organisational and social effects (the processes) that the introduction of technology may have on settings of work (e.g. work practices, job satisfaction, division of labour, re-training, etc.). PD research focuses specifically on these issues including the nature of participation, the politics of design, and techniques used to incorporate end-users in the design process (Kensing & Blomberg, 1998: 168). PD techniques provide an avenue with which to explore worker-related issues through, for example, cooperative prototyping, scenario development and workshops where end-users are given the opportunity to engage directly in the process of design, thus contributing to the identification and definition of requirements.

PD's theoretical and philosophical underpinnings include activity theory (Bødker & Grønbaek, 1991), Heideggerian phenomenology (with its explication of tool use), Marxism (with its notion of worker emancipation) and Wittgenstein's language games (understanding professional languages) (cf. Ehn, 1990). More specifically, 'professional language games' include the ways in which concepts and practices are attended to in a work setting. These games require that members know and follow its rules, however, such rules may not be explicitly stated and may change based upon contingencies. In order for software developers to gain access to this private language an entirely new language-game must be created for the design process itself that fuses both developer and end-user concepts and practices. This shared language-game is developed through co-design activities, that take place between software developers and end-users, such as future workshops, story-telling, and iterative prototyping to name a few (Ehn, 1993; Kensing *et al*, 1998). In this way, PD research is located within the *practice of design* itself.

Through the development of cooperative and participatory design practices the division of labour in software design is reconfigured so that designers and skilled practitioner/end-users work together with the goal of producing technological artefacts that ensure greater job enrichment (Ehn, 1990). PD techniques have been developed to orient the design process towards the local, practice-driven and bottom-up as opposed to centralised, vision-driven and top-down (Ehn & Kyng, 1987). PD sits in stark contrast to traditional HCI approaches and also differs from the ethnographic fieldwork approaches used in CSCW. As Kensing & Blomberg

(1998) note both CSCW and PD share concerns with ‘design-oriented research’, however, the authors note four key differences between the two:

- PD does not single out the notion of requirements for cooperative work but rather focuses on “cooperative strategies for system design”.
- CSCW does not focus on the development of techniques whereas PD is “defined, in part, by the techniques and methods used”.
- CSCW does not require direct end-user participation in design, whereas PD is oriented to developing techniques for direct end-user participation in design.
- CSCW has oriented itself towards enabling the design of novel *technologies* (e.g. collaborative or groupware systems), whereas PD has oriented itself towards a commitment to workplace *democracy* and “worker participation in design”.

Kensing & Blomberg (1998) conclude with a suggestion that further dialogue between CSCW and PD is needed and a recommendation that CSCW researchers include end-users as participants in the design of systems. This recommendation is taken seriously and addressed with the development of the Contextual Evaluation Framework (See chapter 5), which seeks the involvement of potential users of e-Research systems through their evaluation of prototypes, so that any issues which may arise during an evaluation may be attended to within an iterative software development lifecycle.

3.4 Prototype evaluation in HCI & CSCW

With the overview of the analytic orientations of HCI and CSCW discussed in the previous two sections, this section will introduce the primary practical applications of these theoretical and philosophical orientations, specifically as they relate to the evaluation of software prototypes. Software prototype evaluation is conducted to examine whether or not a system performs as specified (Bannon, 1996), not only in terms of the execution of code (cf. Hoare, 1996) but also in terms of its usefulness and usability. Twidale *et al* (1994: 441) define evaluation as “a significant check of a system’s capacity to deliver what is required of it”. Evaluations may be either *summative*, to verify completeness by examining the results of a designed system or *formative*, to identify problems by focusing on improving design within an

iterative software development process (ibid). In addition to this, prototype evaluations can be used to determine both the *usefulness* of a system, by examining how functionality supports work practices and *usability*, by examining the user interface for layout, terminology and navigability (Ovaska, 1991: 47; Twidale *et al*, 1994: 444).

Most often, the evaluation of prototypes within HCI is framed from a developer perspective (Ovaska, 1991: 51) in that the session is set-up to determine how well an end-user can make-sense of the interface and its underlying functionality, on its own terms, without explicitly evaluating it for the ways in which it may support, constrain or transform work practices. Additionally, prototype evaluation frameworks have historically used pre-defined categories to guide an assessment such as ‘visual clarity’, ‘informative feedback’, ‘error prevention and correction’, etc. (ibid: 52). These categories are more than likely not of concern to end-users as they interact with a system, rather, the categories are meaningful to the software development team. Participants are most likely more concerned with evaluating a system for how well it integrates with their current organisation of work, as well as the ways in which it may support the coordination of activities with others.

As an alternative to these formalistic approaches to prototype evaluation, CSCW and PD focus on understanding technology-in-use within its context-of-use; with an interest in understanding *how people adapt* to systems (MacKenzie & Wajcman, 1999; Grint & Woolgar, 1997) and *how systems are adapted* by the people who use them (Büscher *et al*, 2001; Hartswood *et al*, 2002). Even so, evaluation in CSCW has proved to be difficult for three key reasons identified by Grudin (1988: 87): it is difficult to evaluate multi-user collaborative systems where groups of users participate in a single session; observation needs to be extended over a longer period of time to understand a system’s influence on work practices; and challenges to conducting evaluations within distributed environments. Even though these challenges remain and have proved difficult to overcome, CSCW researchers have found that sociologically-informed techniques do provide the resources necessary to evaluate systems that are designed to support cooperative work.

This section will discuss five common *formative* prototype evaluation approaches across HCI, PD and CSCW, respectively. For HCI, this will include the pluralistic walkthrough in section 3.4.1, and the think-aloud protocol in section 3.4.2. Cooperative prototyping, a PD approach, will be discussed in section 3.4.3. After which, two CSCW approaches, situated evaluation in section 3.4.4 and quasi-naturalistic experiments in section 3.4.5 will be presented.

3.4.1 Pluralistic walkthrough

The pluralistic walkthrough is an adaptation of the cognitive walkthrough (Preece, 2002: 422) and is used to evaluate a system in terms of how easy or difficult it may be for end-users to move through a task sequence. It consists of pre-written task scenarios that a group of end-users, developers and usability engineers work through individually and then discuss together as a group (Bias, 1991). The group evaluates low-fidelity paper mock-ups of software screens, which are presented in the same order as they would appear in the actual application. Participants write down the actions they would take on that screen, after which the ‘right answer’ is told to participants (ibid: 95). The session concludes with a discussion of participants responses to the screens and potential usability issues. The process continues until each screen in a task sequence has been reviewed. After the entire sequence of screens is presented, participants suggest changes to task sequences which includes such things as page layout, terms and labels, and adaptations to task flows.

Even though this approach is collaborative its focus remains oriented to evaluating the system against pre-assigned task sequences rather than evaluating the system for its support of work practices within a naturalistic setting of use. As a consequence, the evaluation focuses on the examination problem-solving and learning strategies as participants work within the constraints of a system. The artificial setting of the session is also problematic, as both developers and usability engineers take on the role of end-users and incorporate their viewpoints into the evaluation. However, they are not members of the intended user community and may not fully understand or know of the routine work practices and collaborative work activities of those members. The result is a decontextualised and abstracted

procedure for evaluating pre-defined task sequences that may or may not support the seen-but-unnoticed coordination of activities as they unfold moment-to-moment in the actual achievement of work.

3.4.2 Think-aloud protocol

The think-aloud protocol is the most well-known prototype evaluation technique and is most frequently referred to generically as ‘usability testing’. It utilises the verbal protocol (Ericsson & Simon, 1980) as its analytic orientation and method. The verbal protocol originates in cognitive psychology where it is used to access the introspective thoughts and mental processes of participants (Boren & Ramey, 2000: 262). Utilised in prototype evaluation, individual participants are given a set of tasks to perform and asked to verbalize their thoughts to an observer as they step through each task. The technique is used to gain access to end-user misconceptions of the system (Nielsen, 1993: 195) by focusing the analysis on the dialogue between the end-user and computer (*ibid*). During a session, an observer will continuously prompt the participant with statements such as ‘what are you thinking now’, ‘what do you think this message means’ or ‘keep talking’. While the pluralistic walkthrough is related to task analysis, the think-aloud protocol is associated with the evaluation of end-user conceptual models, that is, how representations are internally constructed (see §3.2). During the session the observer takes notes recording participant’s comments and actions. When the session is completed, the observer will review these in order to identify key events, statements, themes, and patterns of behaviour that may have emerged (Preece, 2002: 381).

Those who use the approach for prototype evaluation have described difficulties with its implementation; including the unnaturalness of asking participants to constantly verbalize their thoughts as they work through tasks (Nielsen, 1993: 196), observers forgetting to or not prompting participants enough (Boren & Ramey, 2000), and participants finding it difficult to verbalize their experiences in general (Smith & Dunckley, 2002). These limitations have been addressed with the introduction of a collaborative think-aloud technique called Constructive Interaction developed by Naomi Miyake (O’Malley *et al*, 1984). Constructive interaction (CI) involves two participants working together in order to work through a shared problem

(Miyake, 1986). The sessions are video recorded as participants work through a broadly defined objective; e.g. “talk together to figure out how a sewing machine makes its stitches” (ibid: 159). After which, participants are encouraged to engage in natural interaction with each other. This interaction between the two collaborators can overcome the unnaturalness of the individual think-aloud protocol technique:

Having two people talk with one another circumvents this problem. The talking is natural. The pair naturally explains not only what they have been thinking about, but why they think it. The situation makes a usually invisible process visible. Moreover, the two participants interact constructively. When a subject working alone claims to have solved a problem, it can be difficult for the subject to come up with a counter example to test it. But this testing process occurs naturally with two persons (ibid: 159).

The original intention of CI was to use it to model the iterative processes of understanding (Miyake, 1986) through the identification of conceptual models held by participants and how new thought processes may develop through interaction (O’Malley *et al*, 1984: 269). The CI approach has been adapted and used in the evaluation of prototypes first by O’Malley *et al* (1984) and subsequently by other HCI researchers (cf. Wildman, 1995; Kahler, 2000; van den Haak *et al*, 2006). However, even though this approach is collaborative, similar to the think-aloud protocol, its focus remains faithful to its orientation within cognitive theory in that it frames participants’ problem-solving and learning strategies as information-processing events that occur within the individual mind (cf. Ericsson & Simon, 1980). Both the think-aloud protocol and CI are used as *elicitation devices* and have been called “the most fundamental usability method and is in some sense irreplaceable, since it provides direct information about how people use computers” (Nielsen, 1993). This direct access, through talk, of how participants interact with computers, specifically within paired evaluations, has given researchers access to some aspects of participants’ tacit knowledge (O’Malley *et al*, 1984: 270). The tacit knowledge that participants rely upon in the accomplishment of work practices is precisely what CSCW researchers have an interest in uncovering and, as will be discussed later, some CSCW researchers have repurposed and modified the CI approach for use in quasi-natural experiments (see §3.4.5). But, before discussing how CI has been modified for CSCW, a PD approach to prototype evaluation, cooperative prototyping will be discussed.

3.4.3 Cooperative prototyping

Cooperative prototyping is a PD approach to prototype evaluation where the aim is to involve participants directly in both the design and evaluation of prototypes (Bødker & Grønbaek, 1991). The prototypes can be either low fidelity (paper sketches) or high fidelity (interactive computer screens). In either case, the preference is to have participants collaborate around a working model that can be used to conduct actual work-related tasks and one that can also be redesigned quickly based upon participant feedback (Grønbaek, 1990). Cooperative prototyping is situated within an iterative software engineering process (Twidale *et al*, 1993; Bowers & Pycock, 1994) where during an evaluation session, the goal is to facilitate mutual learning between developers and end-users regarding what the system might be able to provide and how work practices might be supported. This approach facilitates the development a shared language-game between developers and end-users. During an evaluation session, participants experience what a future-use situation may be like while having an opportunity to participate in modifying the prototype when breakdowns occur, for example, when a prototype behaves differently than anticipated or expected.

This focus on the notion of breakdowns is taken from Winograd & Flores (1987) where the authors argue for the importance of understanding how shared meaning and understanding develop through the course of interaction. Breakdowns are defined as “the interrupted moment of our habitual, standard, comfortable ‘being-in-the-world’” (ibid: 77). Such interruptions make visible the ways in which a prototype might fail to meet the tacit work practices, assumptions, or expectations of end-users. Identifying moments of breakdown during an evaluation session are opportunities in which to identify usability problems. Breakdowns may be best discovered when participants conduct an evaluation of a prototype while performing routine work activities with it. In order to make visible the ways in which breakdowns might occur, cooperative prototyping incorporates simulated work situations into an evaluation session (Bødker & Grønbaek, 1991; Urquijo *et al*, 1993; Twidale *et al*, 1994).

Cooperative prototyping, in contrast to pluralistic walkthroughs and verbal protocols, moves the evaluation session out of a controlled setting where pre-defined tasks are conducted,

into more realistic settings where routine work activities can be conducted within a naturalistic setting. However, cooperative prototyping has its challenges. For instance, Bødker & Grønbaek (1991) note that prototyping activities are technology-centred and that they focus attention on computer interaction rather than the organisation of work. Additionally, the evaluation session requires participants to engage in two activities simultaneously; they are asked to perform routine activities while simultaneously providing a reflective critique and assessment of the prototype in order to contribute to the design activities that are interwoven into the session. Another challenge includes participants' engagement in the conduct of simulated work activities as developers sit beside them in order to identify breakdowns. This kind of evaluation setting may put pressure on participants as developers watch them interact with the system. Finally, cooperative prototyping sessions require that participants play two roles; one of end-user and the other as a designer of the system.

There are other techniques available that overcome these limitations by removing the dual role requirement which may distract participants away from using the prototype unreservedly because they are simultaneously immersed in design activities. If one wishes to eliminate this dual role, participants would need to be able to engage in the evaluation of prototypes without the need for self-reflection mid-activity. The next two sections discuss CSCW evaluation techniques that go some way in allowing for this.

3.4.4 Situated evaluation

As discussed above, cooperative prototyping focuses on facilitating “design-by-doing” (Ehn, 1990: 376), thus enabling end-users to creatively engage, as equal participants, in the design process with developers. Situated evaluation, on the other hand, incorporates a fieldwork approach to the evaluation of prototypes by conducting *in situ* studies of technology-in-use which focus on identifying patterns of interaction that can then be taken into account in the design of prototypes (Crabtree, 2003: 147). In this way, situated evaluation compares known patterns of work against the features and functionality available in the system. By doing so, situated evaluation attempts to move the evaluation process away from being technology-

centred and towards developing an understanding of the ways in which the skilled work practices of potential users are supported *in situ*.

During a situated evaluation, participants work together as they conduct typical activities using a prototype (Crabtree, 2003). The sessions are video recorded and analysed in order to identify breakdowns, and so in this way, its analytical orientation is similar to cooperative prototyping. Situated evaluation has recently been extended to incorporate an analysis of activities that occur solely within a computer. This is achieved through the analysis of digital log files which serve as representations of the actions taken by end-users (Crabtree *et al*, 2006a). The log files provide a more complete picture of the actions of participants particularly as they interact with ubiquitous computing devices (Crabtree *et al*, 2006b). The study of log files to examine the activities of participants is not new and has been used as an analytic tool for many years by website developers to understand the behaviours of website visitors including most visited pages and drop off pages. What is new in this case is that log file data is analysed side-by-side with video recordings of participants using the prototype. Having access to both log file data and video recordings enables CSCW researchers to juxtapose these two data sources in their analysis of participants interactions with and around prototypes.

Situated evaluations can also be used to understand how work practices might be shaped by technology as well as how evolving work practices shape the iterative development of prototypes. These types of evaluations may differ slightly, for example, in the case of ‘co-realisation’ (Hartwood *et al*, 2002) a software developer is embedded long-term into the work setting and modifies the system according to needs that arise in the work setting or in ‘use experiments’ (Büscher *et al*, 2009) which extends PD methods to the co-design of software architectures. In these two approaches a situated evaluation is for ‘co-realisation’ ongoing and embedded in the workplace and for ‘use experiments’ complemented with other PD techniques that facilitate the equal engagement of user-practitioners and software engineers.

The focus of analysis in both cases is to understand how technological assemblies evolve and are combined with existing technologies as a result of their practical use in actual work situations and settings (cf. Büscher *et al*, 2001). Evaluations are also extended over a longer

period of time in order to examine the ways in which embedded prototypes are integrated with the existing workplace technologies. Closely associated with situated evaluations is the quasi-naturalistic experiment. Both these approaches incorporate an EM-informed orientation to understanding how the members of a setting themselves organise activities for each other. Additionally, they both conduct and refer to *in situ* field studies of current work practices in order to inform design and to evaluate prototypes. The key difference between the two is that quasi-naturalistic experiments focus on conducting a detailed analysis of video data using interaction analysis (see §4.3). Additionally, it focuses on an analysis of the development of *mutual understanding*, rather than exclusively looking at breakdowns or the ways in which technical configurations might be re-assembled. The origin and use of quasi-naturalistic experiments, presented in the next section, will conclude this overview of practical approaches to prototype evaluation utilised in HCI, PD and CSCW.

3.4.5 Quasi-naturalistic experiments

Similar to cooperative prototyping and situated evaluations, quasi-naturalistic experiments are formative and their results aim to contribute to the iterative design of software. Detailed in her landmark book *Plans and situated action*, Suchman (1987) conducted one of the first and most well-known EM-informed quasi-natural experiments. The study examined the effectiveness of an ‘intelligent’ photocopier as a kind of conversational partner that would anticipate the needs of users based upon their interactions with the system (Douglas, 1995: 188). Suchman’s study examined the relevance of computer instructions given to participants as they engaged in routine photocopying activities. The evaluation focused on the analysis of ‘conversational turns’ between the machine and human participants. This was achieved by combining (ibid) constructive interaction (discussed in §3.4.2), which engages pairs of end-users in an activity and conversation analysis which examines the sequential production of shared understanding through talk-in-interaction¹². The pairing of participants as they work together to conduct routine work activities using a prototype is central to conducting a quasi-naturalistic evaluation

¹² See chapter 4 for a detailed discussion of conversation analysis.

because it produces a naturally generated protocol in contrast to the artificial verbal protocols produced using the think-aloud technique.

In the interest of collaboration, each makes visible to the other what she believes is going on: what the task is, how it is to be accomplished, what has already been done and what remains, rationales for this way of proceeding over that, and so forth. Through the ways in which each collaborator works to provide her sense of what is going on to the other, she provides that sense to the researcher as well. An artifact of such a collaboration, therefore, is a naturally generated protocol (Suchman, 2007: 123).

In this way, natural conversation can be analysed for the ways in which paired participants develop a shared understanding of a prototype as they engage in collaborative activities using it. A key aspect of Suchman's analysis was to acquire an understanding of the "organization of human-machine communication, including its troubles, in terms of constraints posed by asymmetries in the respective situation resources of human and machine" (ibid: 126). The result of her study showed that the photocopier had an "inability to understand a context of use that is primarily interactive and communicative" (Douglas, 1995: 189). These findings were unsettling within the artificial intelligence community where many believe that it is possible to develop systems that might predict human actions with reference to modelled rule-based notions of activity.

The quasi-natural experiments conducted by Suchman provided researchers with a new analytical focus; the identification of *asymmetries* which may be described as the moments when shared understanding is rendered incomplete because of a lack of mutual shared resources with which to develop such an understanding. This new analytic orientation provided an alternative to breakdown analysis (cf. Winograd & Flores, 1987) in that the resources participants rely upon in the construction of shared understandings could be analysed in order to better understand how technologies might impact upon work practice and communication. More specifically, such an evaluation examines the resources 1) that convey the significance of one's own actions and 2) that enable the interpretation of the actions of others (Suchman, 2007: 125). Quasi-natural experiments have most notably been adopted by Heath & Luff who use the approach to conduct an analysis of asymmetries in order to uncover relationships between work, technology and interaction (Heath & Luff, 1991; Heath & Luff, 1992; Luff *et al*, 2000; Luff *et al*, 2003).

Although, not dissimilar to breakdown analysis, the identification of asymmetries focuses not just on work activities, but attends to the finer details that make collaborative interaction possible. In this way, a prototype is evaluated for its support of interactional accomplishments. Additionally, quasi-naturalistic evaluation, like cooperative prototyping, provides participants with an opportunity to engage in future use situations. However, instead of asking them to engage simultaneously in co-design activities, participants are encouraged to work with the prototype unconditionally in the accomplishment of routine activities. This makes it possible for the CSCW researcher to examine, in detail, the moment-to-moment unfolding of interaction and activities amongst participants as they conduct them without interruption. An approach that combines both cooperative prototyping and quasi-naturalistic experiments is future laboratories (Büscher *et al*, 2004), which engages participants in the experimental use of prototypes alongside cooperative prototyping activities.

This section has provided an overview of the origins and use of quasi-naturalistic experiments as a technique with which researchers may analyse the interpretative work that participants engage in as they use a prototype. For the most part, the technique has been utilised in the assessment of novel technologies such as media spaces and ubiquitous computing. Additionally, quasi-naturalistic evaluations are usually implemented in an *ad hoc* manner by highly skilled researchers within academic settings (cf. Heath *et al*, 1997; Jones *et al*, 2000; Brown *et al*, 2003; Kirk *et al*, 2007). The absence of a practical guide which provides a rigorous and systematic approach to its use has made it difficult for those involved in developing technologies, such as software developers, project managers and requirements engineers, to incorporate the approach into an iterative software development lifecycle. This thesis takes on this challenge by providing a practical guide with which the e-Research community, in the first instance, can utilise this approach and is presented in the case study chapters.

3.5 Summary

This chapter has surveyed the theoretical underpinnings of HCI and CSCW both reviewing and critiquing the main practical approaches to prototype evaluation. Prototype evaluation is a

means with which to provide an opportunity for end-user communities to assess both the usefulness (how functionality supports work practices) and usability (the interface layout, terminology and navigability) of a system. Each of the approaches discussed in this chapter engage participants in the evaluation of prototypes. However, as discussed, the analytical orientation of each approach will determine the kinds of insights derived from an evaluation. It was noted that the two HCI techniques, pluralistic walkthroughs and think-aloud/CI, even though they may be collaborative in nature, nevertheless, frame the analysis of an evaluation within a cognitive construct that aims to identify an individual's information-processing, problem-solving and learning strategies. From this cognitivist perspective results of an evaluation session are usually quantified and reported as task completion, error, success and subjective satisfaction ratings. Additionally, mismatches between an individual's conceptual models and the system's representations are analysed. It should also be emphasized that HCI techniques ask participants to complete pre-assigned task sequences rather than asking them to evaluate prototypes through a range of broad activities that they would normally conduct as part of their routine, day-to-day activities.

The CSCW techniques discussed, on the other hand, seek to engage participants in the evaluation of prototypes in relation to how well it supports 'real world' work practices. This is achieved first by conducting *in situ* ethnographic fieldwork so that the ways in which work practice and collaborative activities are organised may be more accurately understood. Subsequent prototype evaluation activities focus on understanding how a system supports, constrains or transforms the organisation of work practice, based upon the information acquired during *in situ* fieldwork. This is in contrast to HCI with its focus on evaluating systems on its own terms, by how well participants complete pre-assigned tasks. The identification of breakdowns is the analytical orientation used in cooperative prototyping where participants simultaneously use a prototype and contribute design ideas during a session. Both situated evaluations and quasi-naturalistic experiments address the challenges of conducting co-design activities alongside prototype use by engaging participants exclusively in activities with the prototype, thus eliminating the need for self-reflection mid-activity. Both CSCW and PD approaches seek to gain access to the organisation of work by describing not

only the explicit and obvious but also by revealing tacit skills and practices - the 'seen-but-unnoticed' (Suchman, 1995). Within a prototype evaluation session this is achieved by allowing participants to conduct activities as they normally would, without prescribing artificial or pre-determined tasks for them to complete.

With an introduction to e-Research discussed in chapter 2 and a survey of HCI and CSCW accomplished here in chapter 3, the next chapter will go into much greater detail describing the theoretical orientation that the Contextual Evaluation Framework is informed by, which includes ethnomethodology, conversation analysis and interaction analysis.

Chapter 4:

Revealing social order in life and in the work setting

4.1 Introduction

The previous chapter provided an overview of the theoretical orientations of HCI and CSCW and discussed their practical application as found in common techniques used in prototype evaluation. This chapter will build upon the previous two chapters with a discussion of three sociological orientations that provide guidance to CSCW researchers interested in understanding the organisation of social order, human behaviour and practical action; specifically ethnomethodology (EM), conversation analysis (CA) and interaction analysis (IA). Each of these has informed the development of the Contextual Evaluation Framework and so it is essential that their details be discussed.

Indeed, there has been passionate debate and discussion about the differing analytic commitments of each (cf. ten Have, 1990; Lynch & Bogen, 1994; Travers & Calvey, 2001) which, however worthwhile a discussion to pursue, is beyond the scope of this thesis. Rather, this chapter will highlight the methodological commitments and focus of each and point to some of the differences between them. First, section 4.2 will provide a detailed overview of Ethnomethodology as an approach for recovering members' methods of producing, recognising and maintaining social order (Garfinkel, 1967). This is in contrast to the use of a specific theoretical lens (e.g. Feminist Theory, Social Construction, Postcolonial Theory, Critical Theory, Marxism, etc) to interpret phenomena. Next, section 4.3 will discuss Conversation Analysis as its own research program influenced by EM policies and practices, which examines the details of the sequential nature of social interaction (Sacks, 1992). Next

section 4.4 will move to a discussion of the ways in which EM, CA and interaction analysis (IA) differ in their analysis of social order. Finally, section 4.5 will provide a summary of how EM/CA/IA examine the ‘documentary method of interpretation’ as a resource for sociological inquiry.

4.2 Ethnomethodology

In this section, ethnomethodology (EM) will be discussed for the purposes of understanding how its analytic orientation has informed CSCW research. EM holds that members’ methods of producing, maintaining and recognising social order are recoverable for description and analysis. Because its focus is on recovering and describing *members’ methods*, EM does not rely on any sociological theoretical orientation (e.g. Social Construction, Critical Theory, Marxism, Actor-Network Theory and its derivatives (cf. Law & Hassard, 1999), etc) with which to frame the analysis of phenomena. Rather, EM calls for a respecification of social science studies with a move away from constructive analysis of phenomena where the ‘social reality’ of a setting, in particular processes of communication and knowledge production, are scrutinised from an ironic stance (cf. Latour & Woolgar, 1979)¹³ and where the products of social and cultural influence are ultimately interpreted and described from the social scientist’s own perspective. In contrast, EM is committed to producing descriptions of member’s own methods of social order and their own interpretations as “locally produced, naturally organised, reflexively accountable phenomena of order*¹⁴” (Garfinkel, 1991: 15). This respecification of sociological study is expanded in Garfinkel & Wieder (1992: 187) with ‘the rendering theorem’:

$$\{ \} \rightarrow []$$

The theorem can be explained as follows: the curly brackets { } indicate locally produced and recognised actions as lived phenomenon; the arrow → refers to social science analysts’ use of analytical frameworks and methods; and the straight brackets [] refer to written accounts

13 For a critique of the Constructivist approach see: Button, G. (1992). The curious case of the vanishing technology, In In Button, G. (ed.), *Technology in Working Order: Studies of Work, Interaction, and Technology*. London: Routledge: pp. 10-30.

14 The asterisk is used by Garfinkel to indicate a respecification of the word ‘order’ to be read as something that is locally produced and accountable phenomenon and something that cannot become accessible through its description within any chosen theoretical framework.

derived from $\rightarrow$. Using this theorem, Garfinkel & Wieder demonstrate the tenuous relationship between lived phenomena and reconstructed accounts of these activities. EM holds that accounts of social activity imbued with theoretical frameworks, in the form of $\rightarrow$, place an interpretative lens upon actual phenomena{ } resulting in a reconstructed account [] of social activity that more than likely does not reflect the lived experiences of participants and the ways in which social order is produced, maintained and recognised by the members themselves.

In order to recover a description and to conduct a subsequent analysis of social order as produced, maintained and recognised by members of a setting, EM¹⁵ proposes an alternative to the production of reconstructed accounts by setting aside the interpretative lens that theoretical frameworks provide. Instead, the phenomenological philosophy of Husserl and the phenomenological sociology of Schutz¹⁶ inform EM descriptions of lived phenomena (Heritage, 1984: 37-74). In the case of Husserl, his notions of the active constitution of foreground and background (gestalt figures) and of the life-world (Lebenswelt) as the mundane world of immediate, lived experience in which all people are immersed has influenced EM. Schutz's contributions include transforming Husserl's notion of Lebenswelt into a sociological investigation (Lynch, 1993:133) through the development of phenomenological sociology. In this way, Schutz developed concepts that make possible the investigation of *the natural attitude* which he describes as "an on-going world of everyday affairs which is, for the most part, taken for granted in its essential being" (Schutz, 1962:xxvi).

Schutz's phenomenological sociology seeks to understand how a world-known-in-common or *intersubjectivity* (the sharing of subjective states by two or more individuals) is established. He conceptualised it as a practical problem of alignment (Heritage, 1984: 54) which he then

15 When Garfinkel was asked about the origins of the word 'ethnomethodology' he conveys the following story (Garfinkel, 1974): Garfinkel was in a library looking through a list in a cross-cultural index mindful of the research he was in the midst of and sees; 'ethnobotany, ethnophysiology, ethnophysics' in a list. At that time Garfinkel was studying jurors in a criminal case and he wanted to understand 'what makes someone a juror?'. Upon seeing the list Garfinkel made the connection that he too was also looking at 'ethno', that is the study of a particular culture and practices. However, in his case, Garfinkel was trying to understand a culture's, in this case jurors, reasoning practices as produced, maintained and recognised by its members. In other words, the ways in which they engage in methods of producing order, hence the term 'ethnomethodology'.

16 Additional influences to Garfinkel's development of Ethnomethodology include Gurwitsch, Merleau-Ponty and Mannheim (See: Lynch, 1993; 1999).

formalised as two idealizations known as ‘a general theory of reciprocal perspectives’, which hold that there exists (Schutz, 1962:12):

1) *an interchangeability of standpoints*, which states that “I take it for granted - and assume my fellow-man does the same - that if I change places with him so that his ‘here’ becomes mine, I shall be at the same distance from things and see them from the same typicality as he actually does; moreover, the same things would be in my reach which are actually in his. (The reverse is also true)”.

2) *a congruency of relevances*, which states that “until counter evidence I take it for granted - and assume my fellow-man does the same - that the differences in perspectives originating in our unique biographical situations are irrelevant for the purpose at hand of either of us and that he and I, that ‘We’ assume that both of us have selected and interpreted the actually or potentially common objects and their features in an identical manner or at least an ‘empirically identical’ manner, i.e., one sufficient for all practical purposes”.

These two perspectives enable participants in a setting to maintain a world-known-in-common that is taken-for-granted until counter evidence motivates revision. It is “through a continuous process of adjustment ... (that) actors succeed in resolving discrepancies in their perspectives which could otherwise throw doubt on the shared nature of their shared perceptions and cognitions” (Heritage, 1984: 56). This conceptualisation of social order as situated and taken-for-granted led Garfinkel (1967) to develop a set of policies and procedures for revealing the intersubjective procedures employed by members of a setting with “an injunction to treat *methods* as ‘discoverable topics of order’” (Lynch, 1999: 217).

EM policies and practices replace methodological discourse (Lynch, 1993: 13) and in what follows is a discussion of each of these. First the two main policies of *ethnomethodological indifference* and *unique adequacy* will be discussed. Then the set of practices utilised by both members of a setting and ‘professional’ sociologists, for recognising social order will be described. These include *accountability*, *reflexivity*, and *indexicality*.

First, *ethnomethodological indifference* holds that the social scientist disregard theoretical frameworks and instead focus on:

[T]he study of such phenomena, seeking to describe members’ accounts of formal structures wherever and by whomever they are done, while abstaining from all judgements of their adequacy,

value, importance, necessity, practicality, success, or consequentiality. We refer to this procedural policy as ‘ethnomethodological indifference’ (Garfinkel & Sacks, 1970: 166).

The formal structures Garfinkel & Sacks refer to are the practical, situated accomplishments that “exhibit upon analysis the properties of uniformity, reproducibility, repetitiveness, standardization, typicality and so on...” (ibid: 166). These formal structures are situated, local observable orders. The policy of *indifference* allows the EM analyst to focus on developing unbiased descriptions of the ‘cultural machinery’ which displays itself through the practical procedures members employ to produce, recognise and maintain social order.

Next, *unique adequacy* is an EM policy that requires the social analyst to acquire a working knowledge of the endogenous practices and activities of members in a setting. In order to satisfy this policy the EM analyst must display a level of competency regarding the practices of the group which he or she is studying. This policy is also referred to as ‘vulgar competency’ (Lynch, 1993: 146). Both ethnomethodological indifference and unique adequacy combined are seen to be a solution to the rendering theorem described above:

Unique adequacy and indifference together constitute a strong version of the ethnographic requirement to understand local languages and ways of life from a native point of view. However, ethnomethodologists do not elicit this “native point of view” by making use of an independent matrix of social science concepts and methods. The “native point of view” is not represented with a cognitive map or other social science construct. Instead, all that counts as investigations – methods and analyses, and judgements of their adequacy – become endogenous cultural accomplishments (Lynch, 1999: 222).

With these two policies, the EM analyst is tasked with recovering the ‘interactional missing what’:

[T]hat is, to investigate what people are actually doing. This ‘What’ is the topic that should be of principal interest to any social scientist, yet this ‘What’ has gone missing in much social research, in large part because the ‘What’ is an activity and not a thing. Accordingly, the disciplinary task now becomes capturing what methods the people themselves use in their local setting, as a matter of their routine lives, to do the work that is the heart of their life-world. The ethnomethods used by the people who staff the culture being studied must be located, identified, and described in that work’s own identifying details (Liberman, 2004: 33).

These two policies are key to uncovering “how members conduct their ‘methodological’ activities” (Lynch, 1993: 142) and in so doing, EM “respecifies issues such as order and society in the details of their workings” (Button, 1991: 6). Because EM takes social order to be an ongoing achievement it aims to study those concrete *practices* that constitute the

endogenous production, recognition and maintenance of that order. Within this accomplishment it is necessary to understand the ‘sense-making’ practices that make the recognition of social order possible. These sense-making practices, available to both the members of the setting and to the EM analyst, are embedded within the activities themselves and are constituted as *accountability*, *reflexivity*, and *indexicality*.

EM holds that all social order is contingent upon local conditions and self-organising rather than something that is enacted by ‘cultural dopes’ who produce “stable features of society by acting in compliance with preestablished and legitimate alternatives of action that the common culture provides” (Garfinkel, 1967: 68). In this fluid, moment-to-moment unfolding of phenomena, a shared sense of the locally situated actions produced there must be maintained and these actions must be made *accountable* to other members of that setting. In this way, members of a setting continually engage in making visible the rationale behind their actions as ‘observable-reportable’ activities that are made “available to members as situated practices of looking-and-telling” (Garfinkel, 1967: 1). Lynch (1993: 14-15) unpacks the elements of accountable actions to include their *orderliness* in that they are not random but rather repeatable and therefore meaningful and coherent; *observable* in that they are public, can be witnessed and therefore intelligible in some sense; *ordinary* in that they are necessarily witnessed in the ongoing participation of members; *oriented to* in that participants are aware of each others’ activities as they produce their own; *rational* in that they make sense to those who can conduct the same activities themselves; and *describable* in that members can talk about the activities produced. In this way an EM-informed study seeks to identify these accounting practices in order to begin to understand how social order is made visible, through its moment-to-moment production and maintenance.

Next, the EM notion of *reflexivity* refers to the ways in which prior actions inform next actions resulting in the moment-by-moment unfolding of events as they are “recognized for ‘another first time’” (Garfinkel, 1967: 9). Reflexivity is a component of accountability (Lynch, 1993: 16), and according to Heritage (1984: 115) it plays “a central role in the creation and maintenance of the sense of sequence of events and the environing context in which they

occur”. Through reflexivity, it is possible to locate the order and organisation of an activity or scene as “intelligible patterns and their constituent particulars are adjusted to one another” (ibid: 103). This ongoing interpretative practice is an aspect of ‘the documentary method of interpretation’ which results in the emergence of common understanding between members of a setting as each activity elaborates upon the other (Garfinkel, 1967: 77).

Finally indexicality refers to the context bound use of language (Heritage, 1984: 142) within a particular setting. Indexical expressions, or ‘occasioned expressions’ include the use of words such as; ‘here’, ‘now’, ‘this’, ‘that’, ‘it’, ‘I’, ‘he’, ‘you’, ‘there’, ‘then’, ‘soon’, ‘today’, ‘tomorrow’ (Garfinkel & Sacks, 1970: 167). The meaning of such words varies dependent upon the situation in which they are expressed and so may mean different things on different occasions. The vagueness of indexical expressions (both linguistic and interactional - such as pointing), are made comprehensible within the context of their use and result in the achievement of shared understandings that are relevant for the purposes at hand. Glossing practices are also considered here, where a topic of conversation “remains unstated throughout” (ibid: 164) even so, communication and interactional work is achieved unhindered.

Additionally, as noted in Lynch (1993: 18) indexicality encompasses “a way of speaking about the entire field of language use that ethnomethodologists investigate”. In this way, indexicality is embedded within and constitutes the *contextual order* where the meanings of words and phrases are revealed within their settings-of-use. This differs from traditional sociological analysis where there is an attempt to produce concrete definitions and explanations of social action which give primacy to theoretical explanations that are removed from those settings. Rather, for the EM analyst an indexical expression’s “relevance, referential sense, appropriateness and correctness will vary whenever it is used by different speakers, on different occasions, and in different texts” (ibid: 19). As a result, the recognition of indexical expressions is an ongoing achievement that occurs within a setting and so a context-free designation of meaning is viewed as unachievable. In this way, the identification of the ‘formal structures’ of social action differs significantly from traditional sociology where these structures are constructed as abstract concepts that are removed from the specific

contexts of social action. Whereas within EM, structure is identified through the analysis of how members of a setting make sense of and make accountable indexical expressions. In this way, formal structures are studied and analysed as contextual orders of ongoing achievement where these structures are embedded in the routines and patterns of interaction as they are produced, recognised and maintained by the members of a setting themselves.

Both *reflexivity*, and *indexicality* are features of *accountable action* (Heritage, 1984: 156) and paired with ethnomethodological indifference and unique adequacy these policies and practices can be utilised to provide descriptive accounts of the production, recognition and maintenance of social order that the participants of the setting themselves orient to as ordinary, routine, ‘seen-but-unnoticed’ activities and events. To conclude, EM’s program is to understand the achievement of practical action and practical reasoning through their enactment as observable-reportable behaviours (Garfinkel, 1996). From these policies and practices the EM program has produced an alternative response and respecification of classic sociology’s constructive analysis of phenomena to a description of routine sense-making procedures as they are produced, recognised and maintained within and through the very activities themselves. The next section will discuss conversation analysis, influenced by EM, and so is another EM-informed orientation used in CSCW research.

4.3 Conversation analysis

Conversation analysis (CA) emerged as its own research program, although influenced by EM policies and practices, through the work of Sacks (1992) and his colleagues (Sacks *et al*, 1974). The purpose of conversation analysis may be described as follows:

The central goal of conversation analytic research is the description and explication of the competences that ordinary speakers use and rely on in participating in intelligible, socially organized interaction. At its most basic, this objective is one of describing the procedures by which conversationalists produce their own behavior and understand and deal with the behavior of others. A basic assumption throughout is Garfinkel’s (1967: 1) proposal that these activities - producing conduct and understanding and dealing with it - are accomplished as the accountable products of common sets of procedures (Atkinson & Heritage, 1984: 1).

This provides a good example of one of the key differences between CA and EM. In general, the CA analyst is interested in examining *conduct and behaviour*, whereas the EM analyst

places their analytic focus on the examination of *practical action and reasoning*. In CA, the sequential organisation of talk is central to understanding human conduct and behaviour and it is through the sequential organisation of talk that the CA analyst may locate conversationalists' reflexively emerging interaction. Additionally, the sequential nature of talk is publicly observable/hearable and so make available the vocally produced accounting practices of members. An example of the sequential, 'speech exchange system' that CA focuses its analysis upon is demonstrated in the following (Sacks et al, 1974: 702):

A: Oh I have the- I have one class in the e:vening.
 → B: On Mondays?
 A: Y-uh::: Wednesdays.=
 B: =Uh- Wednesday,=
 A: =En it's like a Micky Mouse course.

In the sequence above, both A & B develop a mutual understanding of what is being talked about through the moment-to-moment unfolding of turns of talk within the conversation. Here, the conversationalists are developing an understanding of both the day of the course (Wednesdays, not Mondays) and A's feelings about it (that it is 'micky mouse'). Within this turn-taking system each next action is produced as a response to prior actions allowing for understandings to be achieved reflexively. These reflexive actions are both context-shaped and context-renewing (Heritage, 1984: 242):

A speaker's action is *context-shaped* in that its contribution to an on-going sequence of action cannot adequately be understood except by reference to the context - including, especially, the immediately preceding configuration of actions - in which it participates. This contextualization of utterances is a major, and unavoidable, procedure which hearers use and rely on to interpret conversational contributions and it is also something which speakers pervasively attend to in the design of what they say. The *context-renewing* character of conversational actions is directly related to the fact that they are context-shaped. Since every 'current' action will itself form the immediate context for some 'next' action in a sequence, it will inevitably contribute to the framework in terms of which the next action will be understood. In this sense, the context of a next action is repeatably renewed with every current action. Moreover, each action will, by the same token, function to renew (i.e. maintain, alter or adjust) any more generally prevailing sense of context which is the object of the participants' orientations and actions.

Some CA analysts have extended their analysis beyond talk to include embodied interactional processes as speakers and hearers mutually orient to activities through embodied conduct such as gaze and gestures (cf. Goodwin, 1981; also see section 4.4.3). With the addition of embodied conduct to CA analysis, *talk-in-interaction* is considered to be the true phenomena

of study (Psathas, 1995: 2). For the CA analyst then, embodied interactions and the on-going production of context through talk-in-interaction provides for the “maintenance of a framework of continuously up-dated intersubjective understandings” (Heritage, 1984: 264). CA analyses aim to describe the production of intersubjective order as *behavioural phenomena* which lead to mutual intelligibility amongst co-conversationalists. Here again, CA shares with EM an interest in understanding and describing the ways in which members of a setting attain ‘mutual intelligibility’, which is produced and recognised as a result of the documentary method of interpretation (Garfinkel, 1967: 77).

CA techniques have been used by EM analysts in order to gain access to the content of talk and to the moment-to-moment sense making procedures of members in a setting. However, ‘pure’ CA focuses more on the *context-free mechanics* of conversations such as ‘turn construction units’, ‘adjacency pairs’, ‘repair’, ‘agreement and disagreement’, and ‘conversational openings and closings’, for example (cf. Atkinson & Heritage, 1984; Button & Lee, 1987). CA also relies on the use of audio recordings (and video for research that includes an analysis of embodied conduct). This involves collecting *specimens* (ten Have, 1999) or *instances* of similar types of interactional events which are then compared and analysed in detail. Such an approach has been criticised by ‘purist’ EM analysts (cf. Lynch & Bogen, 1994) because of its microanalytic orientation and its reliance upon a language of professional expertise with which to hypothesize about phenomena from a natural observational science perspective. However, for the CA analyst, recordings allow for precise transcription (Sacks *et al*, 1974) enabling a fine-grained examination of the sequential nature of interactional exchanges in order to reveal what Schegloff (1992) refers to as “the procedural infrastructure of interaction”.

Both EM and CA have influenced researchers in the field of CSCW and are used as approaches to recover both members’ methods of producing, recognising and maintaining social order (Garfinkel, 1967), and the sequential production of mutual understanding (Sacks, 1992) as they relate to the accomplishment of activity in and through computing systems. In the next section the finer-grained differences between EM, CA and interaction analysis (IA)

will be discussed in order to understand the ways in which CSCW researchers have utilised each of these three strands of phenomenological sociology.

4.4 EM/CA/IA approaches to the description of observable order

The CEF is informed by sociological approaches that reveal members' methods of practical action and practical reasoning, specifically through ethnomethodology (EM), conversation analysis (CA) and interaction analysis (IA). This section will provide an overview of each whilst not being diverted by the philosophical controversies which might linger within the research community about them. Rather this section will outline the ways in which each approach uniquely orients to the analysis of 'observable order'. The notion of 'observable order' is invoked across EM, CA and IA as the manner in which members of a setting interpret the meaning of activities and where, for professional sociologists, it is accessed as a topic of research inquiry in order to understand and describe members' sense-making procedures. To provide such accounts, the social scientist attempts to gain access to the very same 'documentary method of interpretation' employed by the members of particular setting. The documentary method of interpretation:

[C]onsists of treating an actual appearance as 'the document of', as 'pointing to', as 'standing on behalf of' a presupposed underlying pattern. Not only is the underlying pattern derived from its individual documentary evidences, but the individual documentary evidences, in turn, are interpreted on the basis of 'what is known' about the underlying pattern. Each is used to elaborate the other (Garfinkel, 1967: 78).

In this way, talk-in-interaction, including the indexical properties of interaction and other 'glossing practices'¹⁷ (Garfinkel & Sacks, 1970), are interpreted by members with reference to prior actions that then inform subsequent ones. The result is a process of interpretation that emerges through the ongoing, moment-to-moment unfolding of activities. Through this process mutual understanding between participants is developed. *Formulations* are also sometimes required to develop mutual understanding between members. These consist of attempts to say 'in so many words' what the participants are doing or talking about (cf. Garfinkel & Sacks, 1970; Maynard & Clayman, 1991). Taken in combination - talk-in-interaction, the interpretation of situated indexical expressions, glossing practices and

¹⁷ This term refers to situations where the meaning of talk or behaviour is not explicitly stated.

formulations, are all applied aspects of the ‘documentary method of interpretation’ and used to attain mutually intelligibility amongst members of a setting. The documentary method of interpretation is a taken-for-granted activity employed by members to make sense of the seen-but-unnoticed accountable actions people produce for each other as ‘observable-reportable’ phenomena. All three orientations - EM, CA and IA examine these observable-reportable phenomena, however, each has a somewhat different focus in its analysis.

The following three sections will describe first, in section 4.4.1 ethnomethodology’s adoption of Wittgenstein’s notion of language games as a means with which to understand the ways in which members of a setting orient to phenomena. Section 4.4.2, will discuss conversation analysis’ focus on understanding the accomplishment of order through the mechanics of talk. Finally, section 4.4.3, will provide an overview of interaction analysis as an approach for understanding embodied conduct.

4.4.1 Ethnomethodology and language games

Wittgenstein’s notion of ‘language-games’ links the rules and meaning of language to its situated use. The most cited example of this concept demonstrates the ways in which language is both used and understood within situated contexts, in this example, between builders:

The language is meant to serve for communication between a builder A and an assistant B. A is building with building stones: there are blocks, pillars, slabs and beams. B has to pass the stones, and that in the order in which A needs them. For this purpose they use a language consisting of the words “block”, “pillar”, “slab”, “beam”. A calls them out:--B brings the stone which he has learnt to bring at such-and-such a call.---Conceive this as a complete primitive language (Wittgenstein, 1953: 3, PI2).

EM employs Wittgenstein’s notion of ‘language-games’ as a way in which to understand how they are applied in their specific situated contexts of use. In this way, the language-games used within a particular community, for example, scientific communities (cf. Lynch, 1985) or international shipping (Crabtree, 2000) are scrutinised. So for example, in scientific communities, the EM analyst would unpack the concepts used in these setting such as ‘measurement’, ‘observation’, ‘discovery’, ‘explanation’ and ‘representation’, etc (Lynch, 1993: 281). And for international shipping concepts such as ‘freight type’, ‘checking space allocation’, ‘negotiating shipping rates’ would be investigated and unpacked (Crabtree, 2000).

In addition to language-games, the *instructed actions* of a community would be examined in detail. These are a kind of rule following and are of particular interest for developing an understanding of the ways in which formal descriptions of work are enacted in actual practice. There are two strands of ethnomethodological studies of work; in the first, the EM analyst focuses on developing an understanding of *practices that organise the production of work*, and in the second, commonly referred to as the sociology of scientific knowledge (SSK), the *practices of knowledge production* as enacted in their situated contexts are described. And because natural language is a vehicle for both the production and display of common-sense reasoning as observable-reportable (cf. Garfinkel & Sacks, 1970; Garfinkel *et al*, 1981; Lynch, 1985), the EM analyst will incorporate the tools of CA, such as its orthography as well as its corpus of typical sequence formulations such as ‘agreement/disagreement’ for instance in its analysis. Using CA transcriptions, an analysis can be conducted of the language-game concepts that appear within conversation and the manner in which they are used. So, like the CA analyst, utterances and turns of talk are investigated, however, the goal is not to come to an understanding of the structure of talk, rather it is to reveal members’ practical reasoning and sense-making procedures.

EM-informed accounts written as part of information system design have also been informed by a Wittgensteinian approach (cf. Hughes *et al*, 1992; Button & Sharrock, 1996; Crabtree, 2000). For the purposes of systems design, particular emphasis is placed upon identifying the skilled practices, artefact use and working division of labour of members (Hughes *et al*, 1992). An EM-informed approach to fieldwork may apply several techniques for gaining access to the details of work practices and the organisation of work. For instance, the EM-analyst will discuss the domain at length with its members through informal interviews and will observe the setting in detail to gain an adequate understanding of the ways in which its language-games, concepts and embodied practices are understood, from the perspective of the members themselves (Crabtree *et al*, 2000). In sum, the EM analyst acquires a working knowledge of the domain by gaining an understanding of the situated use of language and instructed actions.

4.4.2 Conversation analysis and the turn-taking machine

In contrast to EM, conversation analysis is not concerned with understanding members' methods of practical reasoning through the use of situated language-games *per se*. Rather, it is:

primarily concerned with the ways in which utterances accomplish particular actions by virtue of their placement and participation within sequences of actions. It is the sequences and turns-within-sequences which are thus the primary units of analysis (Heritage, 1984: 245).

CA takes a primary interest in the sequential production of talk and with the identification of the 'formal apparatus' of naturally occurring talk (Sacks *et al*, 1974). This formal apparatus, referred to as the 'turn-taking machine' is said to be a *context-free* structure, in that it relies only upon sequentiality for its accomplishment, and operates to either define or constrain how *context-sensitive* responses may be displayed. The 'ordering rules' of talk and its sequential production are oriented to by the CA analyst as context itself. This is in contrast to the EM analyst who will examine turns at talk in relation to how language-games are enacted and the ways in which formal rules are adapted by members of a setting in situated circumstances. In this way, CA's approach to understanding context diverges from the EM orientation in that, for the CA analyst, context is a product exclusively of sequentiality¹⁸. Although, the EM analyst also has an interest in the sequential production of talk it is not the sole purpose of their analysis.

Additionally, CA is a highly 'data-driven' approach (Heritage, 1984) that relies upon audio recordings of conversation in order to conduct a detailed microanalysis of conversational structures and to identify the 'normative expectations' with which the sequential placement of responses restrict or constrain subsequent action (ibid: 249). So, for the CA analyst, structures such as 'turn construction units', 'adjacency pairs', 'turn-allocation techniques', 'repair' (Sack *et al*, 1974) and other context-free 'rules' of interaction are identified in relation to their effects upon communication, mutual intelligibility and interactional behaviour.

¹⁸ With the exception of MCDs which are associated with category-bound activities used in sequentially produced talk (cf. Silverman, 1998).

4.4.3 Interaction analysis and paralinguistic communication

Similar to CA, although quite distinct in its own right, interaction analysis (IA) is concerned with identifying the mechanics of embodied interaction that work as behavioural constraints as well as how stable structures of interaction are enabled. IA's origins can be traced to ethology, which examines behaviour as a biological, evolutionary function and is also informed by kinesics, social anthropology, and experimental psychology (Kendon, 1990). In addition to these, IA has also been informed by the sociologist Irving Goffman's work around the 'interaction order' (Goffman, 1983) where he provides analytic guidance for the identification of social mechanisms that influence conventional practices, ritual interchanges and normative expectations that subsequently place constraints upon behaviour (cf. Goffman, 1981). Additionally, Goffman provides social scientists with guidance in the analysis of the status of participants through an examination of 'participation frameworks' which include ratified participants, unofficial participants and over hearers. Interaction analysis may be informed by a variety of research disciplines from ethology, social psychology, to sociology. The IA utilised in this thesis is informed from within its sociological orientation.

