An Ethnomethodologically-Informed Study of Interaction in the MOOC

Anna Zawilska

St. John's College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy in Computer Science

Acknowledgements

To my family and friends, and especially my parents and Duncan, without whom none of this would be possible.

Thanks to my supervisor, Professor Marina Jirotka, and Dr. Helena Webb who helped me bring my thesis to life. I am also grateful to the entire Human Centred Computing group for their feedback and guidance.

Abstract

The massive open online course (MOOC) is an online educational platform that has garnered significant popularity in recent years and is viewed by some as a transformative force in tertiary education. Originally designed to replicate a classroom setting, social interaction should be a fundamental part of using the MOOC. Most research into MOOCs has focused on using quantitative 'big data' to understand social interaction of large cohorts of students over entire courses. However, recently there has been a turn towards understanding qualitative aspects of social interaction in order to really take account of its features.

A substantial gap in existing research on MOOCs is understanding the fine-grained qualitative features of social interaction in terms of the sequential reportable-observable unfolding of activity. In this thesis, these features are explored in detail in order to unpack and enrich the current understanding of social interaction in MOOCs, and to consider how this might inform improved MOOC interface design that better supports it.

This thesis uses mainly video data of social interaction from quasi-naturalistic studies, that is then subject to ethnomethodologically-informed analysis. The studies involve small groups of participants, both co-present and distributed, collaboratively performing open-ended tasks around MOOC educational content. The analyses reveal recurring patterns in interaction that are largely glossed over by many existing 'big data'-based studies.

I pull out a number of conceptual findings that extend current understanding of interaction in MOOCs. Firstly, I show how an act like a 'click' may have a diverse and

complex role to play in the unfolding of social interaction, at both an individual and group level. Secondly, I explicate some of the key interactional resources that distributed participants may use to transcend boundaries in space and time. I show, for example, how visibility of the immediate context of an act appears to give it a sense of meaning. Thirdly, I discuss how dynamic and complex participation can be; revealing how students can emergently construct and transition between different and multi-faceted modes of participation. Drawing on these findings, I contribute substantively to the MOOC platform by surfacing a number of important issues to inform system requirements.

Methodologically, I reflect on my key learnings taking this approach in terms of advantages and challenges. I find it can be difficult to learn how to undertake an ethnomethodologically-informed study. However, as I show in the thesis, this approach brings to the fore some of the rich seen-but-unnoticed ways participants accomplish social interaction, which other automated methods are not designed to do.

Table of Contents

Chapter 7.

Phase 3 - Exploring Interactional Issues at the Root of Intra-Video Social Interaction for Distributed Groups

Chapter 8. Discussion

Chapter 1.

Introduction

The massive open online course (MOOC) is a video-based online educational platform that has garnered significant popularity in recent years and is seen by some to be a potentially transformative force in tertiary education. Originally intended to replicate a traditional offline classroom, social interaction is an important aspect of the MOOC (Pea and Lindgren 2008). Currently, our conceptual understanding of social interaction in the MOOC is largely limited to what can be gleaned through 'big data': most existing literature uses large-scale quantitative server log data collected remotely, to characterise social interaction by entire enrolled student groups, over the duration of a course. While this provides a view of some statistically significant large-scale features of social interaction, little is known about the detailed mechanics of *how* it unfolds in context, on the scale of individual students, over short periods of time. Understanding this is important because it contributes towards designing a social environment in the MOOC which supports social interaction as well as possible. Studies examining this kind of detail of the sequential unfolding of social interaction in other areas of Computer Science, such as Computer-Supported Collaborative Work, have a substantive and positive effect on computer system design and understanding (Heath et al. 2010). In this thesis, these fine-grained features of social interaction will be the focus of exploration in order to unpack and enrich the current conceptual understanding of social interaction in MOOCs.

Substantively, I will show how this relatively novel view on social interaction in the MOOC can have important implications for the design of the MOOC platform. Through

an analysis of the interface flow of the MOOC, I will show that the interaction between student and educational content is still largely an individual activity. The interaction with the video content on the platform is disconnected from social interaction with other students. My goal is to use a detailed examination of interaction to support intra-video social interaction and work towards creating a more collaborative social environment around the educational content. I will do this by developing a video annotation interface that allows a more seamless integration between viewing video content and interacting with others. The video interface is chosen because, as I will show, video educational content is by far the most popular feature of the platform and is the central delivery mechanism of educational materials to students (Belanger and Thornton 2013).