IA is different from CA in that it is concerned with describing the interactional mechanics of embodied interaction. IA builds upon CA in its analysis of communication with the use of video recordings as data which allows the analyst to not only hear talk but to also view the ways in which embodied actions such as gaze, gesture and bodily orientation are embedded into the production of talk (cf. Goodwin, 1981). Like CA, it relies upon 'data driven' analysis supported by video recordings in order to conduct a detailed microanalysis of not only conversation, but embodied activities. Also similar to CA, IA is used in the analysis of both mundane and task-oriented, institutional interaction where paralinguistic communication devices both vocal, such as tone of voice and non-vocal, such as bodily orientation are examined in detail (cf. Heath, 1986).

Additionally, the IA analyst is interested not only in sequential but *simultaneous* interaction where actions such as smiles, frowns and head nods may take place while talk is occurring and

may shape subsequent actions (Kendon, 1990: 29). Video fragments are examined in detail in order to identify ‘repertoires’ of recurrent forms of behaviour thought to be “patterned according to customary, quasi-grammatical rules” (ibid: 35). Interestingly, the detailed analysis of video and photographic sequences of behaviour first appeared in Bateson & Mead’s (1942) *Balinese character: a photographic analysis* (Fig. 1) and was later developed by Bateson and others as Context Analysis in the 1950’s-60’s in Palo Alto (Kendon, 1990). This later evolved, at both the Institute for Research on Learning and XeroxPARC, into what we now refer to as IA.



Figure 1. The sequential production of a trance state (Bateson & Mead, 1942).

The kinds of IA accounts that have informed information system design rely primarily upon sequential analysis as a methodological resource (Heath & Luff, 1991; Heath *et al*, 1994; Heath *et al*, 2010). It should be noted that this kind of analysis is conducted alongside ethnographic fieldwork and is thus informed by it. These studies often unpack various key concepts as they relate to computer use such as support for awareness, the details of group interaction and the use of gesture and other types of embodied conduct.

4.5 Summary

This chapter has discussed three approaches that provide guidance to CSCW researchers who wish to understand the organisation of social order, human behaviour and practical action; specifically ethnomethodology (EM), conversation analysis (CA) and interaction analysis (IA). For all three *context* exists within the sequential organisation of talk-in-interaction as participants interpret its moment-to-moment production which is at once context-shaped and context-renewing. However, in order to understand the context in which members of a setting orient, one may also examine both the language-games and the use of indexical expressions. Additionally, Cicourel (1992) describes ‘interpenetrating contexts’ which include social and cultural histories, institutional settings and the prior knowledge participants bring to an activity including their expertise and biographies. It is argued that these affect the ‘interpretive frames’ (ibid: 297) each participant may hold as they engage in the immediacy of an ongoing interaction. Also, Schutz (1932) refers to a ‘common stock of knowledge’ as the knowledge gained and shared from past experiences which also informs a participant’s interpretation of activity.

Given the spectrum with which to understand the context of a setting all three orientations frame the notion of *context as an interpretive process* where participants intersubjectively develop mutual understanding moment-to-moment through the unfolding of activity. The task for the analyst then is “showing that some orientation to context is *demonstrably relevant* to the participants” (Schegloff, 1992: 196; italics added). These relevances are accessed through the ‘documentary method of interpretation’ which is used as a *method* by members to achieve both social order and shared understanding and, it is also used as a *resource* by the professional sociologist in the analysis of those very methods for making sense of observable-reportable phenomena. Through adopting the analytic orientations of EM/CA/IA, the analyst must understand and describe the perspectives of participants as they operate in the world using the same inference and common-sense reasoning mechanisms that the participants themselves rely upon (Sacks, 1992/1: 113-125). The development of the Contextual Evaluation Framework has been informed by each of these - ethnomethodology (EM), conversation analysis (CA) and

interaction analysis (IA) with the aim of understanding and describing the moment-to-moment, ongoing production of order, interaction and activity in and through the use of computing systems. From such an analysis, the CEF may be used to understand how software prototypes may support, constrain or transform activity. In the next chapter, the motivation and need for the CEF will be discussed.

Chapter 5:

Developing the Contextual Evaluation Framework

5.1 Introduction

The thesis up to this point has provided the necessary background for understanding first, in chapter 2, that in order for the e-Research vision of distributed collaborative research to be successful it is important to gain an understanding of what researchers do in their scientific work, how they organise activities, what materials and resources they require, and how they collaborate and interact with colleagues. Next, because e-Research is concerned with the design and use of software, chapter 3 discussed the theoretical underpinnings of the two approaches to software design and evaluation; one based in cognitive psychology (HCI) and the other informed by sociological approaches (CSCW). In the previous chapter, the details of three main strands of research associated with conducting ethnomethodological studies of work, an approach commonly used in CSCW research, was discussed. The purpose of this chapter is to tie the previous three chapters together by clarifying the need for an ethnomethodologically-informed approach to e-Research system prototype evaluation that is at once both rigorous and systematic.

First, section 5.2 will provide an overview of the practical guides currently available that focus on contributing to software design through ethnomethodologically-informed studies of work and it will end with a discussion of the identification of a gap in the literature and a need to develop a systematic and rigorous approach to quasi-naturalistic prototype evaluation that is used, in the first instance, amongst the e-Research software development community. Next, section 5.3 presents the notions of relevance and familiarity as a complementary pair that may be used as an analytic resource in the evaluation of prototypes. Finally section 5.4 will

conclude the chapter with an explanation of the initial conceptualisation of the Contextual Evaluation Framework (CEF).

5.2 Ethnomethodologically-informed software design

EM-informed studies of work that are conducted in order to contribute to the software design process received widespread attention with the publication of Lucy Suchman's book *Plans and situated action* in 1987. The book has been referred to as having had the most profound impact on facilitating the emergence EM-informed studies of technology and social action (Heath *et al*, 2000) and that it set the stage for a broader program of the sociological analysis of human interaction and technology (Anderson, 1997; Button, 2000). Since its publication, EM-informed studies of work and technology have investigated coordination in office environments (Bowers & Rodden, 1993; Blythin *et al*, 1997); document repositories (Trigg *et al*, 1999); centres of coordination such as air traffic control (Hughes *et al*, 1994) and the London underground (Heath & Luff, 1991); financial dealing rooms (Heath *et al*, 1994); healthcare (Heath & Luff, 1994); media spaces (Heath *et al*, 1997), ubiquitous computing (Crabtree & Rodden, 2007; Büscher *et al*, 2009) and e-Research (Hartwood *et al*, 2003; Jirotko *et al*, 2005).

Each of these studies, and many more¹⁹, all share a common commitment to developing descriptions and analyses of activities and interaction as they naturally occur *in situ*. The aim of these studies has been to inform the design of software meant to support the skilled practices and cooperative work of professionals. EM-informed fieldwork is utilised to both reveal and describe the ways in which material artefacts, technologies and social interaction are interleaved in the accomplishment of routine practical actions and activities. Through detailed descriptions and analyses, these studies reveal the social organisation of work; tacit, 'seen-but-unnoticed' practices; and the *ad hoc* and contingent procedures and practices that participants rely upon to accomplish activities.

¹⁹ See the proceedings of CSCW and ECSCW conferences and the Journal of Computer Supported Cooperative Work for a large portion of this corpus of research.

As research in this area has advanced, practical explanations for how to conduct EM-informed ethnographic fieldwork began to appear in journals articles and conference papers²⁰ (cf. Bannon, 1991; Bentley *et al*, 1992; Hughes *et al*, 1994; 1995; Anderson, 1994; Blythin *et al*, 1997; Harper, 2000). These articles present general guidelines for how to proceed with an EM-informed study and provide examples using case study material. Additionally, they reflect upon lessons learned concerning the techniques used as well as issues that may arise when embedding ethnographic studies of work within software engineering practice. More recently, practical guides have appeared in book format and provide greater detail regarding the history of ethnographic research, its theoretical underpinnings and general guidelines for its use accompanied with case study examples.

In the following sections an overview of these practical guides will be presented. First, section 5.2.1 will discuss three practical guides for conducting ethnomethodologically-informed fieldwork. After this, section 5.2.2 will describe practical guides to video-based studies of work focusing on the many articles, and one book, devoted to this approach to EM-informed fieldwork. Finally, section 5.2.3 will end with a discussion of the identification of a gap in the literature and a need to develop a systematic and rigorous approach to video-based quasi-naturalistic prototype evaluation that may be used, in the first instance, within the e-Research software development community.

5.2.1 Practical guides for ethnomethodologically-informed fieldwork

In the book *Designing Collaborative Systems: a practical guide to ethnography* (Crabtree, 2003) the need for ethnographic fieldwork in system design is discussed. The book begins with a discussion of the inadequacy of the Waterfall Model in software development and with the limitations of HCI for understanding cooperative work. The organisational context of computer use is then discussed and a brief history of CSCW is provided. The book then presents a practical guide for examining cooperative work using EM-informed ethnographic

20 Although the PD research community has produced numerous practical guides for cooperative design (cf. Ehn, 1990; Greenbaum & Kyng 1991; Bødker *et al*, 2004) these will not be discussed in this chapter. As noted in § 3.3, PD differs from CSCW and WPS (Kensing & Blomberg, 1998) in that the practical guides available focus on ways of redefining the division of labour within the software development process through the introduction of participatory design practices. To generalise, PD focuses more on developing techniques for facilitating *cooperative design sessions* between software engineers and user communities rather than *requirements gathering* techniques that are EM-informed.

fieldwork. The book suggests that a variety of data be gathered through reading job descriptions, conducting informal interviews, drawing out diagrams of the work setting, taking photographs, video recording artefacts in use, writing descriptions of artefacts, and creating workflow diagrams. Once this data is gathered the book suggests an EM-informed approach to data analysis and provides examples of the ways in which conversation may be analysed using CA. Additionally, the book discusses conducting an analysis of natural language by developing an understanding of the language games employed within a setting.

The book attempts to provide strategies for incorporating ethnographic findings into formats that are more easily understood by software engineers, specifically through user scenarios and use cases. The case study examples are derived from an analysis of the use of a electronic library catalogue at a library help desk. The author's advice to readers is to develop 'thick descriptions' of a work setting and from this to gather instances that may be of particular interest to software design. The types of ethnographic studies suggested include 'quick and dirty' ethnographies which allow researchers to get an overview of a work site in a short amount of time; 'concurrent ethnography' where fieldwork and design activities proceed in parallel, and 'evaluative ethnography' which does not require a prolonged period of fieldwork but rather serve as a "sanity-check" (p.91) in relation to design proposals or prototypes. Additionally, it is suggested that the requirements engineer may want to examine other ethnographic studies that have been conducted within the target domain.

The book also provides suggestions for translating ethnographic descriptions of reoccurring patterns of interaction into the Unified Modelling Language (UML)²¹ through the development of 'use case models'. The requirements engineer is advised to describe various attributes of an identified pattern such as its setting, organisational context, and accompanying conversational transcript which provide empirical data. The book then suggests that user scenarios may be written based upon these patterns. Finally, the book concludes with a discussion of how one may conduct a prototype evaluation using 'situated evaluations' (see §3.4.4) and provides an example of its use.

²¹ <http://www.uml.org/>

Another book that serves as a practical guide to EM-informed ethnography is *Fieldwork for Design* (Randall *et al*, 2007). This book differs from the first in that it provides a broad overview of the many ethnographic fieldwork orientations such as Activity Theory, Distributed Cognition, Participatory Design, Grounded Theory and Actor Network Theory. Additionally, the book discusses practical issues around conducting ethnographic studies such as gaining access to a setting and developing trust amongst participants. The book makes a clear distinction between conducting fieldwork and the analytic skills necessary to provide meaningful interpretations of fieldwork data. Like Crabtree (2003), the authors take a Wittgensteinian approach to the analysis of data through the identification of language-games used in a professional setting and, of patterns of interaction that occur within a setting. Additionally, it stresses the need to understand the ‘working division of labour’, which refers to how work is coordinated in actual practice in contrast to formal organisational descriptions of it.

The book also reiterates Crabtree’s (2003) types of ethnographic studies; ‘quick and dirty’, ‘concurrent ethnography’, ‘evaluative ethnography’ and ‘re-examination of previous studies’. Interestingly, the authors use the same diagrammatic descriptions that Crabtree uses which have their origins in Hughes *et al* (1994). The authors suggest that the fieldwork researcher understand the use of artefacts and their ‘lifecycles’ including how they are used, when they are used, why they are used, and when they are not used. Additionally, they suggest that researchers should note ‘troubles’ and ‘work arounds’ that occur in the workplace. The book also suggests that the fieldwork researcher understand the ‘flow of work’ including procedures and its coordination with other members. The authors conclude by stating that “the major analytic focus of fieldwork for design is the deployment of skill, and particularly how skills relate to the business objectives of the organisation. This in turn relates to the effective deployment of technology within that context” (p.239).

The most recent addition to practical guidebooks for EM-informed ethnography is *Studies of work and the workplace in HCI: concepts and techniques* (Button & Sharrock, 2009). The book begins with an overview of the foundations of ethnography from anthropology to the Chicago School of sociology. From here, it discusses how a researcher might approach an EM-

informed ethnography and suggests five general ‘maxims’ (seventeen sub-maxims in total) or plans of action for conducting an EM-informed study. Take for example the maxim ‘keep close to the work’. Within it, an explanation is provided (e.g. ‘actually be with the people who are doing the work’) and then an example of how to proceed (e.g. ‘how report writing is achieved in practice’). This structure is repeated for all maxims where each includes; a statement of the maxim, its explanation and an example of its use. The book concludes with suggested strategies for how to make good observations in the field including; ask questions, follow both the members of the setting and the output of their work, make notes on everything that occurs and not just what participants tell you, for example.

These three practical guidebooks all take a Wittgensteinian approach to the understanding of members’ methods and suggest that a fieldwork researcher come to understand the language-games and working division of labour of members within a setting. Additionally, the advocacy of video-based studies of work does not feature in either Randall *et al* (2007) or Button & Sharrock (2009), however, the use of video does appear in Crabtree (2003). In the next section, practical guides for video-based studies of work will be discussed. These are EM as well as CA/IA-informed studies where the analysis differs in that rather than being oriented to providing descriptions of language-games they are informed primarily by descriptions of the sequential recognition, production and maintenance of both activity and interaction (see §4.4).

5.2.2 Practical guides for video-based studies of work

Video-based studies of work have a long tradition within CSCW²² (cf. Tatar, 1989; Suchman & Trigg, 1991; Tang & Leifer, 1991; Heath & Hindmarsh, 2002; Büscher, 2005; Jirotko & Luff, 2006; Hindmarsh, & Heath, 2007) and have been used in association with conventional fieldwork (e.g. participant-observation, interviews, document analysis) to examine how work is accomplished *in situ*. One of the first practical guides for video-based studies of work was Suchman & Trigg (1991) where video data was used to understand work practices within an airline operations room. In it the authors use this case study to discuss the practicalities of

22 For a review of the use of video in social science research more generally see Knoblauch *et al*, 2006.

video recording *in situ*, approaches to video data analysis and ways of applying video analysis to system design.

A comprehensive follow on to their work can be found in the article *Interaction Analysis: foundations and practice* (Jordan & Henderson, 1995). In it the authors discuss video as a medium that provides empirical data with which to conduct a detailed analysis of organisation of work. They argue that video provides a richer picture of events thus enabling researchers to share, not only notes from the field, but to watch events as they unfold moment-to-moment within a setting. This affordance of video allows empirical evidence to be shared, analysed and verified amongst colleagues. The authors go on to present best practice guidelines for conducting video-based studies. For example, they suggest creating a video content log summarizing the events that occur in the video. Next, they suggest collaborative viewing of video in order to identify patterns of interaction. When interesting patterns are identified the talk within these video fragments is transcribed and the interactions are examined in detail using interaction analysis.

Suggestions for video data analysis include noting the beginning and ending of events, as well as repetitive tasks and routines, identifying the ways in which participants manage contingencies, and the temporal organisation of activity which is identified through participants' turn-taking through talk. Additionally, they suggest that the researcher note the 'participation structures' of a setting which include how individuals share in a common task, their spatial orientation, bodily alignment and gaze. Other suggestions include watching for 'trouble and repair' including how difficulties are confronted and the ways in which they are resolved. Additionally, the researcher should note the use of material artefacts and documents in the setting. The authors also suggest that the film be viewed by the people who appear in the video so that any ambiguities the researchers may have can be clarified.

The key difference between practical guides used in video-based studies of work and those discussed in the previous section is their analytic orientation. The EM-informed guides discussed above take a Wittgensteinian approach to understanding of members' methods whereas the practical guides for video-based studies of work take a CA/IA-informed approach

and focus on developing a detailed understanding of the sequential relevance of prior actions for subsequent ones. Another difference is that video-based studies focus analysis on the embodied interactions of participants in much greater detail than the EM-informed studies discussed in §5.2.1. In this way, video-based studies of work focus the analysis on developing an understanding and descriptions of the emergent structure of ongoing activity as it is revealed through the complex relationship between talk and embodied conduct as participants engage in the accomplishment of activity (Suchman & Trigg, 1991: 83).

In addition to the many articles available, a practical guide in book format has recently been published; *Video in Qualitative Research: Analysing social interaction in everyday life* (Heath *et al*, 2010). The book begins by describing the history of the use of film in social anthropology as well as its use to understand the mechanics of human movement. It then goes on to discuss the ethics of videotaping, project planning and strategies for gaining access to a setting. The authors discuss the importance of conducting preliminary fieldwork to acquire an initial understanding of a work setting. They also provide suggestions for collecting audio-visual data including camera and microphone positioning, and ways to approach circumstances where participants might react to seeing a camera. When discussing data analysis, the authors suggest that the sequentiality of emergent talk-in-interaction should be the main analytic resource. It then provides similar suggestions from the paper by Jordan & Henderson (1995) for the analysis of video data. However, this book goes into far greater detail by providing extended examples and a much clearer explanation of how embodied activities are sequentially organised and presents Goodwin's (1981) notation for transcribing bodily conduct. The book concludes with advice for preparing presentations and publications.

These practical guides for video-based studies of work provide researchers with techniques for the gathering and analysis of embodied activity and interaction within a setting of interest. However, they are broad-based guides for complementing and informing ethnographic fieldwork. The IA-informed guidebooks, although very useful, do not provide the requirements engineering researcher and practitioner with suggestions for conducting video-based evaluations of software prototypes once software design and testing are underway. The next section will discuss the use of video in the evaluation of software prototypes and the need for a

rigorous and systematic technique that may guide software development teams in its implementation.

5.2.3 The need for a practical guide to quasi-naturalistic prototype evaluation

Some researchers have incorporated a quasi-natural approach to the evaluation of prototypes (cf. Tang & Minneman, 1991; Mogensen, & Trigg, 1992; Luff *et al*, 2003; Kuzouka *et al*, 2005; Kirk *et al*, 2005; Tutt *et al*, 2007; de la Flor *et al*, 2010b). Quasi-naturalistic experiments use video data in the same manner as video-based studies of work described in the previous section. However in this case, video recordings are used to analyse, in detail, both conversation and interaction while participants conduct their routine work activities using a software prototype. Video-based quasi-naturalistic experiments provide requirements engineers and researchers with an opportunity to understand how prototypes may effect work practice, interaction, communication and collaboration. However, because prototypes are not yet embedded into work settings, completely ‘naturalistic’ studies may not be conducted and this is when prototype use experiments may be worthwhile (cf. Büscher *et al*, 2004). Suchman (2007) provides an explanation of the quality and purpose of the quasi-naturalistic experiment as it was used in her study of photocopy machines:

This study proceeded, therefore, in a setting where video technology could be used in a sort of uncontrolled experimentation. On the one hand, the situation was constructed so as to make certain issues observable; specifically, the work of using the machine with the assistance of the expert help system. The construction consisted in the selection of tasks observed to pose problems for new users in “the real world”. On the other hand, once given those tasks, the subjects were left entirely on their own. In the analysis, by the same token, the goal was to construct a characterization of the interaction that ensued rather than to apply a predetermined coding scheme. Both predetermined coding schemes and controlled experiments presuppose a characterization of the phenomenon studied, varying only certain parameters to test the characterization. Application of that methodology to the problem of human-machine interaction would be at the least premature. The point of departure for the study was the assumption that we *lack* a description of the structure of situated action. And because the hunch is that the structure lies in a relation between action and its circumstances that we have yet to uncover, we do not want to *presuppose* what are the relevant conditions, or their relationship to the course of the action. We need to begin, therefore, with observations that capture as much of the phenomenon, and presuppose as little, as possible. (pp. 121-122, quotation marks and italics in original).

A quasi-naturalistic experiment is constructed and so in some sense it is artificial. However, because the activities of paired participants are loosely defined it allows them to conduct

activities in a naturalistic manner, thus enabling the researcher to analyse their sense-making procedures as they use the prototype. The above explanation of quasi-naturalistic experiments, however insightful, and similar to the many other descriptions of their use in software design lacks a step-by-step process or a set of best practice guidelines for its implementation by others who may wish to use it.

In light of this, the Contextual Evaluation Framework has been influenced by Participatory Design methods which do provide more detailed explanations of the implementation of prototype evaluation techniques. In particular, the CEF incorporates the PD notion of *collaborative design* through two activities. First, film elicitation techniques are used where selected video fragments, referred to as ‘video collage’ (cf. Karasti, 2001), provide the requirements engineer with an opportunity to ask participants questions in order to understand their work practices as well as the language and concepts used in a setting. Second, CEF prototype evaluation sessions provide participants with an opportunity to influence the software design process through what Mogensen (1992) describes as ‘provocation’ where taken-for-granted practices are investigated as participants use the prototype for the purposes of influencing the reworking of old practices into new ones.

Providing participants an opportunity to evaluate the usefulness and usability of prototypes early on in their development, as part of an iterative software development process, has been recognised within the software engineering community as an activity that produces worthwhile results for software design (cf. Boehm, 1988; Somerville, 1989). However, with no practical guidelines for the use of quasi-naturalistic experiments, like those offered for EM-informed and video-based fieldwork, there remains a gap in the literature and a need to develop a systematic and rigorous approach to quasi-naturalistic prototype evaluation that may be used, in the first instance, within the e-Research software development community. In addition to EM/CA/IA, the next section presents two concepts; relevance and familiarity, discussed in the CSCW literature, which has informed the initial shaping of the CEF as a technique with which to evaluate software prototypes.

5.3 Recognizing a prototype's relevance to work practice

The prototype is a useful resource for system design because of its unique qualities “as an exemplary artefact that is at once intelligibly familiar to the actors involved, and recognizably new” (Suchman *et al*, 2002: 163). The familiar features of a prototype include software tools and processes that resemble, in some manner, the ones which are employed in current practice. Its new or unfamiliar characteristics include the transformation or reconfiguration of those practices as they are mediated through new workflows, methods of accessing data, and processes of communication and data sharing. The prototype's characteristic of being at once familiar and new (unfamiliar) allows for an examination of how well it may support, constrain or transform the taken-for-granted activities and tacit practices of those who might use it as part of their day-to-day work routine.

In order to identify and understand the ways in which a prototype may reconfigure the substantive research practices of those who might wish to use an e-Research system those who build it must understand a particular domain's community of practice (cf. Lave & Wenger, 1991; Wenger, 1998). A community of practice includes a group of people who share the same or similar knowledge and skilled practices. Take for example Structural Engineering researchers (see §7) who study the effects of earthquakes on buildings and bridges. Each member of this specific community shares similar background knowledge including the history of their discipline, how earthquakes are recorded, how experiments are conducted and the ways in which structural damage is assessed. These skilled practices include shared methods, frameworks, and processes, as well as shared material objects and artefacts that enable members of the community to engage in meaningful action. Additionally, members of a community of practice may collaborate with each other in order to share new practices and to extend knowledge. The very details of these work practices, cooperative working and communication define a community of practice and support for these very details must be reflected in the design of prototypes so that they may be recognised as relevant to work practice.

The notion of relevance (or utility) has been discussed in CSCW particularly in Hughes *et al* (1992) with their analysis of the ways in which air traffic controllers conduct activities and use material artefacts such as radar data, paper flight strips and radiotelephone communications in the management of both planes and airspace. Each of these tools, the processes and knowledge of when and why to use them allow members of an air traffic control room setting to engage in meaningful action with each other in the accomplishment of air traffic management. The authors note that developing an understanding of the relevance of both artefact use and interaction as work is carried out *in situ* is necessary in order to inform the design of novel technological approaches to conducting the same activities using, in this case, electronic flight strips. The authors conclude that the success of a system that would incorporate electronic flight strips into the work of air traffic control is dependent upon the understanding that the technology project members have about the relevances of artefact use and interaction.

Building upon this notion of relevance in ethnographic fieldwork, Mogensen & Trigg (1992) discuss the ‘recognizability of relevance’ as a crucial factor in the evaluation of prototypes and the appropriation of technologies:

In order for artifacts to be seen as artifacts-that-matter, they had to be *recognized* as potentially or imaginably relevant for practice (p.7, italics in original).

Conducting a prototype evaluation session provides participants with an opportunity to assess a system’s relevance in its support for work practices. In this way, a prototype may be evaluated for its ‘praxeological validity’ (cf. Garfinkel, 1996) which includes the ways in which topically relevant work is supported through a system. In order to develop a deeper understanding of the notion of relevance and how it might be used as a resource in prototype evaluation it is necessary to discuss its character and qualities with reference to Schutz (1970).

In Schutz (1970) the notion of relevance is presented as a Gestalt process where the act of interpretation involves the selection of a foreground, which is directly attended to, and a background, which lies at the periphery. In this manner, the interpretation of events and arising subsequent actions are filtered through an ongoing process where theme and horizon

oscillate dependent upon one's purposes; e.g. "why I turn to 'this' rather than to 'that' at 'this' time in my life, in the course of 'this' action" (Zaner, 1970). Schutz examines this interpretive process and frames it within a 'system of relevance', which is composed of three components.

First, topical relevance involves an orientation to a *theme* of an activity or a *topic* of thought in, what is most usually, a taken-for-granted manner and with it only becoming an issue when "*an unfamiliar experience imposes itself upon us by its very unfamiliarity*" (ibid: 28, italics in original). Second, interpretive relevance is the process of developing an understanding of a topic and is made up of a participant's background knowledge. This interpretative process occurs "*for the task of recognizing and interpreting the actually experienced segment of the world*" (ibid: 37, italics in original). Interpretative relevance enables participants to create coherent contexts by transforming what are initially unstructurized common fields (e.g. phenomena as it exists) through the selection of a foreground and background, thus attributing appropriate meanings to objects and activities for the purposes at hand. Finally, motivational relevance refers to the biographical situation of a participant and includes the participants goals (e.g. in order to achieve x, y, z) and reasons for action (e.g. because of x,y,z and associated historical reasons for an action).

I would argue that each of these relevances are examined in detail, whether implicitly or explicitly, when conducting EM-informed ethnographic fieldwork. Additionally, these relevances may be examined during prototype evaluation sessions where participants have an opportunity to assess a system according to its familiarity (or recognizability). Here again, Schutz (1970) provides a deeper understanding of the concept of familiarity. Alongside a 'system of relevance', Schutz introduced the notion of 'degrees of familiarity', employed by people for the purposes of both recognizing and orienting to objects in the world. Familiarity is defined as "*being sufficiently conversant with an object of our experience for the actual purpose at hand*" (Schutz, 1970: 61, italics in original). In this way, familiarity enables participants to orient to objects that support skilled practices and activities in an unproblematic manner through their learned meaningfulness. Because they are *already known*, familiar objects are taken-for-granted:

Being already known (so far as my then prevailing purposes required, at least), it is to that extent beyond question and as such it is *taken for granted*, an element of the now unquestioned world (Schutz, 1970: 61, italics in original).

This taken-for-grantedness places the object or activity in a position of needing no further analysis or explanation and can be oriented to from within the participants' natural attitude; in an ordinary, unquestioned manner. From within this taken-for-granted perspective Schutz (1970: 27) identifies degrees of familiarity:

same - previously perceived objects that are recognized.

similar - like those that are identifiable and recognizable.

unfamiliar - possessing a unique individuality.

determinable - known as merely belonging to a familiar type.

habits - referred to in order to recognize, identify and choose actual experiences.

stock of knowledge - acquired in everyday life and within professional settings.

Using Schutz's degrees of familiarity, the unfamiliar may now be described as the ways in which it does not fit participants' expectations and so, the unfamiliar may become problematic. Only when such issues arise, when "*an unfamiliar experience imposes itself upon us by its very unfamiliarity*" (ibid: 28, italics in original), then it takes on what Schutz refers to as 'imposed relevance'. In this way, the concept of familiarity is tightly coupled to the notion of relevance.

These two concepts may inform the development of a sociologically-informed prototype evaluation technique, especially, when used to identify imposed relevances. That is, when the taken-for-granted becomes disrupted and when the focus of attention turns to the unexpected. In this way, unfamiliarity imposes new topical relevances on to a situation by interrupting or modifying the taken-for-granted idealizations of "and so on" and "again and again" (ibid). Taken together and for the purposes of software design, the notions of relevance and familiarity may be referred to when conducting EM-informed fieldwork and prototype evaluation in order to understand the ways in which participants orient to prototypes as both useful (*relevant*) and usable (*familiar*) tools. In this way, the concepts of relevance and familiarity have implicitly informed the analytical orientation of the CEF, in particular the ways in which a prototype may be recognizably relevant to participants as they perform

routine work activities while using it. However, gaining access to the details of these relevances as they are interactionally achieved requires the tools of CA and IA and the final section of this chapter will provide an explanation of the Contextual Evaluation Framework (CEF) as it was initially developed for use in the case studies for this thesis.

5.4 The contextual evaluation framework

The purpose of any prototype evaluation technique is to assess the usefulness and usability, hence the *relevance & familiarity*, of computing systems. Another function is to allow other researchers and practitioners to use the technique by providing guidelines for its implementation. Many CSCW researchers have discussed the difficulties of evaluating multi-user systems (cf. Pinelle & Gutwin, 2000) including difficulties in conducting remote evaluations with distributed users (Neale *et al*, 2004), the time needed to conduct associated fieldwork, and the impossibility of creating laboratory conditions that reflect the actual social organisation of work (Grudin, 1988). The approaches discussed in chapter 3, cooperative prototyping, situated evaluation and quasi-naturalistic experiments all seek to address these difficulties as discussed in that chapter. However, a systematic and rigorous approach to quasi-naturalistic prototype evaluation is currently not available and so it has not been widely adopted outside of a very specific academic research community.

The aim of this thesis is to address the absence of a systematic approach to quasi-naturalistic prototype evaluation which may be useful to a broader community such as requirements engineering practitioners, computer scientists and others not familiar with the details of sociological approaches such as EM/IA/CA. Such an aim is in line with the ways in which prototype evaluation approaches, particularly in HCI, have been successfully disseminated throughout the computer science research community - through the provision of guidelines. Likewise, the CEF is conceived of to be implemented in a similar manner. Informed by the literature in both CSCW and the social sciences the initial concept of the Contextual Evaluation Framework (CEF) has been formulated and a loose structure for its implementation developed.

With the basics of the CEF conceptualised, as a result of this and the preceding two chapters, the e-Research case studies provided an opportunity with which to shape the technique more concretely. In order to begin the process of developing a practical guide to quasi-naturalistic prototype evaluation, a matrix was developed and iterated over which was used as a guide in the analysis of video data. Within CSCW research the design and utilisation of matrices is not uncommon when introducing guidelines for one's own use and when making it available for use by others. For example, Suchman's (1987; 2007: 129) matrix provided the analytic framework with which to study the asymmetries in the resources available between 'intelligent' photocopiers and the people who use them. Also, Tang & Leifer (1988) develop a matrix used to conduct an analysis of design activities. Additionally, Urquijo *et al* (1993) present a matrix for conducting breakdown analysis within synchronous CSCW systems. Finally, the use of matrices has also appeared in EM-informed fieldwork (Martin *et al*, 2001), where it was used to identify patterns of interaction. These matrices function simultaneously as both analytical frameworks and heuristic tools that guide the analysis of participants' interaction within a setting of interest. The CEF matrix is intended to function in a similar manner with its use informed by EM/CA/IA in that the focus of analysis is on the overt behaviours of members as accountable, describable phenomena (Garfinkel, 1967) which are sequentially produced in and through activity and interaction. The matrix will appear within the context of each case study along with my developing reflections and conceptualisations of its use within actual e-Research projects.

In the subsequent three chapters the CEF will be examined as a framework that was iteratively tested and developed through its implementation in three e-Research projects case studies spanning academic disciplines from the humanities, the sciences, and the social sciences (§6,7,8 respectively). Through my experience of its use in the field, knowledge has been acquired regarding the kinds of interactions and issues the framework would need to address. Each chapter will describe how the CEF was used to evaluate e-Research prototypes for their support of substantive research practices. First, chapter 6 presents the use of CEF within the BVREH e-Humanities project case study. Through the use of video observation, both the *in situ* work practices of researchers and their evaluation of the prototype VRE are

framed within the CEF to assess the ways in which the system supported, constrained and transformed the practices and activities associated with deciphering digital images of ancient manuscripts. Second, chapter 7 presents the use of CEF within the UK-NEES e-Research project which developed scientific software to support synchronous and distributed experimentation thus creating new practices where none existed before. Third, chapter 8 explores the use of CEF in the NCeSS project as an approach with which to evaluate a VRE portal interface for its usability with no prior *in situ* fieldwork. The case study was conducted in this manner partly because of the rather broad remit of this aspect of the project, which was to deliver a ‘generic’ web portal for a broad spectrum of social scientists and partly to evaluate the usefulness of the CEF as a stand-alone evaluation framework. Finally, chapter 9 closes with practical guidelines that present the CEF as a robust technique for the evaluation of software prototypes. The guidelines are informed by knowledge gained from the CEF’s implementation within the three e-Research case studies.

Chapter 6:

Seeing through digital images to an analysis of ancient texts

6.1 Introduction

The Building a Virtual Research Environment for the Humanities (BVREH) case study was conducted over a 6 month period during a time when Classics scholars at the University of Oxford were involved in the analysis of an ancient Roman writing tablet, the Tolsum tablet. Whilst involved with this research they were also members of the BVREH project which was developing a proof-of-concept Virtual Research Environment (VRE) for humanities researchers. For this case study, interviews were conducted with project participants (classicists, software developers and the project manager) in order to understand both the e-Research project and the academic research domain. The main aspect of my study however, was to work through how the CEF might be used within the context of an actual e-Research software development project. In order to do this, I introduced the evaluation technique to the project manager and lead software developer and video recordings were taken of the classicists working with two technical interventions. In the first, enhanced digital images of the ancient tablet were projected onto a large screen allowing the classicists to work collaboratively over projected images of the manuscript rather than working with the tablet alone. In the second, the classicists used a prototype e-Research system to analyse the same ancient tablet. In the first intervention, the aim was to understand current work practices of the classicists as they worked towards a re-interpretation of the tablet and so, for the purposes of this project, was considered as a modified version of *in situ* fieldwork. In the second intervention, the understandings gained from this fieldwork were drawn upon to conduct an analysis of how well the prototype supported classicists work practices.

The chapter begins first with section 7.2, which provides background information to the project and an introduction to both the research domain and the purpose of the e-Research system. Next, section 7.3 provides an analysis of the work practices of researchers as they used a small-scale intervention where they worked collaboratively in the analysis of an ancient writing tablet using projected and enhanced digital images. Section 7.4 then provides analysis of a second intervention where classicists analysed the same digital images using a VRE prototype. Section 7.5 will summarize the requirements that emerged from the prototype evaluation session and will end with a discussion of how the CEF was utilised in this case study.

6.2 The study of ancient documents and manuscripts

The Building a Virtual Research Environment for the Humanities (BVREH) is an e-Research pilot project that has developed a proof-of-concept VRE for Humanities researchers²³. Its project members conducted initial requirements gathering activities from 2005-2006 and as a result a proof-of-concept prototype was developed which aims to address the needs of classics scholars. The pilot system studied for this thesis was the VRE for the Study of Ancient Documents (VRE-SDM) project, which ran from 2007 to 2009. The purpose of the system was to enhance collaboration amongst classicists by making available digitized images of ancient documents and manuscripts as well as providing tools for their analysis²⁴. The classicists involved in this project are from the fields of epigraphy (the study of ancient inscriptions) and papyrology (the study of ancient texts written on papyrus or writing-tablets).

As part of the project, an ancient Roman tablet, the Tolsum tablet, was digitized and used to populate the system. The Tolsum tablet was first translated by the scholar Vollgraff in 1917. Typical in this type of tablet, the front (Fig. 1, left) describes the purpose of the tablet (e.g. a loan, will, shopping list, etc.) and includes 10 lines of text. The back (Fig. 1, right) of the tablet includes the signatures of participants and witnesses, the date and consular name; it includes 6

²³ <http://bvreh.humanities.ox.ac.uk/>

²⁴ A related project, e-Science and Ancient Documents (eSAD), extends this work with technologies to support the interpretation of digitized manuscripts using letter recognition technology.

lines of text. In this study the focus is on the activities involved in analysing the front of the tablet. Although the marks on the tablet may be seen, nevertheless, their character is difficult to determine.

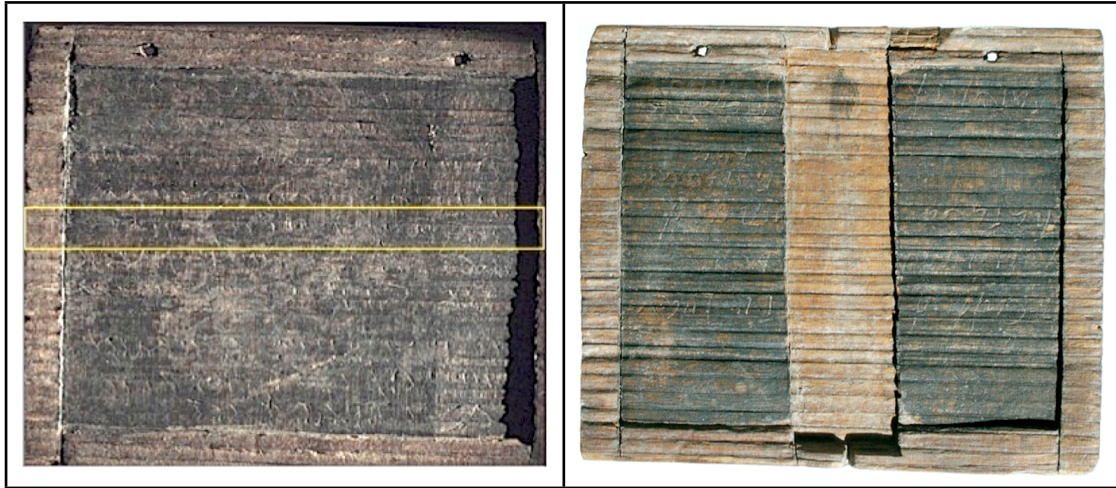


Figure 1. The Tolsum tablet front (left) and back (right).

The use of image processing in both epigraphy and papyrology is not that new. For the past twenty years infrared image processing techniques have been used to improve the legibility of ancient texts incised on wooden surfaces. More recently a new technique has been introduced where digital images of a physical object are taken using variable directional lighting, after which a series of algorithms are applied to the image that analyse the shadows cast across the surface at a low elevation (Molton *et al*, 2003). This enhances edge detection of the markings within the image, making it easier to identify the strokes that constitute each letter. In addition to this, the eSAD project, which ran, in part, parallel to VRE-SDM, is developing algorithms that remove additional 'noise' from an image. This will enhance images further through woodgrain removal and techniques that enable automated letter-stroke detection (Tarte, *forthcoming*). The classicists in the VRE-SDM project had access to both types of digital images; those that use variable directional lighting to improve edge detection and those that remove the woodgrain of the tablet which is thought to expose more of the letter 'underneath' the grain.

These two projects have worked closely together and its members are collocated in the same department in adjacent rooms on the same floor of the building. The eSAD project is

working towards the development of an Interpretation Support System (ISS) for the day-to-day reading of ancient documents and will include functionality that would keep a record of how each digital image of a document is transcribed and interpreted (Roued *et al*, 2009). This system will eventually require the implementation of capabilities (Terras, 2006) such as: databases of standard letter shapes; algorithms for the automatic detection of letters within digital images; letter frequency detection; and character lists aggregated from both ancient manuscripts and the research literature. Through these resources the eSAD system may eventually provide functionality that could estimate the statistical likelihood of certain patterns in the texts matching particular letters and words thus automating some of the analysis.

However, the VRE-SDM pilot system has had much more modest goals. It was developed to demonstrate a virtual workspace that would allow classicists to view, annotate and manipulate enhanced digitized documents and to eventually enable cross-searches of relevant datasets as well as to host virtual meetings where readings of texts and inscriptions could be made with colleagues across distributed locations. These design goals were identified through an initial requirements gathering process made up of in-depth interviews with the participants as they discussed the variety of artefacts and materials they use in their work and through attendance in graduate seminars in papyrology (cf. Bowman *et al*, 2008). Through these initial requirements gathering exercises preliminary features of the prototype were identified and developed into a proof-of-concept system. While the prototype was under development, two small-scale interventions were conducted in parallel. First, classicists worked collaboratively in the analysis of the Tolsum tablet using projected and enhanced digital images and second, the classicists analysed the same digital images of the tablet using the VRE prototype.

6.3 Fieldwork: working with projected digital images

In traditional practice, when working directly with a wooden tablet, classicists will attempt to make the text legible by tilting the tablet in various directions in order to steer the ways in which light and shadow are cast upon the incised surface. This tilting of the tablet in relation to lighting may make letters, words and lines of text legible (Bowman *et al*, 1997). Digital images, to some extent, can simulate this through the application of variable directional

lighting algorithms. Moreover, image processing can also reduce the visibility of the woodgrain of the tablet, making it easier to distinguish the marks that are made on it (Tarte, *forthcoming*). With the introduction of the first intervention, the classicists worked collaboratively over large projected images of the manuscript displayed on a screen, rather than working with the tablet alone. Multiple images of the same manuscript were also available, each differing in the direction of the light source, brightness and contrast. This first intervention provided access to the practices involved in analysing ancient manuscripts through the displays of talk and bodily conduct produced by the participants themselves as they scrutinized the projected digital images

Although the fieldwork undertaken in this intervention did not examine traditional practices it did provide access to those practices which may have been more difficult to reveal if the classicists had worked alone as they traditionally do. Observations of collaborative activities amongst a small group of classicists was video recorded and the video data subsequently analysed. This was the preferred approach over and above using a ‘think-aloud protocol’ which would have been generated if only one participant were working alone. This section presents three key activities identified through this fieldwork that were clarified through interviews and film elicitation techniques with the participants. First, section 6.3.1 examines the reliance on gesture as a communicative resource in letter identification. Second, section 6.3.2 reveals the ways in which a working hypothesis is substantiated through letter comparison and references to related source material, in this case a previous translation of the tablet. Finally, section 6.3.3 examines the use of paper mockups used to draw out more clearly what might lie behind the ambiguous and degraded markings on the tablet.

6.3.1 Analysing the scribe’s hand

In order to locate possible words and phrases, the classicists must first identify the ways in which ambiguous shapes within the manuscript might form letters. To achieve this, iterative readings of the entire text are worked through. These iterative readings transform the interpretation of the manuscript’s purpose and meaning overtime as letters and words are identified, justified as being correct or not, and finally agreed upon. For instance, at the

beginning of the project the manuscript was thought to be a soldier's will, however, this initial reading soon changed.

Letters within the manuscript are illegible or uncertain for many reasons, for example because of damage or because the tablet was used many times for more than one purpose with different writings layered on top of one another. Additionally, although some letters have standard forms, there may be great variety in how letters are written. This variation depends upon the period at which the text was produced; for example, there are significant differences between writing in the 1st century and the 4th century AD (and later). Furthermore, whilst following the conventions of the day, each scribe's personal writing style may vary widely. This makes the identification of both the generic writing style of the time and the particular style of the scribe crucial to the transcription of ancient texts.

As the scholars undertake their work they make use of multiple resources. For instance, in following fragments the classicists rely on tracing as a key resource in their analysis. As the researcher traces letter forms for the others, the group develop a common perceptual field (cf. Goodwin, 1994). Take for example the following fragment where Rupert, Axel and Colin, the three classicists participating in the VRE-SDM project, are sitting at a desk looking at paper copies of new images of the tablet alongside the original 1917 translation completed by the scholar Vollgraff. Each has his own area of expertise in the interpretation of the meanings of ancient texts; Axel specializes in Latin, Roman History and Papyrology (the study of ancient texts written on papyrus or writing-tablets). Rupert in Latin and Palaeography (the study of ancient handwriting) and Colin in Greek and the study of ancient inscriptions etched in stone. While they conduct their analysis of the tablet, high-resolution digital images of it are projected on to a large screen. In their discussion Rupert raises some doubts about Vollgraff's reading of the first two words in line 4 ('EMI TESTE')²⁵.

²⁵ For details regarding transcription notations see Appendix 2.

R: Actually, >I think in the next line< where they read 'TESTE' I was gunna read that as 'RES' (.3) Umm, you got 'E' 'M' (.3) And >thennit< loo:ks (.8) Ah, in line 5 (.3) Umm, and thennit loo:ks (.) >terribly like one< of these 'R's ↑again A: Just point to about where you're talking about	Rupert Colin Axel 
R: =Sorry, I'm at line [A: I'm lost R: I think it's ↓her:e A: =Yea, there	

Figure 2. Re-interpreting Vollgraff's 1917 translation.

Rupert's new reading of 'TESTE' as 'RES' is based upon the identification of a non-standard form of the letter 'R'. As seen in Figure 3, when compared with the standard form of the letter 'R' (Figure 3, on the left) the Tolsum tablet 'R' (Figure 3, on the right) is atypical.

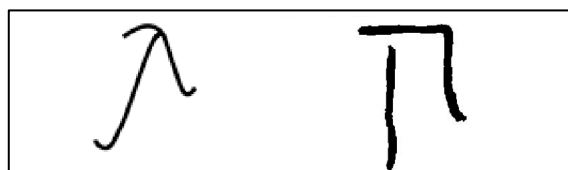


Figure 3. The standard (left) and the Tolsum 'R' (right).

Through the collaborative deciphering process the classicists begin to make sense of the writing produced on the tablet and repeatedly reassess each line of the tablet as they talk and gesture over the projected images. With each iterative pass the focus of analysis may change according to the arising concerns of the classicists as they revisit areas of the tablet in a “first-time-through-again” manner (cf. Garfinkel *et al*, 1981). For over 90 years the text was thought to be the sale of an ox between local residents and a Roman citizen. However, the new reading

of the second word in line 5 of the tablet does not support this. The transformation of the word ‘TESTE’, which translates as witness, into ‘RES’ - a thing, object or event - has implications for the purpose and hence the meaning of other words that appear in relation to it within the tablet.

In their doing of textual analysis, the classicists attempt to retrace the historical context and purpose of the tablet and thus orient to the digital image not simply as a representation of the manuscript, but as a way in which to access the text that resides within it. Significantly, the classicists have identified the letter ‘R’ found within the manuscript to be unique to the scribe’s hand and not following the conventions of the time. This newly identified, atypical letter-form has immediate implications for ways in which other letters and words within the tablet might be re-interpreted.

In addition, because of the vague and ambiguous quality of the text, line location cannot be taken-for-granted but rather is achieved through pointing to particular areas of the tablet. Axel, who is now looking at the projected image, indicates that he has difficulty locating the word in question; (‘Just point to about where you’re talking about’). In response, Rupert leaves his chair and approaches the screen. However, because of his close proximity to the projected image he takes a little time to locate the area of interest within the magnified line of text; (‘I’m at line, I think it’s ↓her:e’). Once he finds the line he begins to point and trace in more detail at the markings he sees within the image. In this fragment Rupert’s gestures reconstruct the sequence and motion of the scribe’s hand in the production of each letter.

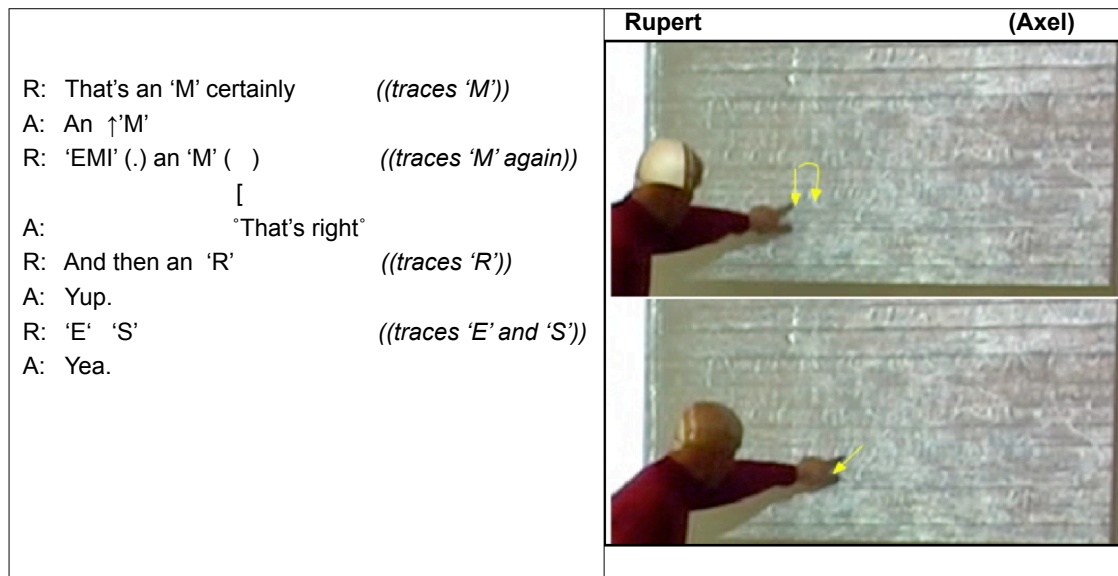


Figure 4. Identifying letters in the image.

Rupert, in some manner, is recreating the order in which the marks might have been produced which include the sequence and motions made with the metal stilus by the scribe. He relies upon the coordination of his talk and gesture to make visible, if only for a moment, the proposed reading of a section of the text. He then steps through the remaining letters (Fig. 5, #1-5) in the new word 'RES'.



Figure 5. Mimicking the scribe's hand (left image: the word 'RES'. right image: the word 'RES' with Rupert's tracings).

In addition to conducting an analysis of the text using digital images, the classicists' also try to understand the material characteristics of how a stilus moves through waxen wooden tablets. Indeed, when analysing tablets some classicists, including Rupert, are known to use wax tablet models in order to recreate the motion of the stilus on a piece of soft wax similar to the kind used for Roman writing. Like other kinds of text analysis, the classicists are attempting to

understand how the text was produced. For the classicists, the material qualities of letter production including the ways in which the scribe may have pressed the stylus across the tablet and the angles used to produce markings in the wax are integral to letter and word recognition. The classicists' gestures suggest how the stylus may have been used and reveals how crucial this activity is to their interpretive practices so that they may develop an understanding of the scribe's embodied production of letters. When working with projected images of the tablet the classicists draw upon an understanding of the complex embodied practices of writing to understand how the texts were produced.

In the next fragment, the classicists return to their analysis of line 5 where Rupert conducts an analysis of the way in which the scribe may have produced the letters:

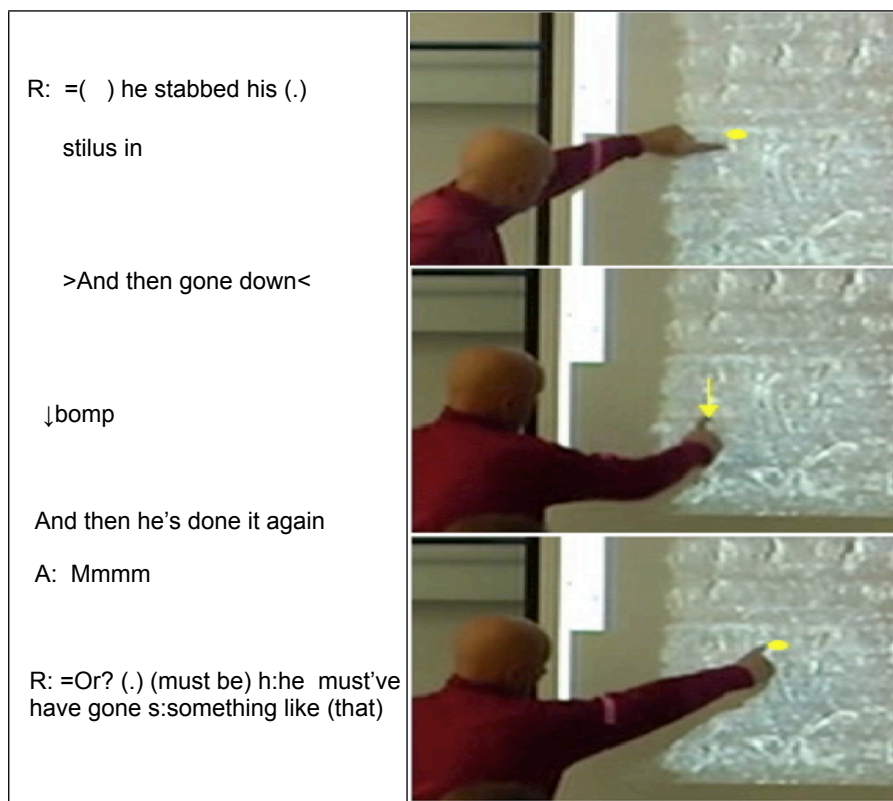


Figure 6. Animating the handling of the stylus.

In this fragment, Rupert establishes how the stylus might have been used and his downward gesture is articulated to mimic not only the movements of the stylus, but also the ways in which the stylus could have been handled to produce the marks by first stabbing the stylus into the wax and then pressing down to create the stroke. In this way, the scholars imagine the ways in

which a scribe may have used a stylus, thus assisting them in the identification of letter-forms within the tablet. Classicists often refer to the perceived motion of the stylus to justify their interpretation of the strokes that make up the shape of a letter. In this example, Rupert is making the claim that, based upon the motion of the stylus, the letters under scrutiny can be made recognizable as ‘E’ and ‘M’.

6.3.2 Comparing letters within the geography of the manuscript

In addition to tracing and gesture, the classicists also rely upon the geography of the tablet to compare letters in relation to one another. In this fragment, the researchers are trying to reconcile what might be multiple occurrences of the letter ‘E’ which have differing stroke formations. In line 4 of the tablet Vollgraff translated the line as ‘RITE UTI LICET BOVEM’, referring to an ox (‘BOVEM’). Hence, Vollgraff interpreted the text as the sale of an ox between a local resident and a Roman citizen. However, the classicists found a discrepancy in his translation where a strange letter form was identified by Vollgraff as an ‘E’.

Whilst the classicists rely upon their expertise to assess a scribe’s writing style when interpreting an ancient text, they also keep in mind the ways in which handwriting conventions evolve over time. Conventions in Roman cursive handwriting are divided into two periods: Old Roman Cursive (ORC), which was used from approximately the 1st century BC to the end of the 3rd century AD, and the later New Roman Cursive (NRC) from approximately the 3rd century to the 7th century AD. With regard to the Tolsum tablet, even given the variation in the literature about its dating, the classicists would expect the tablet to be written in ORC. In all but one instance, Vollgraff had interpreted a recurring combination of marks (Figure 7a) as an ‘E’. However, this was a highly unusual interpretation because the letter ‘E’ is normally identified by two vertical lines (7b). In one instance, on the back of the tablet, Vollgraff read the shape in 7a as an ‘A’ (which would usually be written as in 7c). This one exception of Vollgraff’s, interpreting the shape as ‘A’, led the classicists to reconsider all instances of that letter-form wherever it occurred.

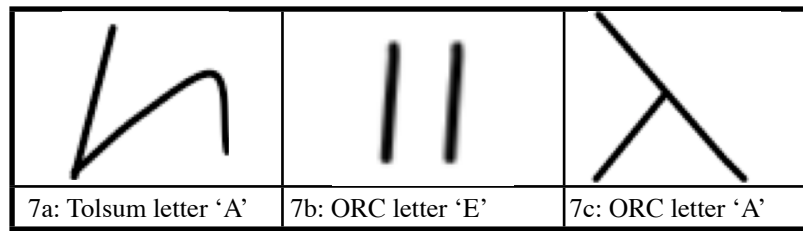


Figure 7. Identifying the 'A'.

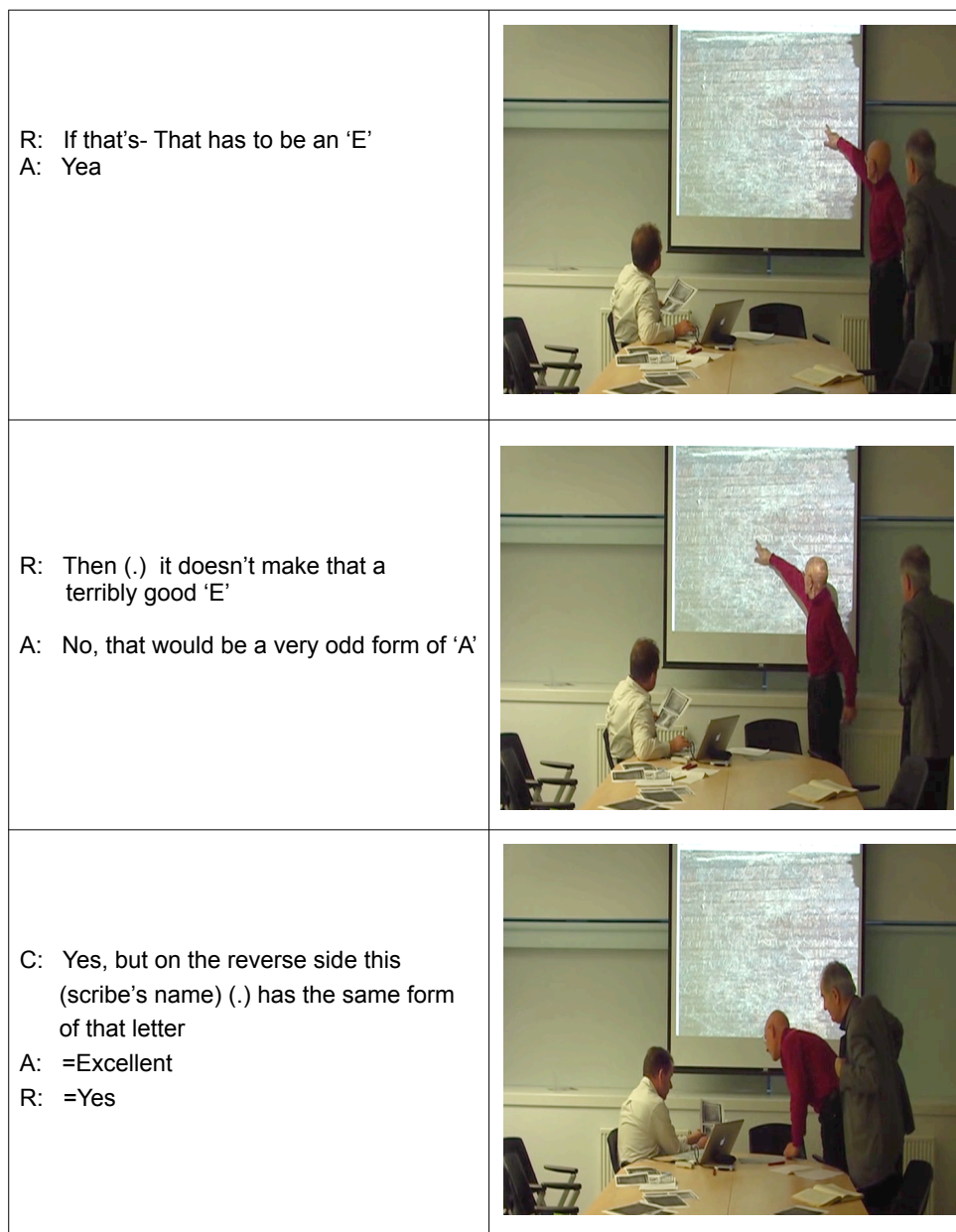


Figure 8. Comparing E's and A's.

Here, the classicists identify what might be two occurrences of the letter 'E' while simultaneously trying to resolve the ('very odd form of 'A)'). To resolve the discrepancy, Rupert works towards matching the 'E' that has previously been agreed upon as being typical

to other instances of the letter which Vollgraff also identified as ‘E’s. To make such a comparison, Rupert locates the certain letter (‘That has to be an ‘E’’) and secures agreement. He then compares it to the letter in question (‘Then it doesn’t make that a terribly good ‘E’’) and secures agreement for a second time. The puzzle of the letter ‘not being an ‘E’ but also a very odd ‘A’’, is resolved when Colin identifies an instance in which Vollgraff interpreted this letter-form in one instance, on the back of the tablet, as an ‘A’.

Here, the classicists rely upon more than the visual appearance of markings on a digital image in order to interpret letters. The work of analysing ancient manuscripts has been described as *not* a performance of “a letter-by-letter transcription like sleepwalkers, but a wakeful testing of possibilities in the light of other knowledge” (Bowman & Tomlin, 2005). Their ability to identify atypical letter-forms and anomalies within the manuscript depends upon their experience and expertise in identifying variations in handwriting along with their ability to read intuitively and imagine how letter combinations may form intelligible words. They also frequently reference other material such as prior translations, dictionaries of Roman names and historical documents, whilst examining a manuscript.

They also rely upon their knowledge of Latin, Roman History, Papyrology, Palaeography and other areas of study within Classics. With this background they are able to make informed guesses about the likely frequency of certain letters appearing in the text, as well as how particular types of text, such as wills, loans, sales receipts, shopping lists, etc. are formatted. For example, in this tablet, the opening line where the date and names appear, provides clues about the purpose of the text and therefore what its meaning might be. As Sharrock & Coulter (1998) suggest rather than considering the interpretation of images with regard to an individual’s perceptual capacities, viewers draw upon their “vernacular intelligibility”, gained through years of training and hands-on-experience, in this case, of ancient manuscripts. The classicists in this study make quite a surprising discovery in their re-reading of the tablet. Through their analysis they develop a distinctive interpretation not only of the text’s meaning, but also of the emergence of conventions in writing within the first century.

The fragments in section 6.3.1 revealed the ways in which the classicists identify strokes that form letters through tracing, motion and gesture. In this section, we can see that in order to reach agreement on working hypotheses letters may be identified through comparison. If a single letter, such as 'E' or 'A', can be identified in multiple locations within the manuscript it brings the classicists another step closer to the identification of similar letter-forms which will result ultimately in the identification of words. As the classicists explained in interviews, their analysis of ancient manuscripts is akin to 'trying to solve a crossword puzzle' and has been compared to detective work and a process of piecing together clues (Terras, 2006). In this way, the classicists work through deciphering the Tolsum tablet in an iterative manner; where the larger context of the manuscript is continually being referred to as the classicists bring to bear their knowledge of known letters, words and lines of text into the deciphering of unknown letters, words and lines of text.

6.3.3 The role of re-inscription in the interpretation of ancient texts

Interpreting texts relies upon being able to shift between different kinds of reading, ranging from the reading of letters and the markings that form them, to reading of the entire text with respect to the historical setting in which it was written. The classicists propose possible interpretations and test their hypotheses by constantly contextualising and re-contextualising the text at different levels of analysis. Through their readings and interpretations, the classicists arrive at a reconstruction of the text within the tablet. In the previous fragments the classicists re-enacted the production of letters, referenced additional resources such as prior translations and compared letters. In the following instance, instead of tracing on top of the projected image Rupert draws the proposed letter on a piece of paper (these paper drawings are referred to as mockups by the classicists). While drawing he expresses his initial doubts about the characteristics of the letter 'A', particularly in relation to differences in writing conventions between the 1st and 4th centuries:

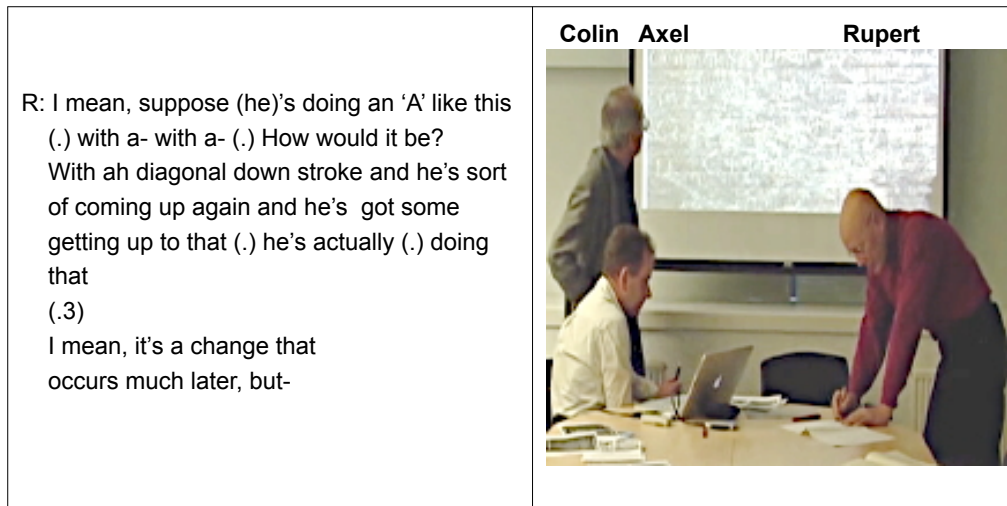


Figure 9. Reconstructing the text within the image.

The 'later' being referred to here is the historical context (e.g. the 4th century). Axel responds to Rupert's initial doubts with ('Well, it's only a late looking 'A' because we think it's a late looking 'A)'). Through his drawing, Rupert produces a re-inscription of the letter and attempts to reveal the hidden details of the object through his graphical representation of the text that resides within the image. Similar to Lynch's (1990) discussion of the ways in which the photographs of ribosomes in a scientific laboratory are interpreted by researchers as somehow imperfect because they are blurry, messy and ambiguous. In this same way, the classicists produce cleaner textual mockups of the ancient manuscript that both reduce and add information. In this way, it can be said that the classicists are attempting to see through the image to the text.

The classicists go on to produce other textual mock-ups of the ancient manuscript that aim to make letter-forms less ambiguous. These paper mock-ups are Latin transliterations of letters and words that reside within the digital image and culminate in a line-by-line re-inscription of the entire text (Fig. 10, left). Added to these mock-ups is commentary which accompanies the re-interpretation, thus providing further historical evidence and references to the literature.

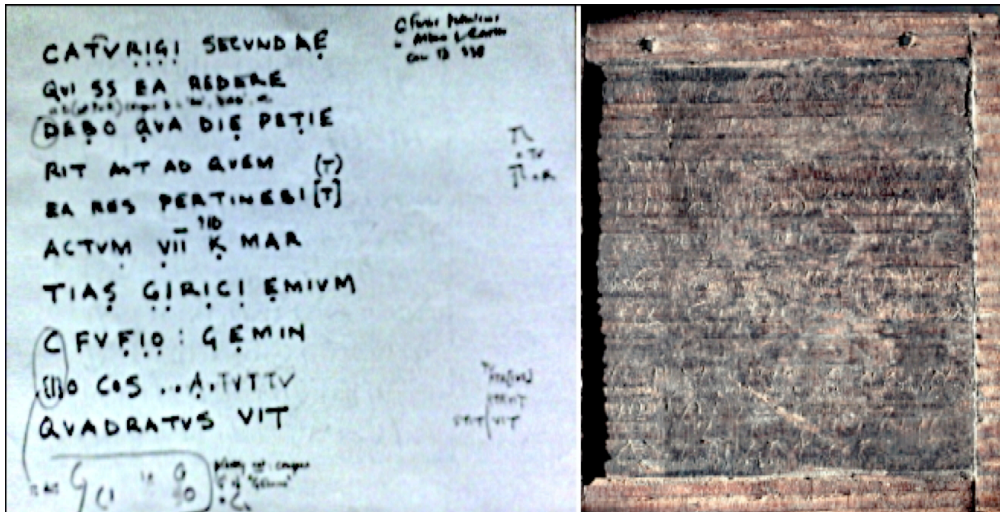


Figure 10. A working re-inscription (left) and the digital image (right).

The juxtaposition of a mock-up of the manuscript's text and the digital image of the manuscript above (Figure 10) demonstrates just how the image has been transformed from an ambiguous object to an ordered and organised piece of textual data. The result of such 'rendering practices' is that the digital image is clarified, completed, extended and identified (cf. Lynch, 1990). Through the activity of re-inscription the classicists arrive at a working re-interpretation of the text and in this way attempt to see through the image to the text.

Through projecting enhanced digital images onto a screen, classicists from different sub-fields are able to work together and discuss different readings of letters and words and their consequences for the interpretation of the meaning of the text. An epigraphist can propose an unexpected form of the letter 'A' for example, as well as discuss the ways in which it may have been written by a scribe using a stylus and as a result, raise issues that such a reading may have for the text as a whole. By working together whilst co-located, differing proposals can be immediately juxtaposed and compared with each other. In this case, the projection of the enhanced images allowed the classicists to work together to identify different readings of letter-forms and to re-interpret the text as a whole. This initial intervention led to a number of requirements that informed the design of a prototype system for VRE-SDM. These being; providing the ability to compare and annotate images, adjust contrast and brightness, as well as provision of a zoom-in feature to view specific areas and a zoom-out to view the entire manuscript.

6.4 Prototype evaluation: working with digital images in the VRE

Drawing from preliminary studies of the work of classicists detailed above and conducted within a broader requirements gathering process (cf. Bowman *et al*, 2008), a prototype system was designed that provided the project team with an opportunity to develop some of its core functionality. It was clear that the functionality the classicists needed was to provide a means with which to display and annotate digital images of ancient manuscripts. From the analysis of the classicists viewing the images *in situ* it also seemed desirable to provide a resource for collaborative pointing and drawing.

For the prototype evaluation session two of the scholars from the first fieldwork intervention, Colin and Axel, were asked to use the prototype system to conduct an analysis of the same ancient tablet using the same digital images that were projected on a large screen in the first intervention. The prototype VRE, at the time of its evaluation, consisted of a mouse and keyboard driven system providing a workspace allowing the classicists to select high-resolution digital images which could be manipulated (brightness & contrast) and viewed alongside other images (Fig. 11). The system could also magnify detailed areas of an image as well as provide an overhead view with which to see an image in its entirety. The system also included an annotation feature, making it possible for the classicists to comment upon or translate letters, words and phrases.

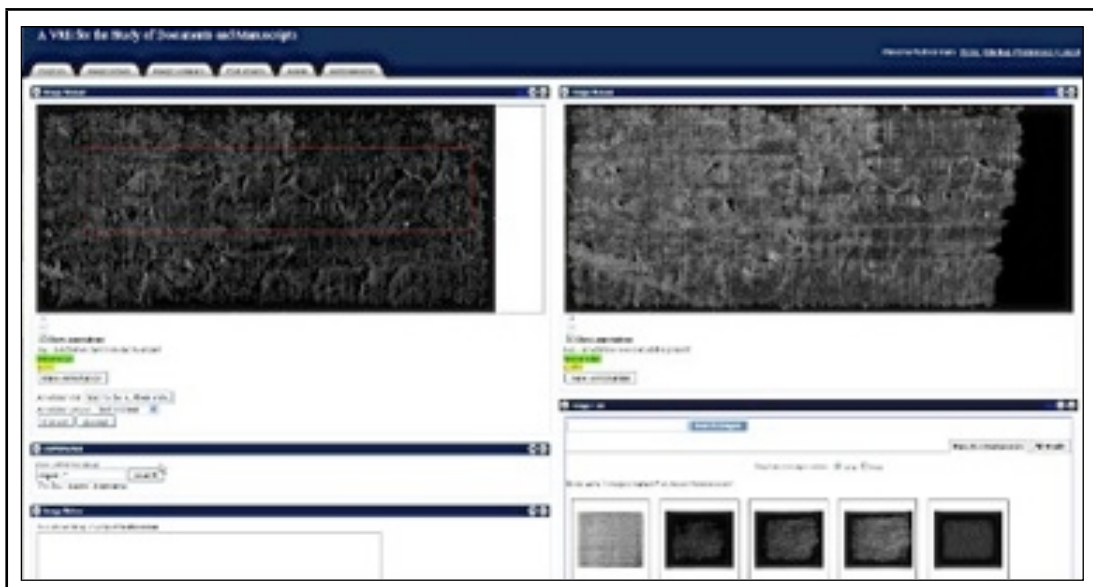


Figure 11: The VRE-SDM prototype system (Image Viewer and Annotation Workspace).

Before the prototype evaluation session began, a software engineer with the project provided a demonstration of the features and functionality available as well as general instructions for how to navigate through the prototype. The evaluation session was videotaped and conducted in a quasi-naturalistic setting where researchers worked in pairs as they performed the usual activities that occur within the deciphering process, as was discussed in the previous section. The evaluation session was conducted in order to assess how the features and functionality of the prototype may support, constrain or enhance researchers substantive research practices. They were not asked to walkthrough a specific set of tasks but rather, the purpose of the session was to allow the researchers to use the system in a way that would be most natural to them. For this reason, the participants were asked to conduct the ordinary activities associated with analysing the tablet. So in this case, the general activity given was to:

“Continue your analysis of the front of the Tolsum tablet using the prototype”

The primary activities included the annotation of four areas of the manuscript²⁶; in line 4 they examined the word ‘QUEM’ with its re-reading from ‘BOEUM’, and the word ‘AD’ with its re-reading from ‘IL’. The two areas in line 8 examined were the consular name (which indicates the date) with the word ‘GEMIN’, including an analysis of the letters just before it in the word ‘C FUF“O’. Through these activities the classicists used the features of the system that were available.

The participants viewed two images side-by-side, each with different contrast and brightness with the settings adjusted to identify letters, words and phrases. As they worked with the VRE, they annotated areas of the manuscript in order to both identify letters and words and provide justifications for their re-readings. When they performed their analysis of the tablet using the VRE they brought with them various mockups and notes from the previous *in situ* sessions. This material was referenced in order to recall and comment upon some of the work that they had done previously.

²⁶ Because the system was still in development some functionality could not be evaluated. This included the brightness & contrast controls and the ability to select a variety of images from the database for comparison. This restriction meant that only the image viewer and annotation feature were available for the evaluation.

This section presents three key activities identified as points of comparison where the prototype could be evaluated from the standpoint of how well it was aligned with the ways in which the participants themselves organise activities, share knowledge and conduct collaborative activities based upon prior fieldwork. This comparison provided a way in which to conduct a qualitative analysis between how these practices were supported, constrained or transformed when using the prototype. First, section 6.4.1 examines the use of gesture around the prototype. Second, section 6.4.2 reveals the ways in which orienting to the geography of the image is transformed. Third, section 6.4.3 examines re-inscription activities at the interface. Finally, section 6.4.4 examines the ways in which ancillary resources are integrated into activities performed around the system.

6.4.1 Re-inscription activities at the interface

In this first fragment, Axel is questioning the working translation of a letter made, not by the original transcriptionist, Vollgraff, but rather by his colleague, Rupert who is not in attendance at the prototype evaluation session. The questioning of his colleague's translation begins with a comment from Axel ('I just wonder about those 'F's'). He goes on to describe to Colin the shape as perhaps being more akin to an 'H'. Axel draws a mockup of the newly proposed letter-form.

A: I'll have to talk to Roger about it but I think it looks- it looks like (2.8) <i>((drawing on paper))</i> a little curl to the left on the tail and then m:ayb:e- maybe li(k)e that! =<i>((A shows C his paper mockup of the letter))</i> C: ↑Hmm	 Drawing  Displaying
---	--

Figure 12: Drawing mockup letters to examine possible letter-forms.

Axel uses paper as a resource with which to develop a counter-claim to Rupert's reading of the letter. This is achieved through the production of a mockup of the letter-form he sees residing within the digital image and is similar to the activity one performed *in situ* (see §6.3.3). When comparing this activity with the fieldwork data it is immediately apparent that there are no mechanisms available for participants to either draw or to display drawings to others using the prototype in its current iteration. Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	- Talk about a letter identified as 'F' perhaps being a different letter perhaps with the likeness of an 'H'. - This talk leads Axel to draw a mockup of how the letter may be constructed.	- One participant is drawing on a sheet of paper and then displays the drawing to the other participant.	<i>Available:</i> Image viewer <i>Constraint:</i> no drawing mechanism.	Provide a drawing mechanism and/or scanning device. The participants can only view the image and are unable to either draw or to upload a scanned version of drawings.
Cooperative work	- Axel says: 'I'll have to talk to Roger about it'.	- Axel displays his drawing to Colin after he has produced it.	<i>Constraint:</i> - There is no feature for participants to share drawings and mockups <i>Constraint:</i> - Communication is limited because there is no feature to communicate asynchronously.	Adding a forum area would be useful for enabling discussion about hypotheses regarding letter formation as well as for developing claims and counterclaims.

Table 1: CEF analysis - Drawing mockup letters to examine possible letter-forms.

This fragment reveals that a substantive work practice is not supported by the system in its current iteration. The activity of drawing along with the ability to display drawings to others remains external to the system. Once a drawing is composed participants orient to it and engage in further discussion. Additionally, Axel says that he would like to speak to Roger about this particular letter. The ability to communicate with colleagues asynchronous within the system is also unavailable.

Using the CEF framework data analysis is focused on what is being talked about and the activities that occur around talk-in-interaction. From this several suggestion for future iterations of the prototype can be made:

Drawing - provide the ability for researchers to produce digital re-inscriptions and freehand drawings of their interpretations of letter-forms. This may be achieved with a drawing tool embedded into the system or provision of a drawing table. Additionally, a scan and upload feature for paper-based drawings could be provided.

Public threads - provide mechanisms where discussion threads can be started. Allow linking to areas within a digital image under discussion.

Private messages - provide mechanisms where researchers can query each other and discuss ideas privately. Include the ability to link to specific areas of the digital image (e.g. the specific letter under investigation, in this instance an 'F' that may indeed be a letter with an 'H' likeness).

Hypothesis folders - provide a mechanism with which researchers can store and track their own claims and those made by colleagues. Allow users to associate specific digital images of a manuscript with specific claims that have been made about it.

Through discussions with the participants they identified an additional requirement associated with drawing:

Layers - the implementation of layers would allow each researcher to draw over the original translation layer and to add additional layers (e.g. Google maps image and satellite views). This would allow them to view the emerging identification of letters as transparencies placed over the object of interest, in this case, the Tolsum tablet.

Using the CEF we see how the analysis of a prototype evaluation session can be conducted which does not merely assess how the current features of the system may be usable but also how routine work practices may be supported, constrained or transformed.

6.4.2 Tracing and gesture activities around the prototype

Upon completing the drawing activity described above, Axel displays the paper mockup with the proposed letter-form to Colin, after which they both shift their orientation from the paper to the digital image displayed within the VRE workspace. This reorientation follows directly after the previous fragment where Axel describes the shape whilst drawing it ('a little curl on the left on the tail and then maybe, maybe, li(k)e that!'). However, after Axel completes the drawing we see in this next fragment that Colin requests further clarification in order to identify the shape when he asks ('So it starts here?') whilst pointing the cursor at the top of the letter. Axel responds by overlapping Colin's question with ('Do you know what I mean?') and begins to trace over the digital image, with his pen functioning as a pointer. While Axel traces he describes the shape of the letter in more detail than he did initially in the previous fragment with ('curling to the left and then back').

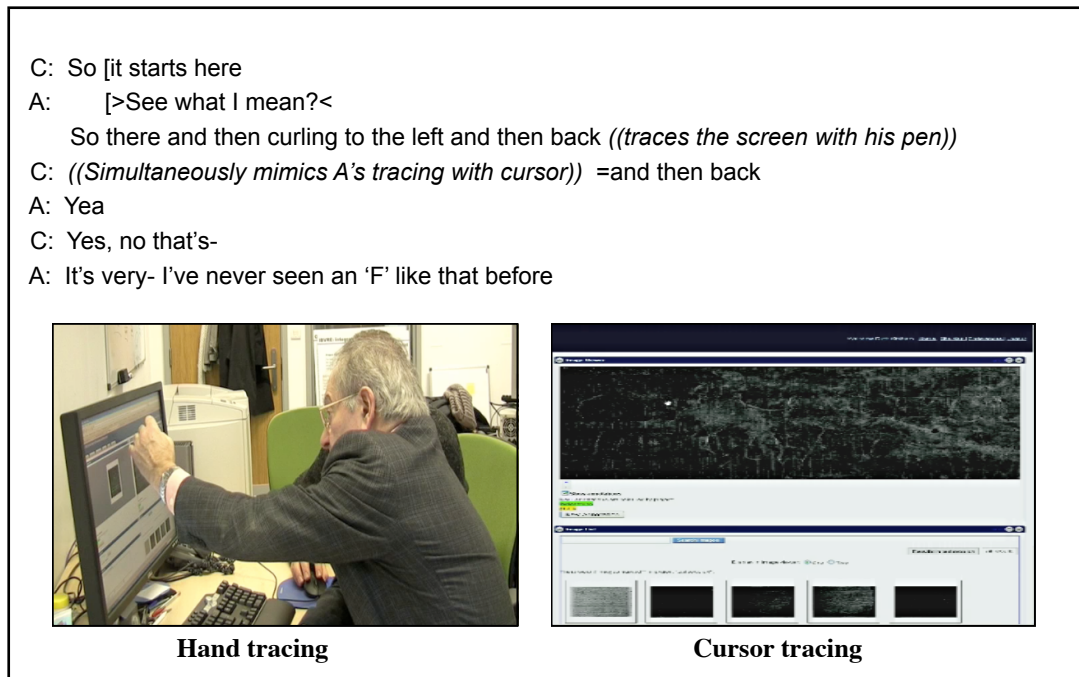


Figure 13: Tracing letters.

Here, Axel continues his analysis of Rupert's proposed letter 'F'. Whilst he traces over the digital image with his pen Colin mimics his colleague's motion simultaneously with the cursor tracing over the letter with the arrow icon. Colin's simultaneous tracing over the letter provides an indication to Axel that a shared perceptual field is emerging. As he follows the motion of the pen with the cursor Colin demonstrates to Axel that he can also identify the same shape. This simultaneous tracing produces an accountable action that confirms to both participants that they see the markings within the digital image in a similar manner. After tracing over the digital image with his pen Axel concludes with ('I've never seen an 'F' like that before'). Through the VRE the participants were able to successfully examine a letter-form previously identified as an 'F' and to then begin to suggest an alternative reading. This is achieved in a similar manner as when analysing the projected image *in situ* discussed in §6.3.2, through gesture and tracing. However, if this activity had been conducted in a distributed manner we see that the system does not offer features to support the production of tracing activities that would allow each participant to come to see the area of an image in a similar manner.

When comparing this activity with the fieldwork data it is apparent the participants are able to successfully identify letter-forms similar to when the image is projected. However, there are

no mechanisms that support tracing either synchronously or asynchronously; individually or collaboratively. Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	- Talk about the letter in question moves to a demonstration of how its shape may appear within the digital image. - The participants identify the letter by locating it within the digital image and then by tracing over it.	Tracing is performed 'outside' of the system using a pen.	<i>Available:</i> The cursor over the image. <i>Constraint:</i> No mechanism for recording and saving traces.	Provide a mechanism that can capture tracings produced with the cursor.
Cooperative work	One participant is tracing over the image with his pen. The other is tracing with the computer's mouse cursor.	<i>Constraint:</i> participants are unable to conduct either individual tracing which could be shared with others asynchronously or collaborative tracing activities which might occur synchronously over distributed locations.	<i>Constraint:</i> No collaborative tracing over digital images. <i>Constraint:</i> - Communication is limited because there is no feature to communicate asynchronously.	Add multiple cursors or some other mechanism to support collaborative tracing.

Table 2: CEF analysis - Tracing letters.

From this several suggestion for future iterations of the prototype can be made:

Tracing - provide the ability to trace over digital images and to save in a layer. This would allow researchers to both display and make counter claims about the letter-forms that appear within an image. Providing the ability to save tracings could allow for asynchronous discussion of hypotheses developed about specific letters.

Multiple tracing - provide the ability for each researcher to collaboratively trace through a multiple tracing mechanisms that could be used in real-time. Each trace could be identified with a different colour associated with each participant. This could provide a mechanism for distributed, synchronous demonstrations and discussion.

Through discussions with the participants another requirement associated with tracing was also identified.

Layers - the implementations of layers would enable participants to trace on to a layer above the original translation layer. A layer might function as a transparency image that could be placed on top of the digital image to produce both ‘mockup’ and ‘tracing’ layers.

Tracings could also be linked to conversations within both public threads and private messages.

Using the CEF, empirical evidence is made available which reveals the ways in which tracing over digital images is a substantive aspect of the collaborative analysis of ancient texts. Such fragments can be shown to software developers and other project stakeholders as a justification for adding new functionality to a system. It also reveals that adding mechanisms to achieve this practice would render the prototype more relevant and familiar to the classicists and thus align it more closely with the practices of deciphering ancient documents.

6.4.3 Referencing and juxtapositioning other resources

As the classicists continue with their analysis of the letter Axel would like to understand what may have prompted Rupert to make his initial reading of the letter when he states (‘anyway, I can discuss that with him’). But for now, Axel refers back to Rupert’s notes (‘actually, I don’t know what he said in his notes’). As he looks towards the notes and begins paging through them Colin asks if a specific drawing (‘that drawing he did on the tablet’) is in the pile of notes. Once the drawing is located Axel inspects the paper mockup (‘Yea, oh I see. He’s done it like that, but I don’t- I don’t entirely believe that either’). After this initial judgement is made both Axel and Colin reorient to the image displayed on the computer screen in order to compare the paper mockup with the digital image.

- 1 C: Do you- do you have that drawing that he did on the tablet?
[
2 A: Yea
I've got it here, Yes, somewhere ((A looks through paper notes))
3 C: 'Ah, there we are'
4 (3.3) ((A & C looking at R's drawing))
5 ↑Mmm (.3)
6 A: Uhh (.8)
7 Yea, oh I see. He's done it like that but I don't- I don't entirely believe that ↑either
8 (.8)
9 'No' (.) I don't like ↑that
10 C: So he's got that, but he's separating that out ((looking to image displayed on screen))
[
11 ((traces R's letter form))
12 A: Yea



Image 16.1 (L.3)



Image 16.2 (L.9)



Image 16.3 (L.10)

Transcript 14: Referencing other resources.

After consulting the paper mockup the letter is examined at the computer screen. Here, Colin describes two distinguishing features of the letter-form ('so he's got that, but he's separating that out'). While he says this he again, as in the previous fragment, simultaneously traces the letter-form with the cursor. Additionally, they refer back and forth from the paper mockup to the computer screen several times in order to locate on the digital image the same letter-form as is seen in the mockup.

The fragment reveals how the classicists coordinate the juxtapositioning of a paper re-inscription taken together with a digital image. After they both agree that the re-inscription of the letter-form may need revising, Axel remarks ('I think I'll have to go and negotiate with him about that'). An analysis of these activities using the CEF may lead to further requirements for the system:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	Referencing other resources, in this case, a paper mockup.	Juxtaposing the paper mock-up with a digital image on the computer screen.	<i>Available:</i> The mouse and cursor. <i>Constraint:</i> comparison activities are conducted 'outside' the system.	In order to compare mockups with digital images, researchers would need to scan and upload them.
Cooperative work	- Comparing a paper mockup with a digital image. - Commenting on a colleague's claims, Axel would like to discuss this particular letter-form with Rupert.	N/A	<i>Constraint:</i> participants are unable to capture within the system how and why a material artefact was referenced in relation to a digital image.	Add a mechanism for linking resources to digital images of the manuscript by line number (in this case the 'F' at this particular line is discussed in association with a specific paper mockup).

Table 3: CEF analysis - Referencing other resources.

As the evaluation continues, the system's accommodation of topically relevant work practice appears minimal. In addition to the suggested changes to the prototype discussed in the previous sections, this fragment demonstrates that the following requirements may be considered for future iterations:

Line numbers - provide researchers the ability to create line numbers that correspond to each line of text within a manuscript.

Linking to other resources - provide a mechanism where links can be made from an annotated region to relevant external resources, such as colleague's notes, dictionary entries (also as seen in §6.4.1).

Side-by-side comparisons - Not only of digital images of the manuscript, which is currently available, but also scanned drawings and other scans that may be uploaded to the system.

Again, the CEF provides empirical evidence in a format that is readily comprehensible for software developers. The combination of video and the CEF matrix can reveal what may otherwise be overlooked or 'seen-but-unnoticed' activities. For instance, the suggestion of

linking to other resources associated with the manuscript, such as mockups that include the ability to associate these with specific line numbers within the image.

6.4.4 Transforming practices with the overhead image view

This final fragment reveals one way in which practices may be enhanced or transformed through the prototype. Axel is finishing his final annotation of the word ‘FUF“O’ where he types into the system a comment about the ambiguous letters that have yet to be agreed upon:

First letter clearly is a ‘C’. ‘FU’ looks good (V’s reading). Three letters of which the middle one looks as if it must be ‘E’. Can just about reconcile with ‘FUFEO’ for ‘FUFIO’ but vowel change needs to be accepted. Looks as if we must have the name of only one of the consuls - why??.

Following Axel’s typing of the final annotation, Colin switches to the overhead view of the digital image and points to each annotation within it using the cursor.

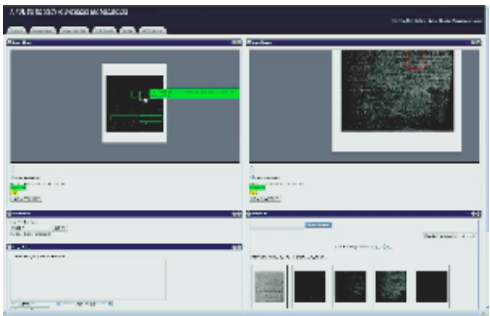
A: Right C: OK A: That'll do	Axel (Colin) 	
C: So now that will zoom out and this is sufficient from there A: Mmm C: to show why you- A: Yea C: How you got to that	((C using the scroll wheel of mouse to switch to overhead view)) ((C moves cursor to annotation of 'GEMIN' in line 8)) ((C moves cursor to 'FUFIO', 'AD' & 'QUEM'))	
A: OK, well that seems to work alright C: Yes		

Figure 15: Viewing the image as a whole.

As Colin hovers the cursor over the words - ‘GEMIN’ and ‘FUFIO’ and then over the words ‘AD’ and ‘QUEM’ he talks through how the overhead feature may be used to (‘show why you-how you got to that’). In other words, how each annotation relates to the other. The scholars may use this feature to view, at a glance, annotations within the manuscript in relation to each other and to the text around them (Figure 19). In this way, researchers can view gaps as well as relationships between words in an emerging re-translation. They can also view how a re-interpreted word in one area may have implications for the re-interpretation of words in other areas of the text.



Figure 19. Overhead view of the annotated digital image.

This fragment reveals one way in which the prototype may transform work practice in a manner that was not possible before. An analysis of this activity using the CEF includes:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	- The final annotation is typed into the system and the image is reduced in size (zooming out). - This enables the participants to examine an overhead view of all four annotations.	N/A	<i>Available:</i> Overhead view of the annotated image. <i>Transformed:</i> This activity appears to provide an entirely new route into information about the deciphering process that was not available before.	N/A

Cooperative work	- The overhead view may provide a quick overall sense of how each scholar is interpreting the manuscript.	N/A	<i>Transformed:</i> The overhead view may potentially benefit collaborative inquiry with the availability of this kind of visualisation tool.	N/A
-------------------------	---	-----	--	-----

Table 4: CEF analysis - Viewing the image as a whole.

The CEF reveals that this new feature allows users to zoom out and provides a high-level view of the annotations that have been made to a digital image. Here, the CEF is used to identify an enhancement and potential transformation of practice that may improve the ways in which research might be conducted. Of course it's transformative potential would need to be proven once the system has been implemented. However, the feature is highlighted through the CEF for its potential benefits which may encourage uptake and use of the system.

6.5 The use of the CEF in the BVREH project

The CEF analysis of the classicists' practices revealed just how complex the annotation functionality would need to be in order to properly address the needs of documentary scholars. Far from the basic, originally envisaged functionality of drawing a rectangular box around an area of an image and writing text that would appear once scrolled over the selected area, it became clear that scholars were working in a much more sophisticated manner and that the requirements for this part of the VRE needed far more consideration than originally planned. However, the complexity of trying to capture embodied action and material conduct that might be displayed across distributed sites is a non-trivial technological challenge and in most cases has been limited to reproducing particular features of gestures, like the movement of a cursor across an image or the real-time production of the mark of a pen (cf. Fraser *et al* , 2006). It was not possible to conduct a distributed evaluation of the prototype due to its level of technical maturity. Even so, this case study brought to the fore the ways in which the CEF could be used to evaluate a prototype intended to be used in a distributed manner within a co-located setting. The collaborative activities and tacit practices displayed during the co-located evaluation

session revealed that the prototype reinforced certain *asymmetries* (cf. Heath & Luff, 1992) that may interfere with the accomplishment of tacit practices and collaborative activities.

For instance, although the need for private workspaces and collaborative work areas were envisaged from the start, differing levels of notes and annotations such as ‘dialogues’ (private conversations), ‘drawing’ (marking areas of the image with hypothesised letter forms), ‘layers’ (overlying tracings, mockups and prior translations over the digital image) only came to light through observing the scholars at work and their discussions with each other. Support for these activities constitute the ordinary, taken-for-granted practices that the classicists rely upon in their analysis of ancient texts. Practices such as drawing, tracing and juxtaposing differing translations provide the classicists with the resources necessary to communicate ideas and to collaborate on a reinterpretation of the meaning of a text. Additionally, tracing and drawing are used to display to other collaborators developing interpretations and to generate a shared perceptual field. Tracing appears to afford scholars with the ability to actually *demonstrate* to each other how letters might be shaped. While drawing provides a platform with which to *convey* new hypotheses regarding the formulation of words as each letter shape is placed in relation to the other.

The CEF was incorporated into the BVREH project by working with the project manager and lead software developer to organise two small-scale technical interventions. In the first, the classicists worked collaboratively with enhanced digital images of the manuscript as they were projected on to a large screen thus providing access to the practices that classicists engage in when working alone as they traditionally do. However, even though classicists may traditionally ‘work alone’ the scholars nevertheless *do* collaborate, perhaps in a less obvious manner, through asynchronous communication and synchronous work-in-progress meetings in order to update each other on their progress, to review each others tentative interpretations, and to reach agreement about the content and meaning of ancient manuscripts. This is achieved through developing consensus of exactly what letters and words strokes and the combination of strokes within the manuscript might represent. In the first intervention, synchronous collaborative interactions between the scholars was observed as they were engaged in the real-

time deciphering and interpretation of the manuscript. In this way, their interpretive practices could be examined in detail for the purposes of supporting those practices through a virtual research environment. Understanding collaborative practices supports a key aim of the e-Research programme which is to provide a shared Internet-enabled resource hub of research applications and communication tools that encourages scholars to collaborate across distributed locations for the purposes of facilitating greater access to digital resources and increased global participation. The hope is that the availability of these systems will result in greater collaboration amongst distributed scholars and quicker turn-around time of research outputs and findings (cf. Atkins *et al*, 2003; Hey & Tefethen, 2003a). In the second intervention classicists used the prototype e-Research system to analyse different aspects of the same manuscript. These two interventions could then be compared so that the prototype could be evaluated from the standpoint of how well it was aligned with the ways in which the participants themselves organise activities, share knowledge and conduct collaborative activities.

The first intervention was conducted at the software development location. A meeting room was used which was transformed into as realistic a work setting as possible. This was achieved primarily by asking the classicists to bring all the materials they would normally access at their own desks. The classicists brought along reference books, drawing paper, laptops and their own wax tablet and stylus used for etching mock-ups with realistic materials. There were four sessions which ran for approximately 2-3 hours each. BVREH project members were directly engaged in these activities and observed the sessions in their entirety. Additionally, the sessions prompted questions from project members and the classicists' responses contributed to their understanding of the domain including the tools and processes relied upon to do the work of deciphering ancient manuscripts.

While the prototype was under development, and when its technical features were still somewhat unstable, project members would conduct demonstrations of its features and the classicists would provide their feedback. These sessions enabled the designers to incorporate changes based upon the classicists responses. After some time, when the prototype was in a much more stable iteration, a prototype evaluation session was conducted. The session was

also conducted at the software development location and again as realistic a setting as possible was provided and the classicists brought along with them the variety of material they had produced (such as mock-ups) and reference books.

The prototype evaluation criteria was drawn from the actual work practices observed *in situ* that the system is meant to support. Using the basic functionality that was available at the time, the classicists used the VRE to analyse the text primarily through the annotation tool. With it, they were able to identify and mark letters and words within a digital image and consider alternative letter-forms. The CEF evaluation proved useful for assessing the ways in which the prototype was relevant and familiar as a tool that would support the work of deciphering ancient manuscripts. It revealed that the current iteration of the prototype does not support researchers activities completely but rather, constrains the deciphering activity to the assignment of bounding boxes around entire letters or words. After the bounding boxes are placed on an area of an image, text may be added and attributed to that area. This feature supports an approximation of the complex activities that were observed during both the ethnographic fieldwork and prototype evaluation session.

In the BVREH case study, the CEF revealed three broad activities as aspects of substantive research practice; the use of gesture and tracing over letters and words, letter comparison activities using the geography of the manuscript, and re-inscription activities where mock-up drawings are produced to identify the form of candidate letters. Additionally, variation of letter-forms are also identified by analysing the scribe's hand and by incorporating knowledge and understanding of the historic conventions of writing. These findings revealed that the skilled practices of classicists, in their identification of letters and words and their interpretation of text, is not based upon their visual capacities alone. Rather, they recognise these shapes by drawing upon what Sharrock & Coulter (1998) refer to as 'vernacular intelligibility' where the recognisable features of work are conveyed socially through both language and, I would add, embodied practices (cf. Goodwin 1994). In the case of the classicists, their ability to identify atypical letter forms and anomalies within the manuscript depends upon the knowledge and expertise they have acquired. This knowledge enables them to recognise, beyond visual perception, the contents of an ancient manuscript. It includes

drawing upon their knowledge of the ways in which particular types of text, such as wills, loans, sales receipts, shopping lists, etc. are formatted. For instance, the wording of a document's opening line, where the date and names appear all provide clues to the classicists about the meaning and purpose of a text. It also includes experience with variations in handwriting, knowledge of Latin and the ability to read intuitively and imagine how letter combinations may form intelligible words.

The project team was very busy and so these results were communicated to them through brief, informal meetings with the project manager where video fragments were viewed and discussed. Additionally, a bullet-pointed list summary of requirements identified using CEF was also provided:

Line numbers - provide researchers the ability to create line numbers to correspond to each line of text within a digital image.

Hypothesis folders - provide a mechanism with which researchers can store and track claims made by both themselves and their colleagues. Allow users to associate specific digital images of a manuscript with specific claims that have been made.

Layers - the implementations of layers would enable users to view a) original translation layer, b) translations-in-progress layer c) counter claims layers, etc. (cf. google maps 'map', 'satellite', 'terrain', 'photos', 'videos' layers).

Private messages - provide mechanisms with which researchers can query each other and discuss ideas privately. Include the ability to link these to hypothesis folders and to public threads.

Public threads - provide mechanisms with which discussion threads can be started and can be linked to hypothesis folders.

Drawing - provide the ability for researchers to produce freehand drawings of letter shapes which can be uploaded and made available within the system.

Tracing - provide the ability to trace mock-up letters on to the digital image. This would enable researchers to comment on claims and to make counter claims.

Linking to other resources - provide a way in which links can be made from an annotated region to relevant external resources, such as colleague's notes, dictionary entries, etc.

These recommendations could not be incorporated into the prototype's design due to project time constraints. However, the results did contribute significantly to facilitating the

development of a sensitivity around the practices and activities within the research domain (cf. Dourish, 2006) and provided project members an opportunity to reflect upon how a future system might support these. For instance, digital tablets for drawing and tracing are now being considered and may meet both the drawing and tracing requirements. Additionally, because of the tight time constraints on the project it was not possible for the project members to evaluate the CEF on their own terms as if they were to utilise the technique independently. This is a well-known issue when developing a methodological thesis. However, I have reflected upon and can offer my own assessment of the evaluation framework.

The CEF was particularly useful in this case study in the evaluation of a prototype intended to be used in a distributed manner within a co-located setting. Asking pairs of participants to perform their routine research activities through the prototype revealed the ways in which both tacit practices and collaborative activities are accomplished and also identified asymmetries within the prototype that may interfere with these activities. Additionally, the CEF was used to facilitate collaborative interaction amongst traditionally solitary scholars through the paired use of two technological interventions. The CEF matrix was used as a way in which to organise and present findings derived from raw video data. The matrix was also used to summarize the activities that were being conducted in each fragment and to show the ways in which material artefacts and system landmarks were utilised to achieve *work practices* (the procedures and resources used to achieve activities) and *cooperative work* (the ways in which participants interact and orient towards one another in the accomplishment of activities). From this, the last column of the matrix provides suggestions for potential requirements.

The BVREH prototype was designed as a proof-of-concept system and gave the classicists an opportunity to begin to envisage how the VRE might support their substantive research practices. The circumstances of this case study did not afford the use of the CEF within a distributed setting but did provide insights into its usefulness as a technique for revealing tacit practices and collaborative activities amongst what are traditionally solitary scholars. The next chapter will look at the use of the CEF in the UK-NEES project where distributed prototype

evaluation was possible and will discuss the use of the CEF in a geographically distributed prototype evaluation.

Chapter 7:

Coordinating local and distributed earthquake experiments

7.1 Introduction

The previous chapter discussed the use of CEF in the BVREH project where the usefulness of the technique was demonstrated particularly for revealing the tacit practices and collaborative activities of traditionally solitary scholars. It was also found through the BVREH case study that a co-located evaluation session could provide insights into the ways in which distributed work may be supported, constrained or transformed as substantive practices are enacted using the system. In this chapter, unlike in the BVREH case study, the CEF is used in the UK-NEES project to conduct distributed prototype evaluations. Having the opportunity to conduct a geographically distributed prototype evaluation allowed me to consider how the CEF might be further refined to accommodate this more complex setting.

The UK Network for Earthquake Engineering Simulation (UK-NEES) case study was conducted over 10 months when structural engineering researchers were in the process of the design and testing of an e-Research proof-of-concept system that would link together three earthquake engineering laboratories at the universities of Bristol, Oxford and Cambridge. During this time, in-depth interviews were conducted with project participants (two co-PIs, a research associate and a doctoral research student) where they described the details of the project and the variety of artefacts and materials used in their work. Similar to the previous case study, I continued to work through how the CEF might be employed within the context of an e-Research software development project where domain researchers worked closely with software developers in the design of a system at the level of code. This is because scientists

need to retain a significant amount of control over the code they develop as part of their research process. For example, in order to conduct a hybrid earthquake experiment the researcher must include both a numerical model of a structure and a computationally produced workflow that will execute and record the experimental procedure. Earthquake experiments tightly couple the computer system with the experimental procedure and so domain researchers expect to be able to modify software features for their research purposes (cf. de la FLor & Jirotko, 2009).

For this case study, a series of video recordings of focused activities were gathered *in situ*, at the earthquake engineering laboratory. First, researchers were videotaped conducting local hybrid earthquake experiments. Second, structural engineers at two locations were observed as they collaboratively conducted distributed hybrid earthquake experiments using the prototype system. Using the same format discussed in the previous chapter, the aim of observing local hybrid earthquake experiments was to understand current work practices through *in situ* fieldwork. The understandings gained from this were drawn upon to conduct an analysis of how the prototype supported, constrained or transformed those practices.

The chapter begins first with section 7.2, providing background information to the project and an introduction to both the research domain and the purpose of the e-Research system. Next, section 7.3 will provide an analysis of the work practices of researchers as they conduct local laboratory experiments. Section 7.4 will discuss the prototype intervention where earthquake engineers conduct a distributed experiment. Finally, section 7.5 will summarize the requirements which emerged from the prototype evaluation session and will end with a discussion of how the CEF was utilised for this case study.