Methodologically, I will draw on a qualitative Sociological approach called 'ethnomethodology'. A key challenge to applying ethnomethodology to the study of massive online systems, is how to capture detailed data of the sequential unfolding of social interaction when all system users are distributed. As I will show, one way to mitigate this challenge is by considering the ways in which studies of social interaction of co-present groups may be used to inform issues around social interaction of distributed groups like those on MOOCs. Over the course of the thesis, I will incorporate an ethnomethodologically-informed approach into a three-phase iterative computer system design process to develop an increasingly fuller conceptual understanding of social interaction, and produce a proof-of-concept video annotation interface.

These conceptual, methodological, and substantive goals will be synthesised to create the central research questions of this thesis, presented in the next section, followed by an outline of the thesis.

1.1. Research Questions

This thesis is driven by the following set of research questions: a primary overarching research question, followed by three sub-questions each referring to the conceptual, substantive, and methodological goals of this thesis.

Primary Research Question:

What are some fine-grained features of social interaction in MOOCs, and how could an investigation into these features be incorporated into the development of the MOOC interface? Given the prominence of video in the MOOC, these questions are explored through the development of a video annotation interface for MOOCs.

Sub-questions:

1. *Conceptually, how may a view of the detailed aspects of social interaction in the MOOC contribute towards a respecification of existing concepts around interaction in MOOCs?*

2. *Substantively, how may the understanding from question 1 be used to inform the requirements for a video annotation interface?*

3. *Methodologically, what are the advantages and challenges of taking an ethnomethodologically-informed approach to the development of an interface to support distributed users?*

1.2. Outline of Thesis

The outline of the chapters of this thesis is as follows:

Chapter 2.

Literature Review: Interaction in the MOOC - Controversies, Limited Visibility of Fine-Grained Features of Interaction, and Interface Flow Issues

In Chapter 2, I will review key issues from existing literature on interaction in MOOCs. I introduce some key platform features, and present an important current design limitation of the MOOC, which forms the basis for the substantive goals in this thesis. Through an analysis of the interface flow, I will show how disjointed viewing video content on the MOOC is, from social interaction on the platform. This is a limitation, in light of literature on interaction of groups on educational platforms that suggests the value of seamlessly integrating viewing video content with social interaction (Pea and Lindgren 2008). Conceptually, I will present the landscape of existing literature on interaction in MOOCs, showing that most work uses 'big data' collected from the educational platform remotely. These data are usually in the form of quantitative statistics based on server logs, such as button clicks and time spent viewing a page by a student. Alternatively, the data is commonly in the form of large quantities of free text discussion forum posts. These data present decontextualised artefacts of interaction and the fine-grained qualities of *how* social interaction unfolds are minimally attended to. Understanding these qualities is important because it contributes towards designing a social environment in the MOOC which supports interaction as well as possible. In this thesis, these features will be explored in order to unpack and enrich the current

understanding of interaction in MOOCs, and to consider how this might inform the design of a video annotation interface in the MOOC that better supports social interaction.

Chapter 3.

Methodology: An Ethnomethodologically-Informed Approach to the Iterative Development of a Video Annotation Interface for MOOCs

Picking up on the discussion for the need for further study of the fine-grained aspects of interaction in the MOOC from Chapter 2, I will review literature that suggests methodological approaches to this. In recent years, there has been substantive methodological innovation in this area although it remains challenging. I will justify using an ethnomethodologically-informed approach, considering advantages and limitations of it. I will suggest taking a three-stage iterative system design approach, with each iteration adding new insight into the details of how social interaction may take place through the MOOC, leading to new conceptual insights and system design requirements. The approach uses quasi-naturalistic studies with both co-present and distributed groups performing open-ended tasks. Participants' activities are video recorded and analysed for patterns of activity. Each iteration will be presented in subsequent chapters.

Chapter 4.