7.2 Laboratory practices in earthquake engineering

The UK Network for Earthquake Engineering Simulation (UK-NEES) project began in 2006 and ran for three-years. The purpose of this e-Research project was to build a proof-of-concept system that would link together three earthquake engineering laboratories at the universities of Bristol, Oxford and Cambridge. The original vision and design goals for the UK-NEES pilot

system²⁷ includes: facility upgrades to improve existing equipment, tele-participation allowing remote participants to observe experiments, tele-operation to include remote operation of certain facilities, open access to consistently formatted data and distributed hybrid testing that would enable synchronous and coupled operation of distributed experimental facilities in a common experiment²⁸. The project draws upon research from the US-NEES project which began in 1999 and that, in 2004, connected fifteen experimental facilities through a shared national network of open source collaborative tools, a centralised data repository, and earthquake simulation software. Similar technology is being developed as part of the UK-NEES project which began in 2006 and developed, in the first instance, a proof-of-concept prototype in 2009.

The relationship between scientific instruments and the technical methods of experimental science is one in which research questions are inextricably linked to the technologies used to observe, produce and capture data from phenomena (de Solla Price, 1984; Rosenberg, 1992). This coupling of science and technology has produced equipment and instruments that facilitate the extension of embodied perception (Lynch & Woolgar, 1990) enabling the acquisition of new kinds of data and new knowledge that were inaccessible prior to the instrument's appearance. The design of experimental instruments and improvements to existing ones are, in most cases, conceptualised and constructed by researchers working in the field within a specific research area (von Hippel, 1976). Because researchers have the in-depth knowledge necessary to recognise what new devices and processes might support and enhance their research; they are more successful than manufacturers in identifying requirements, developing design ideas, building prototypes, testing their applicability in the field and finally diffusing the innovation into the research domain. Described as a 'bottom-up' process, it is considered the most effective way of embedding technological innovations into research communities; through its repeated demonstrable usefulness and through word-of-mouth (ibid).

27 <http://www-civil.eng.ox.ac.uk/research/structdyn/presentations/uknees.html> (last accessed 03 August, 2010).

28 The goals of the project that were stated in the high-level vision were reduced as technical and time constraints emerged so that only the facilities upgrade and development of a proof-of-concept system to support distributed hybrid testing (DHT) were pursued in the first phase. The vision for tele-presence was found to be incompatible with the necessity to conserve all bandwidth for use during experiments; where it is more important for researchers to produce good results than to have the ability to see what is taking place in remote labs.

This case study will present how scientists develop a prototype for a new scientific instrument or meta-instrument that links laboratory equipment across three earthquake simulation laboratories. By linking these laboratories, local capabilities are extended allowing researchers at one location to gain access to and share the resources of other geographically distributed laboratories so that they may conduct joint experiments. Having access to this new meta-instrument may extend the kinds of phenomena available for observation and measurement.

Earthquake engineers study the response of structures subject to seismic action for the purposes of developing technologies that improve seismic resistance. In this domain there remains a pressing need for large-scale experimentation that captures non-linear and rate dependent behaviour of structural components in order to give confidence in new anti-seismic technologies. Various experimental techniques exist, the simplest of which apply ramp or sinusoidal displacements on the test specimen. However, the more realistic the reproduction of true seismic loading on a specimen, the greater the confidence that the findings represent true seismic performance.

Seismic testing of structures is often carried out at large shaking table facilities, where the structural system, or more often a scaled representation, is mounted upon a table and can be actuated in real-time (i.e. where real earthquake duration equals the simulation time) to simulate seismic ground motion. While shaking table testing is arguably the most realistic technique, capturing both global and local seismic response physically, it is expensive to implement at large scale. Additionally, smaller scale tests can be difficult to extrapolate to prototype scale and tests can be difficult to control accurately.

An alternative seismic testing technique, hybrid testing, bridges the gap between fully physical and fully numerical testing. This is achieved by separating the structure of interest into numerical and physical parts or substructures. Here, only key parts of the structure for example, a novel seismic-resistant device and its attachments are tested physically and coupled to an adequately representative numerical model of the rest of the structure (not the main focus of the test). The substructures are coupled together via a transfer system of actuators and

sensors with an earthquake simulated through the entire structural system. The test proceeds in time-stepping fashion and is continuously loaded with typically, interface displacements calculated and derived from the numerical part, applied to the physical part and interface forces measured and fed back at the end of each time-step to allow calculation of the next interface displacement. In this manner the complete structure's seismic response is determined while focusing experimentally on the performance of the physical part, the technology under development (Fig. 1).

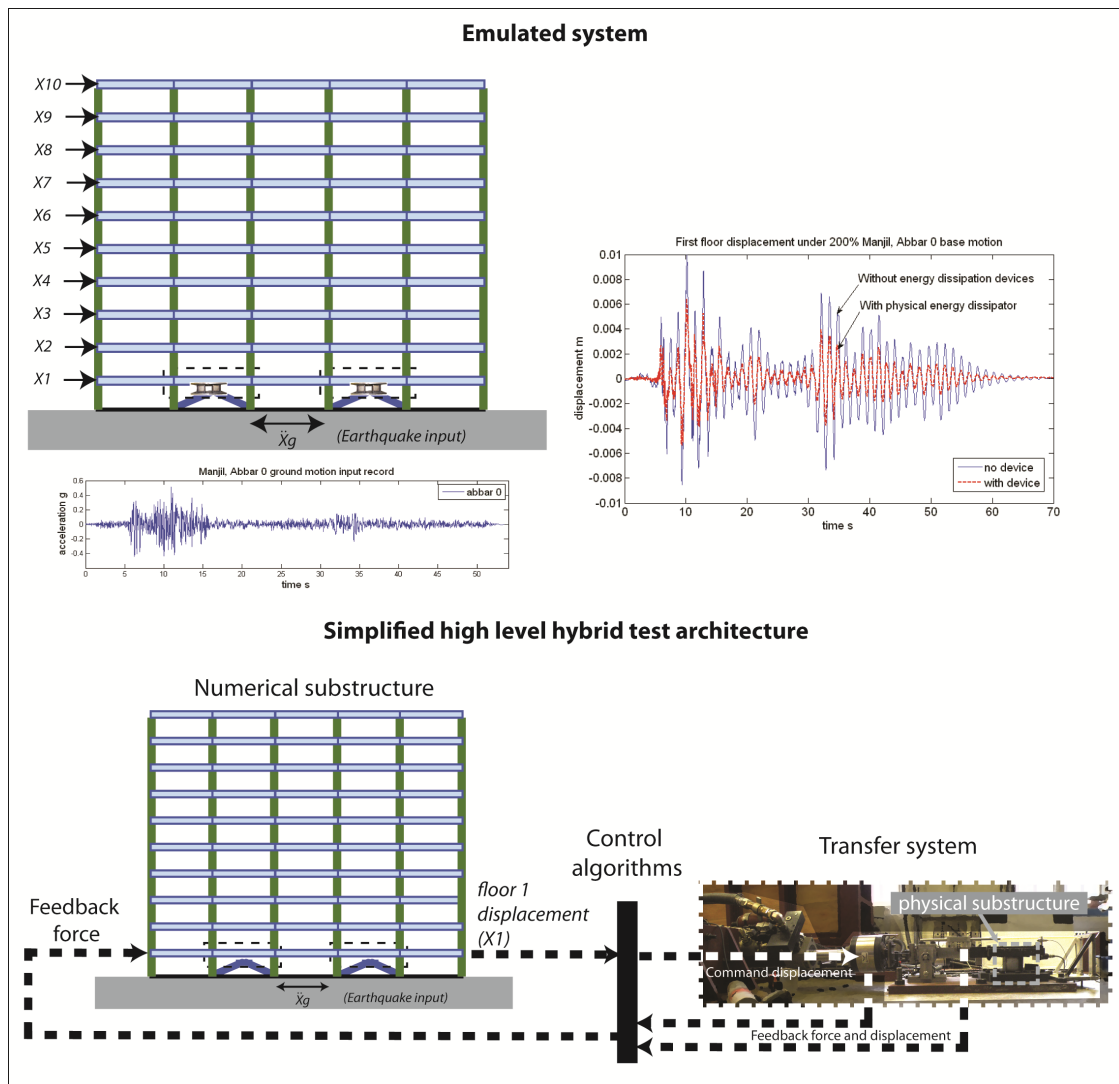


Fig. 1. Hybrid test: a numerical model of a ten-story building coupled with physical substructures (row: XI , top left) and the modelled structure (bottom left) are measured and earthquake forces applied.

Hybrid testing has advanced in recent years from original pseudo-dynamic tests (Mahin & Shing, 1985), where experiments take place over an extended time-scale. With a need to capture rate dependent behaviour, real-time tests have been developed that use dynamic

actuators (Blakeborough *et al*, 2001). Hybrid testing is a more cost effective seismic testing technique and though requiring sophisticated control, is less effected by actuation capacity limits, since only the part of the structure of main interest is tested physically.

In answer to the capacity limits and scaling issues associated with shaking tables, the E-defense table was commissioned in Miki, Japan in 2005. As the largest shaking table in the world, it can subject an entire medium size building to realistic earthquake shaking. However, with a growing requirement for testing ever larger and complex structural systems at full scale (Nakashima, 2008) even E-defense has significant capacity and cost limits in testing larger structures such as represented by Fig 1. As an alternative, hybrid testing particularly in real-time is capable of simulating the seismic response of larger structures. However, in practice, as the number and size of the physical sub-structures increase, it can quickly exceed the testing capability of any one lab. Taking advantage of the principles of hybrid testing and the expansion of networking infrastructure, a new concept, distributed hybrid testing (DHT), can extend the local hybrid testing technique to address many of the limitations within local laboratory settings.

7.2.1 Distributed hybrid testing

DHT aims to address local laboratory capacity issues by allowing the numerical portion of an experiment to be coupled to multiple geographically distributed substructures over a computer network, thus forming a much larger structural system. DHT is a new and novel approach to earthquake experimentation that may eventually make it possible for researchers to take advantage of complementary facilities and specialist technical expertise across distributed locations so that experiments may be carried out that have either never before been possible or are larger than has previously been possible (Ojaghi *et al*, 2007). While extended time scale tests have been demonstrated, particularly by US-NEES (Spencer *et al*, 2004) the technique introduces new and prohibitive obstacles that remained to be solved. One of the main research themes of UK-NEES and the focus of work at the Oxford node has been to develop new algorithms and software enabling distributed hybrid testing at rates up to and including real-time. The project successfully demonstrated its capabilities in a series of two and three site

distributed experiments conducted in 2009. An example of a three site prototype evaluation test is shown in Fig. 2 (adapted from Ojaghi *et al*, 2010).

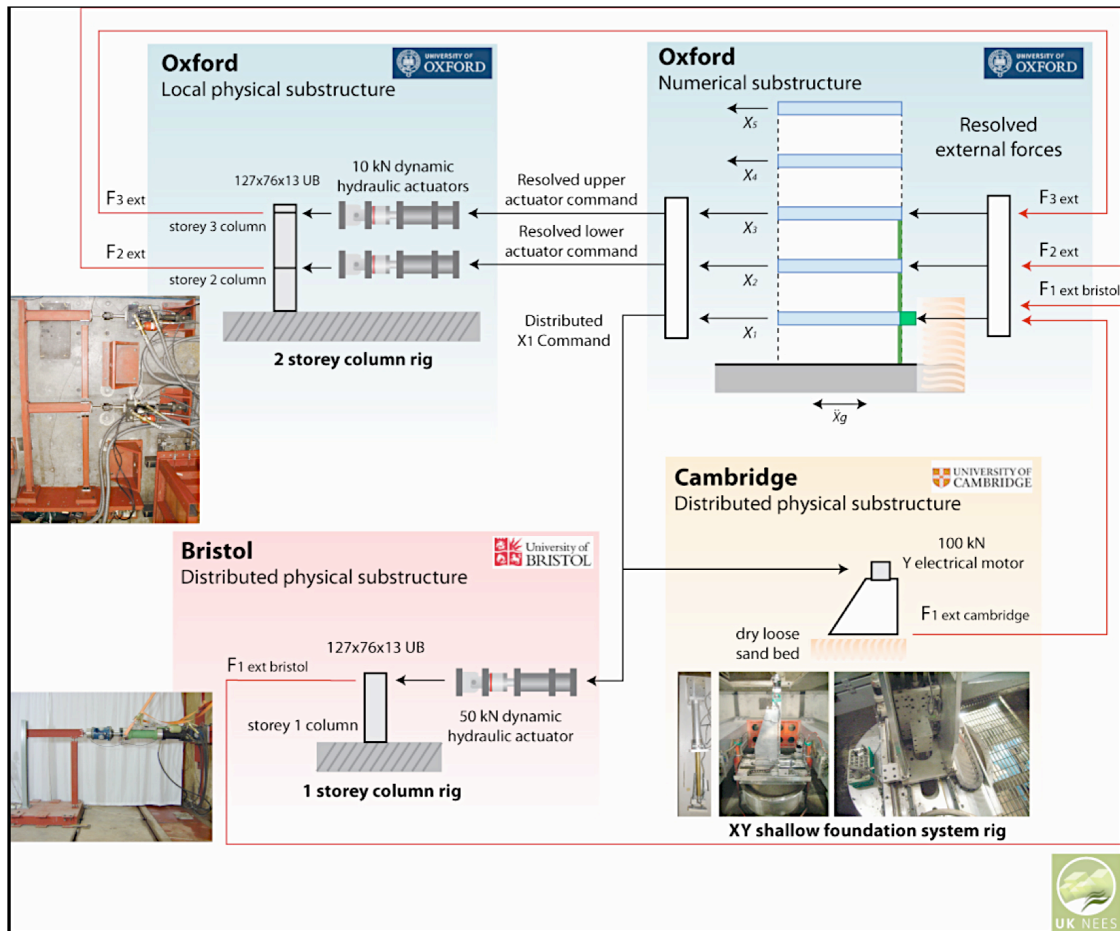


Fig 2. Distributed hybrid test: local physical substructure and numerical model at the Oxford node (top), distributed column at the Bristol node (left), distributed foundation rig at the Cambridge node (right).

Hybrid testing offers a synergy between numerical and physical simulation by combining both numerical and physical components as substructures of an experiment. In this way, a physical model is used to examine the critical structural component of interest and is linked to a numerical model of the remainder of the structure. Distributed hybrid testing extends this procedure by making it possible for researchers to access equipment they may not have in their own labs by providing access to remote facilities. Each of the three laboratories involved in the UK-NEES project contributes its particular area of expertise and accompanying scientific equipment which includes:

Bristol node - a large shaking table with a three-meter square, high load capacity platform. On this platform models can be mounted and shaken realistically using ground motion data captured from previous earthquakes.

Cambridge node - a large centrifuge used in geotechnical engineering research that investigates foundation behaviour. This area of research focuses on the interaction between subsurface conditions, physical materials and soil stresses.

Oxford node - the smallest and most flexible of the three. The lab has highly dynamic actuators and flexible rig set-up capabilities providing the ability to run faster experiments because they do not rely on big equipment but instead focus on the analysis of computationally modelled structures. Most significantly, the Oxford lab is at the forefront of hybrid testing and is the lead node of the UK-NEES project.

In order to extend each laboratory's scientific instruments so that they may be used in a single, distributed experiment it is necessary to be able to both produce phenomena and capture data across remote locations. One of the most challenging aspects of the project is designing a system where scientific instruments within each laboratory can be extended so that they may be used in a single meta-instrument. For this to be possible it becomes necessary to have the ability to produce distributed phenomena that are linked so that the displacements of a physical specimen at one laboratory effect both the numerical model and the displacements of a physical specimen at another laboratory. Additionally, it must be possible to both capture and aggregate the data produced across each location. Although there have been many technical challenges involved in linking together distributed laboratory facilities, such as compensating for time delays over the Internet, and connecting legacy systems, embedding this new infrastructure into research settings has also exposed challenges to the management, including communication and coordination, of distributed experiments as they occur, often in real-time. This chapter will discuss the ways in which the CEF was used to examine the management of experiments, in particular, their coordination and control.

7.3 Fieldwork: local laboratory experiments

Earthquake experiments are stressful not only because researchers are working in a safety critical setting but, also because test procedures can be complicated and unforgiving of error. In such an environment, an incorrect test control or test procedure will destroy unique or

expensive physical specimens. Add to this the likelihood of loud mechanical equipment, high temperatures and the heavy smell of hydraulic fumes. The laboratory in Oxford (Fig. 3) includes a central experimental space cluttered with equipment utilised during experiments such as braces for holding physical specimens, steel columns and actuators which apply forces to specimens. The researcher is located at an elevated area where he has access to an overhead view of the entire experimental space. Additionally, he uses three computers (one for the numerical model, one for the software that controls the actuators and one for conducting general tasks and is connected to the Internet). The elevated area also includes a large wall-mounted display which the researcher uses to project a close-up live video feed of the physical specimen in order to observe its behaviour more closely.

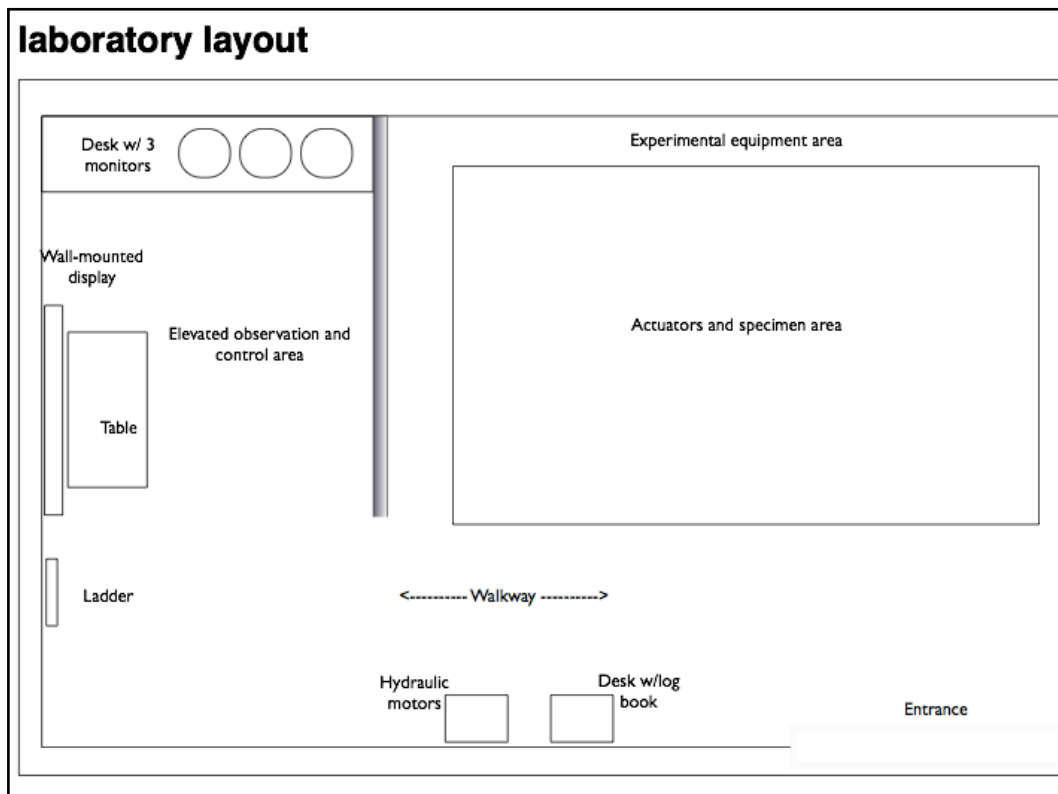


Fig 3. Laboratory layout at the lead node of the UK-NEES project in Oxford.

The main areas of interest in the laboratory during an experiment are highlighted in Fig.4 and include the computer control and display area (L-1), the experimental area (L-2), with a close-up view of the physical specimen (L-2 inset: 1), the actuator (L-2 inset: 2) and video camera (L-2 inset: 3). The wall mounted display (L-3) projects a live video feed from the camera

positioned in front of the specimen. As Lynch (1991) has noted, conducting ethnographic work in a scientific laboratory presents challenges:

Beyond questions on the intelligibility of technical references was the more massive fact that laboratory work seldom provides a very interesting spectacle. For long periods of time one or a few individuals would sit silently, tapping at the keyboard of a computer terminal or scribbling notes while viewing data displays (p.53).

I too found this to be the case when entering into the earthquake laboratory. The technical language within the domain is complex, additionally, the researcher spends a majority of his time looking at computer displays that depict graphs showing actuator inputs and sensor outputs (see ‘Thisisanexperiment.mov’ in Appendix 3). Even so, during *in situ* fieldwork it became immediately apparent that the spatial and ambient properties of the setting were crucial for the start, monitoring and termination of experiments. The purpose of the fieldwork was to observe and understand the traditional practices of earthquake engineers with a particular focus on the resources they rely upon in the coordination and control of local experiments. This remained the focus in order to understand how these activities might be supported through the prototype so that it could enable jointly coordinated distributed experiments.

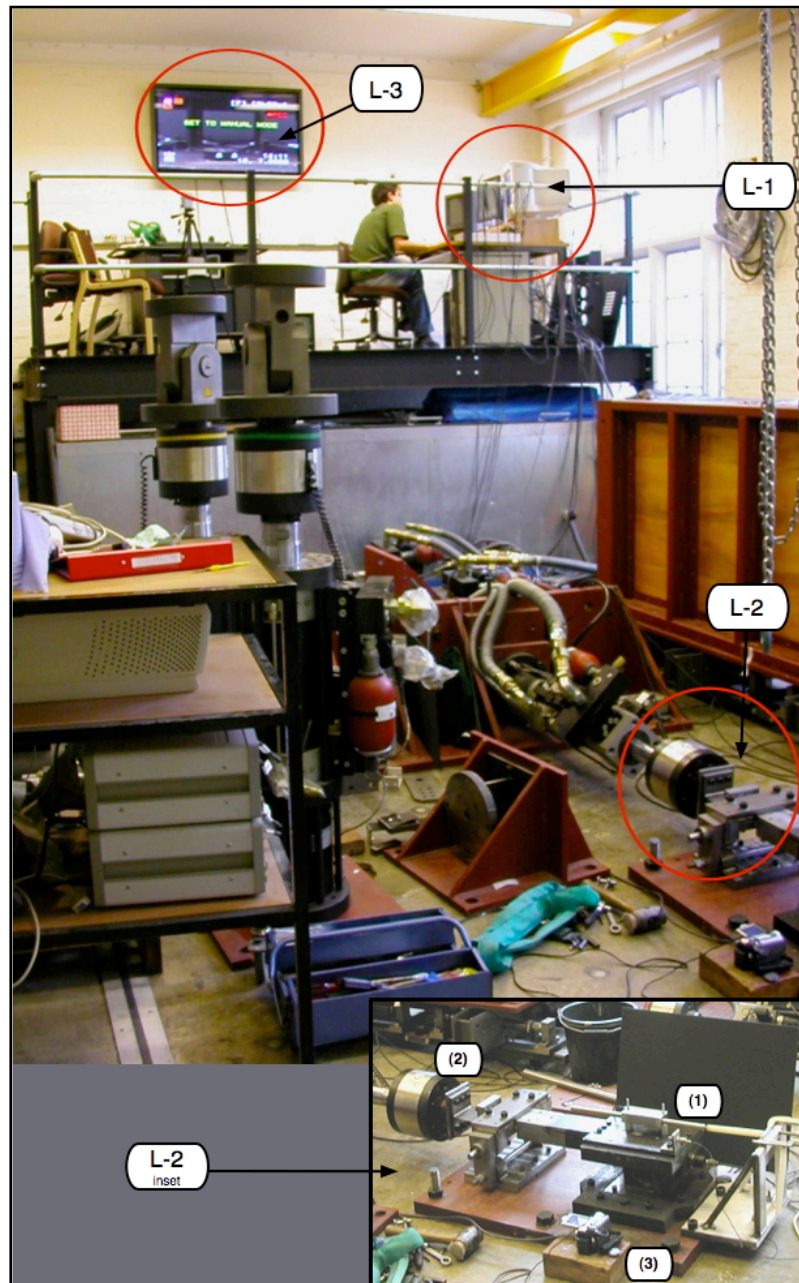


Fig 4. Main areas of the laboratory: computer control area (L-1), the experiment (L-2) with close-up of physical specimen (L-2 inset: 1), actuator (L-2 inset: 2) and video camera (L-2 inset: 3). The wall display (L-3) projects a video feed from the camera positioned in front of the specimen.

Conducting local earthquake experiments is primarily a solitary activity. However, prior to running it the researcher relies on two technicians. The first builds the test specimen and includes soldering together pieces of metal which is meant to replicate an area of a building's structure, usually the brace and joint areas are particularly vulnerable during earthquakes and these are tested in small experiments. The second technician then builds the test rig which will house the test specimen. The test rig is a safety critical element soldered to the floor of the

laboratory and built to ensure that the physical specimen remains in place and will not hurl out into the room which could potentially cause damage to equipment and bodily harm.



Fig 5. Disassembling the frame of a test rig and inserting a dissipative device.

When the rig has been built, it is common to run experiments in succession and in Figure 5 Morgan, the research student involved the UK-NEES project, is opening the test rig to remove one physical specimen and replace it with another. Setting up a test involves dismantling the previous experiment (if conducting more than one), the use of wrenches to loosen and tighten bolts, positioning the specimen into the rig properly, checking that the rig is secure and connecting the actuators to the test rig. These preliminary activities are critical to conducting an experiment. This is so because the researcher is concerned with ensuring that these activities support the experimental process and resulting data analysis through the smooth operation of equipment and the accurate recording of sensor data.

Considering the complexity of experimental science from design and setup to data analysis, the requirements analysis in this chapter focuses specifically on the experimental procedure itself because this is the activity that the prototype is meant to support and enhance. Working with earthquake engineers in a participatory process, they soon began to articulate the importance of apparently ‘minor’ seen-but-unnoticed practices. This section presents four key activities each of which relate to embodied practices rather than talk-in-interaction and so in this sense is more in line with non-verbal interaction analysis (cf. Kendon, 1990). Observations of local earthquake experiments occurred over a period of 3 months and ran in parallel to ongoing prototype development. Because it is common practice to conduct experiments individually, observations were conducted with a lone PhD student in the lab, Morgan, who is also the lead researcher in the development, implementation and testing of the UK-NEES

system. Through *in situ* and video observation and film elicitation techniques with participants the following activities were identified as aspects of substantive research practice. The first three sections examine the ways in which researchers monitor experiments as they attend to sights, sounds and events within the setting, while the final fragment examines specific activities around the initiation and termination of experiments. First, section 7.3.1 examines the researcher's reliance on visual indicators when conducting local laboratory experiments. Second, section 7.3.2 reveals the ways in which audio indicators suggest cracking and breaking of specimens. Third, section 7.3.3 examines the ways in which the researcher observes the physical behaviour of the specimen during experiments. Finally, section 7.3.4 investigates the ways in which an experiment is both initiated and terminated.

7.3.1 Observing an experiment through multiple views

When an experiment is in process the researcher orients to aspects of it that are accessible to visual inspection. These include representational technologies such as computer-generated graphs that record sensor measurements, and observable events, such as looking at the physical specimen from the platform where the researcher is seated. In the Oxford lab three distinct areas provide this type of visual information: the wall mounted display; with a close-up, live video feed of the physical specimen projected onto its screen, the computer monitors; with the graphs displayed within them that measure and record sensor data, and the experimental area; where the physical specimen itself is secured within the test rig. Each area makes available specific resources that allow the researcher to visually inspect the specimen for indications, such as yielding and structural failure, which determine at what point in the experiment it becomes compromised. These visual indications appear as cracks, warps and bends.

In the following fragment Morgan notices something displayed within the graph which captures his attention as an interesting event (Fig 6a). He indicates this through shifting his bodily orientation towards the monitor as he advances forward moving closer towards the computer screen (Fig 6b). He then looks down at the physical specimen located in the experimental area (Fig 6c) and checks the graph a second time moving closer to the computer screen again (Fig 6d). From here Morgan switches to a second graph which provides a

different view of the data (Fig 6e). The first graph consisting of two large bar charts and the second graph displaying a more detailed view that includes 6 areas of information. From this point onwards, the live video stream is interleaved with other situated embodied activities such as looking down at the experimental area and at the graphs within the computer display (Fig 6f). In these sequences of action, assisted through the use of a variety of instruments, observational practices are produced with “the *intertwining* of worldly objects and embodied practices” (cf. Garfinkel *et al*, 1981: 137, italics in original).

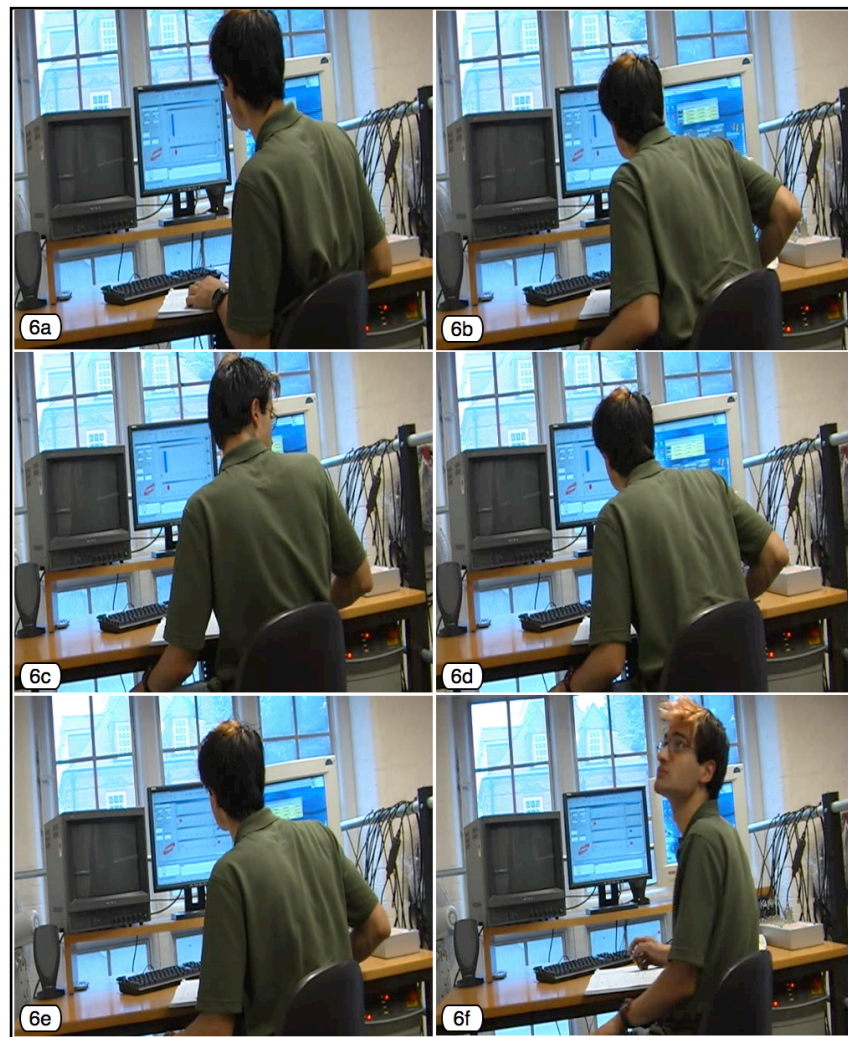


Fig 6. Interest is prompted when Morgan notices a feature on the graph.

Morgan has his pen at the ready so that he is able to make notes immediately upon observing something of relevance which originates either within the graph or through the live video feed. He glances at the computer monitor and jots a brief note into the notebook and again looks up at the live video feed. After an extended examination of the video (3.0) he returns to the

notebook and writes in further information. His note-taking indicates that what was at first an interesting event has now taken on some relevance to the experiment.



Fig 7. Writing down the observation transforms the event from one of interest to a matter of relevance.

As he writes a description of the event down in the lab notebook it is transformed from being something of interest to a matter of relevance. Data from the graph includes the time of the event and measurements such as frequency, amplitude, temperature, acceleration, force and displacement. However, the graphs do not provide all the necessary information and so the notes made in the lab book include descriptions of its physical appearance. During conversation and film elicitation sessions with researchers they verified that although they reference the computer-generated graphs which display sensor measurements, *“they can tell you what’s happening only up to a point”*. While Morgan notices subtle changes that first appear through the graphs where data from sensors located on and around the specimen are recorded, the times and descriptions of cracking behaviours are also important to note. The visual characteristics of the specimen provided through the live video show its behaviour in detail and the ways in which the ‘web’, the internal part of the structure, and the ‘box’, its surrounding external shape, deteriorate as an experiment progresses (Fig. 8). These defining moments are extracted from video recordings taken during experiments, analysed and used in publications to explain device failure.

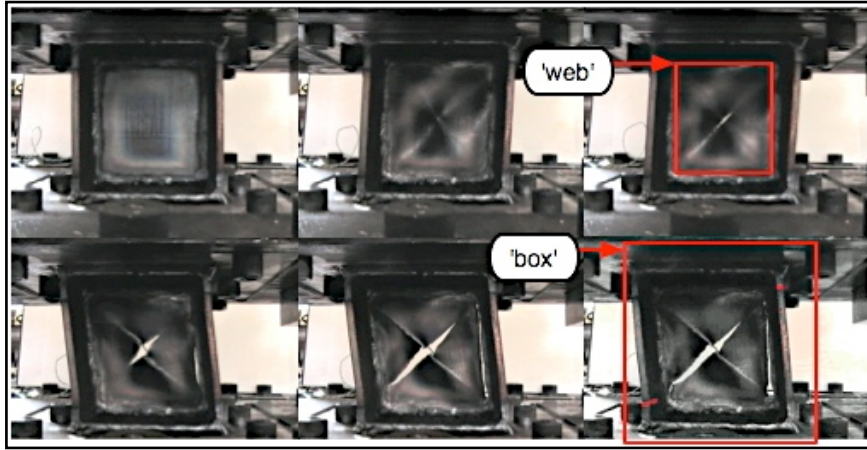


Fig 8. The details of a physical specimen's deterioration (Ojaghi *et al*, 2008).

However, the researchers remark that not all cracks are identified by looking at the video alone and that graphs are used to track crack growth over time with their associated measurement readings. On the other hand, the graphs could give indications that are related to conditions which are outside of experimental parameters such as a loose bolt within the test rig or some other artefact of the experiment. This explains why Morgan looks down from the platform to observe the movements of the test rig and the specimen; he is watching for rig motion to determine if its motion is being recorded as artefactual data within the graphs.

In this way, crack growth is tracked using both the graph (quantitative) and the physical (qualitative) visual indications with each source of information mutually supporting the other. Suchman (1997) describes this type of activity as 'reading a scene' where participants in a setting learn to identify and use information from various sources such as computer displays, the physical objects within a setting, etc., which are juxtaposed and interpreted "in real time, into provisional assessments of an emerging situation". These provisional assessments are made possible by having access to different views of the activities taking place, and in this setting they provide information to researchers about the ways in which an experiment is unfolding in real-time. Through monitoring representations of data (the graphs) and actual observable events (the physical specimen) researchers develop a sense of what is going on in a scene using multiple views.

7.3.2 Interpreting audible behaviours that suggest cracking

In addition to having access to multiple views of the physical specimen earthquake engineers rely upon its audible behaviours to note when cracking and breaking occurs. Audio indicators suggest breaking and observational notes are taken at the time of cracking sounds. Additionally, sounds are monitored for indications of how smoothly the mechanical components of the experiment, such as the actuators, are operating. Because sounds are not recorded by any applications (other than through the video camera) these audible behaviours are noted down by hand when the researcher hears them. The researcher will note down the time the sound is heard and associated descriptions of its appearance as seen through the video feed and the measurements recorded in the graphs when the sound was heard. Interestingly, the researchers reported that there is no ‘categorisation of sounds’ that they reference either explicitly or implicitly which associate specific meanings to sounds such as pops, cracks and pings, etc.

Each fragment reveals the ways in which various embodied practices build upon one another in what Goodwin (1995) refers to as “an architecture of perception”. In the previous fragment the researcher relies upon the availability of multiple views to assess the specimen’s physical behaviour. In the following instance, the availability of audible behaviours also suggests cracking and breaking. The fragment begins with Morgan monitoring the progress of the graphs at the computer display.



Fig 9. The first audible indication of cracking.

However, 3.0 seconds into the sequence the physical specimen makes an audible sound and Morgan immediately attends to his notebook and glances up at the graphs to note down the

measurement readings at the approximate time in which the sound was heard. Another crack is heard 20.0 seconds into the sequence.

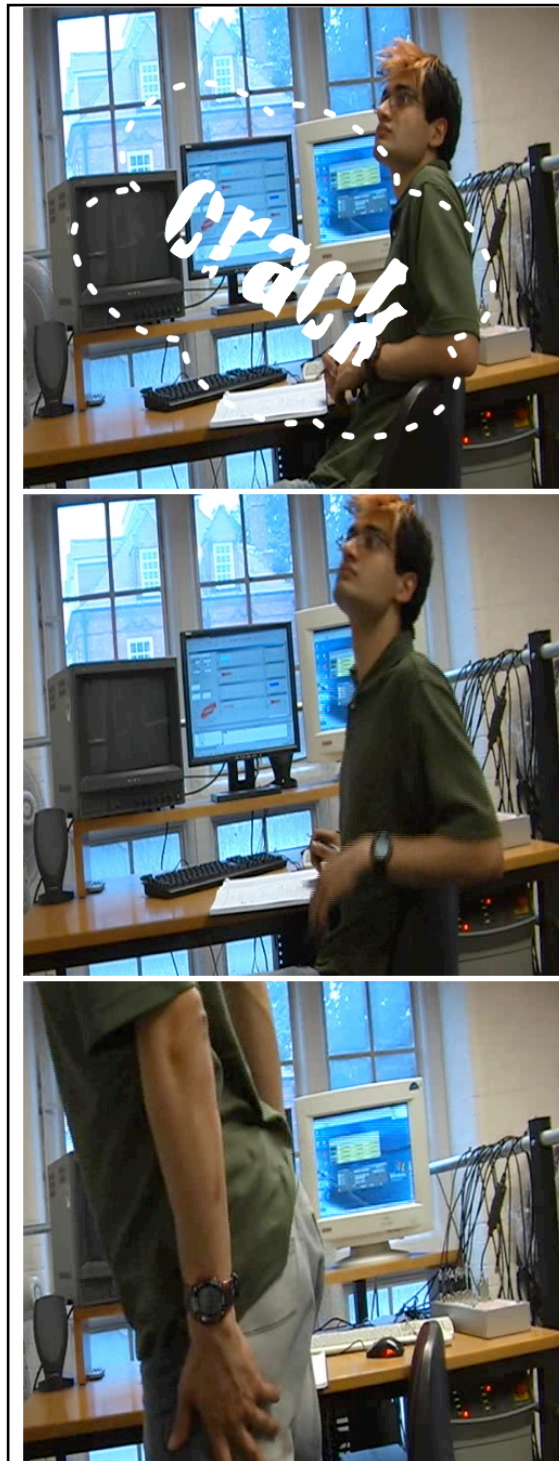


Fig 10. The second audible indication of cracking.

Upon hearing the second crack Morgan had already been watching the physical specimen through the live video feed (see Appendix 3, Figure 8). This second audible behaviour prompts

him to inspect the specimen's physical behaviour more closely. First by orienting his body towards the screen and next getting up from his chair to view the screen more closely. He then returns to his chair and continues to monitor the graphs and take notes.

Each of these two instances reveals how the audible behaviour of the specimen prompts a response from the researcher. In the first instance, Morgan orients to the measurements indicated within the graphs and takes notes. In the second, he orients to the video monitor and observes the physical behaviour of the specimen. On each occasion, the sounds that occur within the setting provide the resources necessary for the researcher to make relevant observations which result in noting down activities that may be important in the data analysis phase after the experiment is completed.

7.3.3 Ambulatory checks around the test rig

During experiments having access to the physical specimen is important in order to check both the stability of the test rig itself and the secureness of the specimen held within it. In the following fragment, towards the end of the experiment, Morgan leaves the platform to check on the test rig and is, as in the previous fragment, prompted by the audible behaviours of the specimen. At this point, the specimen is not longer making intermittent cracking sounds. Rather, because it is at the end of its lifespan and structural integrity it is producing a constant 'back and forth' sound as it is pushed by the actuator.

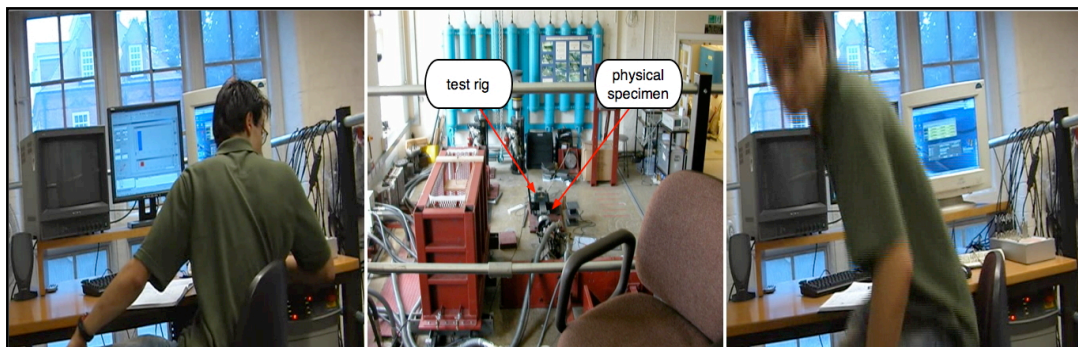


Fig 11. Leaving the platform to perform ambulatory checks around the test rig and physical specimen.

As can be seen in Figure 11, having ambulatory access to the test rig allows the researcher to attain a 360 degree view of the entire event. Having such access, the researcher can survey the area for loose bolts, unwanted vibrations and other effects that may skew the sensor data.

Although, the live video feed is useful it presents only a single, restricted close-up view of the specimen making it impossible to assess the complete scene.

7.3.4 Initiating and terminating an experiment

The fragments above reveal that the earthquake laboratory is a space where multiple embodied activities are relied upon during an experiment, each informing the other. While the computer records sensor data into digital format, the researcher nonetheless requires access to other types of information that are accessible through embodied sensory channels such as auditory and visual cues. The earthquake engineering laboratory is also a space where the researcher is engaged in peripheral monitoring, or what Heath & Luff (1991) describe as ‘overseeing the local environment’, as an ongoing activity. In this way, digital representations in the form of graphs are verified or triangulated with embodied sensory information. This last section discusses what the researchers themselves reported may be the two most important activities of an experiment, its initiation; without an accurate initiation of time-steps and synchronisation of equipment the resulting data would be inaccurate and invalid and its termination which is equally important; critical for safety reasons in an effort to prevent equipment from becoming damaged as well as the safety hazards involved in stopping moving hydraulic equipment and the physical specimen, if this is done incorrectly, the data may be corrupted near the experiment’s end.

It is commonplace in most laboratories that legacy systems, that were not originally intended to interoperate, are adapted so that software systems (e.g. the numerical model of the building, the actuator control software, the time-stepping algorithm, the earthquake simulation) and hardware systems (e.g. hydraulic actuators, sensors, test rigs and the physical specimen) work together. Developing technical workarounds is expected in most laboratories with each designing its own solutions. Workarounds ensure interoperability, however, sometimes systems may not be optimally integrated. In one case, the procedure for local testing required the use of two interaction devices (computer mice) simultaneously or in quick succession to start an experiment. The start of an experiment begins by first turning on hydraulics and placing them on stand-by and then synchronising the actuator’s time-steps with the numerical model’s time-

steps. Synchronisation is related to the alignment, measured in time, of the numerical model and the physical specimen as commands are sent to the specimen and responses received back to the numerical model. The synchronisation of time-steps is critical for conducting a reliable experiment because hybrid tests are achieved by separating the structure of interest into numerical and physical parts or substructures. Here, only key parts of the structure, for example a novel seismic-resistant device and its attachments, are tested physically and coupled to an adequately representative numerical model of the rest of the structure (not the main focus of the test). Substructures are coupled together via a transfer system of actuators and sensors with an earthquake simulated through the entire structural system. The experiment proceeds in time-stepping fashion and is continuously loaded with, typically, interface displacements calculated and derived from the numerical part, applied to the physical part and interface forces measured and fed back at the end of each time step to allow calculation of the next interface displacement. In this manner, the complete structure's seismic response is determined while focusing experimentally on the performance of the physical part, the technology under development (Fig. 12).

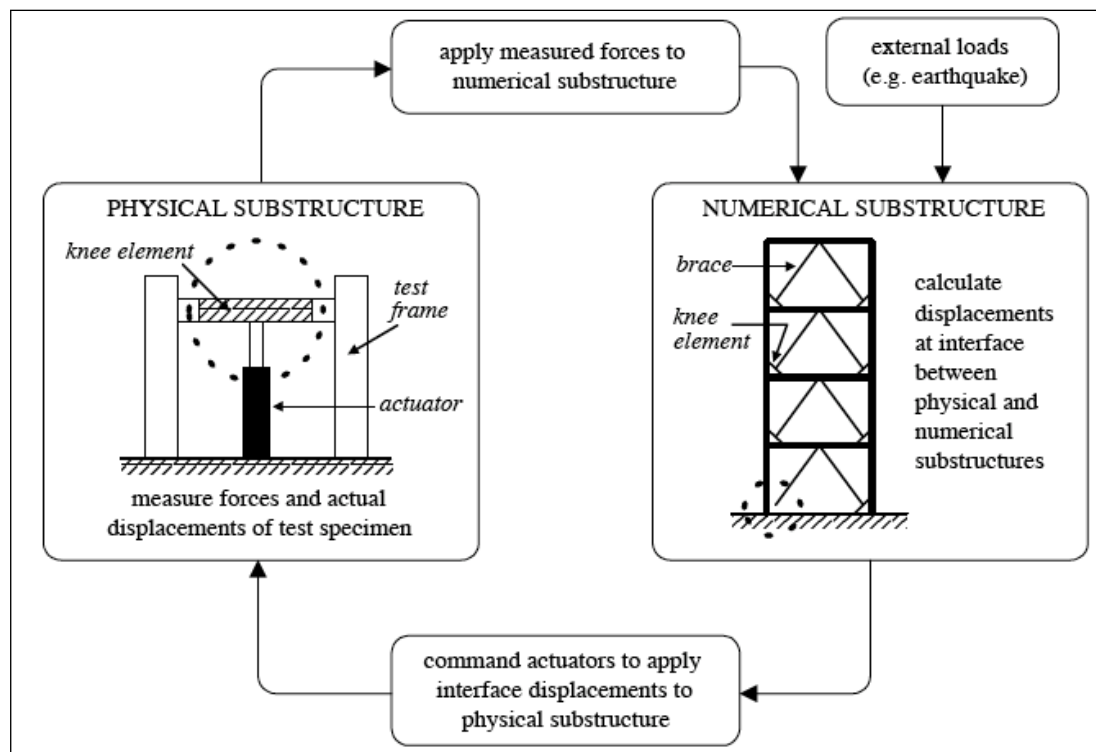


Fig 12. The substructure test loop (Williams & Blakeborough, 2001).

The procedure of using two computer mice simultaneously or in quick succession to start an experiment allows for an acceptable synchronisation of time-steps with no conflict between commands from the time-stepping test controller and the actuation system, and so was considered good enough for local testing even though it was awkward for the user. The local workaround for integrating the two systems is seen in Figure 13. The machine running the numerical model and data capture software (Computer A) and the machine for synchronising the mechanical testing equipment (Computer B) are side-by-side on the desk area. Additionally, there is a third system involved (Actuator control) which turns the mechanical equipment on and places it on stand-by.

Because the two systems (the numerical model and the mechanical equipment) were not optimally integrated, synchronising the time-steps between the mechanical hydraulics and the numerical model required handling two mouse controls simultaneously or in quick succession in order to synchronise them. This lack of integration and control of time-steps between the two systems, although good enough for a local experiment, would not be robust enough in a distributed environment. Early on in the design of the UK-NEES distributed system, it became clear that the synchronisation of remote time-steps would need to be more accurately controlled to ensure both reliability and repeatability. As a result, complex technical solutions were developed and an accompanying new experimental procedure developed (see §7.4).

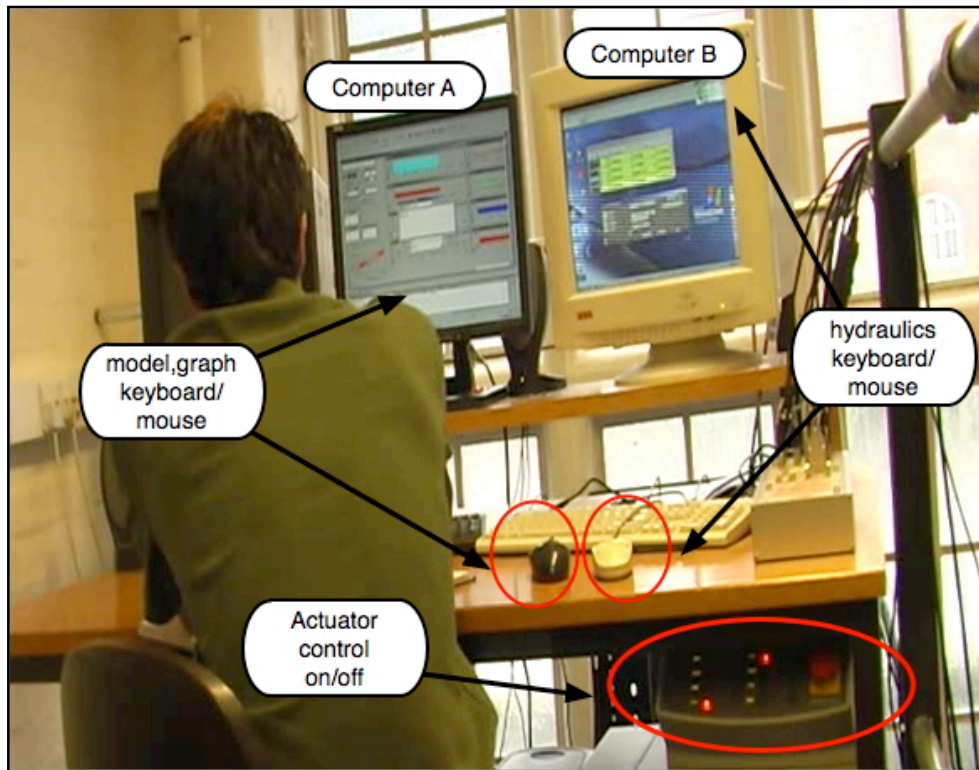


Fig 13. Coordinating the use and monitoring of two computer systems and the actuator control.

In addition to the complexities of initiating an experiment through the synchronisation of time-steps, the termination of an experiment is important from a safety perspective and there are two ways to terminate them. The first is through the emergency stop button located on the actuator control (Fig. 14, left) and the second is through Computer B where the machinery is switched off using the hydraulics software (Fig. 14, right).



Fig 14. Terminating an experiment: emergency stop (left) or through Computer B (right).

So, initiating experiments require the synchronisation of time-steps and their termination requires the winding down of hydraulic equipment. These technical challenges were

considered in the design of the prototype, however, a key usability challenge was also discovered; supporting the coordination and control of distributed experiments as they occur often in real time across laboratory locations. Transitioning experiments into a distributed setting requires that information be conveyed and activities coordinated across remote laboratories. Several requirements were identified during the fieldwork phase which the researchers concurred were essential to consider. These include providing mechanisms for:

Experimental set-up - provide researchers with the ability to communicate how each location is progressing with the set-up of physical specimens and actuator controls.

Initiate - provide a mechanism where researchers can coordinate the start of software and equipment.

Monitor - allow researchers at each location to have access to the types of activities occurring, and to query each other about events in local laboratories as each monitors their own setting.

Terminate - provide mechanisms that indicate to researchers that the single laboratory controlling the distributed experiment is in the process of terminating it.

These requirements correspond to the main activities performed during an earthquake experiment and the details of these types of activities, as they were conducted in a local laboratory setting, were described in this section.