Method:

A Three Stage Iterative Study of Social Interaction, and Design of Proof-of-Concept Video Annotation Interface

This chapter will describe the key facets of the three stage iterative approach that this thesis will use to develop a video annotation interface for the MOOC. In addition to describing the overarching issues around recruiting participants, ethical approval, and the practicalities of setting up and running studies, this chapter will walk the reader through each phase, giving snippets of data and signposting the contents of the following chapters.

Chapter 5.

Phase 1:

Broad and Narrow Context of Social Interactions in MOOCs

Using a qualitative and ethnomethodologically-informed approach, I will take the first steps towards extending our understanding of interaction around MOOCs. Using a survey and quasi-naturalistic experimental studies, I will show the ways in which students interact with MOOC video lectures, which are not captured in typical 'big data' approaches. Rather, students may actively engage in evaluating video content with respect to their existing knowledge and experiences. During social interaction around educational video content, the data surface the seen-but-unnoticed ways participants use gesture and speech to collaborate, including when transitioning

between video-viewing and discussion. This transition is not constituted by any one act, including pausing the video, but is accomplished by a complex sequence of interplaying gestures and speech by both participants. Over the course of their collaboration, the participants implicitly and seamlessly transition into and out of asymmetrical roles. At times, participants are able to unproblematically focus on different artefacts in the environment, while maintaining an ongoing conversation about educational content. Surfacing these features of social interaction and the competencies that the participants bring to the collaboration, will extend the concepts of 'video viewing' and 'collaboration' around MOOC educational video. The ethnomethodologically-informed approach adds depth and new complexity to how social interaction is actually accomplished in a quasi-naturalistic setting. By considering those patterns of interaction that appear most central to the accomplishment of collaboration, a first set of requirements for supporting intra-video social interaction online will be created.

Chapter 6.

Phase 2:

The Individual and the Collaborative - Intra-Video Social Interaction of a Co-Present Group

Going deeper into the way that groups interact around MOOC educational video, in this chapter I will study the fine-grained aspects of how larger co-present groups of participants collaboratively accomplish the annotation of a video. The participants use a low-fidelity prototype informed by the results from Chapter 5. The analysis will

illustrate how even seemingly individual activities by participants have a collaborative nature. For example, the pressing of a button to play or pause the video, by an individual, alone does not accomplish the transition by the group to or from video viewing. Rather, the interplay between all the participants' gestures and speech leading up to, during, and after the act, is the foundation for how the group accomplishes the transition. Furthermore, participants seamlessly move into and out of peripheral roles, accomplished through a delicate and complex sequence of gestures and speech. The analysis will serve to extend our understanding of the mechanics of how social interaction in a quasi-naturalistic setting is accomplished. This understanding will inform a set of requirements for a prototype to support these kinds of mechanics of social interaction in a distributed group.

Chapter 7.

Phase 3:

Exploring Interactional Issues at the Root of Intra-Video Social Interaction for Distributed Groups

Using a prototype of an online system informed by previous chapters, I study the social interaction of a group of distributed participants, in this chapter. The study involves collecting data remotely from the platform as well as video-recording the local activity of each participant. All video collected is subjected to ethnomethodologically-informed analysis. Collecting both types of data allows a contrast to be created between the kind of insight about social interaction that can be gleaned using online data (or 'big data' only) and an ethnomethodologically-informed approach. The study also involves using

both the existing MOOC platform as well as the prototype, creating a view of how the two systems compare. Among the more important insights, the analysis will reveal new aspects of the role of the 'lurker': a student who views content in the MOOC but doesn't actively participate in social interaction. I will find patterns in the sequential unfolding of interaction that provides important opportunities for interventions to shift students out of 'lurking' into more active modes of participation. Another important insight surfaced from analysis concerns communication through forums. By studying the way a misunderstanding between participants persists, I will show how the lack of availability of peripheral awareness of the gestures and speech of others in the distributed group provide conditions for potentially ineffective communication. The analysis extends our conceptual analysis of aspects of social interaction in the MOOC like stepwise development of contributions, details of interactional shifts between viewing video and discussion forums, and root causes around breakdown in communication. Methodologically, this study requires a complex setup and offers some important learnings around incorporating an ethnomethodologically-informed approach into the study of the MOOC. The new insights will be used to derive requirements for future prototypes of the MOOC.