In order to provide support for the distributed coordination of experiments, researchers designed communication software for the purposes of allowing distributed researchers to maintain awareness of concurrent multiple activities as they occur across remote locations. These software tools may potentially provide researchers with an understanding of the overall context of distributed activities which are crucial for the synchronisation and monitoring of experiments. The greatest technical challenges for the project have been the synchronisation of time-steps and compensating for delays across the high-bandwidth connection. Developing solutions to those technical challenges has been the main goal of the project. However, researchers' practices as it relates to experimental procedure would also need to be reconfigured. In the next section, the prototype evaluation focused specifically on the ways in

which the practices of initiating, monitoring and terminating experiments is supported, constrained or transformed through the communication software developed for the project.

7.4 Prototype evaluation: distributed earthquake experiments

A local hybrid test has certain key procedural components that ensure its successful execution, monitoring and termination. When transitioning to a distributed environment, it is critical to ensure that these procedures are properly reconfigured. For instance, in a distributed setting it is no longer possible for one researcher to oversee every aspect of an experiment. Instead, a team of researchers must rely upon one another; sharing their expertise to ensure that experiments are set-up properly and that observations are relayed to distributed partners and recorded into lab notebooks. In a local experiment, the most critical aspect is the start, where incorrect settings, time-step synchronisation, or incorrectly followed procedure will result in test failure. During an experiment, researchers typically make observations to determine its quality and the significance of its results. At the end, researchers are concerned with where data will be saved for later analysis and making sure that systems are shutdown safely.

Coordinating multiple activities with multiple distributed participants is a complex activity where researchers must coordinate both individual and joint activities. To support these activities, project members designed two applications that link the technical aspects of an experiment, such as control algorithms and hardware (the background process), with the ability to coordinate activities and to ensure that the correct procedure is followed (the foreground process). The goal for these two applications has been to remove as much complexity as possible from the experimental procedure, so that researchers can concentrate on the earthquake engineering aspects of an experiment. Unlike local experiments, a distributed experiment does not provide a viewpoint from within a single location, rather a range of partial information resources must be assembled to produce a coherent view (cf. Kuzouka *et al*, 2004).

In response to the loss of ambient properties at remote locations during the initiation, duration and termination of a distributed experiment, two new software interfaces have been

developed that guide researchers through coordinated procedures: a command line interface designed to direct activities, the foreground process, (Figure 15a) and a visual ‘traffic light’ designed to display the state of an experiment, the background process (Figure 15b). These new applications were designed to support the coordination of distributed events by both *directing the activities* of researchers and *displaying the state* of sensor inputs and hence the experiment. They were designed to provide better integration between the two systems described above (see Fig. 13) so that both the synchronisation of machines and the coordination of human activities might be supported in distributed experiments.

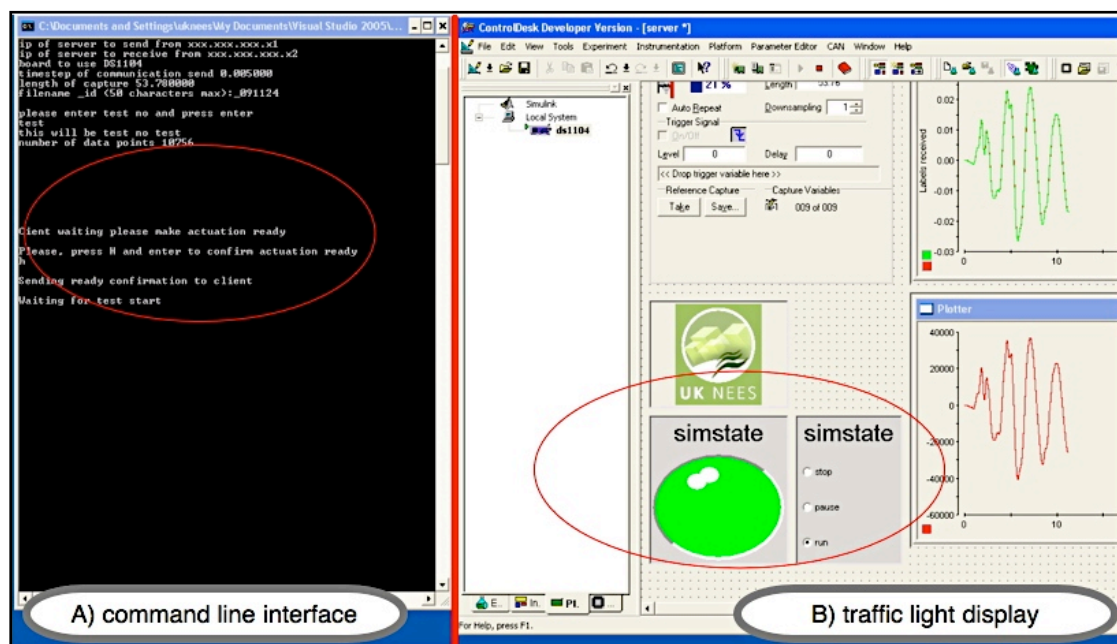


Fig 15. The two communication applications: a) Directing the activities of remote researchers with the command line interface and b) displaying the state of an experiment with the traffic light display.

The prototype evaluation sessions were conducted to evaluate highly complex engineering challenges and so did not model any real engineering problem. That meant that during the prototype evaluation sessions only numerical hybrid experiments were conducted and so it was not possible to observe how physical specimens may have been incorporated. Rather the sessions were conducted to “*prove the connection*” between remote laboratories, to synchronise distributed time-steps, to identify other technical issues, and to ultimately “*generate stable tests*”. Because of the highly technical nature of the prototype, the earthquake engineering researchers were primarily focused on understanding the technical challenges of conducting such tests and solving problems with communication between distributed time-

stepping programs. For this reason, the participants were asked to conduct the ordinary activities associated with testing the proof-of-concept system and in this case, the general activity given was to:

“Conduct a distributed numerical experiment”

As part of the prototype’s design, a client/server architecture was created, assigning the lead laboratory as the client which allowed that location to maintain control of both the numerical model and the overall test (see §7.2.1, Fig. 2). For the duration of the project, the lead laboratory was assigned to Oxford because of their expertise in hybrid testing and because they were also the lead in the design and development of the prototype. Laboratories supplying additional physical specimens are assigned as servers that share local resources yet maintain control over their own local equipment (Ojaghi *et al*, 2010). This client/server configuration has enabled tighter coordination and control of distributed experiments. In what follows, the reconfiguration of practices in distributed experiments between Oxford and Bristol are discussed in order to reveal its impact on coordination, control and cooperative work.

The following fragments reveal how earthquake engineers, Morgan a doctoral research student in Oxford who specialises in numerical modelling and hybrid testing, and Magnus a post-doctoral research associate in Bristol who specialises in structures collaboratively start, monitor and end distributed experiments. The researchers typically run up to 10 numerical tests per prototype evaluation session in quick succession. The following fragments provide insights into the ways in which mutual intelligibility is achieved through a reliance, not on the communication software developed, but rather telephone communication.

This section presents an analysis of how new practices and activities are accomplished using the prototype, specifically, coordinating multiple activities with multiple distributed participants. In this way, the prototype evaluation was not used to compare how current practices were supported, constrained or transformed as was done in the BVREH case study. Rather, entirely new practices were being conducted with the prototype and so needed to be assessed independently of how work is currently conducted. First, section 7.4.1 examines the collaborative monitoring of experiments using the prototype. Next, section 7.4.2 reveals the

ways in which distributed experiments are initiated. Finally, section 7.4.3 examines how experiments are ended and results saved.

7.4.1 Collaborative monitoring using local graphical displays

Even though remote researchers followed the prompts given by the command line interface and monitored the traffic light visual display, they remained in constant telephone communication while monitoring the experiment's progress. In this fragment, as the first distributed test is running both Morgan (Mo) in Oxford and Magnus (Ma) in Bristol discuss the behaviour of their local graphs. In this particular instance Magnus is receiving a ('nice command signal') however, Morgan is not receiving any input.

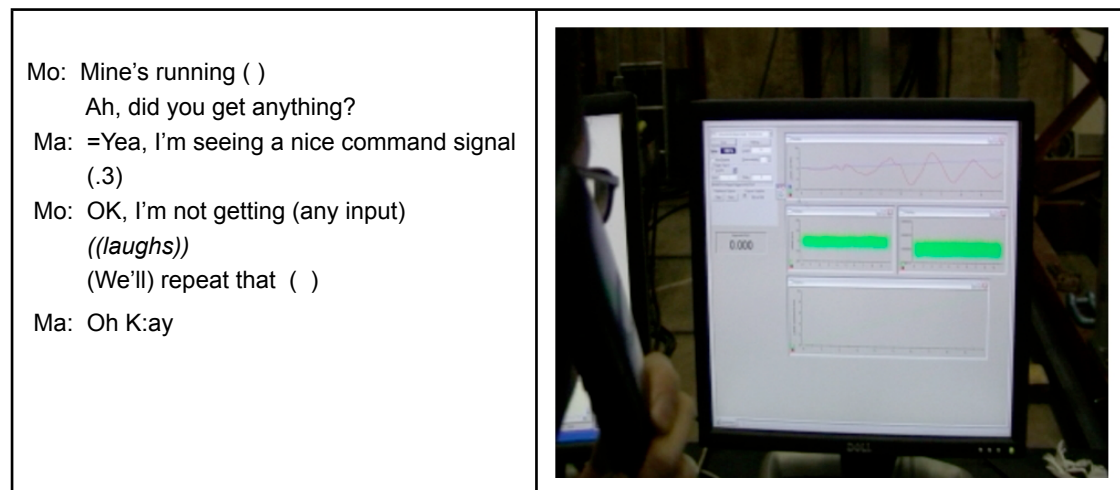


Figure 16: Communicating success at one location and failure at the other.

With only the local graphical display available to each participant there is a discrepancy in the information available between the participants, and so telephone communication is used to fill in the gaps. First, Morgan let's Magnus know that the experiment is running and asks Magnus if he is receiving ('anything'), a command signal. Magnus tells Morgan that he is receiving a ('nice command signal') indicating that the connection has been made at his location and that the experiment is running as expected. However, Morgan responds that he is not receiving input from the graph and that the test will need to be repeated. In this fragment, the graph was animating in Bristol, indicating a command signal is being received, but not in Oxford. If Magnus in Bristol were to have relied on the graphical display alone he might have assumed that the experiment was proceeding in the same or similar manner Oxford. However, this was

not the case and so activity at the remote location was not represented within the prototype. As it turns out, the mismatch in graphical input occurred because Morgan forgot to select ‘animation mode’ within the graphical display’s program.

Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The researchers are checking with each other that if connection between locations has been achieved.	Telephone	<i>Available:</i> Local graphical display <i>Constraint:</i> No feature that allows participants to view the inputs received at remote locations.	Constrain activities to follow the ‘required’ procedure at critical points (e.g. using a bank machine requires a prescriptive order to be followed in order to receive money). In a similar manner, requiring ‘animation mode’ to be selected before an experiment can begin may prevent errors in its initiation.
Cooperative work	One participant assumes that the connection has been made and that the test is running, the other participant indicates that he is not getting input.	Telephone	<i>Constraint:</i> No feature that allows participants to view the inputs received at remote locations.	Design a common graphical display so that local graph data from each participating location may be viewed.

Table 1: CEF analysis - Communicating success at one location and failure at the other.

From this the following suggestion for future iterations of the prototype can be made:

Required fields - design a mechanism to check that the ‘animation mode’ is switched on in order for the experiment to begin. This would be a required field where ‘animation mode’ would need to be selected before any subsequent activities can occur, such as sending the first command signal out to a remote partner.

Common graphical display - design a feature that presents data from the local graphs of distributed participants. This would enable a view into the types of information remote participants are receiving.

Although, this type of sequential ‘mistake’ (possibly indicated by Morgan’s laughter) in the procedure may not occur often, adding a required field may prevent costly mistakes that could occur if a hybrid test is conducted which incorporates actual physical specimens. Additionally, having access to a common graphical display may allow distributed researchers to monitor and comment on each others’ graphical inputs in order to identify problems and to comment on other behaviours represented by the graphs such as specimen’s temperature and other variables monitored by sensors.

In the next fragment, the connection between the two locations has been successfully established and the researchers are in the process of determining if the command signals are being sent and received at the appropriate rate of speed. Morgan describes it as (‘a bit jerky’) and Magnus agrees, elaborating on this description with (‘seems yea, very step-like’). In what follows, Morgan continues with a ‘play-by-play’ commentary on how the graph is behaving and from this they both draw the conclusion that they are looking at (‘very high delays’) sending and receiving data through the network.

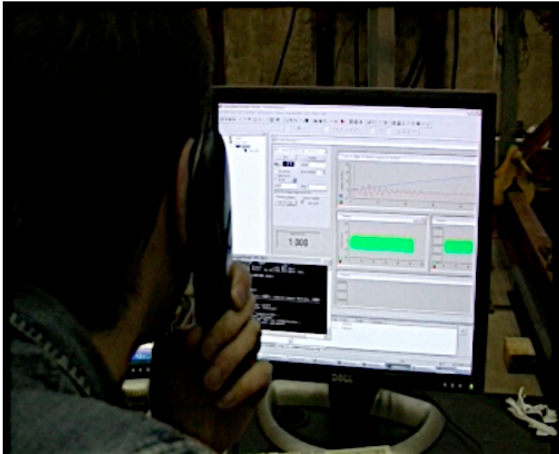
Ma: I'm receiving a command signal Mo: Hold on (3.3) it's a bit jerky Ma: Seems very (.3) yea step-like (1.0) Mo: () shooting up a little bit (2.3) still manageable but () Ma: Yea Mo: Millisecond delays it goes up to () drops down again, went up again Ma: ((sighs)) (2.0) Mo: Yea, we're looking at (seven) millisecond delays () Ma: Very high delays	
---	--

Figure 17: Monitoring the connection speed between laboratories.

Both Morgan and Magnus refer to their local graphical displays to discuss the progress of the signal and as in the previous fragment, the telephone is used to communicate. They comment

on the shape of the graph as it appears at their individual location and from this are able to combine the two observations in order to make an assessment of the stability of the connection.

However, unlike the previous fragment where the participants are updating each other on the initiation of the experiment at each location, in this fragment, they are observing and commenting on its progress based upon the behaviour of their local graphs. These two separate graphs are now synchronising with one another so that the delays in Oxford have an effect on the transmission of signals to Bristol and so are reflected in Magnus's graph. In this way, the distributed experiment is represented through the two separate graphs as they output data at each location. However, each graph may only be interpreted and understood in relation to the other as the time-step data between the two laboratories is integrated so that the separate signals being sent and received are combined into a single distributed event.

From these observations they are able to comment on the quality of the connection speed and with the identification of high delays, the numerical experiment's results would not be not valid and therefore the experiment is terminated. Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The researchers are monitoring and commenting upon the connection speed between the two locations.	Telephone	<i>Available:</i> Local graphical display <i>Support:</i> The graphical display provides an indication of connection speed between the two laboratories.	Provide a bandwidth display mechanism where bandwidth quality can be tested independently of the experiment's graphical display.

Cooperative work	The participants are in the process of determining if the connection speed is appropriate for the experiment to proceed.	Telephone: used to convey commentary and observational judgements.	The graphical display is intended to represent the outputs from an experiment. However, in this instance it is used to determine bandwidth quality.	(see above)
-------------------------	--	--	---	-------------

Table 2: CEF analysis -Monitoring the connection speed between laboratories..

From this the following suggestion for future iterations of the prototype can be made:

Bandwidth display - design a mechanism where bandwidth quality can be tested and monitored independently of the experiment's graphical display. The sending and receiving of signals could be tested first through such a mechanism and would allow researchers to determine whether the connection's expected speeds were regular enough to go ahead with an actual experiment. Such a feature could provide the following: network service quality, expected quality, event analysis, bandwidth sizing, etc (Cisco Bandwidth Quality Manager 4.0²⁹ is such a network performance management tool).

Such a mechanism may be valuable when actual distributed hybrid experiments are performed. As with the previous requirement, it may prevent costly mistakes from occurring when using physical specimens. If the connection speeds are unreliable a hybrid test could be postponed until the bandwidth quality checks showed a consistent connection between sites.

In the final fragment of this section, a numerical experiment is running as expected and nine seconds into the test Morgan responds to a comment by Magnus in Bristol³⁰. He tells Magnus that he'll ('have to check on that') and confirms that he can see what Magnus is describing ('yea, I can see that'). After which, he tells Magnus that the outputs being displayed on the graph are correct ('yes, it will do that'). Here, even when things are going as expected questions may arise about a graph's behaviour and so having access to the expertise of remote partners can assist others in both understanding and interpreting a graph's activity.

²⁹ http://www.cisco.com/en/US/prod/collateral/netmgtsw/ps6341/ps6385/ps8796/prod_bulletin0900aecd806d3acb.html

³⁰ Magnus cannot be heard because Oxford does not have speaker phone capability.

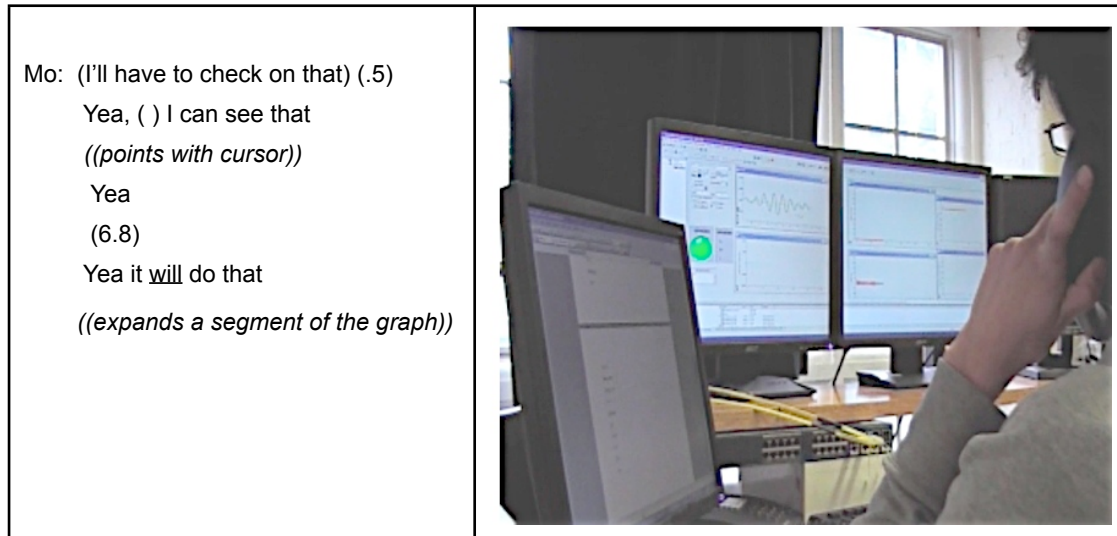


Figure 18: Interpreting the lines of a graph through telephone communication.

This fragment reveals how the prototype supports distributed experimentation. However, the participants' interaction with it may still be improved. While Morgan attends to Magnus's question he expands the graph to view its details. However, because he is constantly holding the phone to his ear, he lacks full mobility and cannot easily move around his work area. The same is true for Magnus in Bristol. This fragment may be used as a representative of a reoccurring pattern; the awkward positioning of the phone between the participants' head and shoulders along with the engagement of arm which hand decreases the participants' overall mobility. This fragment along with the others may be compiled to document the pattern. Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The researchers are analysing the behaviour of a graph during an experiment.	Telephone: held in between the head and shoulders with the participant's arm and hand.	<i>Available:</i> Local graphical display <i>Support:</i> Provides an indication of how the numerical experiment is progressing.	Providing a telephone headset or speaker phone would significantly increase the mobility of each researcher.

Cooperative work	The participants are observing the behaviour of the graph as an experiment is underway.	Telephone: used to asks questions during an experiment and to respond to them.	N/A	(see above)
-------------------------	---	--	-----	-------------

Table 3: CEF analysis - Interpreting the lines of a graph through telephone communication.

From this the following suggestion for future iterations of the prototype can be made:

Telephone headset or Speaker phone - providing each location with a telephone headset or speaker phone may significantly increase the mobility of each researcher. Currently, they must either hold the phone with their arm and hand or place it between the side of the head and the shoulder so that both hands are free. Both the headset and speaker could be trialled during distributed testing to determine which is preferred by the researchers.

Each local graphical display synchronizes with the other and provides the details of how the experiment is progressing. However, the researchers still rely on vocal communication to understand and interpret the graph's behaviour. For actual distributed hybrid experiments where physical specimens are used, a wireless telephone headset or speaker phone may be ideal. This would allow researchers to move around the laboratory freely as they do under normal local laboratory circumstances.

The CEF was used to provide empirical evidence of what may otherwise be overlooked or 'seen-but-unnoticed' activities. While the participants did monitor graphs they also conducted other essential activities in conjunction with graph use. For instance, the procedure for switching on the 'animation mode' needs to be adhered to, bandwidth quality needs to be determined before conducting an experiment and interpreting and understanding the behaviour of graphs requires the sharing of expertise. The next section will examine the coordination of the initiation of experiments in greater detail.

7.4.2 Coordinating the initiation of experiments

During the prototype evaluation sessions, numerical experiments occurred in quick succession one after the other. In order to initiate new experiments the animation mode trigger needs to be

selected, the server sites must be switched to ‘listening’ mode in order to receive the first command signal from the client site, and the test parameters for each experiment have to be changed accordingly at each location. In the following fragment both Morgan and Magnus coordinate the sequence of events necessary in order to initiate an experiment. First, Morgan tells Magnus which activities he is performing as he does them (‘I’m pressing play, I’m going into animation mode’) and Magnus confirms that he is following along at his location with the same activities (‘me too’). Then Morgan tells Magnus to start his server. Magnus does this and confirms with (‘server is listening’). Upon receiving confirmation, Morgan can begin to send command signals from Oxford to Bristol.

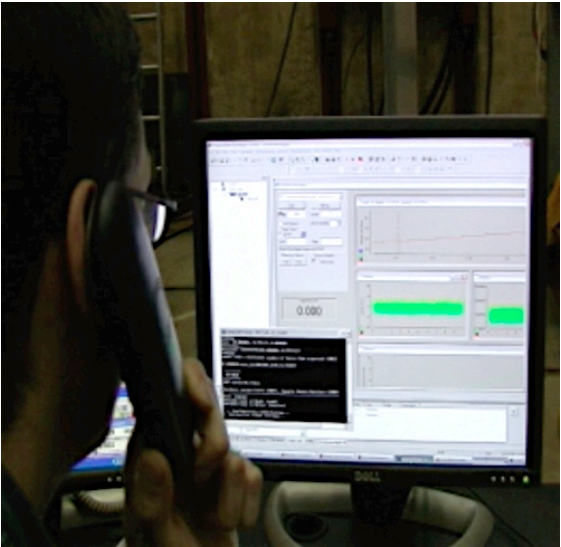
Mo: Um, can we do a quick test (while)- if you get the new sample going, get the new () file I sent-cha Ma: Yea Mo: I’m pressing play (.3) (Alright) I’m going into animation mode Ma: Me too Mo: Start your server Ma: Server is listening Mo: Line is connected. uhh (.3) I’ll send stuff ↑now (.3) uhh, I don’t seem (to be on) <i>((laughs))</i> Ma: Uh::h, yea somethings happened Mo: (>You’ve<) massive, massive- Ma: =There was a massive delay though	 "Server is listening"
---	---

Figure 19: Coordinating a distributed experiment (first prototype iteration).

Morgan notifies Magnus when the line is connected and that he will send a command signal (‘Line is connected ... I’ll send stuff now’). After which Morgan says that he doesn’t (‘seem to be on’) based upon his observations of the movement from the graph. However, Magnus indicates that something is happening at his location and that the graph is animating. Morgan responds with an observation that there are (‘massive’) with Magnus interrupting (‘there was a massive delay though’).

The fragment reveals how the researchers coordinate the start of an experiment by letting the other know exactly what activities are being performed and when (e.g. ‘I’m pressing play’,

‘me too’, ‘start your server’, ‘server started’). Such communication is necessary so that each location may conduct the distributed procedure using the correct sequence of events. For the distributed experiment to run correctly the client location (Oxford) must know when certain activities have occurred at the server location (Bristol). Receiving confirmation that both locations have ‘pressed play’ and are ‘going into animation mode’ lets each participants know that their local graphical displays should reflect the effects of command signals being sent and received across the two labs. Additionally, letting the remote laboratory know that the line between the two has been established, and that the client will send the first command signal gives the server location an implicit instruction to pay particularly close attention to their graphical display so that the connection speed may be assessed. Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The researchers are initiating an experiment and communicate local set-up activities to each other.	Telephone: used to convey when each activity is undertaken within a local setting.	N/A <i>Constraint:</i> During the set-up and initiation of an experiment there is no shared graphical display and activities are performed on local computers.	Design a common graphical display for experimental set-up and initiation.
Cooperative work	The participants are in the process of initiating an experiment which requires that activities take place in a prescribed order. For instance, the server must be switched to listen before the client can begin to send command signals.	Telephone: used to coordinate the progress of activity across locations.	N/A	(see above)

Table 4: CEF analysis - Coordinating a distributed experiment (first prototype iteration).

From this the following suggestion for future iterations of the prototype can be made:

Common display for experimental set-up & initiation - currently researchers rely upon their local graphical displays to monitor experimental outputs as they occur in real-time (see §7.4.1). Providing a common display for the set-up and initiation of experiments may provide remote participants with a more complete indication of the types of activities occurring elsewhere which may be relevant to the next actions that need to be taken by those in participating locations.

In order to initiate a test, participating laboratories must conduct activities in a pre-determined order. The requirement for a common display for experimental set-up and initiation was identified during early prototype evaluation where the video data was analysed to understand how distributed coordination takes place. As a response, the researchers developed a shared display providing an indication of when each laboratory has completed required activities. In the second iteration of the prototype, the ‘command line prompts’ work in conjunction with the ‘traffic light display’ to provide a visual indication of the status of each laboratory.

The following fragment reveals how the new software is used to coordinate the initiation of an experiment between laboratories. The traffic light display provides a visual indication of the status at the remote location. Once each location is ready and on stand-by, the visual display changes on all systems from red (stop) to amber (pause). Upon connection (for hybrid tests that incorporate physical specimens) the server sites are prompted through the command line interface to place their hydraulic actuation systems on standby and to confirm laboratory readiness. Upon receipt of remote server confirmations, the client (Oxford) is prompted to place its local actuation systems on standby and, when ready, to start the distributed experiment with a single command. Upon receiving the start signal, the experiment is initiated setting the state of each local system to green (run), indicating that the experiment is active. At the end of the test, the state of all sites returns to red (stop) automatically confirming that the experiment has terminated.

This second iteration of software was evaluated through conducting distributed numerical experiments with physical specimens used. However, even with the visual indicator in place, the researchers still remain in constant telephone contact. In the following fragment, Morgan indicates that he is (‘good to go’), however, the traffic light display in Oxford remains red (experiment not running). Even so, the display has changed to amber (experiment on stand-by)

in Bristol, indicating that Oxford is ready. After an unusually long delay in responding, Morgan asks if everything is OK ('you alright there') and vocalises this reminder ('yea we're good to go now') which prompts the Bristol researcher to initiate the experiment at his location, where Magnus then sets his server on 'listen'. Once the initiation of the experiment is coordinated, Morgan starts the experiment, sending the first command signal and commenting on the first displays of data from the graph.

Mo: Good (.) good to go

(31.0)

Mo: >You alright<
You alright (in) there?

(0.5)

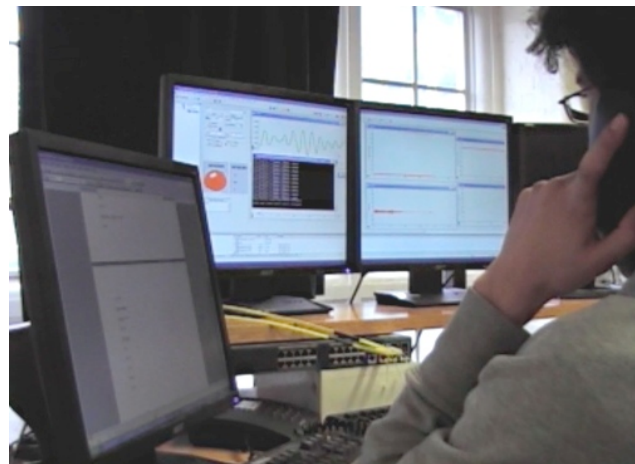
Mo: Yea, yea we're good to go now

(0.5)

Mo: Yea, yea



*((traffic light turns amber indicating that
the Bristol server is ready to receive))*



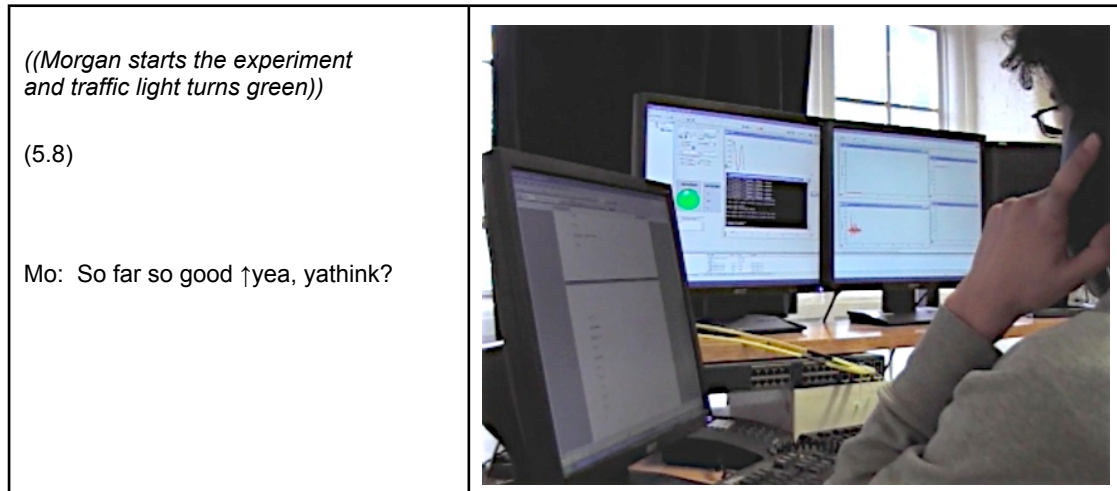


Figure 20: Coordinating a distributed experiment (second iteration of the prototype).

The ‘traffic light display’ software supports the initiation of the new distributed procedure where remote researchers maintain awareness of each location through a colour-coded visual display. In this instance, Morgan was alerted to the non-responsiveness of Bristol because the visual display did not change colour from red to amber. On the other hand, even though Bristol receives visual indications through the traffic light display that indicate Oxford’s readiness (by displaying its state with the colour amber), the researcher there does not notice the visual cue and must be alerted vocally by Morgan. Additional vocalisations include, Morgan usually beginning each experiment with an audible notification (‘good to go’). Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The researchers coordinate the start of an experiment.	Telephone: used to indicate the status of the client to the server location.	<i>Available:</i> Traffic light display <i>Support:</i> Provides a visual indication of remote location's status.	Provide audible alert sounds that coincide with the change in visual indications of the traffic light display.

Cooperative work	The researcher must coordinate their activities in the correct sequence.	- Telephone: used to alert a participant that the experiment may begin.. - A separate computer is used to take notes (replacing the paper notebook)	<i>Available:</i> Traffic light display <i>Constraint:</i> Visual indications may not be sufficient because they are not always noticed.	(see above)
-------------------------	--	--	--	-------------

Table 5: CEF analysis - Coordinating a distributed experiment (second prototype iteration).

These audible vocal cues suggest that perhaps an alert sound could be incorporated into the visual display thus augmenting visual indications with audible ones. This may prove to be a helpful feature when actual distributed hybrid experiments are conducted which incorporate physical specimens. For those kinds of experiments researchers will be attending to their local environment and may not always be present at the computer display. Making available only a visual indicator requires researchers to pay close attention to the computer monitor and may restrict their movement around the laboratory during actual hybrid experiments. From this the following suggestion for future iterations of the prototype can be made:

Audible alerts - currently the traffic light is the only shared display and is used to monitor the status of each location. Providing audible alerts in addition to visual cues may increase the likelihood that system prompts are recognised by distributed partners. A sound indication could prompt researchers when they are looking away from the computer's display.

In the last fragment of this section examining the coordination of an experiment's initiation, the numerical model in Bristol needs to be modified before the test can begin. Because Morgan possesses the expertise in numerical modelling, he has developed them for both laboratories and so knows exactly which parameters need to be modified. Here, Morgan asks Magnus to change some parameters in the model from one to two milliseconds. Morgan names the file which prompts Magnus to locate it. After which, Magnus proceeds to locate the section of the file containing the parameters and confirms with Morgan which should be changed ('so all three to two milliseconds') while he types into the document. Morgan confirms with ('all three to two milliseconds').

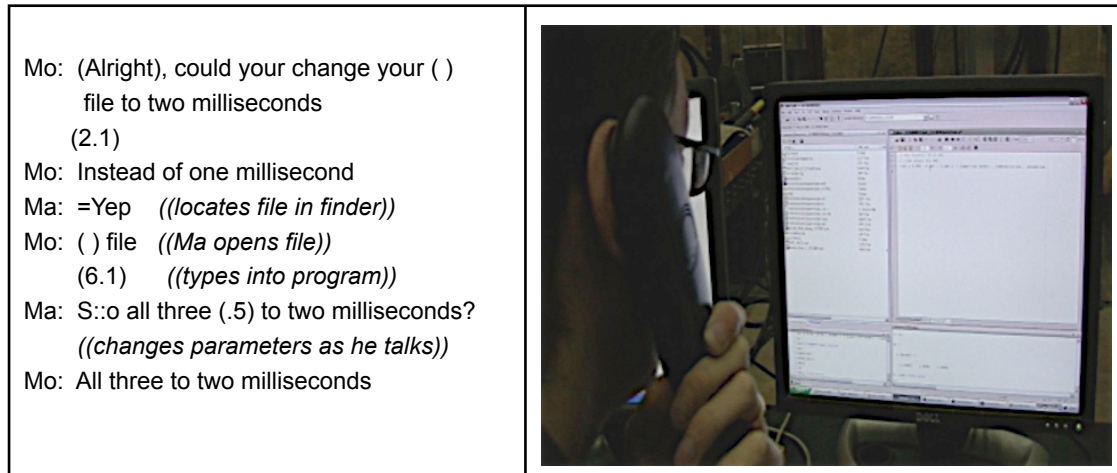


Figure 21: Changing experimental parameters through file editing.

Here, the telephone is used to communicate instruction to the Bristol researcher, who has less experience and expertise in numerical modelling. Even though this activity is successfully accomplished, activities around file editing may be supported through a networked file storage system which could provide each location access to the remote files relevant for conducting experiments. Such an option would allow the researcher leading the experiment to change parameters across locations if required. It would also allow remote participants to diagnose code written at other locations. Currently files used in experiments are emailed (see next fragment) to distributed partners with emails sometimes being sent in the set-up and termination phases of experiments. Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	Participants are discussing the need to change parameters of a file located in Bristol and related to an aspect of the experiment in Bristol.	Telephone: used to communicate a change in experimental parameters.	<i>Available:</i> Local file system <i>Constraint:</i> Files cannot be accessed across locations.	Provide a shared network file storage area that would give distributed participants access to files relevant to the experiment.

Cooperative work	Morgan instructs Magnus to change specific parameters.	Telephone: used to instruct the Bristol researcher.	N/A	(see above)
-------------------------	--	---	-----	-------------

Table 6: CEF analysis - Changing experimental parameters through file editing.

From this the following suggestion for future iterations of the prototype can be made:

Shared file storage - Create a networked drive or web portal feature that allows all participants access to files that will be used during experiments.

The three fragments above reveal the ways in which distributed numerical tests are initiated. The first two fragments discuss distributed coordination using different iterations of the prototype. The first fragment (Fig. 19) was from an evaluation session conducted early on in the system's development and the second fragment (Fig. 20) was from an evaluation session conducted after the second iteration of software was designed to support distributed activities. In each fragment, the suggested requirements discussed could lead to a more usable system when conducting actual hybrid experiments that use physical specimens which is the ultimate goal of the UK-NEES system.

7.4.3 Terminating distributed experiments and saving data

In this final section, the termination of experiments will be discussed. Experiments are terminated automatically through the numerical model, however once completed files need to be saved at each location and then shared. In this fragment, Morgan instructs Magnus to save the file which has recorded the data outputs from the experiment. Additionally, Morgan tells Magnus how he should name the file and Magnus responds by asking if he should put ('a date reference on the end'). Finally, the file is saved and the email client is opened so that Magnus can email Morgan the Bristol result.

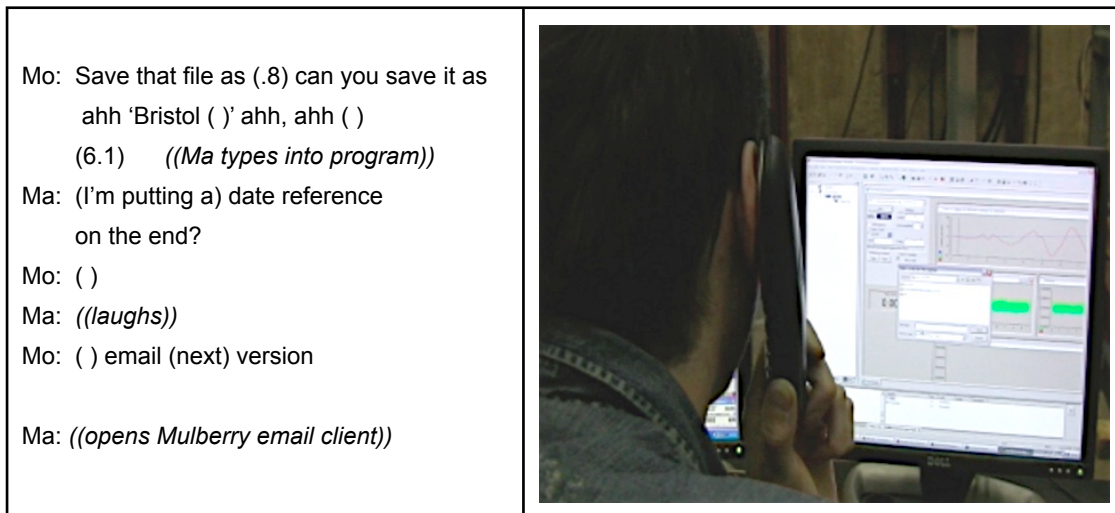


Figure 22: Activities around terminating a distributed experiment.

The termination of experiments involves the saving and sharing of experimental results, such as graph outputs and numerical data. These are in electronic file formats and in order to share them the researchers email them to each other. The requirement here is identical to the one above (shared file storage), however for two completing different purposes. The requirement in the previous fragment is related to the initiation of experiments (editing files), while this fragment is related to the termination of experiments (sharing files). Because the UK-NEES project's initial focus was on developing solutions to highly technical challenges a web-based portal or other mechanism for sharing data could not be implemented at the time.

During prototype evaluation the researchers were involved in activities around the testing and debugging of distributed numerical experiments. The prototype evaluation sessions revealed the ways in which distributed researchers rely upon the telephone to coordinate activities. They also revealed how the communication software tools were used in actual practice and so identified potential requirements that may be considered for future iterations of the prototype. Importantly, the prototype evaluations revealed the need for project members to consider the ways in which non-computational mechanisms, such as person-to-person telephone communication, may be crucial to the coordination of highly complex activities undertaken around computational systems intended to support distributed work.

7.5 The use of the CEF in the UK-NEES project

The CEF analysis of the earthquake engineers' practices revealed that making observations in a local laboratory setting involves a variety of activities including maintaining an awareness of the ambient properties of the laboratory which are crucial for the start, monitoring and termination of local hybrid experiments. During distributed evaluation, the experiments conducted were constrained to numerical testing due to the level of technical maturity of the prototype and so a direct comparison of how these ambient properties were represented through the system could not be assessed. Even so, this case study allowed me to use the CEF to evaluate a prototype within a distributed setting.

I found that there were significant challenges in conducting distributed prototype evaluations. For instance, it was not possible to place a video camera at the location where I was not present (whether it be Oxford or Bristol) because neither I nor the researchers had spare video cameras. This prevented me from having access to the activities as they occurred in the remote setting. However, when I was in Bristol, the Oxford researcher was placed on speaker phone and this did provide access to that researcher's talk as they both coordinated activities, although it was very difficult to hear Morgan's talk clearly. This was resolved somewhat by processing the video file using audio restoration software³¹. Another software tool I considered incorporating into distributed evaluations was the Digital Replay System (DRS)³² which allows for the synchronisation between system log files which record users activities within the system and video recordings which capture the embodied activities of participants in a setting. This novel technology, unfortunately, could not be used because the DRS tool, at the time, required substantial customisation and a level of software coding expertise that was beyond my own abilities.

Even with these constraints, the coordinating activities of multiple distributed participants could be observed from within a single location. A single-sited approach to understanding the coordinated activities of multiple distributed participants has been referred to as 'strategically

³¹ <http://www.bias-inc.com/products/soundSoap2>

³² <http://www.ncess.ac.uk/tools/drs>

situated single-site ethnography' (Marcus, 1995). This approach involves focusing upon the ways in which local activities are calibrated in response to the activities of remote partners. For the UK-NEES project, this meant observing and describing the coordination work of the researchers as they attempted to develop a coherent understanding of the activities occurring at each laboratory. What I found was that even though the researchers had developed two software applications specifically to support the coordination of activities: one to direct the activities of remote researchers through command line prompts and the other to display the state of experiments through a traffic light display, that nevertheless, the researchers organised their activities primarily through talk on the telephone, where the actual use of these two applications was guided by that talk.

It has been noted in the CSCW literature that technological support for synchronous and distributed (as well as asynchronous and co-located) interaction should include a shared workspace where participants may communicate and interact with one another through the system (cf. Schmidt & Bannon, 1992; Bannon & Bodker, 1997). The two applications developed by the UK-NEES researchers did provide some level of shared communication and interaction support. However, it did so at a very high-level and so did not provide the richness of information they require, which is afforded by talk. The prototype could not provide this richness primarily because its purpose was to solve technical challenges; specifically the synchronisation of time-steps across geographically distributed locations. Because of this constraint and the fact that in order to conduct real-time distributed experiments all bandwidth must be allocated exclusively to the experiment³³ the workspace that the researchers developed could not include rich features like audio and video.

In terms of presenting the CEF video data and preliminary findings to UK-NEES project members, key video fragments taken during fieldwork were discussed with the lead researcher using film elicitation. Early on in the project we identified that the ambient properties of the setting were crucial in all aspects of experimental set-up and initiation, monitoring and termination. The lead researcher was surprised that the project team had not considered these

³³ The project did eventually acquire a dedicated Lightpath <<http://www.ja.net/services/lightpath>> line which provides end-to-end connectivity between equipment across laboratories. Having this dedicated line in place could allow for a rich workspace to be developed to include audio/video links between locations which would run on the 'public' JANET Internet alongside the Lightpath connection.

features of experimental practice when discussing the design of the prototype with each other. These however could not be incorporated into the design of the system (see above), even so, the findings have provided the researchers with data that they may draw upon when there is an opportunity to expand upon the prototype.

The prototype evaluation criteria were drawn from the goals set out by the researchers themselves which was to coordinate and conduct a numerical distributed experiment. In this way, the prototype evaluation was not used to compare how current practices were supported, constrained or transformed as was done in the BVREH case study. Rather, entirely new practices were being conducted with the prototype and so these activities were assessed independently of how work is currently conducted. It was discovered from this that the CEF is a technique that can be used for the evaluation of prototypes in the absence of practice (cf. Crabtree, 2004).

This case study has shown that the CEF can be used to evaluate prototypes in a synchronous and distributed setting as well as a technique that can be used for evaluating prototypes in the absence of practice. The case study has also shown that paying particular attention to non-verbal, embodied conduct and interaction may reveal tacit and substantive research practices of laboratory scientists. When I asked the researchers how they would resolve their lack of access to the ambient properties of distributed locations they replied that they would leave those observations to the local experts who have knowledge of the operation of both equipment and physical specimens in their specific local settings. Whilst verbal communication is extensively used to assist in the coordination of activities at a fine-grained level of detail, researchers have expressed that they would like to reduce the need for telephone communication by improving software functionality. It is unclear how or if software can replace talk in the coordination of activities and it should not be considered an inadequacy of the system if talk persists when distributed experiments are conducted. Many other highly technologically complex distributed activities rely on talk in conjunction with the use of software tools in order to coordinate activities (e.g. London Underground, air traffic control, police and fire operations). However, a higher degree of automation is preferred by earthquake engineers, particularly as the number of sites participating in distributed experiments increases

it is believed that verbal communication across each location might become more difficult. If talk is to be replaced with software cues or automation it will require a rethinking of just how the substantive practices of researchers such as changing parameters, monitoring graphs and equipment, and the saving and sharing experimental results might be coordinated without it.

In addition to communicating and discussing work practices through brief, informal meetings with the lead researcher where video fragments were viewed and discussed, a bullet-pointed list summary of requirements identified using CEF was also provided:

Required fields - design a mechanism to check that the ‘animation mode’ is switched on in order for the experiment to begin. This would be a required field where ‘animation mode’ would need to be selected before any subsequent activities can occur, such as sending the first command signal out to a remote partner.

Common graphical display - design a feature that presents data from the local graphs of distributed participants. This would enable a view into the types of information remote participants are receiving.

Bandwidth display - design a mechanism where bandwidth quality can be tested and monitored independently of the experiment’s graphical display. The sending and receiving of signals could be tested first through such a mechanism and would allow researchers to determine whether the connection’s expected speeds were regular enough to go ahead with an actual experiment. Such a feature could provide the following: network service quality, expected quality, event analysis, bandwidth sizing, etc (Cisco Bandwidth Quality Manager³⁴ is such a network performance management tool).

Telephone headset or Speaker phone - providing each location with a telephone headset or speaker phone may significantly increase the mobility of each researcher. Currently, they must either hold the phone with their arm and hand or place it between the side of the head and the shoulder so that both hands are free. Both the headset and speaker could be trialled during distributed testing to determine which is preferred by the researchers.

Common display for experimental set-up & initiation - currently researchers rely upon their local graphical displays to monitor experimental outputs as they occur in real-time (see §7.4.1). Providing a common display for the set-up and initiation of experiments may provide remote participants with a more complete indication of the types of activities occurring elsewhere which may be relevant to the next actions that need to be taken by those in participating locations.

Audible alerts - currently the traffic light is the only shared display and is used to monitor the status of each location. Providing audible alerts in addition to visual cues may

³⁴ <http://www.cisco.com/en/US/products/ps6385/index.html>

increase the likelihood that system prompts are recognised by distributed partners. A sound indication could prompt researchers when they are looking away from the computer's display.

Shared file storage - Create a networked drive or web portal feature that allows all participants access to files that will be used during experiments.

Providing support for these requirements could improve the prototype's capacity to enable the coordination of distributed experiments amongst earthquake engineers. Importantly, the ability to communicate vocally appears to be central to the coordination of distributed experiments. This activity currently takes place outside of the prototype system boundaries, nonetheless, it is an aspect of the larger socio-technical system and should be taken into account when designing future iterations so that the complimentary artefacts of work practice may be better integrated into the primary system under development. The next chapter and final case study will look at the use of the CEF in the NCeSS e-Research project where no prior fieldwork was conducted. This was done to examine how the CEF might be used as a stand-alone prototype evaluation framework.

Chapter 8:

Web portal for the social sciences

8.1 Introduction

The the National Centre for e-Social Science (NCeSS) case study concludes the use of the CEF within an e-Research software development project. It was conducted over a 6 month period and similar to the previous two case studies, I continued to work through how the CEF might be utilised within the context of this particular e-Research software development project. The case study focuses on the evaluation of the e-Infrastructure web portal prototype designed to bring together many different applications that social scientists might use as part of their day-to-day research. It differs from the other two case studies in that no prior fieldwork was conducted with which to provide a comparative basis that might assist in the assessment of how well activities are supported, constrained or transformed when using the prototype. Rather, this chapter explores the use of the CEF as a stand-alone technique in the evaluation of an application's interface for 'usability', thus, issues around its 'usefulness' were not addressed.

The NCeSS case study was conducted in this manner partly because of the rather broad remit of the e-Infrastructure project, which was to deliver a 'generic' web portal for a broad spectrum of social scientists and partly to evaluate the usefulness of the CEF as a stand-alone evaluation framework. Utilising the CEF in this manner is important to consider because many e-Research projects share similar characteristics to other academic and commercial software development projects, in particular, a lack of time and/or resources (cf. Lloyd & Simpson, 2005) necessary to conduct ethnographic studies of current work practice *in situ*.

In this case study, the CEF was used as a stand-alone framework to evaluate the NCeSS VRE web portal. The chapter begins first in section 8.2, with background information about the NCeSS project. Next, section 8.3 will introduce the e-Infrastructure project in order to develop an understanding of the purposes of the prototype system and will subsequently provide an analysis of video recorded prototype evaluation sessions using the CEF. Finally, section 8.4 will summarize the requirements that emerged from the prototype evaluation sessions and will end with a discussion of how the CEF was utilised for this case study.

8.2 Developing an e-infrastructure

The National Centre for e-Social Science (NCeSS) began in 2004, operating for five years until 2009³⁵. The goals of NCeSS were to develop tools, techniques and services that would support a range of social scientists who use quantitative, qualitative and mixed methods approaches in their research through the development of software that might support these research practices³⁶. NCeSS was coordinated and managed centrally through its hub in Manchester which was responsible for designing and managing the research programme, funding and developing cooperative links across six distinct software development projects (referred to as nodes), research dissemination, and strategic planning (Halfpenny & Procter, 2010). Six of the seven nodes were funded to develop proof-of-concept prototypes that would aim to support work practice within specific areas of social science research. This case study took place in between the first and second phases of the NCeSS programme and at that time the seven nodes included:

Geographic Virtual Urban Environments (GeoVUE) - a grid-enabled virtual urban environment (VUE) that would link spatial data to geographic information systems (GIS)-related software relevant to urban planning and other scientific and design problems.

Modelling and Simulation for e-Social Science (MoSeS) - a representation of the entire UK population as individuals and households, together with a package of modeling tools that would enable research and policy questions to be addressed.

35 Although a new 'National Strategic Directorate' <see: <http://www.ereseearchsouth.ac.uk/uk-e-social-science>> has been formed to replace NCeSS the details of this change is outside of the scope of this chapter and so will not be discussed.

36 http://www.ncess.ac.uk/about_us/

PolicyGrid - an application for rural policy development and appraisal using the Grid and Semantic Web technologies to explore how Semantic Grid tools can be created to support social scientists and policy makers.

Digital Replay System (DReSS) - a Computer Aided Qualitative Data Analysis (CAQDAS) tool to enable the synchronization, replay, and analysis of audio and video recordings in conjunction with system log file data.

Mixed Media Grid (MiMeG) - aiming to generate tools and techniques for social scientists to analyse qualitative audio-visual data and related materials collaboratively.

Collaboratory for Quantitative e-Social Science (CQeSS) - developed tools for social scientists using large numerical data sets, in support of social scientists interested in using e-Science methods to support their research.

Oxford e-Social Science (OeSS) - the only non-technical node which examines the 'social shaping' of technologies through the availability of technologies, institutional relationships and researcher work practices.

I attended many of the NCeSS internal meetings and workshops throughout as a participant/observer. Notably, members of a second phase node were also involved in this case study:

Obesity e-Lab - a VRE for producing, sharing, communicating and finding obesity research between epidemiologists, public health researchers and social scientists.

Each node developed prototypes independently of each other, however, in an effort to increase communication and to showcase the developments of each, in 2007 the e-Infrastructure project was funded, initially for three years. Its goal was to support the exchange of information and knowledge across nodes as well as to integrate the tools and data sets that were being developed into a single web portal. It was thought that by developing a single e-Infrastructure web portal that it would allow nodes to disseminate the benefits of e-Social Science both within the NCeSS community and to the wider audience of social science researchers (Daw *et al*, 2007). From an initial set of requirements gathered by the NCeSS team (cf. Allan, *et al*, 2009) a preliminary prototype was developed which included, in the first instance, generic web portlets that would enable communication across work sites.