Chapter 8. Discussion

This final chapter of the thesis will synthesise the themes and issues raised through data analysis from the previous chapters to draw out the main conceptual, methodological, and substantive contributions of this thesis. Conceptually, this thesis will reconceptualise the role of 'big data' events like clicks, how distributed participants transcend boundaries in space and time, and how 'participation' is much more

complex than most current studies suggest. The thesis will speak to the complexities around these aspects that are revealed through an analysis of the fine-grained details of the social interaction into which they are embedded. These complexities have tended to be glossed over in traditional literature using the 'big data' approach. Methodologically, I will discuss the lessons learned through undertaking the various studies in this thesis. A key advantage of incorporating an ethnomethodologically-informed approach is the ability to surface key seen-but-unnoticed aspects of how participants accomplish social interaction around educational video. By their nature, these aspects may be difficult to surface without an ethnomethodologically-informed approach. A difficulty of this approach is the common logistical challenges around arranging participants and arrangement the technical setup to execute the quasi-naturalistic studies. Using the understanding of social interaction gleaned in previous chapters, I will briefly touch on what they suggest about the MOOC as a learning device. Finally, I will present a set of qualitative requirements for the MOOC which can contribute to creating a social environment on the MOOC that may better support interaction.

Chapter 2.

Literature Review:

Interaction in the MOOC - Controversies, Limited Visibility of Fine-Grained Features of Interaction, and Interface Flow Issues

2.1. Introduction

This chapter will introduce the MOOC platform 'Coursera', and explore interaction in the MOOC from both a conceptual and design point of view. The MOOC, while proving extremely popular with high enrolment rates, has been the subject of substantial debate and controversy, particularly around student participation and low completion rates. Existing data into interaction on the platform (both between the student and the platform, as well as social interaction) suggest participation by students is different in nature to traditional offline classroom settings, and further research is required to understand interaction in the MOOC (Chengjie 2015).

This chapter will explore interaction in the MOOC from primarily two perspectives. Firstly, existing literature will be reviewed to illustrate the main ways that interaction in the MOOC has been studied and conceptualised. I will show that our lens on interaction is, so far, mainly based on inferences made from the mainly automated analysis of large quantities of quantitative data collected remotely. These data are primarily from server logs and reflects artefacts of interaction such as website button clicks and web page views. As a result, our understanding of the details of how interaction unfolds is minimal. For example, we don't understand well the resources a

participant draws on to accomplish activities and the reasons *why* interaction unfolds the way it does. Secondly, I will undertake an analysis of the current system interface flow, to reveal a concrete platform design opportunity. The platform interface flow tends to dis-aggregate interaction by a student with educational content, from social interaction with other students. Existing literature suggests it would be better to integrate the two types of interaction since contributions by a community to a piece of content tend to counter any bias in that content and increase the quality of it (Pea and Lindgren 2008). In this chapter, I will suggest creating a new video annotation interface for the MOOC to support this integration. I will put forward the view that extending our understanding of interaction to include the fine-grained features of *how* it unfolds, should be made part of developing requirements for a new video annotation interface for the MOOC.

2.2. The Rise of Coursera MOOCs: Controversies around Participation

While formal educational systems are expected to be slow to change (Kolowich 2013), new opportunities for informal learning online are appearing and attracting considerable popularity. One such opportunity is the proliferation of the online platform called the 'massive open online course' (MOOC). The acronym 'MOOC' was first used in 2008 by Stephen Downes, George Siemens and Dave Corimer to describe their online course 'Connectivism' which had around 2000 registered students. Only a small portion of these students paid a fee and had a more directed course to follow, while most students signed up for free and could explore various educational resources provided online. The only pre-requisite for registration was that each student have a

valid email address through which to receive a daily course email. This email contained links to online content that was intended to be a seed for discussion and collaboration amongst students.