8.3 Prototype evaluation: a web portal for collaborative work

Although the purpose of the NCeSS portal was to integrate the software developed by each node within a single collaborative working environment its early iterations could not facilitate such integration. This was due to several factors which included the technical challenges associated with encasing stand-alone software into portlets and concerns about software being prematurely released for public consumption. In light of these circumstances, the first release of the e-Infrastructure web portal would focus on developing tools that would allow distributed researchers to exchange information through features such as an announcement tool, calendar, discussion, chat, wiki, and file management facility (Fig. 1).

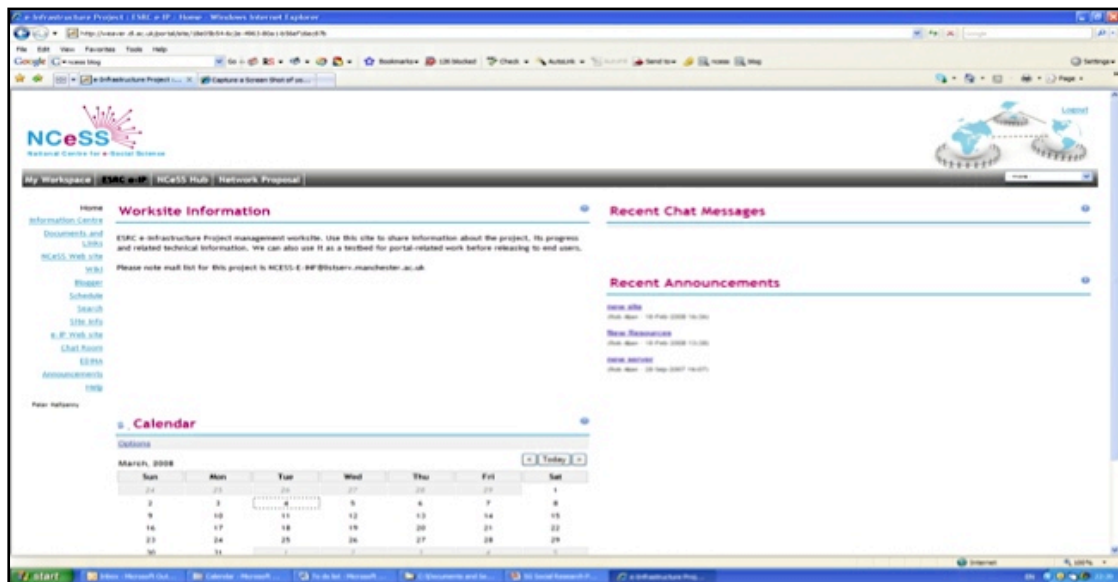


Fig. 1. The NCeSS prototype web portal.

The portal was developed using Sakai³⁷, an open source software development environment and while it was under development, two user evaluation sessions were conducted and the findings shared with the project manager and lead software developer. Given that the evaluation sessions were conducted during the early phase of the prototype's design, only basic features could be evaluated including: uploading, sharing and editing a document, creating a wiki subject area, creating a workspace for a project and adding an event. Two evaluation sessions were conducted. In the first session a social scientist and a requirements engineer

³⁷ <http://sakaiproject.org/>

directly involved in the NCeSS project worked together and in the second session the paired participants included a web developer for the e-Infrastructure project and an epidemiologist from Obesity e-lab. This section will present findings from the two evaluation sessions. First, in section 8.3.1 how participants come to an understanding of the ways in which they can control documents through sharing and editing functionality is discussed. Second, section 8.3.2 reveals the ways in which participants interpret system messages as they create a wiki subject area. Finally, section 8.3.3 discusses how the post-evaluation interview was used to review with participants what they found to be difficult or challenging during the evaluation sessions.

8.3.1 Encountering system limitations

Participants were asked to upload a document into the web portal then share and edit it. Uploading documents was a straightforward process, however once the document was in the web portal two concerns arose within the sharing and editing features. First, participants discussed with each other the insufficiency of the level of access control and second they discussed uncertainties about which document version was in the system once it had been edited.

In the following fragment Phil and Marc would like to control access to the document by assigning rights to specific people. Phil notices that ('there's only one way of controlling sharing') and reads aloud the options that appear on the screen ('members of this site or publicly viewable'). The participants then embark on a search looking for another way to control sharing. Phil's first guess is to control sharing through the copyright. Marc looks at the options in the copyright pulldown menu and then clicks on the 'more info' link which then takes them out of the portal and to an information page at Stanford University. Phil laughs ('ahh I don't want a lesson in copyright, can we go back'). When they return to the portal Marc looks at the other details of the document; when it was 'uploaded', 'last changed' and the 'web address'. However, Phil is still concerned with assigning file access controls and asks if 'optional properties' might contain finer-grained control features for document sharing. Marc clicks the link only to find options for entering more metadata.

Phil: But we only have (2.3) Oh that's f- (.8)
 Well, I suppose two ways (then)-
 No, there's only one way of controlling sharing
 (Nothing else) seems to be available in the access

Marc: =Yea

Phil: 'Members of this site' or 'publicly viewable'

Marc: =Yea

Phil: Is there anything else? I mean (.) could you control
 the sharing by the copyright (.3) would be awkward
 (.8)

Marc: Hmmm, probably ((looks at pulldown options))
 What's on 'more info', let's have a look ((clicks more info))
 (.5)
 Stanford (.8)

Phil: ((laughs))

Marc: OK, so we get quite complex resources about copyright
 Overview from Stanford University

Phil: =Ahh I don't want a lesson in copyright

[

Marc: Noo

Phil: Can we go back ((laughs))

Marc: Yea, close that window. So um (.3) Oh and we get now
 some more details here, when we- when it was uploaded,
 last changed (2.0) web address (.5) OK

Phil: 'Optional properties', does that allow us t:o control sharing?
 Uhm, just more metadata

Marc: =No:o just more metadata

Phil: So the only sharing- the only controls on sharing are (.3)
 Uhehm (available to access)

Marc: Yea

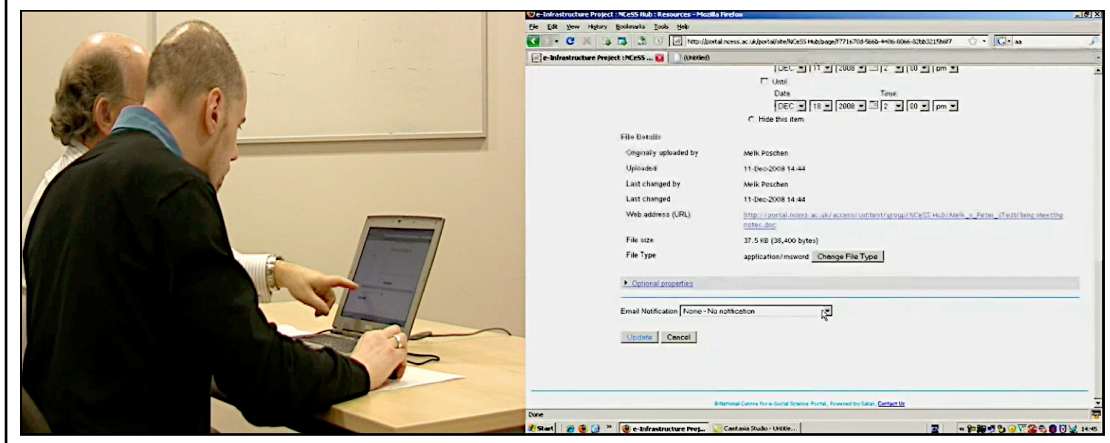


Fig. 2. Document sharing and access controls.

When examining the fragment using the CEF, it is through the sequential analysis of the participants talk-in-interaction which makes visible the troubles participants encounter. For instance, the fragment begins with Phil's identification of a constraint *we only have one way of controlling sharing*. From this point onwards the participants engage in activities that might provide workarounds such as controlling sharing through the 'copyright' or 'optional

properties', however these do not provide a solution. Finally they conclude that there is *only one way of sharing* with that way being 'available to access' which refers to the 'availability and access' section of the web page. The video fragment, the transcribed talk and the CEF matrix were all utilised as resources for the analysis of troubles which in turn were helped to identify potential requirements for the prototype. Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The participants have uploaded a document and are trying to configure its access control.	N/A	<i>Available:</i> The 'edit details' page of the file uploaded. <i>Constraint:</i> There is no feature that allows participants to control the visibility of documents for specific members.	Having two broad options 'members of this site' and 'publicly viewable' may not provide the right level of granularity or security for users if they wish to share documents with a few specific members rather than the whole group.
Cooperative work	Sharing documents may be a sensitive issue and may require fine-grained access controls.	N/A	<i>Constraint:</i> There is no feature for selecting specific member to share and collaborate over files.	(see above)

Table 1: CEF analysis - Document sharing and access controls.

From this the following suggestion for future iterations of the prototype can be made:

Fine-grained access control - design a feature that would allow users greater control over which individual members of the portal have access to files. At the moment, the options of 'members of this site' and 'publicly viewable' are too broad. Providing users the option to control access at a fine-grained level may increase the use of the file sharing feature.

By focusing video analysis on both the topic of conversation and the activities performed simultaneously around talk the troubles participants encounter and the ways in which they address them are identified and analysed in detail using the CEF.

While participants conduct this same activity other system limitations are encountered and in both evaluation sessions (Figures 3a & 3b) the prototype's version control features are identified as potentially problematic. In the following fragment Marc is uploading a new version of a document to the 'resources' area of the portal. Once uploaded he remarks that the file name has remained the same and that ('this a good thing'). He then clicks 'edit details' to look at the file's details and immediately remarks ('okay, we've got last changed date, but there's no versioning'). Phil also remarks on the information displayed ('last changed you've got the new time') and concludes ('right, so you would guess from that that it had been changed'). He then comments that on the page ('it does say last changed by') indicating that the feature does provide useful information about the document's provenance.

Marc: Okay, this was the document
 We say open (.5) file notification *((looks at email notification pulldown))*
 Upload new version *((clicks the 'upload new version now' button))*
 (1.0)
 Okay! (.8) so the good thing is it has uploaded it but
 it hasn't changed the name *((looks at Phil))*

Phil: No

Marc: This is a- this a good thing (.5) Um but, of course,
 Well if you se- if you look at the details let's see what we have down here
 (.8)
 Okay, we've got 'last changed' date
 But there's no versioning *((looks at Phil))*

Phil: Last changed you've got the new time

Marc: Yea, we've got the new time

Phil: Ahum, (1.3) right, so (.8) you would guess from that
 that it had been changed

Marc: Yea

Phil: Yes, it does say last changed by, OK

Marc: Last changed, last changed by whom, yea

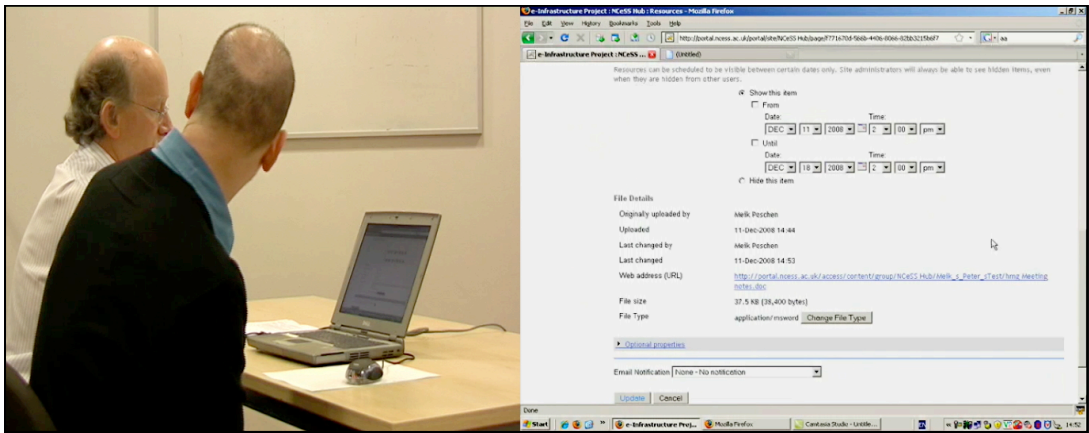


Fig. 3a. Version control.

The feature as it currently functions might require collaborators to mark down the times they changed a document. Even though the upload button is labelled with ‘upload new version now’ the notion of uploading a *new version* is not followed through in the system. Rather the system only displays the *time* that the document was ‘last changed’ and not its version number. So, although information about the document’s provenance does appear by displaying the name of the person and the time the document was last changed, it does not provide information about its particular state within the editing process itself (e.g. it’s version). Because the participants have initiated a discussion of *the absence of versioning as a topic* it takes on relevance because the topic of talk identifies a kind of *imposed relevance* which is identified by the participants as the absence of version numbers.

Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The participants have uploaded a new version of a document and comment on the absence of versioning information.	N/A	<i>Available:</i> The ‘edit details’ page of the resource uploaded. <i>Constraint:</i> There is no indication of the version of the document.	Create a feature that enables automatic versioning. For instance, if the new document has the same title, append the ‘edit details’ page with a version number.
Cooperative work	When distributed participants collaborate on document writing they are not concerned about the time it was uploaded, they are more concerned with knowing which version it is.	N/A	<i>Constraint:</i> When collaborating remotely with documents, participants will need to know which is the most recent version to prevent overwriting newer material.	(see above)

Table 2: CEF analysis - Version control.

From this the following suggestion for future iterations of the prototype can be made:

Version control - provide a mechanism for document version control. This might be achieved through the introduction of a version control system similar to those used in software development environments such as Concurrent Version System (CVS).

Additionally, the feature may also retain old versions of files on the systems (with an option to delete) and mark in sequential order according to the time new files with the same name are uploaded (v.1,2,3, etc).

The fragment above reveals that the prototype may not provide the necessary details that researchers require when collaborating around documents with multiple participants. Similar to this, in the next fragment Derek comments on the absence of versioning within the prototype. However, his remarks about it are somewhat different to Marc's when he states that ideally the system ('will just rename or at least tells you that this is the newer version'). Lisa counters this suggestion with ('you've got the date modified') but Derek refuses the suggestion when he shakes his head slightly 'no' and responds that such a feature may be good if it is used amongst a group of people who are ('working every hour') on editing documents. After justifying his reasons he concludes with ('it's just odd').

Derek: The ideal thing I think is, even if its uploaded a new version
It will just rename (.3) Or at least tells you that this is the newer version
Rather than working-
Lisa: =()You've got the date modified ()
Derek: Yea *((slightly shakes his head no))*
That might work when you- It automatically change dates
For example like in JISC, I could imagine 'cause their working
(with those) every hour *((laughs))*
Lisa: "Yea"
....
Derek: In this case you have to go through (.) or duplicate, it's just odd



Fig. 3b. Version control.

In this fragment the more experienced participant Lisa, who is also a member of the development team suggests that the 'date modified' indicator should be enough. But Derek does not agree, indicated by his head shaking 'no' and a further comment on when the 'date modified' feature may work; within high turnaround environments, such as 'JISC', where people are modifying documents constantly throughout the day and even ('every hour'). In this session, one participant suggested that the system change the file name of the new version whereas in the previous fragment a participant indicated that not changing the file name was a good thing. However, in both sessions each remark that the system should indicate the document's version number suggesting that the date modified feature may not provide the sufficient level of information needed when collaborating over documents through the portal.

Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The participants discuss the uploading of a new version of a file on to the system.	N/A	<i>Available:</i> The 'edit details' page of the resource uploaded. <i>Constraint:</i> There is no indication of the document's version.	Same as Fig 3a with the following addition: Offer an option to change the name of newer versions of the same file uploaded to the system.
Cooperative work	It is suggested that perhaps in fast-paced work environments where documents are changed every hour the 'date modified' feature may be useful.	N/A	<i>Constraint:</i> When collaborating remotely with documents, participants will need to know which is the most recent version to prevent overwriting newer material.	N/A

Table 3: CEF analysis - Version control.

This fragment identifies an additional feature that could be incorporated into a version control component:

Append file name - provide an option as part of a version control feature for those who prefer to let the system change the name of newer versions of documents.

In the two fragments above, two participants in different evaluation sessions had suggestions about how the system could work (to change the file name) or (not change the file name). These differences are important to identify and analyse using the CEF where they can be documented through the conversational manuscript, the video fragment and the CEF matrix. These can then be shown to project members and design decisions made. Even though each participant may have had a slightly different technical solution they both remarked that the system should at least indicate the version of the file. A Google Docs-type solution was discussed during the post-evaluation interviews which would allow all documents to be created, shared and edited using the online system without having to download and upload files from a personal PC. Providing robust file management, specifically version control may be beyond the scope of what is available in the Sakai platform. If this is the case, the limitation of open source software would need to be highlighted when it comes time to discuss the level of investment projects commit to when purchasing software including the pros and cons of using either open source tools or purchasing more advanced file management systems from Sun or Oracle, for example, which do provide many advance file management features.

In both fragments the participants encountered limitations when conducting activities using the system. The first involved the lack of fine-grained access controls for sharing and the second involved the lack of information available about a document's version. In each case, the fragments were identified when watching the video during the analysis phase. A textual summary of the identified troubles was then emailed to two participants who volunteered their time (Marc and Derek) to review them. This gave participants of each session an opportunity to provide comments about the analysis. Unfortunately, it was not possible to go through the video fragments with each of them because of their prior time commitments and their, in general, very busy schedules.

8.3.2 Getting to grips with system idiosyncrasies

The previous section discussed the ways in which participants worked through system limitations and in the process they suggested technical solutions, which were identified as potential requirements for the next iteration of the prototype. In this section, participants work through some of the prototype's idiosyncrasies. Specifically, the workflow for creating a wiki subject area and system messages encountered by participants once they've created the area. Both fragments reveal how many of the customisable features of Sakai (the software used to develop the web portal) could be modified further to accommodate the expectations of people who may want to use the system.

In the following fragment Lisa has just explained to Derek the workflow for creating a wiki page. She asks him if he understands ('why you create the link like that'). Derek begins his reply as a sort of technical response ('to link it to say to the document') but interrupts himself and instead asks Lisa a question ('is it a link to another part of the system?'). In response Lisa again explains how a wiki page needs to be created in Sakai ('what you do in Sakai is you basically create a link and it creates a page for you') and Derek responds with a vocalisation and exaggerated gestures which suggest that the system's workflow may be confusing but that now he 'gets it' or understands. Although Lisa explains the system's 'logic' to Derek several times he does not fully grasp the way in which it functions ('when you say *link* it's a new wiki page link to the prior whatever'), Lisa nods her head in agreement.

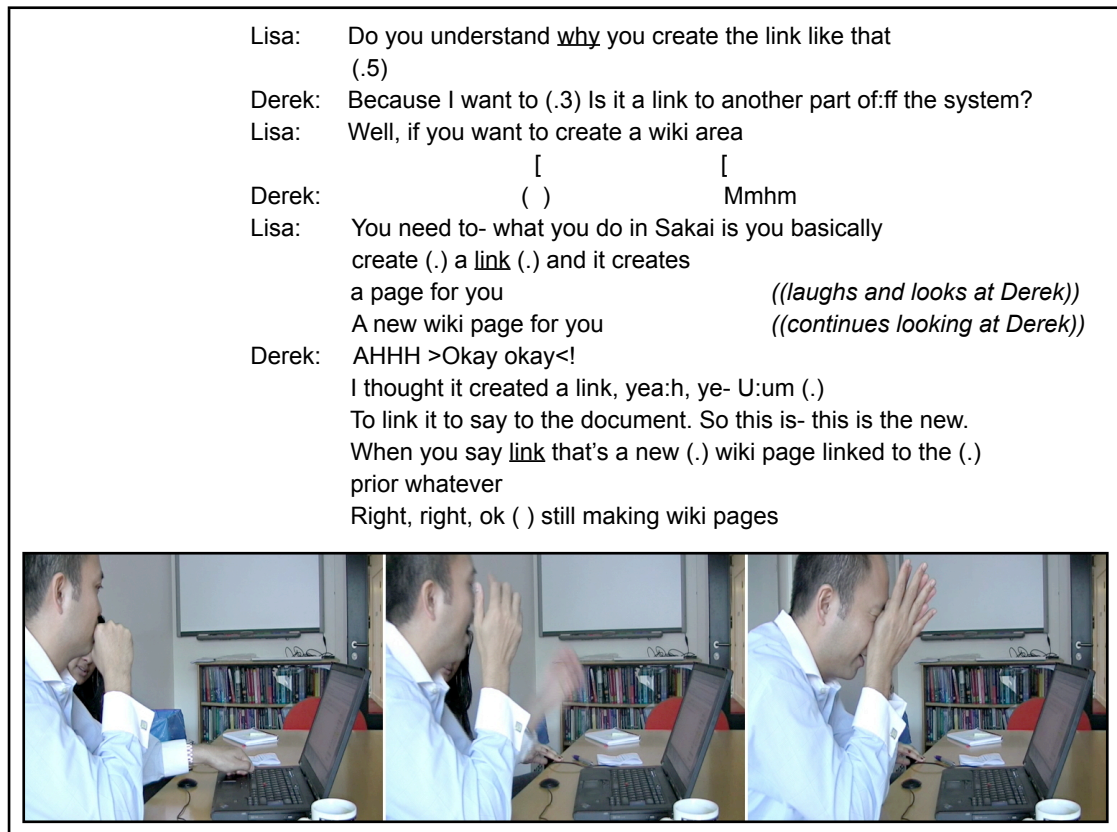


Fig. 4. Understanding the system's workflow.

Similar to Figure 3b, the more experienced participant Lisa, who is a member of the portal development team tries to explain the system's 'logic' to Derek, an epidemiologist working in the Obesity e-Lab project. Lisa's laugh as she explains the procedure may indicate the randomness of the system's 'logic'. In most applications when a user wants to create a new page or a 'new' *anything* there is more than likely a 'new' feature available (e.g. new document, new mail message, new image, etc). In this case, the workflow is unfamiliar to the ways in which most software supports this type of activity. In order to create a 'new' wiki page first a link to that page must be created and then by clicking on that link it will take the user to the newly created wiki page. The epidemiologist tries to understand the system's workflow and this requires some explanation because it is so different from other familiar software application workflows.

Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The participants are discussing the procedure for creating a new wiki page.	N/A	<i>Available:</i> Wiki homepage <i>Constraint:</i> The process for creating a new wiki page is not familiar and not presented in the usual manner expected to create a 'new' file.	Change the workflow so that in order to create a new wiki page, the user can select 'new' from the main edit page.
Cooperative work	N/A	N/A	N/A	Providing an understandable procedure (e.g. create a feature - 'new wiki page') might increase its use allowing social scientists to share information using the wiki.

Table 4: CEF analysis - Understanding the system's workflow.

From this, the following suggestion for future iterations of the prototype can be made:

System workflow - The workflow for creating a new wiki subject area is unfamiliar and does not function in an expected manner. Most, if not all software applications provide users with a 'new' feature (e.g. new document, new mail message, new image, etc). Provide a 'new' feature that would allow members to create new wiki pages in an familiar manner. Because the untrained or unassisted user may not understand the process, use of this feature may not be as high as the project would hope if it remains as it is.

Once the process is explained, inexperienced users face another challenge. In the next fragment Phil is trying to understand the process of creating a wiki page. The pair have just completed the procedure and he repeats back to Marc his understanding of the system's workflow ('editing the homepage of the wiki to add another heading') Marc overlaps this talk with a laugh and replies ('yea'). Phil displays his lack of understanding of the system's workflow with ('but, you tell me, we still haven't got a space'). The only item the participants have created is the page title, which is now a link on the wiki homepage. Marc explains to Phil that when the title link is clicked ('we now get this information'). However, when the link is

clicked the system's message indicates that the 'page does not exist'. Both participants then discuss the implications of presenting this kind of information to users.

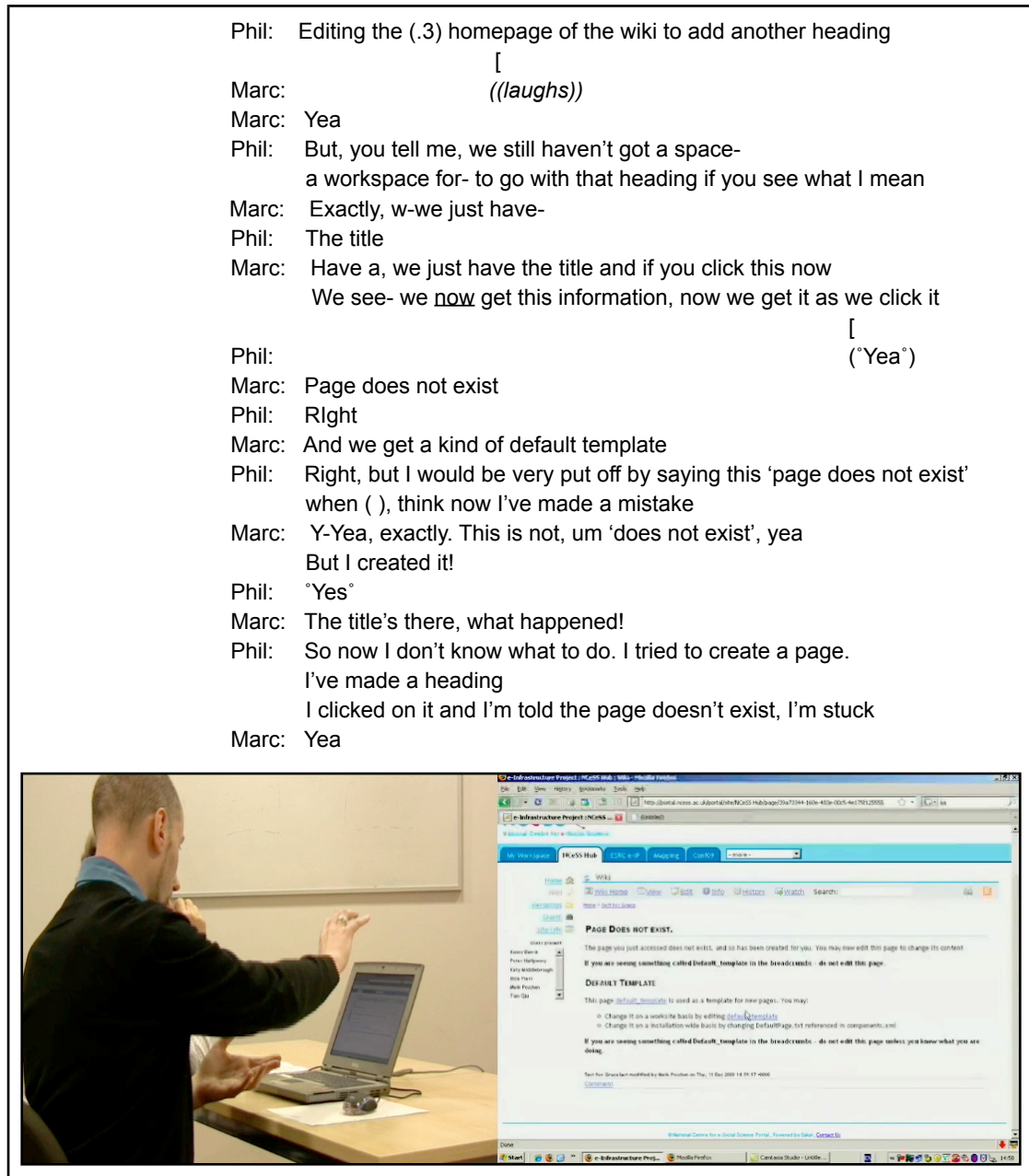


Fig. 5. This page does not exist.

Here, the system message is unfamiliar and unexpected and both participants agree that it is easy to get 'stuck' when confronted with the message 'page does not exist'. Later in the evaluation session they decide that the message 'page does not exist' should be replaced with 'this is a template - you can now edit it'. This simple change to the system message provides a

more accurate description of the page and its purpose (an editable template page), thus increasing its understandability.

Applying the CEF as a guide to video analysis the fragment may be analysed as follows:

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
Work practices	The participants create a link to a wiki page and then click on the link to view the page.	N/A	<i>Available:</i> Newly created wiki page <i>Constraint:</i> The system message 'page does not exist' does not explain the purpose of the page.	Rename this system message to something that describes the purpose of the page, one suggestion provided by participants: 'this is a template - you can now edit it'.
Cooperative work	N/A	N/A	N/A	Displaying understandable system messages may facilitate more widespread use of the system.

Table 5: CEF analysis - Page does not exist.

From this the following suggestion for future iterations of the prototype can be made:

System messages - The 'new' wiki page system message should convey the purpose of the page as well as provide prompts for further action. Messages such as 'page does not exist' may confuse users. Additionally, conducting a review of all system messages would be useful to ensure that each is understandable (e.g. replacing 'page does not exist' with 'this is a template - you can now edit it').

Both fragments in this section reveal the activities of participants as they encounter unfamiliar workflows and system messages. The CEF has revealed that many of the 'out-of-the-box' features of Sakai have been identified as problematic for participants as they work through activities. In section 8.3.1, the participants' discussions with each other about the lack of fine-grained access controls for sharing and the lack of information available about a document's version provide insights for further requirements that would give users greater control. In this

section, the unfamiliar workflow and system messages present obstacles as participants worked towards the completion of activities.

For the prototype evaluation sessions activities were conducted that were not integrated into the day-to-day activities of the social science researchers who participated. Rather, two inexperienced users of the system, Phil and Derek were paired with members of the NCeSS e-Infrastructure team, Marc and Lisa. The value of conducting these sessions was to develop an understanding of where the system failed to make sense to, in this case, novice users. Additionally, the feedback provided by participants after the evaluation helped to clarify exactly where troubles had arisen. The post-evaluation interviews will be discussed in the next section.

8.3.3 Post-evaluation feedback

At the end of each evaluation session I asked participants to describe which activities they felt were successful and which were unsuccessful. However, in both sessions the discussion focused primarily on the difficulties encountered when using the prototype. In the first session with Marc and Phil the discussion began with their difficulties in understanding how to use the wiki. They suggested that if more information was provided for users on each page that it may improve its understandability. We discussed the current workflow where users must press ‘edit’ first to create a *new* wiki page and at this point I made a suggestion (‘then it should be something like *new*’) and Marc and Phil agreed. The feedback session proceeded in this manner where Marc and Phil would identify issues, they would provide suggestions or I would make a suggestion to them and note the opinions of the participants as potential requirements.

In addition to discussing their difficulties, Marc also walked me through several of them. We started with the wiki where he stepped through the workflow and messages. He pointed out that the system appends a newly created page title with a question mark (indicating that the page has been made, but not been edited). He also noted the ‘this page does not exist’ message as potentially problematic. We then discussed the ‘sharing a document’ activity where the participants commented on the lack of access controls, the lack of clarity regarding the purpose

and functionality of the 'hide' button and the 'show item' features. Phil commented that the whole page, 'edit details', was confusing. We then discussed the 'edit document' activity where the participants discussed the inconvenience of having to download the document to their PC preferring a Google Docs-type feature where documents creation and editing could be integrated into the system. After the activities were discussed we ended the session and it was agreed that I would write a list of potential requirements and send them in an email for confirmation and amendments.

In the second evaluation with Derek and Lisa the discussion began in a similar manner. First, I asked participants to describe which activities they felt were successful and which were unsuccessful. Derek began the discussion with his observation that it may be just a matter of familiarising oneself with the system features, implying that not knowing how to use the prototype is not so much a 'usability' issue as a training one. I then asked them if there were any specific things that they would like to flag as 'potential usability issues'. As with the first session, the discussion began with difficulties associated with using the wiki. Lisa began by remarking upon ('the navigation on the top') noting that the mapping of the breadcrumbs was inaccurate. Interestingly, this was not discussed amongst the participants during the evaluation session. Soon after this Derek remarks that ('it also takes a bit of wiki experience to make the most of this. it took me awhile to get the hang of the concept again') implying that 'lack of experience' and familiarity was an issue. When discussing the 'upload file activity' Derek noted that the system does not indicate the file's version and that it also retains the same document title. In the 'edit activity' Derek added that it wasn't clear if it was possible to edit online (e.g. Google Docs) suggesting that it would be a good feature. Additionally, when looking at a file's details we discuss the label 'edit details' as possibly confusing and that it could be changed to 'details'. Towards the end of the feedback session Derek concludes with the same sentiment he expressed at the beginning ('it's all there, it's just a matter of familiarising yourself with the features'). Similar to the first evaluation, after the participants gave their feedback about each activity we ended the session and it was agreed that I would write a list of potential requirements and send them in an email for confirmation and amendments.

Through a combination of the evaluation session and the post-evaluation feedback a summary of requirements were identified using the CEF:

Fine-grained access control - design a feature that would allow users greater control over which individual members of the portal have access to files. At the moment the options of 'members of this site' and 'publicly viewable' are too broad. Providing users the option to control access at a fine-grained level may increase the use of the file sharing feature.

Version control - provide a mechanism for document version control. This might be achieved through the introduction of a version control system similar to those used in software development environments such as Concurrent Version System (CVS). Additionally, the feature may also retain old versions of files on the systems (with an option to delete) and mark in sequential order according to the time new files with the same name are uploaded (v.1,2,3, etc).

Append file name - provide an option as part of a version control feature for those who prefer to allow the system change the name of newer versions of documents.

System workflow - The workflow for creating a new wiki subject area is unfamiliar and does not function in an expected manner. Most, if not all software applications provide users with a 'new' feature (e.g. new document, new mail message, new image, etc). Provide a 'new' feature that would allow members to create new wiki pages in an familiar manner. Because the untrained or unassisted user may not understand the process, use of this feature may not be as high as the project would hope if it remains as it is.

System messages - The 'new' wiki page system message should convey the purpose of the page as well as provide prompts for further action. Messages such as 'page does not exist' may confuse users. Additionally, conducting a review of all system messages would be useful to ensure that each is understandable (e.g. replacing 'page does not exist' with 'this is a template - you can now edit it').

From the post-evaluation feedback the following were also identified:

Unknown symbols - Placing a question mark after a page title link is created (e.g. 'wiki page ?') is a system message that causes confusion. This could be changed to (e.g. 'wiki page - click to edit').

Hide & show item - There is no explanation about the purpose or functionality of these features. Provide explanatory text in close proximity with each.

Web-based documents - Provide a feature where documents can be created, shared and edited directly within the system (e.g. Google Docs).

Details page - Change the name 'edit details' to 'details'. It is not clear that within 'edit details' users can also 'view details'

Internal help pages - Provide 'help' documentation as well as indications on each page describing what types of activities are possible. The 'more info' link from the copyright feature which links to Stanford should be removed because it does not provide immediate help but is rather related to a copyright tutorial.

Incorporating each of these potential requirements³⁸ could improve the system's capacity to support distributed file and resource sharing amongst social scientist who choose to use the NCeSS portal. Even so, it must be stressed that these evaluation sessions examined the prototype's usability at the interface and not its usefulness as a tool that might be integrated into the day-to-day practices of social science researchers. Although this was the case, the evaluation and feedback sessions did uncover substantive requirements and gave participants an opportunity to participate in the prototype's design both through using the system and in providing suggestions for its improvement based upon their reflections. The final section of this chapter will discuss the use of the CEF in the NCeSS e-Infrastructure project.

8.4 The use of the CEF in the NCeSS project

This chapter has described the use of the CEF in what may be described as a technology-driven project in that the software used to develop the portal, Sakai, had pre-determined or default workflows and system messages that were implemented in the NCeSS portal as they functioned 'out-of-the-box'. This meant that unless considerable time was allocated to re-designing those default settings social scientists interested in using the system would need to be trained in its functionality in order to conform their activities to what could be achieved by the system. As a result, ordinary activities such as sharing documents, uploading files and communicating information through the creation of a wiki page were all rendered in some way problematic when using the portal. In this case study the use of the CEF was limited in that the evaluation sessions investigated the *usability* of the system at the interface, and not its *usefulness* as an embedded system that would support routine work practices. As Bloomberg *et al* (1996: 239) note; *usability*, refers to the system's general intelligibility particularly at the

³⁸ The inaccurate breadcrumb trail which Lisa remarked upon was due to technical 'work-in-progress' and so was a temporary software bug rather than a usability issue related to the intended functionality of the prototype.

interface and *usefulness*, refers to the ways in which a system's functionality actually makes sense and adds value in relation to a particular work setting. Even though the evaluation sessions were conducted in this way they did reveal that the purpose and functionality of some of the prototype's features were not immediately understood by participants. These difficulties could be identified and workflows potentially redesigned before the system becomes publicly available which may ensure greater adoption when it eventually is embedded into actual work settings.

The CEF proved useful for revealing participants' moment-to-moment sense-making procedures as they worked through each activity. In this case, longer video fragments are presented in order to reveal the unfolding mutual understanding between participants as they work through activities. As a result, their common-sense understandings of the prototype could be observed, which provided access to features of the web portal that presented specific troubles as they worked through activities. Obtaining access to developing common-sense understandings of the prototype could only be achieved by engaging pairs of participants in collaborative activities thus enabling discussions around both the activity and how the system may or may not support its accomplishment. In this way, these sessions functioned similar to collaborative design activities used in PD (e.g. cooperative prototyping) where participants engage in each activity in a reflective manner for the purposes of making design suggestions to improve the prototype's interface and functionality. This was made possible through the configuration of the participant pairs as developer/social scientist pairs. Especially during the post-evaluation sessions, the developers of NCeSS (whom were paired with social scientists) could observe difficulties first-hand and discuss them in detail. These sessions also differed from the other two case studies (BVREH and UK-NEES) in that they incorporated a post-evaluation feedback session. This provided participants with an opportunity to walk through the details of specific troubles in a reflective manner in order to inform design so that the understandability of both workflows and system messages (e.g. version control, wiki workflow, 'page does not exist') could be improved upon.

When setting up the sessions, the participant pairs were unusual in that for each an experienced member of the portal software development team was paired with, in one case, a social scientist (Phil) and in the other an epidemiologist (Derek). This was due to difficulties in finding social scientists willing to participate. In an effort to recruit participants I emailed the lead researcher of each of the NCeSS nodes explaining the details of my project and asking if they would be willing to participate. All nodes stated they were too busy, with the exception of Derek from Obesity e-Lab and so this resulted in the particular pairings used in this case study. Securing participation is a key challenge and difficulty that may be faced when trying to engage domain researchers in the software development process and prototype evaluation. This is primarily due to the heavy workloads of those who may want to participate and obstacles to finding time where they can take a few hours in the course of a project to provide feedback. This may be remedied by explicitly allocating prototype evaluation activities into a project proposal, budgeted and included in a project's Gantt chart.

When conducting the sessions, I asked each participating pair to install Camtasia³⁹ on to their computers so that screen activities could be recorded. Unfortunately in the second session, Derek checked his email before the start of the session and inadvertently turned off the screen recorder. I had already left the room at this point so it was impossible to identify until I had returned to the room at the end of the session and was beginning to pack up. Even though this was the case, it was still possible to assess the video based upon my knowledge of the system and the activities that they were performing for the session. The pairs were given the same activities to conduct which provided an opportunity to compare and contrast the two sessions. For instance, in the first session Marc stated that it was a good idea not to change the file name when a new version had been uploaded, whereas in contrast to this, Derek remarked that such an option would be useful. Also, in both sessions each remarked on the usefulness of providing 'web-based' documents with the ability to create and edit online similar to Google Docs. Whether feedback from participants is in some way contradictory or shows consensus, the CEF provides a mechanism to identify these so that they may be documented and eventually prioritised as 'essential', 'optional', etc when documenting requirements.

³⁹ <http://www.techsmith.com/camtasia.asp>

The use of video, which was analysed in detail after the sessions, in conjunction with post-evaluation interviews complimented the requirements identification process. For instance, in the first session, Marc and Phil did not mention version control in the post-evaluation feedback session although during the evaluation it was discussed by the pair in some detail. This is a good example of how discussions alone may not reveal all the difficulties participants may encounter. I emailed participants a list usability issues they reported that they experienced during the session and potential requirements that may solve them. The participants then made amendments if necessary and then I emailed the list to the project manager. There was no opportunity to show the video to either the project manager or the lead software developer because of tight constraints on their time. In any case, they indicated that they were primarily interested in the bullet-point list and not so much in the video data. Unfortunately, it was also not possible to go through the video fragments with the participants because of their prior time commitments and very busy schedules. Similar to securing initial participation, this might be remedied by explicitly allocating post-evaluation activities such as video data sessions into a project proposal, budgeted and included in a project's Gantt chart. When video data sessions cannot be held because of time constraints both software developers and participants miss an opportunity to develop a shared understanding of the organisation of actual activities as they unfold when using the prototype. Additionally, they miss an opportunity to share design ideas and solutions as a result of viewing selected video fragments. Because video is qualitative data it may be seen by some as unnecessary for the design process, hence the request - "just show me the bullet-point list". Here, the challenge remains to demonstrate to software engineers and others that video data sessions can produce meaningful, and at times unexpected, design solutions.

When using the CEF matrix for video analysis, there were no 'material artefacts' used making that column of the matrix redundant for this case study. Additionally, the 'cooperative work' row of the matrix did not capture *actual* cooperative working but rather was used to suggest how the identified requirement *might* support cooperative work. In general, when novice participants were trying to make sense of the prototype, having a more experienced person paired with a less experience one made visible their sense-making procedures through

their conversations with each other. Again, this is similar to the PD cooperative prototyping approach where mutual learning takes place during the session (developers understand difficulties using the system and participants understand what the system is capable of doing). One of the participants, Derek, summarized the experience as needing to familiarise himself with the system and it appears that offering training may address some issues. However, the *usefulness* of the NCeSS e-Infrastructure web portal can only be understood by conducting an evaluation of technology-in-use as it is embedded in actual work settings where social scientists may use the prototype as part of their day-to-day work routines.

Creating an easily understandable system workflow and messages might increase the likelihood that those social scientists who are interested would take the time to learn the system. On the other hand, if the system is incomprehensible from the start it is less likely that busy researchers will take time to work through getting to grips with how it operates. In sum, the post-evaluation interviews were a key feature in these evaluation sessions. Because there was no *in situ* fieldwork with which to compare the ways in which the prototype might have supported, constrained or transformed day-to-day work practice, the post-evaluation discussions provided access to the participants reflections of their experience using the system and provided them with an opportunity to make suggestions for its improvement.

For this case study the Contextual Evaluation Framework was utilised in a similar manner to the PD notion of collaborative design especially as implemented in cooperative prototyping where participants conduct activities in the presence of software developers and where they also, at times, take a reflective stance in order to offer design solutions during the session. In this manner, the CEF was adapted to the circumstances of the NCeSS software development project so that participants' mutual understanding of system features could be revealed. The video data of participants' talk-in-interaction was also analysed in detail as was done in the previous two case studies and provided access to the details of troubles that may not have been reported during the post-evaluation session (e.g. Marc and Phil did not mention version control in the post-evaluation feedback session although during the evaluation it was discussed by the pair in some detail).

This, and the previous two chapters presented the ways in which each case study shaped the development of the CEF which was loosely envisaged and described when this research began. The next chapter will present the CEF as a systematic and rigorous technique, developed as a result of my experiences in the three case studies detailed above, providing detailed instructions for e-Research project members interested in using the approach to evaluate prototypes.

Chapter 9:

A practical guide to the Contextual Evaluation Framework

9.1 Introduction

The three case studies detailed above have provided me with both the experience and knowledge with which to explore concepts, ideas and issues related to the development of the CEF through its utilisation in three e-Research projects. Each case study presented its own unique challenges to incorporating the technique that, in turn, has informed the design of a systemized approach which may be used as a practical guide for conducting sociologically-informed prototype evaluation. In each case study, the CEF was used to identify emerging requirements related to system functionality and interface usability thus contributing to each project's software development process through sharing those findings and through facilitating the development of a sensitivity around the practices and activities that people rely upon within each research domain. The case studies also contributed to my own conceptualisation of the CEF's usefulness in a range of evaluation settings within the humanities, the sciences, and the social sciences.

For instance, in the BVREH case study, a co-located evaluation session was conducted which demonstrated the CEF's usefulness in revealing the ways in which both collaborative activities and tacit practices were supported through the prototype. Additionally, the CEF was used to facilitate collaborative interaction amongst traditionally solitary scholars. In the UK-NEES case study, the CEF was implemented to evaluate a prototype in a synchronous and distributed setting. The case study also demonstrated the usefulness of the CEF in evaluating prototypes in the absence of practice and that a single-sited approach to understanding the

coordinated activities of multiple distributed participants could be used to acquire meaningful results in synchronous and distributed prototype evaluation. Finally in the NCeSS case study, the CEF proved useful for revealing participants' moment-to-moment sense-making procedures as they worked through each activity during the prototype evaluation session. These sense-making practices are identified through observations of people as they *interpret the meaning of appearances* (cf. Garfinkel, 1967: 78), which is achieved through the analysis of talk-in-interaction, and in this case, was revealed within the context of conducting activities using a web-based virtual research environment. Such practices are important to understand because participants' interpretations inform the production of their subsequent actions (discussed in §4) and so, in this case, may be understood in terms of assessing a software prototype's usability and its subsequent support for action.

The aim of this chapter is to provide a practical guide with which interested members of the e-Research community may incorporate the CEF prototype evaluation technique within their own projects. The CEF is a sociologically-informed quasi-naturalistic technique composed of five elements:

Pairs of participants - Two participants are asked to evaluate a prototype together. This is done in order to reveal, through an analysis of talk-in-interaction, both the ways in which participants make sense of the system and how they may use its features and functionality to accomplish routine and cooperative work.

Routine work activities - The pairs are asked to work through a selection of routine work activities that are identified through prior fieldwork (e.g. EM-informed *in situ* observation, interviews, document analysis) and through consultation with the participants. Selecting routine activities will reveal the ways in which a prototype might support, constrain or transform these day-to-day practices.

Naturalistic setting - The session may take place in the participants' actual work setting or in one that simulates that setting. In some cases it may be possible and even necessary to embed a prototype directly into a work setting alongside already existing technical infrastructures as was the case for the UK-NEES project.

Video record - The video recording of each prototype evaluation session is done in order to capture participants talk and interaction as they work through routine activities using the prototype. Video data allows for the analysis of interactions in detail so that themes and patterns of work and interaction may be identified as well as to identify participants difficulties when using the prototype.

Ethical approval - Before conducting any fieldwork ethical approval from within the organisation is required. If the work is undertaken at a university or research institute there is usually an application process where the purpose and duration of the project is presented. If the work is undertaken within a commercial setting check with the management team and human resources. In either case, include the rationale for project and the purposes for recording participants. Also provide written assurances to follow Data Protection guidelines⁴⁰ when recording audio-visual materials which includes the details of storage, distribution or the future publication of materials. Added to this, each participant must read and sign a consent form where they formally agree to be video recorded and where they may also give consent for either the audio and/or video to be used in presentations and publications. It is important to emphasize to the participant that they may request the recording to be stopped at anytime and that, upon their request, footage already recorded will be destroyed.

The CEF is a highly empirical approach to prototype evaluation which emphasizes conducting observations of a prototype-in-use within a naturalistic setting and situation of use. This is in contrast to traditional HCI evaluation techniques, which examine how a system operates as defined from a high-level requirements specification, thus constraining participants to the evaluation of tasks that are pre-assigned by a software development team and aimed at validating functionality already implemented.

Through the analysis of video data informed by EM/CA/IA an understanding can be developed of how a prototype supports, constrains or transforms routine work. Additionally, video fragments provide evidence, through the identification of particular instances, which may reveal the difficulties participants encounter and patterns of interaction that might be transformed. Instances where a prototype takes on an *imposed relevance* are signposts that may reveal where functionality is absent or configured in an unexpected manner, or where the interface is not easily understood. The following sections provide general guidelines for conducting a CEF prototype evaluation with examples taken from the three e-Research case studies.

⁴⁰ http://www.ico.gov.uk/for_organisations/data_protection/the_guide.aspx

9.2 Get to know the domain: conducting initial fieldwork

Developing an understanding of the activities and interactions within a research community as they exist prior to a prototype's design and introduction into a work setting will allow e-Research project members to familiarise themselves with the domain for which the system is being developed. A good first step in this process is to make contact with a lead researcher who is involved in the e-Research project who could also introduce you to the particular research area and to other researchers. From this, initial interviews may be conducted in order to understand the nature of the domain and as well as the purposes and aims of the e-Research system. When conducting interviews it is useful to have questions written out beforehand. Prepared questions can work as a device to get the conversation going and to steer its direction. However, certainly do not constrain yourself to the question list because many unanticipated topics may arise during the conversation that you may want to pursue. A few sample questions include: 'Provide an overview of your research', 'What are the aims of the system under development', 'How do you envisage it supporting your research', 'What kinds of tools do you currently use in your research process', 'How do you collaborate in your research'. During the interview also ask for a good introductory paper in the research subject area that you could read to familiarise yourself with the field of study.

In addition to interviews and informal conversations, the next step would be to ask if you could observe the researchers as they conduct their routine work activities. Before conducting *in situ* fieldwork and video-based studies, you may need to acquire ethical approval from either your own or the case study organisation. In all cases, informed consent should be obtained through a signed form which explains the purposes of the study, confidentiality, data protection, as well as asking explicit permission to share video, audio and other data acquired as part of the case study. Observation within the research setting will provide an opportunity to see how the work that you discussed during the interview is actually carried out in practice. Additionally, it may also provide an opportunity to ask questions as they go about their work. Pay particular attention to the procedures they follow, the material artefacts they use and the collaborative interactions that occur within the setting.

It is important that your case study participants build a sense of familiarity and trust with you so that it may be possible to begin video recording their work practices. Some participants may be reluctant at first, however, it is important to stress that any video taken would be strictly confidential and used only to understand work practices for the purposes of informing the design and evaluation of the system under development. It is also important to explain the benefits of video recording as a means to acquiring a detailed understanding of tacit, seen-but-unnoticed practices that may go undocumented. These undocumented practices could have consequences for the usefulness and usability of the system if support for them is not specified.

Conducting initial fieldwork will also provide a comparison between how the practices observed prior to a prototype's evaluation are supported, constrained or transformed. From this, a prototype may be assessed for how well it is aligned to the ways in which researchers themselves organise activities. In this way, it may be evaluated for its *relevance* or *praxeological validity* (its value to the working practices of those who might use the system). From this initial fieldwork, potential requirements for the system may be identified. For example, in the UK-NEES case study having access to the ambient properties of the setting (e.g. multiple views that enable visual inspection, monitoring cracking sounds and having ambulatory access around the experimental area) were identified as essential to making observations during an experiment. Also, in the BVREH case study the use of gestures, tracing and drawing of candidate letters were identified as aspects of substantive research practice. In both of these case studies however, these complex activities could not be supported in the proof-of-concept prototypes that were developed. However, having access to these descriptions of work practice may inform future iterations of both systems.

9.3 The equipment: preparing for an evaluation session

When the initial prototype is ready for evaluation it is important to have the right equipment available which can capture the interactions between participants, a view of the computer display and the conversation between the pair as they conduct activities. First, a reliable video camera is needed. This would ideally be a high-definition camera that records onto tape. The video recorded on to tapes can be transferred to a computer desktop and the tapes stored in an

archival library. Recording in high-definition mode will also allow you to select and view close-up areas that may be of interest such as the ways in which artefacts are handled. For these case studies a Canon MiniDV Camera HV-30 was used.

In addition to a camera which records the participants' activities within a setting, the computer screen should also be recorded. Tools for screen recording are available for Windows machines using Camtasia⁴¹ and for Apple machines using Silverback⁴². For these case studies Camtasia was used in the BVREH and NCeSS evaluations because they were conducted on Windows machines. It is important to note that screen recording may not always be possible, as was the case in the UK-NEES case study. In this instance every bit of computer processing speed was required to perform experiments. Although this was a limitation, having a good angle on both the participant and the screen provided adequate results. Even when using screen recording software ensure that the video camera is angled in such a way as to record the participant *and* the computer screen. In cases where it may not be possible to install screen recording software you may choose to use two video cameras, one focused on the participants and the other on the computer screen.