Even though the term existed, 'MOOC' remained largely excluded from the public vocabulary until 2012 when the online platform *Coursera*[1] launched. The founders, Stanford Computer Science Professors Andrew Ng and Daphne Koller, were driven by the idea of creating a social intervention in response to the rising cost of tertiary education and social disparities associated with higher education both in the US and internationally (Gannes 2012). Coursera provides the framework of functionality for online courses, with the course content supplied by academics at prominent universities. The Coursera MOOCs consist of video lectures, a discussion forum, and online assessments. The platform was developed without a clear set of requirements, in the spirit of experimentation, largely intended initially to reproduce some interactional aspects of a traditional classroom (Don 2013). Anyone who has an Internet connection with which to access the platform may search and enroll in as many courses as they wish. There are no requirements on the student to participate in the course in any particular way to continue to be enrolled, although students are required to watch all lectures and pass assessments if they wish to get a certificate of accomplishment. Since the launch of Coursera and as of August 2016, over 140 top universities and educational organisations have created Coursera courses.

One of the strengths of Coursera MOOCs is their popularity and ability to reach students worldwide. MOOCs have an average enrolment rate in the order of 10,000s of mostly adults (Jordan 2013). Demographically, available data suggest that most

[1] http://www.coursera.org

students come from first-world countries (Jordan 2013) and that a substantive proportion of enrolled students in MOOCs is educated to at least a tertiary level. To illustrate typical proportions, over 70% of those that took the MOOC 'Software as a Service' were educated to at least tertiary level, where over 30% of those that took 'Circuits and Electronics' reported having graduate degrees (Pappano 2012).

MOOCs are the subject of debate in the mainstream news, particularly around the role they could play in revolutionising the traditional educational system. Mainstream news has portrayed MOOCs in contradictory ways across a number of key issues, as summarised in Table 2.1 below (Selwyn and Bulfin 2014).

Table 2.1. Summary of conflicting assertions in the coverage of MOOCs in news media (Selwyn and Bulfin 2014)

Issue	Dominant assertion	Counter-assertion
Large class sizes	Students are motivated and excited by the crowded stadium-like space.	Students are disconnected from the intimacies of learning.
Higher education 'status quo'	MOOCs circumvent traditional higher education systems.	The legitimacy of the MOOC as an educational form appears to derive primarily from their association with high status, elite universities.
The role of teachers	Courses are driven by renowned professors enjoying the cachet of celebrity and status.	MOOCs are promoting the mass of non-elite educators and institutions worry about their futures.
Pedagogy & learning	Pedagogy and learning is heralded as innovative and '2.0'.	Pedagogy and learning is derided as replicating the passive instructionism of twentieth century higher education.
Revenue	MOOCs present a new and uncertain phase of the fee-paying structure of university tuition.	MOOCs are bolstering the stock of established university 'brands' and their revenue streams.

Alternate mode of provision	MOOCs offer an alternate means of studying at a university-level of education.	MOOCs are reinforcing the established *status quo* in higher education – offering an alternative 'way in' to later study for 'proper' courses at 'proper', 'face to face' universities.
Technology development	MOOCs have developed as part of a natural evolution of technology.	MOOCs have developed as part of a deliberate process of scientific innovation and Silicon Valley entrepreneurship.

The themes of course participation and completion are particularly contentious. MOOCs usually have very low completion rates of under 5%, even though enrolment numbers are high. This is sometimes considered a sign that MOOCs are not successful or worthwhile (Konnikova 2014). While several hundreds to several thousands of students enrol, numbers of engaged students tend to drop substantially with each new learning activity in the course. Balch quotes the percentage of students who watched the first video of Coursera's 'Computational Investing, Part I' as 50% of the total enrolled (Balch 2013). Online blogger Sue Gee highlighted the dropout rate through her post 'MITx – The Fallout Rate' (Gee 2012). With 150,000 sign-ups, only a 5% pass rate was achieved. The numbers drop significantly with each new activity: of the 150,000 sign-ups, 69,221 people looked at the first problem set, 26,349 earned at least one point in the problem set, 13,569 looked at the midterm, 10,547 people got at least one point on the midterm and 9,318 passed it. 10,262 people looked at the final exam, 8,240 people got at least one point on the exam and ultimately 5,800 passed the exam. As another example, Levine points out that for the 'Social Network Analysis' class he took on Coursera, 61,285 people registered but only 1,303 (2%) earned the certificate (Levine 2012).

At the same time, there exists substantive debate over whether completion rate is an adequate indicator of the success of a MOOC (e.g. (ECampus 2013; Kolowich 2013; Lucena 2013; Spaniel 2013)). Firstly, the completion rates are not surprising when compared to other indicators of Internet engagement; for example, Google advertisements have a click rate of about 5% (Gabbert 2010). Secondly, Pappano (2012) puts forward the view that the concept of 'dropping out' for a MOOC is necessarily different to the concept in traditional, formal educational systems. The open nature of the MOOC means students make various levels of commitment to the course and sometimes are only interested in portions of the course content. This view is supported by studies that suggest that even though students participate in the MOOC in various ways over time, they report a similar quality of experience. For example, in one case it was found that students who do not participate in the MOOC assessment procedures but who view the lecture videos, report a similar level of satisfaction with the course experience to others who take all assessments (Kizilcec et al. 2013). This was measured by administering post-course surveys to students across several MOOCs in which students rated their "overall experience with the course" on a 7-point Likert scale from 'Poor' to 'Excellent'.

A main reason why so much controversy and debate exist, is the lack of foundational understanding of social interaction in the MOOC. While it's clear that social interaction in the MOOC is different from more traditional educational settings like a classroom, not enough is known yet to be able to fully unpack and resolve the debates around participation and course completion. As I will discuss in the next section, social interaction is important to consider when thinking about how the MOOC may or may not foster learning.

2.3. The Role of Social Interaction in MOOCs as a Learning Device

As the MOOC rises in popularity, the pedagogy behind the social interaction fostered on the platform, has increasingly become the subject of research (e.g. Bali 2014; Ross et al. 2014). The research questions in this thesis are focused on understanding reportable-observable aspects of social interaction, rather than trying to infer what 'learning' means in this context. A fuller discussion on pedagogy is outside of the scope of this thesis. However, a brief discussion on some issues from pedagogy is useful to understand why social interaction is important to the MOOC as a learning device. In this section, some of these are touched upon to give this context, before moving on in the next section to understanding how social interaction has mainly been studied in the MOOC so far.

Among the key learning theories that exist, there are at least the following two broad views on what constitutes learning: behaviourist learning theory (Watson 1930) and social constructivism (Kim 2001). Behaviourist learning theory is based on the idea that knowledge is acquired through controlled stimulus/response conditioning (Watson 1930). The role of the instructor here is to provide factual knowledge and to then observe, measure, and modify behavioural changes of the student. This is associated with rote memorisation of facts, assertions, rules, laws, and terminology. The role of behaviourism is generally understood to be at the lowest order of learning: that of factual knowledge and basic standard skills development. Behaviourism is based on the idea that concepts of 'learning' should not be judged or theorised using subjective data, but based on measurable behaviours so as to make them more 'scientific'.

A view of learning from a behaviourist point of view may have particular implications for interface design and the metrics used to measure 'success' of a MOOC, such as: measurement of 'correct' participation might involve administering standard tests with distinct correct and incorrect answers, along with the monitoring and assessment of response time. Additionally, the MOOC may use incentives and rewards to condition the learner to behave in a standard, pre-determined way with respect to problems with established singular solutions. The use of incentives and rewards for provoking particular standard tasks and behaviours has received criticism from within the educational community (Resnick 2012). The primary argument against it focuses on the ways behaviourist learning is limited in its ability to teach students more complex problem-solving skills beyond fact memorisation. Additionally, using incentives and rewards to encourage fixed behaviours may detract from encouraging genuine interest and exploration of a subject matter by the student.

In the alternative approach, informed by social constructivism, learning is inherently a social process. It doesn't occur within an individual through use of external forces to shape behaviours (McMahon 1997). The approach espouses the view that learning occurs meaningfully when students are engaged in social activities. Social constructivism may include a more adaptive view of knowledge which focuses on active discovery, reflection, and use of personal experience. It favours hands-on, project-based, and group-based learning rather than only rote learning of factual information. Some of the other key characteristics may be: provision of multiple representations of reality, provision of complexity of the real world, a reflection on experiences, context-focus, and collaborative knowledge construction rather than competition (Duffy & Jonassen 1992). Learning which incorporates these

characteristics may also sometimes be referred to as ‘deep learning’ as opposed to the ‘surface learning’ focus of behaviourism (Lublin 2003).