When recording conversations between participants the built-in microphone from the video camera works well. Additionally, you may wish to attach a uni-directional boundary microphone. For these case studies both the video camera microphone and an external boundary microphone were used. Ideally, a pair of wireless microphones would provide the best sound quality. If the sound needs enhancing SoundSoap⁴³ is a good tool for removing unwanted noise and for enhancing vocal sound. These key pieces of equipment; video camera, microphone, screen recording and audio restoration software were used in the case studies and are necessary pieces of equipment for conducting a CEF session. For a detailed exposition of the more technical aspects of equipment and its set-up, which is beyond the scope of this thesis, see Heath *et al*, 2010.

41 <http://www.techsmith.com/camtasia.asp>

42 <http://www.silverbackapp.com>

43 <http://www.bias-inc.com/>

9.4 The naturalistic setting: identify the participants and place

When a prototype is at the stage where it may be assessed it is important to conduct evaluation sessions within naturalistic settings. First, the appropriate participants need to be identified and these are usually the very same participants who were observed when preliminary fieldwork was conducted. It is important that the target end-users of the system evaluate the system because it is their work that it will need to support. Additionally, the pairs that work together would also ideally be the same researchers who are working together on the research project that the prototype is meant to support. However, because of scheduling conflicts it may be necessary to allocate stand-ins. This is not recommended or ideal, however, it was resorted to for the NCeSS evaluation sessions where a social science researcher was paired with a member of the NCeSS software development team. With these pairings many usability issues were identified but even so, it should only be considered a pairing of last resort. The conversations between the participants will be analysed after the session to gain access to their practical understandings of the prototype as they conduct activities relevant to their research, which is why pairing appropriate participants is essential for producing relevant findings and potential system requirements. Also, if time permits, it is ideal to conduct two to three sessions with different participant pairs. This may provide an opportunity for comparison between the sessions. For example, in the NCeSS case study two pairs were given the same activities to conduct which revealed similarities and differences in their interpretation of the system. These can be discussed with the project team where eventually, design decisions would need to be made.

After the participant pairs are identified then the setting should be agreed upon with input from the participants and based upon the practicalities of the prototypes' capabilities. For instance, in both the BVREH and the NCeSS projects, the prototypes were web-based which meant that the sessions could take place either at the location of work or close to the software development team. For the BVREH session, we decided to conduct it at the software development location because new functionality had been implemented since it was last shown to the researchers and so needed to be demonstrated to them. Also, the prototype was still in a

very fragile state and could require on-the-spot bug fixing. If you decide to conduct the session in a simulated setting ask participants to bring associated artefacts that are easily transportable as was done in the BVREH session where participants brought with them drawing material, reference books and notebooks.

In the NCeSS case study, the location of work was used to conduct the sessions. This was very convenient because the social scientists and the software development team are all affiliated in some way with the larger Manchester eResearch Centre and so work in the same building. Because the offices were open plan it was decided to conduct the sessions in a meeting room so as not to disturb colleagues. Using a meeting room also provided a quiet place with which to video record the sessions and ensured that the conversations were audible.

In some cases the prototype may be embedded into the work setting at the start of a project and remain there as it is being developed. For example, in the UK-NEES project the prototype had to be tested with the actuators, sensors and numerical modelling software. These tools are integrated into the laboratory and so it was necessary that the prototype interoperate with them in order to build a system to support distributed experimentation. For this reason, the evaluation sessions were conducted in the actual setting of work.

Wherever the location, it is important to prepare and test any equipment you decide to use during the session. If possible go to the location a day or a few hours beforehand to locate the best positioning for the video camera so that both participants and the computer display are recorded. Also, testing the room's acoustic levels and how well conversation may be picked up by either the camera's internal microphone or an external one is equally important. Finally, install and test screen recording software on the machines used for the session.

Whether the sessions take place in the actual setting of work or in a simulated setting, remember that the purpose of an evaluation session is to provide participants with an opportunity to use the prototype to accomplish routine activities it is meant to support. Keeping this in mind, as long as the place where the evaluation takes place allows them to conduct these activities in as routine a manner as possible then the setting can be considered sufficiently quasi-naturalistic.

9.5 The evaluation session: identify its purpose and focus

At this point the e-Research development team should have a good working knowledge of the research domain for which the system is being designed, which would have been acquired through initial fieldwork. As mentioned above, the findings from fieldwork may provide an opportunity for comparisons to be made between how the practices observed *in situ* are supported, constrained or transformed when using the prototype. The purpose of conducting a CEF evaluation is to assess how well the prototype's features and functionality are aligned to the ways in which researchers themselves organise and conduct activities.

Identifying the scope of an evaluation session will depend upon the features and functionality available and how well the interface provides access to these. A good starting point for deciding upon the evaluation's focus is to revisit the project proposal document which will include a description of the purpose and goals of the prototype. Identify the work practices that the prototype is meant to support and from this identify activities for the prototype evaluation session. So for instance, the BVREH project broad objectives were to make available through a web-based interface digitized images of ancient documents and manuscripts and to provide tools for their distributed analysis. The aim of the UK-NEES project was to provide a system that would link together three earthquake engineering laboratories so that remote equipment could be shared in order to conduct distributed experiments. In the NCeSS project, for its first release, the aim was to provide tools that would provide distributed researchers with the ability to exchange information through a web portal.

When an e-Research system is meant to support broad research practices then researchers should be able to perform those very same practices using the prototype. So when making a determination about which activities to include in an evaluation session it is recommended that participants are asked to conduct their routine work and research activities. For example, in the BVREH session the participants were asked to “continue your analysis of the front of the Tolsum tablet using the prototype” and for the UK-NEES sessions they were asked to “conduct a distributed numerical experiment”. However, for the NCeSS sessions the participants were

asked to conduct much more focused activities such as “upload, share and edit a document and create a wiki subject area”.

Constraining participants to specific tasks that are pre-assigned by the software development team is not recommended because it may simply serve to validate functionality already implemented and may not validate its actual usefulness to researchers who might use the system. Rather, participants in a CEF evaluation are asked to conduct routine activities for the purposes of assessing the prototype for how well it supports work practices and not for how well it performs as defined from a high-level requirements specification. Even though evaluation sessions will undoubtedly be constrained contingent upon the features available within a prototype, providing participants with an opportunity to use the features that are available to conduct routine work will identify usefulness and usability issues that may not have even been considered by the e-Research project. For instance, in the BVREH sessions support for gesture, tracing and drawing practices were not considered initially. Additionally, in the UK-NEES project support for access to the ambient properties of a setting such as sights, sounds and events were not specified in the first phase of the project. Also, in the NCeSS session, even though the activities conducted were more focused, the session still provided participants with an opportunity to make suggested improvements regarding the prototype’s usability.

In terms of strategies for the selection of evaluation criteria, this must be done on a case-by-case basis. Generally, the evaluation criteria should be focused on achieving an understanding of the ways in which the specifics of work practice might be supported through the system. So for example, in the BVREH case study the prototype evaluation criteria was drawn from the actual work practices observed *in situ* which the system was meant to support and extend. In this case, the evaluation criteria focused on how well the system supported the broad activities of ‘deciphering ancient manuscripts through a VRE’. In the UK-NEES case study the prototype evaluation criteria was drawn from the goals set out by the researchers themselves which was to coordinate and conduct numerical distributed experiments. In this way, the prototype evaluation was not used to compare how current practices were supported, constrained or transformed as was done in the BVREH case study. Rather, entirely new

practices, “conducting distributed numerical experiments” were performed with the prototype and so these activities were assessed independently of how work is currently conducted. For the NCeSS case study, there was no criteria *per se*, the sessions were more of an exploratory activity to scrutinize the use of the CEF as a stand-alone technique and so investigated the *usability* of the system at the interface, and not its *usefulness* as an embedded system that would support work practices.

Before you begin a session it is important to provide an introduction to its purpose and to introduce the participants to the functionality available in the current iteration of the prototype. Additionally, the activities to be performed during the session should have been agreed beforehand and would need to be reiterated at the beginning. You will also want the participants to sign an informed consent form, if they have not done so already. This is where they formally agree to be video recorded and where they may also give consent for either the audio and/or video to be used in presentations and publications related to the project. After these preliminaries have been completed you may choose to either leave the room or stay and observe the session. I’ve found that it depends upon the comfort level of the participants and the logistics of the setting. For instance, during the BVREH session I remained in the back of the room as an observer. For the UK-NEES sessions I remained in the laboratory and during one session found it necessary to hold the camera and stand behind the participant because of the noise levels were so high. For the NCeSS sessions I found that the participants were very aware of my presence to the point that they appeared to be conducting a paired think-aloud session rather than interacting with each other! In both cases, I left the room and returned when they had completed the activities. The duration of each evaluation session will vary and may run from anywhere between 30 - 90 minutes dependent upon the complexity of the activities performed.

After the evaluation session is completed a post-evaluation feedback session will provide participants with an opportunity to tell you about their experiences. If possible, it is recommended to keep the video recorder switched on. However, some participants may not express all of their opinions if the recorder is on and in many cases I’ve noticed that after turning it off participants express their opinions in a more forthright manner. During the post-

evaluation feedback, participants usually identify difficulties they faced when using the prototype and provide suggestions for its improvement. A list of recommendations should be made and incorporated into a ‘potential requirements’ document which is verified with participants and then made available to e-Research project members.

9.6 Data analysis: identify fragments of work and interaction

After the evaluation session is completed you will have an opportunity to review the video data in order to understand the details of the ways in which activities and cooperative work were accomplished during the session. Here the CEF matrix is used to guide the analysis of the sequential organisation of activities. Remember, when conducting an analysis of video data acquired during an evaluation session the purpose is to understand the ways in which work and interaction is supported, constrained or transformed when using the prototype. This is done in order to gain an understanding of the prototype’s *relevance* and *praxeological validity* in relation to the work it is meant to support.

The analytic focus rests on understanding and describing the sequential production, recognition and maintenance of routine order through activity and interaction. However, the CEF matrix provides guidance to those whom are interested in using it and who have little experience with using EM/CA/IA approaches to understanding interaction. In this way, the CEF guidelines support video analysis by focusing on key analytic themes.

Coordination of activities	Topic & activity	Material artefacts	System landmarks	Potential requirement
<i>Work practices</i>				
<i>Cooperative work</i>				

Table 1: The Contextual Evaluation Framework.

Using the CEF framework as an analytic focus, the sequential organisation of activities and interaction of participants taken from video recorded during an evaluation session may be analysed. Each cell of the matrix and its purposes include:

Work practices - What types of activities are participants engaged in doing? Note how they initiate activities. What is the sequence of an activity? Note how they work through contingencies that may arise. How is an activity terminated? What is the transition to a next activity?

Cooperative work - In what ways are the participants interacting with each other? Note how each participant orients towards the activity. How are the specialized skills of each participant contributing the accomplishment of activities?

While the focus is on understanding the details of work practice and cooperative work, these details may be more finely attended to through noting the following details of participants' interaction with each other and around the prototype:

Topic and activity - What are participants talking about? Note the activities that are enacted while speaking. How might the topic of talk be related to the activity?

Material artefacts - What types of material artefacts do participants engage with during the activity? This may include pens, paper or other objects such as microscopes, books, or computing systems not linked with the prototype.

System Landmarks - How do participants engage with the system to conduct activities? Note the features & functionality selected and used with an orientation to the topic and activity. Note the material artefacts that are used in conjunction with the prototype.

The outcome of this detailed analysis will reveal potential requirements which may then be presented to the participants with the CEF matrix and corresponding video fragments:

Potential requirement - Which system features are available or absent for this activity. How does the system support, constrain or transform the activity. Identify potential requirements to support work practice and cooperative work.

Each of these criteria may be used as a guide to focus the analysis of quasi-naturalistic prototype evaluation video data particularly on the ways in which material artefacts and the prototype are used to accomplish routine activities and cooperative work. Pay particular attention to the problems participants may encounter when trying to make sense of the

prototype. The CEF may also facilitate a comparison (based upon prior fieldwork if available) between how practices are supported, constrained or transformed when using a prototype system and from this assess how well a system is aligned with the ways in which researchers themselves organise activities.

When viewing the evaluation session video it is important to keep an open mind and simply make notes of what the participants are doing and taking about. Focus particularly on turns of talk which will reveal the sequential production of both the topic and activity and their mutual understanding of the ways in which the system may be supporting that activity moment-to-moment. Also, note the embodied practices of participants such as gesturing, bodily orientation and gaze direction. Pay particular attention to the other material artefacts the participants may turn to as resources in the accomplishment of work. Note the difficulties and troubles participants may have when using the prototype. Finally, watch for activities that are repeated, which will help you to identify reoccurring patterns.

Create a log file of the video that provides a summary of its content and highlights. Log files should include the time that a particular fragment of interest begins and ends, a description of the activities occurring, and the characteristics of the constraint or pattern of interaction. Include with this a list of questions you may have for the participants so that you may better understand their activities. In sum, the video log file may be used to raise issues, to indicate events of interest, to point to places where more data might be needed, and to suggest themes, categories, and strategies for analysis (Suchman & Trigg, 1991: 292).

9.7 Film elicitation: confirm preliminary analysis with participants

After video fragments of interest are identified your analysis can be refined using film elicitation (Banks, 2001: 96) techniques in order to gain a better understanding of the activities participants are conducting and the difficulties they may have encountered. Film elicitation involves gathering together one or several video fragments of interest showing these to participants and asking questions about the details of the activities within a scene. For example, in the UK-NEES case study video fragments were identified where the researcher

was relying on access to the ambient properties of the laboratory setting. In this case, I collected a few instances together and played the video fragments with the lead researcher and asked him to provide feedback regarding my initial analysis. He agreed that the ambient properties of the laboratory are indeed seen-but-unnoticed aspects of experimental practice and was amazed that the e-Research team hadn't thought about how they would support access to these properties across distributed laboratories. Here, film elicitation provided an opportunity with which to confirm seen-but-unnoticed patterns of interaction.

Additionally, film elicitation can be used if you are unsure about the purpose and meaning of activities and may ask participants: "in this fragment I wasn't sure about x,y,z - what is going on here"? Film elicitation may also be used to understand domain-specific terminology used by participants during a prototype evaluation session. It is also useful for understanding *the details* of the difficulties participants encounter which are observed on video but still require further understanding from the members own perspectives. Ask questions such as 'what was the difficulty you encountered here', 'how should the prototype support this activity', and 'what should the prototype be supporting at this point in this activity'.

After film elicitation is conducted, it is suggested that video fragments also be viewed and discussed amongst e-Research project members. Known as 'data sessions' (Heath *et al*, 2010: 102) these internal project group meetings are a matter of course in both CA and IA research. They provide an opportunity for ethnographers to discuss audio-video fragments in detail and when used within the software development process these sessions can be extended to include software engineers, requirement engineers and project managers. Data sessions provide an opportunity for the project team to discuss particular instances where difficulties have occurred or where patterns of interaction emerge and allow the team to work through potential solutions and design ideas.

Film elicitation, taken in combination with the other materials acquired during requirements gathering activities will include field notes from interviews, video from *in situ* observations, prototype evaluation video data, post-evaluation feedback and data session feedback. Outcomes from each of these activities will assist in refining the project team's understanding

of the domain and in turn inform the ways in which prototypes might support work practice and cooperative interaction. From each of these sources the e-Research project team may assemble potential requirements in order to inform future iterations of the prototype's design and the CEF may be used as one tool for presenting potential requirements.

9.8 Assemble potential requirements: informing software design

The CEF matrix along with associated video fragments can be used as tool with which to share prototype evaluation analysis with software developers and other members of an e-Research project. In addition to the matrix and its associated video fragment, CEF outputs also provide recommendations based upon analysis that are listed as potential requirements. In each of the case studies however, the software development team was not interested in accessing either the video fragments or the matrix. The most common response was 'just give me the bullet point list'. Even so, the researchers and software developers were interested in the results of the evaluation but felt pressed for time and so could not commit to viewing them. Upon reflection, it may be useful to allocate time for such activities into e-Research project proposals, where within the work plan software engineers would be scheduled to participate in data sessions so that they may better understand the relationship between software functionality and its impact upon work practices. Even though software engineers could not view the video data because of time constraints, having video fragments available provides both empirical evidence for a potential requirement and traceability to the requirement after it has been agreed and implemented into the prototype's design.

Potential requirements were emailed to participants as bullet point lists where they confirmed or clarified their specific concerns. The prioritisation of requirements including which to include in the prototype's next iteration was agreed amongst the software developers, the project managers and the participants. The CEF is meant to be used within an iterative and incremental software development process and for this reason it is recommended that with each major iteration of the prototype a CEF evaluation be conducted if time permits.

9.9 Summary

The CEF can be used as a guide in the analysis of quasi-naturalistic prototype evaluations and offers a systematic and rigorous approach that includes a practical guide for its use detailed in this chapter. Identifying and documenting potential requirements using the CEF provides empirical evidence of both the difficulties participants encounter and patterns of work and interaction. Using the CEF, the sequential analysis of video data is the analytic focus and the matrix provides a guide to those who may wish to use it for understanding participants' activities. In particular, the ways in which the participants' topic of talk and the ongoing activity is supported with a focus on understanding how material artefacts and the prototype are interleaved as resources in the accomplishment of work practice and cooperative work.

Requirements analysts wishing to use the CEF would need a working knowledge of the concepts of EM/CA/IA and most especially the notion of the sequential organisation of work and interaction as a primary analytic resource. Ideally, the CEF would be introduced to e-Research project teams through tutorial workshops where they could familiarise themselves with EM concepts, quasi-naturalistic evaluation and the ways in which the CEF can be used to guide the analysis of video data. Finally, although the CEF was designed within the e-Research domain it is suitable for use in any iterative and incremental software development project that wishes to apply qualitative methods for system requirements definition and refinement. The final chapter of this thesis will discuss the successes and challenges of using the CEF as well as further work that is planned as part of the EPSRC funded Phd Plus fellowship programme.

Chapter 10:

Discussion and further work

10.1 Introduction

The CEF guidelines are designed to be used by any member of an e-Research project or any other software development project interested in applying qualitative methods for system requirements definition and refinement, including project managers, software engineers and requirements analysts, without necessarily having expert knowledge in EM/CA/IA. The aim for developing this prototype evaluation framework is to provide guidance to researchers and practitioners wishing to evaluate the usefulness and usability of software prototypes using a sociologically-informed approach. Additionally, the CEF facilitates a sensitivity towards the practices and activities relied upon by professionals within the domain a system is designed for and provides project members an opportunity to reflect upon how future iterations of a prototype might support work practices.

The CEF is a very specific contribution to requirements engineering focused on the evaluation of prototypes in quasi-naturalistic settings, although it is recommended that it be used in conjunction with ethnomethodologically-informed fieldwork so that an adequate understanding of the domain in which the technology is being designed may be obtained. EM-informed studies of work provide access to recognisable orders that are produced and maintained by the members of a setting. However, when evaluating prototypes, the recognisable order is in some way breached either because the practice does not exist at all or ordinary, routine practices may be transformed in some manner. To identify these breaches, CSCW researchers have evaluated prototypes for ‘breakdowns’ which identify when routine work practices are interrupted (Winograd & Flores, 1987) and ‘asymmetries’ which identify

when shared understanding is incomplete due to the lack of mutual resources that convey the significance of one's own actions and enable the interpretation of the actions of others (Suchman, 1987). The outcomes of this research contributes to requirements engineering research by providing an addition to the two main analytic frameworks currently used to evaluate systems in CSCW - 'relevance'.

Using the CEF, a prototype's 'relevance' as a meaningful tool is assessed by conducting initial fieldwork that will provide a comparison between how the practices observed prior to a prototype's evaluation are supported, constrained or transformed. From this, a prototype may be assessed for how well it is aligned to the ways in which skilled practitioners themselves organise their activities. This qualitative analysis between how practices are supported, constrained or transformed when using a prototype makes possible an assessment of the ways in which participants interpret usefulness and usability, which map on to concepts in phenomenological sociology of *relevance*, *familiarity* and *praxeological validity*. In this way, the CEF's focus is on the analysis of an e-Research system's relevance as a tool that is in some manner familiar to researchers. This familiarity does not need to map exactly to the ways in which work is currently conducted. However, some degree of familiarity is necessary in order for researchers to make sense of what the system has to offer and the ways in which it would support their routine work practices and collaborative activities so that they may confidently appropriate it into their work settings.

The CEF is a generic framework that may be used not only within an e-Research context but also for the evaluation of prototypes developed in other domains such as healthcare, mobile and pervasive computing. In short, any domain where day-to-day practices need to be understood in detail so that relevant systems may be designed to support those practices. The CEF may be embedded within the most commonly used iterative and incremental software development methodologies that fully support and incorporate mutual learning in design practice such as Agile (Cockburn, 2007), Spiral (Boehm, 1988), Rapid Prototyping and Evolutionary (Sommerville, 1989) approaches.

Additionally, the CEF may be used in a similar manner as ‘quick and dirty’ ethnographies (cf. Hughes *et al*, 1994; Crabtree, 2003; Randall *et al*, 2007) where as an alternative to conducting fieldwork, brief prototype evaluation sessions may be conducted and quickly evaluated by project members to get a general sense of how the system might support work practice and collaboration. This type of quick session might be conducted in a single day where a video-based quasi-naturalistic evaluation would be followed by a post-evaluation session in order to gather both quick observations in the first activity and participant feedback in the second. The NCeSS case study came close to this type of evaluation where both a prototype evaluation and feedback session were conducted as paired activities which resulted in the identification of requirements at the end of the two-hour session.

The CEF is a formative prototype evaluation technique and may be classified as a resource for *evolutionary* prototyping (cf. Floyd, 1984) where there is a possibility for a system to be adapted based upon findings as part of an iterative software design lifecycle. The outputs of a CEF prototype evaluation are not intended to provide metrics (the main output of traditional HCI evaluation) of participant success rates in performing tasks, the time a task requires, error rates, or the quantification of users’ subjective satisfaction using Likert scales. Rather, it is a qualitative approach designed to provide *ecological validity*. In contrast to metrics, ecological validity provides results that may demonstrate whether an effect observed “is representative of what happens everyday life” and that activities observed have “been demonstrated to occur under conditions that are typical” (Brewer, 2000: 12). It is important to note that the concept of ‘validity’ is taken from the natural sciences and so its definition as accuracy of measurement which may be used to replicate phenomena cannot be transferred over easily to the assessment of social phenomena (Hammersley, 1998). First, the notion of replication is difficult, if not impossible, when examining social phenomena because human behaviour cannot be controlled as can physical objects or animals in a laboratory setting. Secondly, the notion of ‘naive realism’ - that there is a single world independent of us where absolute truth exists, has been rejected in postmodern thought and replaced with the notion of the social construction of reality where knowledge acquisition is not considered to be a process of

discovering an absolute truth but rather as an interpretive process that occurs within and is influenced by social contexts (cf. Berger & Luckman, 1966).

Rather, the notion of validity in relation to the study of social phenomena should be evaluated using the following criteria (Hammersley, 1998: 67); *plausibility* (how likely is the account to be true from what we take to be well-established knowledge and through the demonstration of evidence), and *credibility* (are the researcher's judgements accurate based upon the phenomena presented). Using these criteria, studies of social phenomena may be assessed for the extent to which an account accurately represents the social phenomena to which it refers (cf. Silverman, 2006). And so, ecological validity is met by the CEF firstly, by asking participants to conduct their typical work activities using the prototype, thus providing representative instances of what occurs in everyday life. Secondly, conducting an evaluation session within a quasi-naturalistic setting provides a close approximation to typical conditions. Additionally, video recordings and transcripts are gathered as evidence that can be reviewed by others as a way of validating claims made by the requirements analyst which may address issues of validity and *reliability* (cf. Peräkylä, 2004).

Finally, *generalizability* can be described as the wider significance results may have which derive from a small sample size. This may be achieved by taking a 'comparative approach' (ibid: 296) where similarities and differences across a number of settings or situations are identified. However, even when identifying patterns of interaction it is less important to count the number of instances but rather to focus on the quality and meaning of the interaction observed. Because of its qualitative orientation, findings may be generalised on a case-by-case analysis and this is recognized as an appropriate approach to take within the social sciences (ibid: 299). And so, ecological validity, reliability and generalizability are addressed within the CEF. The next section will discuss the boundaries of the CEF including the successes and challenges of implementing it.

10.2 Successes and Challenges

Because of tight time constraints within each of the e-Research projects it was not possible for the project members to evaluate the CEF on their own terms as if they were to utilise the

technique independently. This is a well-known issue when developing a methodological thesis. However, I will address this by conducting a brief case study early next year as part of an award our project received under the EPSRC funded Phd Plus fellowship programme. For this, I will engage in knowledge transfer activities with selected software development projects with the aim of validating the CEF with people who will use it independently and in their own setting. However, I have reflected upon and can offer my own assessment of the evaluation framework.

Successes with CEF include firstly, its ability to be used as a resource in the software development process, which provides empirical evidence that may be examined in detail in order to reveal what may otherwise be overlooked or ‘seen-but-unnoticed’ activities and patterns of interaction. Second, the video data can be used for film elicitation with participants so that a more complete understanding of activities and interactions may be gained. Additionally, the matrix provides guidance in the evaluation of video data for assessing a system’s relevance. From a technical perspective, there is good software support for video data analysis, including video editing tools, audio restoration and screen recording software that are simple to use and relatively inexpensive. Also, the video fragments coupled with the CEF matrix provide for requirements traceability after these are agreed upon and implemented into the prototype’s design. Finally, the CEF provides participants with an opportunity to participate in the software development process as well as to reflect upon their own practices.

Challenges with the CEF include its labour intensive nature where schedules need to be coordinated and where travel to different work settings needs to take place. Additionally, conducting evaluations across multiple settings may be difficult if there is not another camera available to be positioned at the remote location. Also, video data analysis is a very time intensive activity especially when fragments of interest are identified for conversational transcription and detailed scrutiny. The CEF may never provide enough domain knowledge and so it needs to be used in an iterative fashion and those using it must commit to acquiring a level of competency about the domain for which the system is being developed. In this way, the software development project needs to build in the use of the CEF as part of the design

process and not simply add it in if it is convenient. Other limitations of the CEF may be the audio-video equipment investment necessary as some projects may lack the funds to purchase such equipment. Another limitation in its use may be the amount of time needed to set-up and conduct an evaluation session and to then analyse the video data. However, video analysis times may be shortened if the project team commits to conducting group data sessions. This however brings to light challenges with conducting interdisciplinary work where not all members of a project may want to participate in the same activities (e.g. ‘just show me the bullet-point list’). Even so, it may even be sufficient to have a number of project members sit and watch the evaluation session recording while filling in the CEF matrix; where they may at any time stop the recording to discuss specifics or they may discuss their observations at the end of the viewing.

A prototype evaluation session may run from between 30 to 90 minutes. Therefore, the time required to review and discuss the evaluation session data would take no more than half a day. After the data session is complete, the list of requirements generated can then be prioritized amongst the developers and participants. Further potential uses of the CEF include the possibility of conducting distributed requirements data sessions through Access Grid⁴⁴ or using some other Internet-enabled video software tool.

10.3 Conclusion

The goal of e-Research systems is to support collaborative research across geographically distributed settings. However, developing an understanding of the details of researchers’ scientific work including the materials and resources they rely upon and the ways in which they collaborate and interact with colleagues is a non-trivial matter. This thesis has addressed this challenge with the development and use of a sociologically-informed requirements engineering technique for prototype evaluation. Its purpose is to facilitate an understanding of the complex relationship between work practice and technology as well as the ways in which new distributed work configurations may transform work practice or create entirely new ones.

⁴⁴ <http://www.ja.net/services/video/agsc/AGSCHome>

In each case study, the CEF was used to identify emerging requirements related to system functionality and interface usability in a variety of academic disciplines from the humanities, the sciences, to the social sciences. In the BVREH case study, a co-located evaluation session was conducted which demonstrated the CEF's usefulness in revealing the ways in which both collaborative activities and tacit practices were supported through the prototype. These practices traditionally occur through asynchronous communication and synchronous work-in-progress meetings in order to update each other on their progress, to review each others tentative interpretations, and to reach agreement about the content and meaning of ancient manuscripts. The CEF sessions were used to observe synchronous collaborative interactions between the scholars as they engaged in the real-time deciphering and interpretation of the manuscript. In this way, their interpretive practices could be examined in detail for the purposes of supporting those practices through a virtual research environment. In the UK-NEES case study, the CEF was implemented to evaluate a prototype in a synchronous and distributed setting. The case study also demonstrated the usefulness of the CEF in evaluating prototypes in the absence of practice and that a single-sited approach to understanding the coordinated activities of multiple distributed participants could be used to acquire meaningful results in synchronous and distributed prototype evaluation. Finally in the NCeSS case study, the CEF proved useful for revealing participants' moment-to-moment sense-making procedures as they worked through each activity in the prototype evaluation session.

In each case, the CEF provided an alternative to the limitations of traditional requirements engineering techniques that rely upon gathering retrospective user feedback. The CEF extends and builds upon other CSCW prototype evaluation techniques by assessing systems for their relevance as meaningful tools. This orientation corresponds to the two key attributes associated with the design of successful systems (Bloomberg *et al*, 1996: 239); their *usability*, which refers to the system's general intelligibility particularly at the interface and *usefulness*, referring to the ways in which a system's functionality actually makes sense and adds value in relation to a particular work setting. More specifically, within any academic discipline complex skills and knowledge are shared across its community. Additionally, researchers learn to use the tools, instruments and processes necessary for research through engagement in long-

term apprenticeships within laboratories and research institutes. Implicit in these social practices, researchers go through a process of learning whereby they gradually acquire the knowledge and skills necessary to become full members of a community of practice. In this way, the processes, objects and artefacts of practice may come to possess specific meaning and significance for participants. The CEF addresses this by providing a new analytical framework focused on understanding, ‘relevance’, which identifies how well a system is aligned with the endogenous organisation of work within a community of practice through an analysis of its usefulness as an emergent cultural object and collaborative tool. Assessing how well a system supports the knowledge, skills and practices that already exist within a community is equally as important as developing solutions that will eventually reconfigure those practices, create new ones and extend modes of collaboration. The aim of the CEF is to contribute to the successful uptake and use of new technologies by providing a sociologically-informed technique that can be used to evaluate prototypes within quasi-naturalistic settings of work so that a more detailed understanding of the disciplinary concerns of potential users and how they collaboratively use technologies in actual practice may be understood.

Appendix 1:

UK e-Research user studies

Completed studies

Three short-term studies have been completed thus far. The earliest was completed in June 2005 and the most recent in February 2007.

Security and Usability of Grid Projects: Implications for e-Science (Craft & Sasse, 2005):

A four-month study which addressed the usability of security. Recommendations include; 1) a need to improve documentation, 2) require the use digital signatures only when necessary and provide alternative authentication options, 3) provide outreach, education and training, 4) give clear guidance to developers, 5) consider security at the beginning of a project and 6) plan for long-term adoption and use.

JISC human factors audit report (Kalawsky et al, 2006):

The report includes a summary of eight e-Science projects with a detailed description of the technical aspects of each system. It then summarises a list of human factors best practices. Recommendations include; 1) support project management with easy to access online documentation, 2) manage expectations by providing clear project goals, 3) provide ethics management by addressing issues at the proposal stage, 4) create USE cases to support an understanding of user needs, 5) provide a workflow audit trail of project deliverables, 6) take complexity out of user interaction and interface, 7) increase grid awareness with user pull being more effective than developer push, 8) provide training, 9) provide good documentation, 10) support interdisciplinary working with frequent communication.

OMII study of user priorities for e-infrastructure for e-research (Newhouse et al, 2007):

A six-month study to understand user priorities for e-Research and to identify short-term needs of projects developing e-Infrastructures. Recommendations include; 1) software to support file sharing and the creation of VOs, 2) policy to develop standards for data curation, VO construction and annotation schemas and 3) support for best practice guidelines, self-help training, outreach and education.

Current projects

There are four main research projects currently being undertaken to examine UK e-Science usability. The first project began in 2004 and the most recent projects began in April 2007.

Usability task force (UTF):

The UTF was created as part of the e-Science community to formally address usability challenges, their report entitled *Usability Research Challenges in e-Science* describes four main challenges; 1) the need for new methods and techniques to understand global scientific communities and how to foster new forms, 2) support of trust and ethics within systems, 3) understanding the knowledge production lifecycle and the representation of information and expertise and 4) design, evaluation and management of global e-Science grids.

e-Science Institute theme: Adoption of e-Research Technologies:

This programme theme is funded by the e-Science Institute <<http://www.nesc.ac.uk/esi/>> and its purpose is to study factors that inhibit the uptake of e-Research tools with an emphasis on the social sciences, arts and humanities. The theme has hosted three of workshops <http://www.nesc.ac.uk/esi/themes/theme_03/community.htm> which provided an arena for researchers to discuss issues specific to the uptake of e-Research tools.

JISC e-Infrastructure – Barriers to Uptake & Service Usage Models:

The JISC Research Programme funded two projects that began in April 2007 which are investigating two specific issues. First, Enabling Uptake <<http://www.e-researchcommunity.org/projects/e-uptake>> will focus on producing a series of case studies that will report on specific barriers to uptake and provide recommendations for training

programmes. Second, the Service Usage Models project <<http://www.eius.ac.uk>> will focus on describing how researchers are currently using e-Infrastructures. Based upon case study findings it will then develop scenarios and use cases that will describe research workflows.

EPSRC Research – Usability Challenges to Emerge from e-Science:

The EPSRC put out a call in 2005 for proposals that focused specifically on research into e-Science usability <<http://www.epsrc.ac.uk/CMSWeb/Downloads%5CCalls%5CUsabilityChallengesFrome-ScienceCall.doc>>. Five projects have been funded within this call and include:

- *Particle-physics Engagement with the Grid: A Socio-technical Usability Study* <<http://www.ucl.ac.uk/research-computing/research/e-science/pegasus.html>> investigating both the political and social aspects that influence the construction of technology design within the domain of particle physics
- *Usable Image Project* <<http://www2.idl.dundee.ac.uk/usableimage>> is engaged in developing novel methods for gathering data about tool use to further define user needs specific to the Open Microscopy Environment
- *ADVISES: Adaptive Visualisation for e-Science: Requirements analysis approach and experience* <<http://www.informatics.manchester.ac.uk/research/groups/isd/projects/advises>> is investigating the development of 'sub-languages' that express clinicians' high-level requirements for visualisation tools
- *Easier Specification of Authorisations Policies* <<http://www.cs.kent.ac.uk/news/2005/EasyGrant>> is conducting research into the development of a 'policy tool', which will make it easier for end-users to create authorisation policies for their e-science applications.
- *Embedding e-Science Applications: Designing and Managing for Usability* <<http://www.oerc.ox.ac.uk/research/embedding-e-science>>, investigating novel methods for both the design and management of e-Science projects that specifically take into account the distributed, collaborative nature of research practice, software development and project management.

Appendix 2:

Transcription notation

The transcription notation system used in conversational data segments is an adaptation of developed by Gail Jefferson in Atkinson & Heritage, 1984.

1. Simultaneous utterances

Utterances starting simultaneously are linked together with single left-hand brackets :

T: [I used to smoke a lot when I was young
B: [I used to smoke Camels

2. Contiguous utterances

When there is no interval between adjacent utterances, the second being latched immediately to the first (without overlapping it), the utterances are linked together with equal signs:

T: I used to smoke a lot=
B: =He thinks he's real tough

3. Intervals within and between utterances

When intervals in the stream of talk occur, they are timed in tenths of a second and inserted within parentheses, either within an utterance:

L: When I was (0.6) oh nine or ten

or between utterances :

H: Step right up
(1.3)
H: I said step right up
(0.8)
J: Are you talking to me

4. Characteristics of speech delivery

Punctuation may be used to capture characteristics of speech delivery. For example, a colon indicates an extension of the sound or syllable it follows:

R: What ha:ppened to you

and more colons prolong the stretch :

M: I ju::ss can't come

T: I'm so::: sorry re:::ally I am

The other punctuation marks are used as follows:

- . A period indicates a stopping fall in tone, not necessarily the end of a sentence.
- , A comma indicates a continuing intonation
- ? A question mark indicates a rising inflection, not necessarily a question.
- ! An exclamation point indicates an animated tone
- A single dash indicates a halting, abrupt cutoff,

Marked rising and falling shifts in intonation are indicated by upward and downward pointing arrows immediately prior to the rise or fall :

T: I am however (0.2) very ↑fortunate (0.4) in having (0.6) a
↑mar:vlous dep↓uty

Emphasis is indicated by underlining:

A: It happens to be mine

Capital letters are used to indicate an utterance, or part thereof, that is spoken much louder than the surrounding talk:

A : an the winner : ↓iz:s (1.4) RACHEL ROBERTS

A degree sign is used to indicate a passage of talk which is quieter than the surrounding talk :

M: (.) °Um::° 'Ow is yih mother by: th'wa:y

Audible aspirations (hhh) and inhalations ('hhh) are inserted in the speech where they occur :

P: An thi(hh)s is for you hhh
D: 'hhh O(hh) tha(h)nk you rea(hh)lly

Double parentheses are used to enclose a description of some phenomenon the transcriptionist does not want to deal with:

T: I used to ((cough)) smoke a lot
B: ((sniff)) He thinks he's tough
A: ((snorts))

When part of an utterance is delivered at a pace quicker than the surrounding talk, it is indicated by being enclosed between '>' signs:

J: the Guar:dian newspaper looked through >the manifestoes< la:st
↑week

5. Transcription doubt

Items enclosed in single parentheses are in doubt:

T: I (spose I'm not)

When single parentheses are empty, no hearing could be achieved for the item in question:

T: My () catching

6. Marking the interesting parts of the transcript

An arrow (→) in the left-hand margin can be used to mark the parts of the transcript that are being considered at a particular stage in a document.

Appendix 3:

Video fragments

See attached DVD.

Chapter 6 - BVREH:	Chapter 7 - UK-NEES:	Chapter 8 - NCeSS:
Fig2.mov	Thisisanexperiment.mov	Fig2.mov
Fig4.mov	Fig6/7.mov	Fig3a.mov
Fig6.mov	Fig8.mov	Fig3b.mov
Fig8.mov	Fig9/10.mov	Fig4.mov
Fig9.mov	Fig11.mov	Fig5.mov
Fig12.mov	Fig16.mov	-
Fig13.mov	Fig17.mov	-
Fig14.mov	Fig18.mov	-
Fig15.mov	Fig19.mov	-
-	Fig20.mov	-
-	Fig21.mov	-
-	Fig22.mov	-

Appendix 4:

List of publications

Journal articles:

de la Flor, G., Jirotko, M., Luff, P., Pybus, J., Kirkham, R. (2010): Transforming scholarly practice: embedding technological interventions to support the collaborative analysis of ancient texts. *Journal of Computer Supported Cooperative Work*, 19(3/4): pp. 309-334.

de la Flor, G., Ojaghi, M., Martínez, I., Jirotko, M., Williams, M., Blakeborough, T. (2010). Reconfiguring practice: the interdependence of experimental procedure and computing infrastructure in distributed earthquake engineering. *Philosophical Transactions of the Royal Society, A*: pp. 4073-4088.

Conference papers:

de la Flor, G., Luff, P., Jirotko, M., Pybus, J., Kirkham, R., Carusi, A. (2010). The Case of the Disappearing Ox: Seeing Through Digital Images to an Analysis of Ancient Texts, In *Procs. 28th ACM Conference on Human Factors in Computing Systems*, CHI 2010, Atlanta GA: pp. 473-482.

de la Flor, G. & Jirotko, M. (2009). Blurring the distinction between software design and work practice. In *Proc. of the 2nd Int. Symp. on End User Development, Work-in-Progress Session*, Siegen, Germany, 2-4 March, pp. 24-31, (available at: [<http://sozio-informatik.org/fileadmin/IISI/upload/IRSI/IRSIV611.pdf#page=24>]; last accessed 01 June 2010).

de la Flor, G., Meyer, E.T. (2008). Talking 'bout a revolution: Framing e-Research as a computerization movement. In *The Oxford eResearch Conference 2008*. 11-13 September 2008. Oxford, UK.

de la Flor, G., Jirotko, M., Spencer, D., Lloyd, S. (2008). Communicating requirements in multidisciplinary teams. In *The Oxford eResearch Conference 2008*. 11-13 September 2008. Oxford, UK.

de la Flor, G., Lloyd, S., Jirotko, M., Warr, A. (2007). Designing software in support of workplace activities-embedding e-science applications. In *Procs. of 3rd International Conference on e-Social Science*, Ann Arbor, MI, (available at: [<http://ess.si.umich.edu/papers/paper189.pdf>]; last accessed 07 September, 2008).

Daw, M., Procter, R., Lin, Y., Hewitt, T. Jie, W., Baird, K., Turner, A., Birkin, M., Miller, K., Jirotko, M., Schroeder, R., de la Flor, G., Allan, R., Yang, X., Crouchley, R. (2007). Developing an e-infrastructure for social science. In: *Procs. of 3rd International Conference on e-Social Science*, Ann Arbor, MI, (available at: [www.ncess.ac.uk/events/conference/2007/papers/paper127.pdf]; last accessed 16 August, 2010).

Warr, A., de la Flor, G., Jirotko, M. & Lloyd, S. (2007). Usability in e-Science: The eDiaMoND Case Study. In *CHI International Workshop on Increasing the Impact of Usability Work in Software Development*. San Jose, California. 2007.

Book chapters:

de la Flor, G., Jirotko, M., Lloyd, S. & Warr, A. (2010). Embedding e-Science applications: designing for usability. In *World Wide Research: Reshaping the Sciences and Humanities* (eds W. H. Dutton & P. W. Jeffreys). Cambridge, MA: MIT Press: pp. 135-152.

References

- Allan, R., Yang, X., Crouchley, R. (2009). ESRC e-Infrastructure: Portal Requirements. *NCeSS Technical report*, (available at: [http://epubs.cclrc.ac.uk/bitstream/3635/e-IP_portal_report.pdf]; last accessed 16 August, 2010).
- American Council of Learned Societies, (ACL). (2006). *Our Cultural Commonwealth*. Technical Report of the Commission on Cyberinfrastructure for the Humanities and Social Sciences, (available at: [<http://www.acls.org/cyberinfrastructure/OurCulturalCommonwealth.pdf>]; last accessed 07 September, 2010).
- Anderson, R. (1997). Work, ethnography and system design. In Kent, A., Williams, J. (eds.) *The Encyclopaedia of MicroComputing*, 20. New York: Marcel Dekker: pp. 159-183.
- Anderson, R. (1994). Representations and requirements: the value of ethnography in system design. *Human-Computer Interaction*. 9(2): pp. 151-182.
- Anderson, R. & Sharrock, W. (1993). Can organizations afford knowledge?. *Journal of Computer Supported Cooperative Work*, 1: pp. 143-161.
- Aoki, P.M., Grinter, R., Hurst, A., Szymanski, M., Thornton, J., Woodruff, A. (2002). Sotto Voce: Exploring the Interplay of Conversation and Mobile Audio Spaces. In *Procs of CHI 2002*, Minneapolis, Minnesota. ACM Press.
- Arminen, I. (2008). Scientific and “Radical” Ethnomethodology: From Incompatible Paradigms to Ethnomethodological Sociology. *Philosophy of the Social Sciences*, 38(2): pp. 167-191.
- Atkins D., et al. (2010). *Building a UK foundation for the transformative enhancement of research and innovation*. Technical report of the international panel for the 2009 review of the UK research councils e-science programme. Swindon, UK:RCUK, (available at: [<http://www.epsrc.ac.uk/research/intreivs/escience/Pages/default.aspx>]; last accessed 07 September, 2010).
- Atkins, D. et al. (2003). Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. National Science Foundation Report, (available at: [<http://www.nsf.gov/od/oci/reports/toc.jsp>]; last accessed 01 July, 2008).
- Atkinson, M. J., & Heritage, J. (eds.). (1984). *Structures of social action: Studies in conversation analysis*. Cambridge, Cambridge University Press.
- Banks, M. (2001). *Visual methods in social research*. London: Sage.
- Bannon, L. & Bodker, S. (1997). Constructing Common Information Space. In *Procs. of the European Conference on Computer-Supported Cooperative Work ECSCW'97*.(Lancaster, UK), Dordrecht: Kluwer: pp. 81-96.

- Bannon, L. (1996). Use, Design and Evaluation: Steps towards an Integration. In Shapiro, D., Tauber, M., Traunmueller, R. (eds.) *The Design of Computer-Supported Cooperative Work and Groupware Systems* (series 'Human Factors in Information Systems' volume 12), Amsterdam: Elsevier, (available at: [<http://www.ul.ie/~idc/library/papersreports/LiamBannon/18/COMIC.html>]; last accessed 29 August, 2010).
- Bannon, L. (1991). From human factors to human actors: the role of psychology and human-computer interaction studies in system design. In Greenbaum, J., Kyng, M. (eds.) *Design at Work: Cooperative Design of Computer Systems*. Lawrence Erlbaum Associates: pp. 25-44.
- Bannon, L. & Bødker, S. (1991). Beyond the interface: encountering artifacts in use. In Carroll, J. (ed.) *Designing interaction: psychology at the human-computer interface*. Cambridge: Cambridge University Press: pp. 227-253.
- Bateson, G. (1972). *Steps to an ecology of mind*. Chicago: University of Chicago Press.
- Bateson, G. & Mead, M. (1942). *Balinese character: a photographic analysis*. New York: New York Academy of Sciences.
- Bentley, R., Rodden, T., Sawyer, P., Sommerville, I., Hughes, J., Randall, D., Shapiro, D. (1992). Ethnographically informed systems design for air traffic control, In *Procs of CSCW'92*, Toronto: pp. 123–129.
- Berger, P. & Luckmann, T. (1966). *The social construction of reality: a treatise in the sociology of knowledge*. London: Penguin.
- Bergman, M., Leslie King, J. and Lyytinen, K. (2002): Large Scale Requirements Analysis as Heterogeneous Engineering. *Scandinavian Journal of Information Systems* 14: pp. 37-55.
- Berman, F. & Brady, H. (2005). Final report NSF SBE-CISE workshop on cyberinfrastructure and the social sciences. National Science Foundation, (available at: [<http://vis.sdsc.edu/sbe/reports/SBE-CISE-FINAL.pdf>]; last accessed 05 September, 2008)
- Berman, F., Fox, G. & Hey, A. (2003). The Grid: past, present, future. In Berman, F., Fox, G. & Hey, A. (eds.) *Grid Computing: Making The Global Infrastructure a Reality*. Chichester: Wiley: pp.9-50.
- Bias, R. (1991). Interface-Walkthroughs: efficient collaborative testing. *IEEE Software*, 8(5): pp. 94-95.
- Bikerton, M, & Siddiqi, J. (1993). The classification of requirements engineering methods. In *Procs of IEEE International Symposium on Requirements Engineering*, 4-6 Jan 1993, San Diego, CA: pp. 182-186.
- Birnholtz, J. P. & Horn, D. B. 2007 Shake, rattle and roles: Lessons from experimental earthquake engineering for incorporating remote users in large-scale e-science experiments. *Journal of Computer-Mediated Communication*, 12(2), (available at: [<http://jcmc.indiana.edu/vol12/issue2/birnholtz.html>]; last accessed 28 December 2009).
- Blakeborough, A., Williams, M.S., Darby, A.P., Williams, D.M. (2001). The Development of Real-Time Substructure Testing. *Philosophical Transactions of the Royal Society A*, 359: pp. 1869-1891.
- Blomberg, J. L., Suchman, L., Trigg, R. (1996). Reflections on a work-oriented design project. *Human-Computer Interaction*, 11: pp. 237–265.

- Blythin, S., Rouncefield, M., Hughes, J. (1997). Never mind the ethno' stuff, what does all this mean and what do we do now: ethnography in the commercial world. *Interactions*, 4(3): pp. 38-47.
- Bødker, K., Kensing, F., Simonsen, J. (2004). *Participatory IT design: designing for business and workplace realities*. Cambridge: MIT Press.
- Bødker, S., Ehn P. Sjögren, D., & Sundblad, Y. (2000). Cooperative Design - perspectives on 20 years with 'the Scandinavian IT Design Model'. Keynote, In *Proc. of NordiCHI*, Stockholm, October 2000.
- Bødker, S. & Grønbaek, K. (1991). Cooperative prototyping: users and designers in mutual activity. *Int. J. Man-Machine Studies*, 34: pp. 453-478.
- Boehm, B. (2006). A view of 20th and 21st century software engineering. In *Procs. of the 28th international conference on Software engineering*, (Shanghai, China): pp. 12-29.
- Boehm, B. (1988). A Spiral Model of Software Development and Enhancement. *IEEE Computer*, 21(5): pp. 61-72.
- Boren, T., & Ramey, J. (2000). Thinking Aloud: Reconciling Theory and Practice. *IEEE Trans. on Professional Communication*, 43(3): pp. 261-278.
- Borgman, C. L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press.
- Bos, N., Zimmerman, A., Olson, J., Yew, J., Yerkie, J., Dahl, E., Olson, G. (2007). From shared databases to communities of practice: A taxonomy of collaboratories. *Journal of Computer-Mediated Communication*, 12(2).
- Bowers, J. & Pycock, J. (1994). Talking through design: requirements and resistance in cooperative prototyping. In *Procs of CHI'94 Human Factors in Computing Systems*, April 24-28th, Boston, MA: pp. 299-305.
- Bowers, J. & Rodden, T. (1993). Exploding the interface: experiences of a CSCW network. *Proc. of the SIGCHI conference on Human factors in computing systems*: pp. 255-262.
- Bowman, A.K., Crowther C.V., Kirkham, R., Pybus J. (2008). A Virtual Research Environment for the Study of Documents and Manuscripts. In *The Oxford eResearch Conference 2008*. 11-13 September 2008. Oxford, UK.
- Bowman, A. & Tomlin, R. (2005): Wooden stilus tablets from Roman Britain. In Bowman, A. & Brady, M. (eds.) *Images and artefacts of the ancient world*. Oxford UK: Oxford University Press: pp. 7-14.
- Bowman, A., Brady, M. & Tomlin, R. (1997). Imaging incised documents. *Literary and linguistic computing*, 12, 3: pp. 169-176.
- Brewer, M. (2000). Research Design and Issues of Validity. In Reis, H. and Judd, C. (eds) *Handbook of Research Methods in Social and Personality Psychology*. Cambridge: Cambridge University Press: pp. 3-16.
- Brown, J. S. & P. Duguid (1991). Organizational Learning and Communities of Practice: Towards a Unified View of Working, Learning and Organization. *Organization Science*, 2(1): pp. 40-57.

- Brown, B., MacColl, I., Chalmers, M., Galani, A., Randell, C., Steed, A. (2003). Lessons from the Lighthouse: Collaboration in a shared mixed reality system. In *Procs of CHI 2003*, Fort Lauderdale. ACM Press: pp. 577–584.
- Büscher, M., Christensen, M., Hansen, K.M., Mogensen, P., Shapiro, D. (2009). Bottom-up, top-down? Connecting software architecture design with use. In Voss, A., Hartswood, M., Ho, K., Procter, R., Rouncefield, M., Slack, R., Büscher, M. (eds.) *Configuring user-designer relations: Interdisciplinary perspectives*. London: Springer: pp. 157-192.
- Büscher, M. (2005). Social life under the microscope?. *Sociological Research Online*. 1, (available at: [<http://www.socresonline.org.uk/10/1/buscher.html>]; last accessed 04 October, 2010).
- Büscher, M., Eriksen, M.A., Kristensen, J.F., Mogensen, P.H. (2004). Ways of grounding imagination. In *Procs of 8th conference on participatory design*: pp 193-203.
- Büscher, M., Gill, S., Mogensen, P., Shapiro, D. (2001). Landscapes of Practice: Bricolage as a Method for Situated Design. *Journal of Computer Supported Cooperative Work*, 10(1) pp. 1-28.
- Button, G. & Sharrock, W. (2009). *Studies of work and the workplace in HCI: concepts and techniques*. Morgan and Claypool Publishers.
- Button, G. (2000). The ethnographic tradition and design. *Design studies*, 21: pp. 319-332.
- Button, G. & Sharrock, W. (1996). Project work: The organisation of collaborative design and development in software engineering. *Journal of Computer Supported Cooperative Work*, 5(4): pp. 369-386.
- Button, G., Coulter, J., Lee, J.R.E., Sharrock, W. (1995). *Computers, Minds and Conduct*. Cambridge: Polity Press.
- Button, G. (ed.) (1992). *Technology in Working Order: Studies of Work, Interaction, and Technology*. London: Routledge.
- Button, G. (1992). The curious case of the vanishing technology, In In Button, G. (ed.), *Technology in Working Order: Studies of Work, Interaction, and Technology*. London: Routledge: pp. 10-30.
- Button, G. (1991). Introduction: ethnomethodology and the foundational respecification of the human sciences. In Button, G. (ed.), *Ethnomethodology and the human sciences*. London: Routledge & Kegan Paul: pp. 1-9.
- Button, G. & Lee, J.R.E. (eds.) (1987). *Talk and social organisation*. Clevedon: Multilingual Matters.
- Buzbee, B.L., & Sharp, D.H. (1985). Perspectives in supercomputing. *Science*, 227(4687): pp. 591-597.
- Capshew, C., Rader, K. (1992). Big science: Price to present. *Oriis*, 2(7): pp. 2-25.
- Card, S., Moran, T., and Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, NJ.