In the case that MOOCs be driven by some principles of social constructivism, the platform may not just be a medium for broadcasting content along with set student assessments and tests, rather it may become a medium focused on creating a social environment; supporting social interaction and conversation around the educational content. Social constructivism may be associated with higher order learning of more mature learners such as adult ones, although the ideas may be successfully implemented at earlier stages of learning too (Brend and Naylor 1996).

In general, there is a substantial group of pedagogical researchers who align with the view that social interaction is critical to learning, drawing on principles from social constructivism (Reigeluth 2013; Siemens 2014). This thesis aligns with this view, and aims to contribute towards future work in pedagogy by offering a lens into the details of how social interaction unfolds in practice in the MOOC. As I will discuss in more details later in this thesis, an approach informed by ethnomethodology seems to be particularly well suited to this, and helps us expand what we understand about social interaction on the platform. In the next section, I discuss how, for the moment, most of what we understand about social interaction in the MOOC are broad patterns in interaction based on large quantities of primarily quantitative server log data.

2.4. Social Interaction in the MOOC: Focus on Server Log Data, and Lack of Visibility of the Detailed Mechanics of Interaction

In this section, I will show how existing research into interaction in the MOOC has given concepts around interaction (both between the student and the platform, as well as

social interaction) a largely quantitative treatment. Studies have focused on the use of large amounts of primarily quantitative data collected remotely online. These data are typically based on server logs that reflect the artefacts of interaction such as, clicks on a website button (like the play/pause button on a video), time viewing a page or educational video, or the frequency of contribution to a discussion forum. There are three main genres of research into interaction in MOOC, which I will now discuss in turn.

One primary genre of literature uses statistical analysis over the server log data to formulate and define categories of MOOC students. For example, Kizilcec et al. (2013) use statistics gathered remotely about MOOC students (such as, the number of videos at least partially viewed by students, the number of discussion forums viewed by students, and the number of contributions to discussion forums made by students) to create categories of participation into which various cohorts of students fit – such as 'Active' and 'Auditing' student categories. The student category of the 'lurker' has been a particular focus across all literature. 'Lurkers' are students that mainly view video content and discussion forums, without actively participating in social interaction. The main driver of lurker-focused research has been to find ways to encourage them to take on more actives roles of participation, or to prevent active students from becoming lurkers (Sinha et al. 2014; Khalil and Ebner 2013). Often, the direction of this genre of research appears to be towards building increasingly granular categories of students based on quantitative server log data.

A second genre of research focuses on identifying large-scale, statistically-significant patterns in interaction with educational video or discussion forums, which is used to create suggestions for video creation or discussion forum design. Guo et al. (2014)

examine aspects of interaction between students and video to understand the features in a video that increase the chances of a student completing viewing a video and performing a post-video assessment. From this, they are able to generate some recommendations for video creators for MOOCs. For example, they find that statistically, shorter and more informal videos with the speaker talking directly at the camera, perform very well. Kim et al. (2014) analyse the moments at which students are highly likely to stop watching a video, or rewind a video to rewatch a fragment. They cross-reference these moments with the contents of the video to create conclusions around the kind of video content that's likely to engage students. For example, non-visual explanations in videos tend to cause students to rewind the videos, suggesting potentially interest and/or problems in understanding. The paper provides a set of implications for video creators to increase engagement by students, such as creating shorter videos, and providing interactive links in videos to help students skip to particular segments of a video. Other studies consider patterns in interaction in discussion forums. Several studies highlight that the overwhelming majority of students across MOOCs don't appear to participate actively in any substantive way with forums, that interaction among students in the forums is inconsistent and bursty, and that forum posts are mostly by small minorities of students who pass the course with either very high or very low course marks (Eynon et al. 2014).