- Carstensen, P., Schmidt, K. (2003). Computer supported cooperative work: New challenges to systems design. In Itoh, K. (ed.) *Handbook of Human Factors/Ergonomics*, (available at: [http://www.itu.dk/~schmidt/papers/cscw_intro.pdf]; last accessed 29 September, 2010). Asakura Publishing: pp. 619-636.
- Carusi, A. & Jirotko, M. (2009). From data archive to ethical labyrinth. *Qualitative Research*, 9(3): pp. 285-298.
- Cicourel, A. (1992). The interpenetration of communicative contexts: examples from medical encounters. In *Rethinking Context: language as an interactive phenomenon*, Goodwin, C. & Duranti, A. (eds.). Cambridge: Cambridge University Press: pp. 291-310.
- Cockburn, A. (2007). *Agile software development: the cooperative game*, 2nd Edition. London: Addison-Wesley.
- Coulter, J. (1979). *The Social Construction of the Mind: Studies in Ethnomethodology and Linguistic Philosophy*. London: Macmillan Press.
- Cox, A. (2005). What Are Communities of Practice? A Comparative Review of Four Seminal Works. *Journal of Information Science*, 31(6): pp. 527-540.
- Crabtree, A. & Rodden, T. (2007). Hybrid ecologies: understanding cooperative interaction in emerging physical-digital environments. *Personal and Ubiquitous Computing*: pp. 481-493.
- Crabtree, A., French, A., Greenhalgh, C., Benford, S., Cheverst, K., Fitton, D., Rouncefield, M., Graham, C. (2006a): Developing Digital Records: Early Experiences of Record and Replay. *Journal of Computer Supported Cooperative Work*, 15(4): pp. 281-319.
- Crabtree, A., Benford, S., Greenhalgh, C., Tennent, P., Chalmers, M., Brown, B. (2006b) Supporting ethnographic studies of ubiquitous computing in the wild. In *Procs of the 6th conference on Designing Interactive systems, DIS '06*, NY, NY: ACM: pp. 60-69.
- Crabtree, A. (2004). Design in the absence of practice: breaching experiments. In *DIS2004 - Designing Interactive Systems: Across the Spectrum*: pp. 59-68
- Crabtree, A. (2003). *Designing Collaborative Systems: a practical guide to ethnography*. Springer.
- Crabtree, A. (2001). On the practical availability of work practice. In *Ethnographic Studies: Issue No. 6, Special issue: Workplace Studies*, pp. 17-30. (available at: [<http://www.socialsciences.manchester.ac.uk/disciplines/sociology/about/events/ethnography/journal/issue6/>]; last accessed 28 June, 2010).
- Crabtree, A. (2000). Talking work: language games, organisations and computer supported cooperative work. *Journal of Computer Supported Cooperative Work*, 9(2): pp. 215-237.
- Crabtree, A., Nichols, D.M., O'Brien, J., Rouncefield, M., Twidale, M.B. (2000). Ethnomethodologically Informed Ethnography and Information System Design, *Journal of the American Society for Information Science*, 51(7): pp. 666-682.
- Craft, B. & Sasse, A. (2005). Security and Usability of Grid Projects: Implications for e-Science. E-Science workshop position paper. <http://www.brock.craft.org/docs/Craft,%20Sasse,%20e-Science%20Workshop%202005.pdf>

- Daw, M., Procter, R., Lin, Y., Hewitt, T. Jie, W., Baird, K., Turner, A., Birkin, M., Miller, K., Jirotko, M., Schroeder, R., de la Flor, G., Allan, R., Yang, X., Crouchley, R. (2007). Developing an e-infrastructure for social science. In: *Procs. of the 3rd e-Social Science Conf.*, Ann Arbor, MI, October. (available at: [www.ncess.ac.uk/events/conference/2007/papers/paper127.pdf]; last accessed 16 August, 2010).
- de la Flor, G., Luff, P., Jirotko, M., Pybus, J., Kirkham, R., Carusi, A. (2010a). The Case of the Disappearing Ox: Seeing Through Digital Images to an Analysis of Ancient Texts, In *Procs. 28th ACM Conference on Human Factors in Computing Systems*, CHI 2010, Atlanta GA: pp. 473-482.
- de la Flor, G., Jirotko, M., Luff, P., Pybus, J., Kirkham, R. (2010b). Transforming scholarly practice: embedding technological interventions to support the collaborative analysis of ancient texts. *Journal of Computer Supported Cooperative Work*, Special Issue: Supporting Scientific Collaboration Through Cyberinfrastructure and e-Science.
- de la Flor, G., Ojaghi, M., Martínez, I., Jirotko, M., Williams, M., Blakeborough, T. (2010c). Reconfiguring practice: the interdependence of experimental procedure and computing infrastructure in distributed earthquake engineering. *Philosophical Transactions of the Royal Society, A*: pp. 4073-4088.
- de la Flor, G. & Jirotko, M. (2009). Blurring the distinction between software design and work practice. In *Proc. of the 2nd Int. Symp. on End User Development, Work-in-Progress Session*, Siegen, Germany, 2-4 March, pp. 24-31, (available at: [<http://sozio-informatik.org/fileadmin/IISI/upload/IRSI/IRSIV6I1.pdf#page=24>]; last accessed 01 June 2010).
- de la Flor, G., Meyer, E.T. (2008). Talking 'bout a revolution: Framing e-Research as a computerization movement. In *The Oxford eResearch Conference 2008*. 11-13 September 2008. Oxford, UK.
- de la Flor, G., Lloyd, S., Jirotko, M., Warr, A. (2007). Designing software in support of workplace activities-embedding e-science applications. *3rd International Conference on e-Social Science*, (available at: [<http://ess.si.umich.edu/papers/paper189.pdf>]; last accessed 07 September, 2008).
- De Roure, D., et al. (2003). The evolution of the grid. In Berman, F., Fox, G. & Hey, A. (eds.) *Grid Computing: Making The Global Infrastructure a Reality*. Chichester: Wiley: pp.65-100.
- Douglas, S. (1995). Conversation Analysis and Human-Computer Interaction Design. In Thomas, P. (ed.) *The Social and Interactional Dimensions of Human-Computer Interfaces*. Cambridge: Cambridge University Press: pp. 184-203.
- Dourish, P. (2006). Implications for design. In *Procs. Conference on Human Factors in Computing Systems*, Montreal CA: pp. 541-550.
- Dourish, P. (2004). What We Talk About When We Talk About Context. *Personal and Ubiquitous Computing*, 8(1): pp. 19-30.
- Dourish, P. (2001). *Where the action is: The foundations of embodied interaction*. Cambridge, MA: MIT Press.
- Dutton, W. H., & Jeffreys, P. W. (eds.) (2010). *World Wide Research: Reshaping the Sciences and Humanities* (eds). Cambridge, MA: MIT Press.

- Dutton, W. H., & Meyer, E. T. (2008). The Diffusion of e-Research: The Use and Non-Use of Advances in Information and Communication Technologies across the Social Sciences. *4th International Conference on e-Social Science*. 18-20 June, Manchester, UK. (available at: [<http://ssrn.com/abstract=1150422>]; last accessed 28 June, 2010).
- Edwards, W.K., Newman, M.W., Shehan Poole E. (2010). The infrastructure problem in HCI, In *Procs. 28th ACM Conference on Human Factors in Computing Systems, CHI 2010*, Atlanta GA, USA: pp. 423-432.
- Ehn, P. (1991). The art and science of designing computer artifacts. In Dunlop, C. & Kling, R. (eds.) *Computerization and controversy: value conflicts and social choices*. London: Academic Press: pp. 288-307.
- Ehn, P. (1990). *Work-Oriented Design of Computer Artifacts*. Lawrence Erlbaum Associates.
- Ehn, P. & Kyng, M. (1987). The collective resource approach to systems design. In Bjerknes, G., Ehn, P., Kyng, M (eds.) *Computers and democracy: a Scandinavian challenge*. Aldershot: Avebury: pp. 17-57.
- Ellis, D., Oldridge, R. and Vasconcelos A. (2004). Community and virtual community, *Annual Review of Information Science and Technology*, 38: pp. 145-186.
- Ericsson, K.A. & Simon, H.A. (1980). Verbal protocol as data. *Psychological Review*, 87(3): pp. 215-251.
- Eriksson, J., Ek, A., Johansson, G. (2000). Design and evaluation of a software prototype for participatory planning of environmental adaptations. In *IEEE Transactions on Rehabilitation Engineering* 8(1): pp. 94-106.
- Ewusi-Mensah, K. (2003). *Software Development Failures: Anatomy of Abandoned Projects*. Cambridge, MA: MIT Press.
- Finholt, T. (2003). Collaboratories as a new form of scientific organization. *Economics of Innovation and New Technology*, 12(1): pp. 5-25.
- Fitzpatrick, G. (2003). *The Locales Framework: Understanding and Designing for Wicked Problems*. Kluwer Academic Publishers.
- Floyd, C. (1984). A systematic look at prototyping. In Budde, R., Kuhlenkamp, K., Mathiassen, L., Züllighoven, H. (eds.), *Approaches to Prototyping*, London: Springer Verlag: pp. 1-18.
- Foster, A. (2008). New Ways to Connect Data, Computers, and People. *The Chronical of Higher Education*, (available at: [<http://chronicle.com/wiredcampus/article/3146/new-ways-to-connect-data-computers-and-people>]; last accessed 06 September, 2008).
- Foster, I., Kesselman, C. & Tuecke, S. (2001). The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *International Journal of Supercomputer Applications*, 15(3): pp. 200-222.
- Foster, I. & Kesselman, C. (1997). Globus: A metacomputing infrastructure toolkit. *International Journal of High Performance Computing Applications*, 11(2): pp.115-128.

- Fraser, M., Hindmarsh, J., Best, K., Heath, C., Biegel, G., Greenhalgh, C. & Reeves, S. (2006). Remote Collaboration over Video Data: Towards Real-Time e-Social Science. *JCSCW, Special Issue on Collaboration in e-Research*, 15(4): pp. 257-279.
- Freeman, P. et al. (2005). Cyberinfrastructure for science and engineering: promises and challenges. *Proc. of the IEEE*, 93(3): pp. 682-691.
- Garfinkel, H. (1996). Ethnomethodology's Programme. *Social Psychology Quarterly*, 59(1): pp. 5-21.
- Garfinkel, H. & Wieder, D.L. (1992). Two incommensurable, asymmetrically alternate technologies of social analysis. In Watson, G. & Seiler, R. M. (eds.), *Text in Context: Contributions to Ethnomethodology*. London: Sage: pp. 175–206.
- Garfinkel, H. (1991). Respecification: Evidence for locally produced, naturally accountable phenomena of order, logic, reason, meaning, method, etc. in and as of the essential haecceity of immortal ordinary society (I)-An announcement of studies. In Button, G. (ed.), *Ethnomethodology and the human sciences*. London: Routledge & Kegan Paul: pp. 10-19.
- Garfinkel, H., Lynch, M. and Livingston, E. (1981). The Work of Discovering Science Construed with Materials from the Optically Discovered Pulsar. *Philosophy of the Social Sciences*, 11(2): pp. 131-158.
- Garfinkel, H. (1974). The origins of the term 'Ethnomethodology', in Turner, R. (ed.) *Ethnomethodology: selected readings*. Harmondsworth: Penguin: pp. 15-18.
- Garfinkel, H., & Sacks, H. (1970). On formal structures of practical actions. Reprint in *Ethnomethodological studies of work*. (1986), London: Routledge: pp. 160–193, (original work published in McKinney, J. & Tiryakian, A. (1970) Theoretical sociology: perspectives and developments)
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Galison, P. (1992). The many faces of big science. In Galison, P., Hevly, B. (eds.) *Big science: the growth of large-scale research*. Stanford, CA: Stanford University Press: pp. 1–17.
- Gibson, J.J. (1986). *The ecological approach to visual perception*. New York: Psychology Press.
- Goffman, E. (1983). The interaction order: American Sociological Association, 1982 Presidential Address, *American Sociological Review*, 48(1): pp. 1–17.
- Goffman, E. (1981). *Forms of talk*. Oxford: Basil Blackwell.
- Goguen, J. & Linde, C. (1993). Techniques for requirements elicitation. In *Procs of RE '93*, Jan. 4-6, San Diego California, IEEE: pp. 152-164.
- Goodwin, C. (1995). Seeing in depth. *Social studies of science*, 25(2) pp. 237-274.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3) pp. 606-633.
- Goodwin, C. & Duranti, A. (eds.) (1992). *Rethinking Context: language as an interactive phenomenon*. Cambridge: Cambridge University Press.

- Goodwin, C. (1981). *Conversational organization: interaction between speakers and hearers*. London: Academic Press.
- Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating GOMS for predicting and explaining real-world task performance. *Human Computer Interaction*, 8(3): pp. 237-309.
- Gray, W. & Salzman, M. (1998). Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods. *Human-Computer Interaction*, 13(3): pp. 203-261.
- Greenbaum, J., Kyng, M. (eds.) (1991). *Design at Work: Cooperative Design of Computer Systems*. Lawrence Erlbaum Associates.
- Greenberg, S. & Buxton, B. (2008). Usability evaluation considered harmful (some of the time). In *Proc of the twenty-sixth annual SIGCHI conference on Human factors in computing system*: pp. 111-120.
- Greif, I. (ed.) (1988). *Computer-supported cooperative work: a book of readings*. SF: Morgan Kaufmann.
- Grint, K. & Woolgar, S. (1997). *The machine at work: technology, work and organization*. Cambridge: Polity Press.
- Grønbaek, K. (1990). Supporting active user involvement in prototyping. *Scandinavian Journal of Information Systems*, 2: pp. 3-24.
- Grudin, J. (1994). Computer-supported cooperative work: history and focus. *Computer*, 27(5): pp. 19-26.
- Grudin, J. (1990). The computer reaches out: the historical continuity of interface design. *Proc of the SIGCHI conference on Human factors in computing systems: Empowering people*: pp. 261-268.
- Grudin, J. (1988). Why CSCW Applications Fail: Problems in the Design and Evaluation of Organizational Interfaces. In *Procs of CSCW 88*, (Portland, Oregon): pp. 85-93.
- Halfpenny, P. & Procter, R. (2010). The e-Social Science research agenda. *Philosophical Transactions of the Royal Society A*, 368: pp. 3761–3778.
- Hammersley, M. (1998). *Reading ethnographic research: A critical guide*, 2nd Edition. London: Longman.
- Harper, R. H. R. (2000). The organisation of ethnography: A discussion of ethnographic fieldwork programs in CSCW. *Computer Supported Cooperative Work*, 9(2): pp. 239-264.
- Hartwood, M., Procter, R., Rouncefield, M., Slack, R., Soutter, J., Voss, A. (2003). 'Repairing' the Machine: A Case Study of the Evaluation of Computer-Aided Detection Tools in Breast Screening. In *Procs of the 8th European Conference on Computer-Supported Cooperative Work, ECSCW '03*, Helsinki, Finland. Springer: pp. 375-394.
- Hartwood, M., Procter, R., Slack, R., Voss, A., Büscher, M., Rouncefield, M., Rouchy, P. (2002). Co-realisation, Towards a principled synthesis of ethnomethodology and participatory design. *Scandinavian Journal of Information Systems*, 14(2) pp. 9-30.

- Heath, C., Hindmarsh, J., Luff, P. (2010). *Video in Qualitative Research: Analysing social interaction in everyday life*. London: Sage.
- Heath, C. & Hindmarsh, J. (2002). Analysing Interaction: Video, ethnography and situated conduct. In May, T. (ed.) *Qualitative Research in Action*. London: Sage: pp. 99-121.
- Heath, C. & Button, G. (eds.) (2002). Special issue on workplace studies. *British Journal of Sociology*, 53(2): pp. 157-317.
- Heath, C. & Button, G. (2002). Editorial Introduction. Special issue on workplace studies. *British Journal of Sociology*, 53(2): pp. 157-161.
- Heath, C. & Luff, P. (2000). *Technology in Action*. Cambridge: Cambridge University Press.
- Heath, C. Knoblauch, H., Luff, P. (2000): Technology and social interaction: the emergence of 'workplace studies'. *British Journal of Sociology*, 51(2): pp. 299-320.
- Heath, C., Luff, P., & Sellen, A.J. (1997). Reconfiguring media space: Supporting collaborative work. In K. Finn, A. Sellen, & S. Wilbur (eds.), *Video-mediated Communication*. New Jersey: Erlbaum: pp. 323-349.
- Heath, C. & Luff, P. (1996). Documents and professional practice: "Bad" organisational reasons for "good" clinical records. In *Procs of Computer Supported Cooperative Work '96*. Cambridge, MA: pp. 354-363.
- Heath, C., Jirotko, M., Luff, P., Hindmarsh, J. (1994). Unpacking collaboration: the interactional organisation of trading in a city dealing room. *Journal of Computer Supported Cooperative Work*, 3(2): pp. 147-165.
- Heath, C. & Luff, P. (1992). Media space and communicative asymmetries. *Human Computer Interaction*, 7(3): pp. 315-346.
- Heath, C. & Luff, P. (1991). Collaborative activity and technological design: task coordination in London underground control rooms. In *Proc. of the 2nd European Conference on Computer Supported Cooperative Work*: pp. 65-80.
- Heath, C., Luff, P., Sellen, A. (1995). Reconsidering the Virtual Workplace: Flexible Support for Collaborative Activity. In *Proc of the 1995 Fourth European Conference on Computer Supported Cooperative Work*: pp. 83-99.
- Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge: Cambridge University Press.
- Heidegger, M. (1977). *The question concerning technology, and other essays*. Translated by William Lovitt. New York: Harper & Row.
- Heritage, J. (2004). Conversation analysis and institutional talk. In Silverman, D (ed). *Qualitative research: theory, method and practice*, 2nd ed. London: Sage.
- Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity Press.
- Hewett, T. et al. (1996). ACM SIGCHI Curricula for Human-Computer Interaction, (available at: [<http://old.sigchi.org/cdg/>]; last accessed 28 June, 2010).

- Hey, T. & Tefethen, A. (2005). The e-Science Challenge: Creating a Reusable e-Infrastructure for Collaborative Multidisciplinary Science. *CTWatch Quarterly*, 1(4), (available at:[<http://www.ctwatch.org/quarterly/articles/2005/11/the-e-science-challenge-creating-a-reusable-e-infrastructure-for-collaborative-multidisciplinary-science/index.html>]; last accessed 07 September, 2010).
- Hey, T. & Trefethen, A. (2003a). e-Science and its Implications. *Philosophical Transactions of the Royal Society A*, 361: pp. 1809-1825.
- Hey, T. & Tefethen, A. (2003b). The data deluge: an e-Science perspective. In Berman, F., Fox, G. & Hey, A. (Eds.) *Grid Computing: Making The Global Infrastructure a Reality*. Chichester: Wiley, p.809-824.
- Hey, T. & Trefethen, A. (2002). The UK e-Science core programme and the Grid. *Future Generation Computing Systems*, 18(8): pp. 1017-1031.
- Hindmarsh, J. & Heath, C. (2007). Video-based studies of work. *Sociology Compass*, 1(1): pp. 156-173.
- Hinds, C. (2008). The case against a positivist philosophy of requirements engineering. *Requirements Engineering*, 13(4): pp. 315-328.
- Hoare, C. A. R. (1996). How did software get so reliable without proof? *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer: pp. 1-17.
- Hollan, J., Hutchins, E., Kirsh, D. (2000). Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research. *ACM Transactions on Computer-Human Interaction*, 7(2): pp. 174-196.
- Hughes, J.A., King, V., Rodden, T., Andersen, H., (1995). The role of ethnography in interactive systems design. *Interactions*, 2(2): pp.56-65.
- Hughes, J., King, V., Rodden, T., Andersen, H. (1994). Moving out from the control room: ethnography in system design. In *Proc. of the 1994 ACM conference on Computer supported cooperative work*: pp. 429-439.
- Hughes, J.A., Somerville, I., Bentley, R., Randall, D. (1993). Designing with ethnography: making work visible. *Interacting with Computers*, 15(2): pp.239-233.
- Hughes, J. A., Randall, D., & Shapiro, D. (1992). Faltering From Ethnography to Design. In *Procs of the Conference on Computer Supported Cooperative Work*. (Toronto, CA): pp. 115-122.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge MA: MIT Press.
- Jirotko, M., Proctor, R., Rodden, T., Bowker, G. (2006). Special issue: Collaboration in e-Research. *International Journal of Computer Supported Cooperative Work*, 15(4): pp. 251-255.
- Jirotko, M. & Luff, P. (2006). Supporting Requirements with Video-Based Analysis. *IEEE Software*, 23(3): pp. 42-44.

- Jirotko, M., Procter, R., Hartswood, M., Slack, R., Simpson, A., Coopmans, C., Hinds, C. and Voss, A. (2005). Requirements for Collaboration and Trust in Healthcare Innovation: The eDiaMoND Case Study. *International Journal of Computer Supported Cooperative Work*, 14(3): pp. 369-398.
- Jirotko, M. (2003). eDiaMoND Technical Report. <http://www.ediamond.ox.ac.uk/>.
- Jirotko, M. & Goguen, J. (1994). *Requirements Engineering: Social and Technical Issues*. London: Academic Press.
- Johansen, R. (1988). *Groupware. Computer support for business teams*. New York: Free Press.
- John, B., & Keiras, D. (1996). "Using GOMS for User Interface Design and Evaluation: Which Technique?". *Transactions on Computer-Human Interaction*, 3: pp. 287-319.
- Johnson, P. (1989). Supporting system design by analyzing current task knowledge. In Diaper, D. (Ed.). *Task Analysis for Human-Computer Interaction*. Chichester: Ellis Horwood: pp. 160-185.
- Jones, M., Rieger, R., Treadwell, P., Gay, G. (2000). Live from the Stacks: User Feedback on Mobile Computers and Wireless Tools for Library Patrons. *International Conf on Digital Libraries*, ACM: pp. 95-102.
- Jordan, B. (1996). Ethnographic Workplace Studies and Computer Supported Cooperative Work. In: *The Design of Computer-Supported Cooperative Work and Groupware Systems*. Shapiro, D., Tauber, M. & Traunmüller, R. (eds.) Amsterdam, The Netherlands: North Holland/Elsevier Science: pp. 17-42, (available at:[http://www.lifescapes.org/Papers/ethnographic_workplace_studies.htm]; last accessed 12 July, 2010).
- Jordan, B. & Henderson, A. (1995). Interaction analysis: foundations and practice. *The Journal of the Learning Sciences*, 4(1): pp. 39-103.
- Kahler, H. (2000). Constructive interaction and collaborative work: introducing a method for testing collaborative systems. *Interactions*, 7(3): pp. 27-34.
- Kalawsky, R.S., Holmes, I.R., O'Brien, J., Lewin, M., Armitage, S., Nee, S.P. (2006). JISC Human Factors Audit of Selected e-Science Projects. Technical report, (available at:[<http://syseng.lboro.ac.uk/jisc/humanfactors>]; last accessed 28 June, 2010).
- Karasti, H. (2001). Bridging Work Practice and System Design: Integrating Systemic Analysis, Appreciative Intervention and Practitioner Participation. *Journal of Computer Supported Cooperative Work*, 10(2): pp. 211-246.
- Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- Kensing, F., Simonsen, J., Bødker, K. (1998). MUST – A Method for Participatory Design. *Human-Computer Interaction*, 13(2): pp. 167-198.
- Kensing F, & Blomberg J. (1998). Participatory design: issues and concerns. *Journal of Computer Supported Cooperative Work*, 7: pp. 167-185.

- Kirk, D. S., Rodden, T. & Stanton Fraser, D. (2007). Turn It This Way: Grounding Collaborative Action with Remote Gestures. In *Proc. of CHI Conference on Human Factors in Computing Systems*, 28th April-3rd May, San Jose, CA, ACM: pp. 1039-1048.
- Kirk, D., Crabtree, A. & Rodden, T. (2005). Ways of the Hands. In *Procs of the 9th European Conference on Computer-Supported Cooperative Work, ECSCW '05*, Paris, France. Springer: pp. 1-22.
- Kling, R. & Iacono, S. (1994). Computerization Movements and the Mobilization of Support for Computerization. In Starr, L. (Ed.) *Ecologies of Knowledge: work and politics in science and technology*. Albany: SUNY Press: pp. 119-153.
- Knoblauch, H., Schnettler, B., Raab, J., Soeffner, H.G. (2006). *Video Analysis: Methodology and Methods*. Frankfurt: Philr Lang.
- Kuzouka, H., Kosaka, J., Yamazaki, K., Suga, S., Yamazaki, A., Luff, P., Heath, C. (2004). Mediating dual ecologies. *Procs. of CSCW '04*, Chicago: ACM Press: pp. 477-486,
- Latour, B. & Woolgar, S. (1979). *Laboratory life: the construction of scientific facts*. Beverly Hills: Sage.
- Lave, J & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Law, J. & Hassard, J. (eds.) (1999). *Actor network theory and after*. Oxford: Blackwell.
- Liberman, K (2004). *Dialectical Practice in Tibetan Philosophical Culture*. Oxford: Rowman & Littlefield.
- Lloyd, S. & Simpson, A.C. (2005). Project management in multi-disciplinary collaborative research. In *Procs of International Professional Communication Conference (IPCC)*. Limerick.
- Luff, P., Heath, C., Kuzuoka, H., Yamazaki, K., & Yamashita, J. (2006): Handling Documents and Discriminating Objects in Hybrid Spaces. In *Proc. of ACM CHI 2006, (CHI '06)* (Montreal, Ontario, April 22-27, 2006): pp. 561-570.
- Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., Oyama, S. (2003). Fractured ecologies: creating environments for collaboration. *Human-Computer Interaction*, 18(1): pp. 51-84.
- Luff, P., Hindmarsh, J., Heath, C. (eds.) (2000). *Workplace Studies Recovering Work Practice and Informing System Design*. Cambridge UK: Cambridge University Press.
- Luff, P. (1997). Computers and interaction: the social organisation of human-computer interaction in the workplace. *PhD thesis*. University of Surrey.
- Lynch, M. (2006). Cognitive activities without cognition? ethnomethodological investigations of selected 'cognitive' topics. *Discourse Studies*, 8(5): pp.95-104.
- Lynch, M. (1999). Silence in context: Ethnomethodology and social theory. *Human Studies*, 22(2-4): pp.211-233.
- Lynch, M. & Bogen, D. (1994). Harvey Sack's primitive natural science. *Theory, Culture & Society*, 11(4): pp.65-104.

- Lynch, M. (1993). *Scientific practice and ordinary action: Ethnomethodology and social studies of science*. Cambridge: Cambridge University Press.
- Lynch, M. (1991). Laboratory space and the technological complex: an investigation of topical contextures. *Science in context*, 4(1): pp.51-78.
- Lynch, M. & Woolgar, S. (eds.) (1990). *Representation in Scientific Practice*. Cambridge, MA: MIT Press.
- Lynch, M. (1990). The externalized retina: Selection and mathematization in the visual documentation of objects in the life sciences. In Lynch, M. & Woolgar, S. (eds.) *Representation in Scientific Practice*. Cambridge, MA: MIT Press: pp. 153-186.
- Lynch, M. (1985). *Art and artefact in laboratory science: A study of shop work and shop talk in a research laboratory*. London/Boston: Routledge Kegan & Paul.
- MacKenzie, D. & Wajcman, J. (1999). *The social shaping of technology*, 2nd edition. Buckingham: Open University Press.
- Mahin, S.A. & Shing, P.B. (1985). Pseudodynamic Method for Seismic Testing. *ASCE Journal of Structural Engineering*, 111(7): pp.1482-1503.
- Marcus, G. (1995). Ethnography In/Of the World System: the Emergence of Multi-Sited Ethnography. *Annual Review of Anthropology*, 24: pp. 95-117.
- Martin, D., Rodden, T., Rouncefield, M., Sommerville, I., Viller, S. (2001). Finding Patterns in Fieldwork. In *Procs of the 7th European Conference on Computer-Supported Cooperative Work, ECSCW '01*, Bonn, Germany, Springer: pp. 39-58.
- Maynard, D.W., & Clayman, S.E. (1991). The diversity of ethnomethodology. *Annual Review of Sociology* 17(1): pp. 385-418.
- Meyer, E. T., & Dutton, W. H. (2008). Top-Down e-Infrastructure Meets Bottom-Up Research Innovation: Fitting e-Social Science Visions to the Realities. *UK e-Science All Hands Meeting*. 8-11 Sept, Edinburgh, UK, (available at: [<http://ssrn.com/abstract=1262211>]; last accessed 28 June, 2010).
- Meyer, E., Schroeder, R., Dutton, W. (2008). The Role of e-Infrastructures in the Transformation of Research Practices and Outcomes. *iConference 2008*, (available at:[http://www.ischools.org/oc/conference08/pc/PA10-2_iconf08.pdf]; last accessed 5 September, 2008)
- Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive science*, 10(2): pp. 151-177.
- Mogensen, P. (1992). Towards a Prototyping Approach in Systems Development. *Scandinavian Journal of Information Systems*, 4: pp.31-53.
- Mogensen, P. & Trigg, R.H. (1992). Using Artifacts as Triggers for Participatory Analysis. In M. J. Muller, S. Kuhn, & J. A. Meskill (Eds.), *Procs. of the PDC'92*. (available at:[<http://www.daimi.au.dk/PB/413/PB-413.ps.gz>]; last accessed 25 August, 2010).
- Möllering, G. (2006). *Trust: reason, routine, reflexivity*. Amsterdam; Oxford: Elsevier.

- Molton, N., Pan, X., Brady, M., Bowman, A., Crowther, C., & Tomlin, R. (2003). Visual enhancement of incised text. *Pattern Recognition*, 36: pp. 1031-1043.
- Muller, M. (2002). Participatory design: the third space in HCI. In Jacko, J., Sears, A., (eds.). *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*. Lawrence Erlbaum Associates: pp. 1051-1068.
- Nakashima, M. (2008). Roles of Large Structural Testing for the Advancement of Earthquake Engineering. *The 14th World Conference on Earthquake Engineering*, (Beijing, China), Keynote address.
- Nardi, B. & Redmiles, D. (2002). Activity Theory and the Practice of Design. *A special issue of The Journal of Computer-supported Cooperative Work*, 11 (1-2).
- National Science Foundation, (NSF). (2007). Cyberinfrastructure Vision for 21st Century Discovery, (available at: [<http://www.nsf.gov/pubs/2007/nsf0728/nsf0728.pdf>]; last accessed 07 September, 2010).
- Neale, D.C., Carroll, J.M., Rosson, M.B. (2004). Evaluating Computer-Supported Cooperative Work: Models and Frameworks, *Proceedings of CSCW'04*, (Chicago): pp. 112-121.
- Newhouse, S., Schopf, J., Richards, A., Atkinson, M. (2007). Study of Users' Priorities for E-Infrastructure for Research (SUPER). OMII Draft report, February 2007, (available at: [<http://www.nesc.ac.uk/action/esi/download.cfm?index=3368>]; last accessed 28 June, 2010).
- Nielsen, J. (1993). *Usability Engineering*. Academic Press.
- Nielsen, J. (1992). The usability engineering life cycle, *IEEE Computer* 25(3): pp. 12–22.
- Norman, D. (1988). *The Design of Everyday Things*. Basic Books.
- Norman, D and Draper S. (eds.) (1986) *User Centred System Design: New Perspectives on Human Computer Interaction*. Lawrence Erlbaum.
- Nuseibeh, B., Easterbrook, S., (2000). Requirements engineering: a roadmap, *Proc of the Conference on The Future of Software Engineering*, June 04-11, 2000, (Limerick, Ireland): pp. 35-46.
- Ojaghi, M., Lamata Martínez, I., Dietz, M., Williams, M.S., Blakeborough, A., Crewe, A., Taylor, C., Madabhushi, G., Haigh, S., Ali, A. (2010). UK-NEES - distributed hybrid testing between Bristol, Cambridge and Oxford universities: connecting structural dynamics labs to a geotechnical centrifuge. In *Proc. 9th U.S. National and 10th Canadian Conf. on Earthquake Engineering*, Toronto, Ontario, 25–29 July 2010, paper 1024. Oakland, CA: Earthquake Engineering Research Institute, [<http://2010eqconf.org>].
- Ojaghi, M., Williams, M.S., Blakeborough, A. (2008). Development of Energy Dissipating Devices Using Real Time Hybrid Testing. In *14th World Conference on Earthquake Engineering: Innovation Practice Safety*, October 12-17, 2008, Beijing, China, (available at: [<http://www.14wcee.org/Proceedings/files/11-0076.PDF>]; last accessed 04 August, 2010)
- Ojaghi, M., Parra-Fuente, J., Saleem, K., Blakeborough, A., Williams M.S. (2007). Grid Based Distributed Hybrid Testing. In *Proc. of the UK e-Science All Hands Meeting*, (Nottingham, UK): pp. 190-196.

- O'Keefe, D. J. (1979). Ethnomethodology. *Journal for the Theory of Social Behavior*, 9(2): pp. 187-219.
- O'Malley, C.E., Draper, S., Riley, M. (1984). Constructive interaction: A method for studying human-computer-human interaction. In *Proc. IFIP Conf. Human-Computer Interaction: Interact'84*, (Amsterdam): pp. 269-274.
- Orr, J.E. (1996). *Talking about machines: an ethnography of a modern job*. Ithaca, NY: Cornell University Press.
- Ovaska, S., (1991). Usability as a Goal for the Design of Computer Systems. *Scandinavian Journal of Information Systems*, 3: pp. 47-62.
- Peräkylä A. (2004). Reliability and validity in research based on naturally occurring social interaction. In Silverman, D. (ed.) *Qualitative research: Theory, method and practice*. London: Sage: pp. 283–304.
- Pinelle, D. & Gutwin, C. (2000). A review of groupware evaluations. *Proc 9th IEEE WETICE Infrastructure for Collaborative Enterprises*: pp. 86-91.
- Pitt-Francis, J., Bernabeu, M.O., Cooper, J., Garny, A., Momtahan, L., Osborne, J., Pathmanathan, P., Rodriguez, B., Whiteley, J.P., Gavaghan, D.J. (2008). Chaste: using agile programming techniques to develop computational biology software. *Philosophical Transactions of the Royal Society A*, 366: pp. 3111–3136.
- Polson, P.G., Lewis, C., Rieman, J., Wharton, C. (1992). Cognitive walkthroughs: A method for theory- based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36: pp. 741-773.
- Potts, C. (1997). Requirements models in context. In *Proc. 3rd IEEE International Symposium on Requirements Engineering*. 6-10 January 1997, Annapolis, MD: pp. 102-104.
- Preece, J., Rogers, Y., Sharp, H. (2002). *Interaction design: Beyond human-computer interaction*. John Wiley & Sons.
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-computer interaction*. Addison-Wesley.
- Price, D. J. d. S. (1984). The science/technology relationship, the craft of experimental science, and policy for the improvement of high technology innovation. *Research Policy*, 13(1), pp. 3-20.
- Price, D.J.d.S. (1963). *Little Science, Big Science*. Columbia University Press.
- Procter, R., Borgman, C., Bowker, G., Jirotko, M., Olsen, G., Pancake, C., Rodden, T., Schraefel, M.C. (2006a). Usability research challenges for cyberinfrastructure and tools. *Proc. of ACM CHI 2006, Workshop on Human Factors in Computing Systems*, 2: pp. 1675-1678.
- Procter, R., Batty, M., Birkin, M., Crouchley, R., Dutton, W., Edwards, P., Fraser, M., Halfpenny, P., and Rodden, T. (2006b). The National Centre for e-Social Science'. In: *Procs of the UK e-Science All Hands Meeting (AHM'06)*, National e-Science Centre (NeSC), September: pp. 542-549.
- Psathas, G. (1995). *Conversation analysis: The study of talk-in-interaction*. London: Sage.

- Rana, M Cirstea, M., Reynolds, T. (2008). Developing a prototype using mobile devices to assist visually impaired users. In *IEEE International Symposium on Industrial Electronics*: pp. 1826-1830.
- Randall, D., Harper, R., Rouncefield, M. (2007). *Fieldwork for Design*. London: Springer-Verlag.
- Randall, D., Marr, L., Rouncefield, M. (2001). Ethnography, ethnomethodology and interaction analysis. In *Ethnographic Studies: Issue No. 6, Special issue: Workplace Studies*, pp. 31-44. (available at: [<http://www.socialsciences.manchester.ac.uk/disciplines/sociology/about/events/ethnography/journal/issue6/>]; last accessed 28 June, 2010).
- Rawls, A. (2008). Harold Garfinkel, Ethnomethodology and Workplace Studies. *Organization Studies*, 29(5): pp. 701-732.
- Rawls, A. (2003). Harold Garfinkel. In Ritzer, G. (ed) *The Blackwell Companion to Major Contemporary Social Theorists*, 2nd Ed. Oxford: Blackwell.
- Rawls, A. (2001). Durkheim's Treatment of Practice: Concrete Practice vs Representations as the Foundation of Reason. *Journal of Classical Sociology*, 1(33): pp. 33-68.
- Rodden, T. (2007). Unknown presentation title. Oxford: Computing Laboratory, University of Oxford.
- Rogers, Y. (2000). Recent Theoretical Developments in HCI; Their Value for Informing Systems Design. Working paper. <http://www.cogs.sussex.ac.uk/>.
- Rosenberg, N. (1992). Scientific instrumentation and university research. *Research Policy*, 21: pp. 381-390.
- Roued Olsen, H., Tarte, S, Terras, M., Brady, M. & Bowman, A. (2009). Towards an interpretation support system for reading ancient documents. *Digital Humanites'09*: pp. 237-39.
- Sacks, H. (1992). *Lectures in conversation: Volumes I and II*. Oxford: Blackwell.
- Sacks, H. (1984). Notes on methodology. In Atkinson, M. J., & Heritage, J. (eds.). *Structures of social action: Studies in conversation analysis*. Cambridge, Cambridge University Press: pp. 21-27.
- Sacks, H. (1972). An initial investigation of the usability of conversational data for doing Sociology. In: Sudnow, D.N. (ed.), *Studies in Social Interaction*. Free Press, New York: pp. 31-74.
- Sacks, H., Schegloff, E., Jefferson, G. (1974). A simplest systematics of the organization of turn-taking in conversation. *Language*, 50(4): pp. 696-735.
- Schmidt, K., Simone, C. (1996). Coordinating mechanisms: Towards a conceptual foundation of CSCW systems design. *Journal of Computer Supported Cooperative Work*, 5: pp. 155-200.
- Schmidt, K., Bannon, L. (1992): Taking CSCW Seriously: Supporting Articulation Work. *Journal of Computer Supported Cooperative Work*, 1 (1/2): pp. 7-40.
- Schegloff, E. (2007). A tutorial on membership categorization. *Journal of Pragmatics*, 39(3): pp. 462-482.

- Schegloff, E. (2001). Accounts of Conduct in Interaction: Interruption, Overlap, and Turn-Taking. In *Handbook of Sociological Theory*, Turner, J.T. (ed.). New York: Springer: pp. 287-321.
- Schegloff, E. (1992). In another context. In *Rethinking Context: language as an interactive phenomenon*, Goodwin, C. & Duranti, A. (eds.). Cambridge: Cambridge University Press: pp. 191-227.
- Schegloff, E. (1987). Between micro and macro: contexts and other connections. In Alexander, J. et al (eds.) *The micro-macro link*. Berkeley: University of California Press: pp. 207-234.
- Schegloff, E. (1972). Notes on a conversational practice: formulating place. In Sudnow, D (ed.) *Studies in social interaction*. New York: Free press: pp. 75-119.
- schraefel, m. c. & Dix, A. (2009). Within Bounds and Between Domains: Reflecting on Making Tea within the Context of Design Elicitation Methods. *International Journal of Human Computer Studies*, 67(4): pp. 313-323.
- Schroeder, R. & Fry, J. (2007). Social science approaches to e-Science: Framing an agenda. *Journal of Computer-Mediated Communication*, 12: pp. 563-582.
- Schutz, A. & Luckmann, T. (1973). *The Structures of the Life-World, Vol. 1*. Northwestern University Press.
- Schutz, A. (1970). *Reflections on the Problem of Relevance*. New Haven: Yale University Press.
- Schutz, A. (1962). *Collected papers 1: the problem of social reality*. Natanson, M. (ed). London: Kluwer Academic.
- Schutz, A. (1932). *The phenomenology of the social world*. Translated by Walsh, G. & Lehnert, F., 1967. Evanston: Northwestern University Press.
- Shapiro, D. (1994). The limits of ethnography: combining social sciences for CSCW. In *Proc.ACM Conference on Computer-Supported Cooperative Work CSCW'94*, Chapel Hill, North Carolina: pp. 417-428.
- Sharrock, W. & Coulter, J. (2003). Dissolving the "projection problem". *Visual studies*, 18(1): pp. 74-82.
- Sharrock, W. & Coulter, J. (1998). On What We Can See. *Theory and Psychology*, 8(2): pp. 147-164.
- Shepherd, A. (2000). HTA as a framework for task analysis. In Annett, J., Stanton, N., (eds.). *Task analysis*. Taylor & Francis: pp. 9-24.
- Shneiderman, B. (1998). *Designing the user interface*. Addison Wesley.
- Siddiqi, J. Shekaran, M. (1996). Requirements Engineering: The Emerging Wisdom. *IEEE Software*, 13(2): pp. 15-19.
- Silverman, D. (2006). *Interpreting qualitative data: Methods for analyzing talk, text, and interaction*, 3rd Edition. London: Sage.

- Silverman, D. (1998). *Harvey Sacks: social science and conversation analysis*. Cambridge: Polity Press.
- Smarr, L. & Catlett, C. (1992). Metacomputing. In Berman, F., Fox, G. & Hey, A. (eds.) *Grid Computing: Making The Global Infrastructure a Reality*. Chichester: Wiley, 2003: pp. 825-835.
- Smith, A. & Dunckley, L. (2002). Prototype Evaluation and Redesign: Structuring the Design Space Through Contextual Techniques. *Interacting with Computers*, 14(6): pp. 821-843.
- Sommerville, I. (2004). Making ethnography accessible. Proc. Workshop on Software Engineering/HCI, Edinburgh, May 2004, *Institute of Electrical Engineers*: pp. 36-40.
- Somerville, I. (1989). *Software Engineering*, 3rd Edition. Addison-Wesley.
- Spencer B. F., Elnashai A., Nakata N., Saliem H., Yang G., Futrelle J., Glick W., Marcusiu D., Ricker K., Finholt T., Horn D., Hubbard P., Keahey K., Liming L., Zaluzec N., Pearlman L., Stauffer E., 2004. *The MOST Experiment: Earthquake Engineering on the Grid, NEES grid whitepaper 1.0*. (available at: [http://neesgrid.ncsa.illinois.edu/documents/TR_2004_41.pdf]; last accessed 25 March 2010).
- Standish Group (1994). The Chaos Report. http://www.standishgroup.com/sample_research/chaos_1994_1.php.
- Star, S. L. & Ruhleder, K. (1994). Steps Towards an Ecology of Infrastructure: Complex Problems in Design and Access for Large-Scale Collaborative Systems. In *Proceedings of CSCW'94*, (Chapel Hill, NC, 1994), ACM Press: pp. 253-264.
- Suchman, L. (2007). *Human-machine reconfigurations: plans and situated actions, 2nd edition*. Cambridge University Press.
- Suchman, L., Trigg, R. and Blomberg, J. (2002). Working artefacts: ethnomethods of the prototype. *British Journal of Sociology*, 53(2): pp. 163-179.
- Suchman, L. (1997). Centers of Coordination: A case and some themes. In Resnick, L. B., Säljö, R., Pontecorvo, C., & Burge, B. (eds.) *Discourse, Tools, and Reasoning: Essays on Situated Cognition*. Berlin: Springer-Verlag: pp. 41-62. (available at: [http://www.lancs.ac.uk/fass/doc_library/sociology/Suchman_Centers_1997.pdf]; last accessed 04 August 2010).
- Suchman, L. (1995). Making work visible. *Communications of the ACM*, 38(9): pp. 56-64.
- Suchman, L. (1994). Do Categories Have Politics? The language~action perspective reconsidered. *Journal of Computer Supported Cooperative Work*, 2(3): pp. 177-190.
- Suchman, L. (1993). Technologies of Accountability: Of Lizards and Aeroplanes. In Button, G. (ed.) *Technology in Working Order*. London: Routledge: pp. 113-126.
- Suchman, L. & Trigg, R. (1991). Understanding Practice: Video as a medium for reflection and design. In Greenbaum, J., Kyng, M. (eds.) *Design at Work: Cooperative Design of Computer Systems*. Lawrence Erlbaum Associates: pp. 65-89.
- Suchman, L. (1987): *Plans and situated actions: the problem of human-machine communication*. Cambridge University Press.

- Tang, J.C. and Leifer, L.J. (1991). An Observational Methodology for Studying Group Design Activity. *Research in Engineering Design*, 2: pp. 209-219.
- Tang, J. C., & Minneman, S. L. (1991): VideoDraw: A Video Interface for Collaborative Drawing. *ACM Transactions on Information Systems*, 9(2): pp. 170-184.
- Tang, J.C. and Leifer, L.J. (1988). A framework for understanding the workspace activity of design teams. In *Proc. of the 1988 ACM conference on Computer-supported cooperative work*: pp. 244-249.
- Tarte, S. (Forthcoming). Papyrological Investigations: Transferring Perception and Interpretation into the Digital World. *Literary and Linguistics Computing*.
- Tarte, S. M., Brady, J. M., Roued Olsen, H., Terras, M., Bowman, A. K. (2008) Image Acquisition and Analysis to Enhance the Legibility of Ancient Texts. In *UK e-Science Programme All Hands Meeting 2008 (AHM2008)*, September 2008, Edinburgh.
- Tatar, D. (1989). Using video-based observation to shape the design of new technology. *SIGCHI Bulletin*, 21(2): pp. 108-111.
- ten Have, P. (1999). *Doing conversation analysis: a practical guide*. London: Sage.
- ten Have, P. (1990). Methodological Issues in Conversation Analysis. University of Amsterdam, (available at: [<http://www2.fmg.uva.nl/emca/mica.htm>]; last accessed 28 June, 2010).
- Terras, M.M. (2006). *Image to interpretation: towards an intelligent system to aid historians in the reading of the Vindolanda texts*. Oxford UK: Oxford University Press.
- Travers, M. & Calvey, D. (eds.) (2001). *Ethnographic Studies: Issue No. 6, Special issue: Workplace Studies*, (available at: [<http://www.socialsciences.manchester.ac.uk/disciplines/sociology/about/events/ethnography/journal/issue6/>]; last accessed 28 June, 2010).
- Trigg, R., Blomberg, J., Suchman, L. (1999). Moving Document Collections Online: The Evolution of a Shared Repository. In *Procs of the 6th European Conference on Computer-Supported Cooperative Work, ECSCW '93*, Copenhagen, Denmark, Springer: pp. 331-350.
- Trigg, R., Bødker, S., Grønæk, K. (1991). Open-ended interaction in cooperative prototyping: a video-based analysis. *Scandinavian Journal of Information Systems*, 3: pp. 63-86.
- Tutt, D., Hindmarsh, J., Shaukat, M., Fraser, M. (2007). The Distributed Work of Local Action: Interaction amongst virtually collocated research teams. In *Procs of the 10th European Conference on Computer-Supported Cooperative Work, ECSCW '07*, Limerick, Ireland, London: Springer: pp.199-218.
- Twidale, M., Randall, D., Bentley, R. (1994). Situated evaluation for cooperative systems. In *Proc. of the Conference on Computer-Supported Cooperative Work*, Oct. 22–26, (Chapel Hill, NC): pp. 441–452.
- Twidale, M., Rodden, T., Sommerville, I. (1993). The designers' notepad: supporting and understanding cooperative design. In *Procs of the 3rd European Conference on Computer-Supported Cooperative Work, ECSCW '93*, Milan, Italy. Springer: pp. 93-108.

- Urquijo, S., Scrivener, S., Palmen, H. (1993). The Use of Breakdown Analysis in Synchronous CSCW System Design". In *Procs of the 3rd European Conference on Computer-Supported Cooperative Work, ECSCW '93*, Milan, Italy. Springer: pp. 281-294.
- Usability Task Force. (2004). Usability Research Challenges in e-Science. Technical paper, (available at: [<http://www.cs.nott.ac.uk/~tar/UTF.pdf>]; last accessed 28 June, 2010).
- Usability Task Force. (2004). Usability Surgeries with EPSRC pilot projects. *Summary paper*.
- van den Haak, M.J., de Jong, M.D.T., Schellens, P.J. (2006). Constructive interaction. An analysis of verbal interaction in a usability setting. *IEEE Transactions on Professional Communication*, 49(4): pp. 311-324.
- von Hippel, E. (1976). The dominant role of users in the scientific instrument innovation process. *Research Policy*, 5: pp. 212-239.
- Warr, A., de la Flor, G., Jirotko, M. & Lloyd, S. (2007). Usability in e-Science: The eDiaMoND Case Study. In *Increasing the impact of usability work in software development workshop, CHI 2007*. 28 April-03 May 2007. San Jose, CA.
- Warr, W. A. (1998). Communication and communities of chemists. *Journal of Chemical Information and Computer Sciences*, 38(6): pp. 966-975.
- Williams, M.S., & Blakeborough, A. (2001). Laboratory testing of structures under dynamic loads: an introductory review. 359: pp. 1651-1669.
- Winograd, T. & Flores, F. (1987). *Understanding computers and cognition*. Reading, Mass: Addison-Wesley.
- Weiser, M. (1994). The world is not a desktop. *Interactions*: pp. 7-8.
- Weiser, M. (1993). Some Computer Science Problems in Ubiquitous Computing. *Communications of the ACM*: pp. 75-83.
- Weiser, M. (1991). The Computer for the Twenty-First Century. *Scientific American*: pp. 94-104.
- Welsh, E., Jirotko, M., Gavagan, D. (2006). Post-genomic science: cross-disciplinary and large-scale collaborative research and its organizational and technological challenges for the scientific research process. *Philosophical Transactions of the Royal Society A*, 364: pp. 1533-1549.
- Wenger, E. (1998). *Communities of Practice*. Cambridge: Cambridge University Press.
- Wertsch, J. (1998). *Mind as Action*. Oxford: Oxford University Press.
- Weinberg, A. (1961). Impact of large-scale science on the United States, *Science*, 134: pp. 161-164.
- Wildman, D. (1995). Getting the most from paired-user testing. *Interactions*, 2(3): pp. 21-27.
- Wilson, P. (1991). *Computer supported cooperative work: an introduction*. Oxford: Intellect.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.

Woodruff, A., Szymanski, M., Grinter, R. and Aoki, P. (2002). Practical strategies for integrating a conversation analyst in an iterative design process. In *Proc. of DIS '02*. London, ACM: pp. 255-264.

Woodward, P. (1996). Perspectives on supercomputing: three decades of change. *Computer*, 29(10): pp. 99-111.

Yang, X., Allan, R., Finch, J. (2008). The NCeSS Portal - A Web 2.0 Enabled Collaborative Virtual Research Environment for Social Scientists. In: *Procs of the 2008 Fourth International Conference on Semantics, Knowledge and Grid*, IEEE: pp. 507-508.

Zaner, R. (1970). Editor's Introduction. In Schutz, A. *Reflections on the Problem of Relevance*. New Haven: Yale University Press.