The third main genre of research focuses on using the large amount of server log data available to do statistical work to create metrics that capture the attributes of the MOOC, and to train Machine Learning systems. One area of research focuses on creating predictors of student behaviour using a variety of Machine Learning approaches. For example, Sinha et al. (2014) use an analysis of large-scale data around

video lecture viewing and discussion forum behaviours (e.g. video playing/pausing, forum posts, and forum comments) to create a predictor for student attrition that achieves a ~65% accuracy rate. Weller (2014) defines and explains various more and less active forms of participation, in terms of correlation to other course- and student-related properties such as, course length, number of enrolled students, and the ranking or prestige of the university offering the course. Other studies use discussion forum text as the basis for analysis, using automatic Natural Language Processing approaches to surface themes and issues that then form the basis for predictors to student behaviour. For example, Wen et al. (2014) apply an automated Sentiment Analysis algorithm to discussion forum contents, to find a correlation between the level of sentiment expressed daily in a forum and the level of student attrition in the MOOC on that day. Other studies focus on developing rigorous quantitative metrics of a MOOC that can be used to train Machine Learning networks. Yang et al. (2014) focus on developing statistical analysis methods for forums whose results encapsulate features around how sustained discussion in a forum is. This could be taken forward to be used to train other predictors on interaction in the MOOC.

The strength of most existing studies on interaction in MOOCs is they exploit the large amounts of data that can be collected remotely and processed automatically by machines. This kind of data collection, management, and processing would be impossible to do at the same scale through human processing. The insights gleaned are also then statistically-valid because they are powered by a large and well-structured set of data.

However, there are several challenges associated with a study of the MOOC based on large-scale server log data collected remotely and analysed, including:

1. The large amount of data collected, places limitations on the **kinds of analysis techniques** that could be applied to them. Importantly, it usually means that most manual techniques of analysis, particularly those applied in traditional Social Science research (such as varieties of Thematic Analysis (Boyatzis 1998)), cannot feasibly be used, as it would take an unreasonable amount of human time and other resources.

2. The **type of data** that can be collected remotely is limited in nature, as it's usually collected from servers. The data is commonly in the form of numerical server log data or text.

3. The **understanding we can gain about the 'social'** are limited, because only the artefacts of interaction are available through server log data (such as button clicks or page views). The data do not give direct access to the sequence of interaction which gave rise to these artefacts, and therefore tends to gloss over the detail associated with *how* and *why* those artefacts were created in the first place.

By exploiting large amounts of server log data to understanding the broad patterns in interaction, most existing studies provide a limited view of the details of *how* interaction unfolds moment-by-moment. We don't yet have a substantive understanding of the underlying structures through which a students' interaction is organised, and its possibly associated complexity and delicacy. This requires an analysis of the fine-grained features of interaction. It should include, for example, a view of a students' emergent negotiation of a setting and the inter-relationship between unfolding interaction and its context. This cannot be accomplished through analysis of

large amounts of server log data, as it requires data that captures a view of interaction at a higher level of granularity over narrow periods of time.

The value of looking at the details of interaction is supported by a large corpus of research in Sociology that has shown the important understanding of the mechanics of interaction that can be gleaned this way. Often, this work is performed in reaction to existing research that provides an over-simplistic, glossed-over view of interaction. For example, in the area of social interaction, Goffman is central in providing an empirically-grounded and rich understanding of how participants in a social encounter collaborate which he formalised into the concept of a *Participation Framework* (Goffman 1981). The development of Participation Frameworks is founded on a critique of the model of a simple dyadic relationship of speaker to hearer in conversation. In this simple model, only two individuals are involved and the role of speaker and hearer are interchanged; passed back and forth between the individuals.

Goffman argues that this model is insufficient for the analysis of social encounters[2] involving both pairs and larger groups of individuals. For example, he decomposes the notion of the 'hearer' and 'speaker'. Goffman suggests the role of 'speaker' may be decomposed into three further categories: (1) the 'animator' who produces the utterance, (2) the 'author' who selects "the sentiments that are being expressed and the words in which they are encoded" (Goffman 1981), and (3) the 'principal' who is committed to the words being spoken. Goffman discusses how these three roles may be taken by one person simultaneously, and alternatively how the roles may be split

[2] Goffman describes an encounter as involving, for participants "a single visual and cognitive focus of attention; a mutual and preferential openness to verbal communication; a heightened mutual relevance of acts; an eye-to-eye ecological huddle that maximises each participant's opportunity to perceive the other participants' monitoring of him" (Goffman 1972 p.19).

among people such as when an individual recites words written by someone else. Additionally, members in a social encounter can switch between these roles. Goffman reveals how participants display and manage their different participant status to one another using gestures, gaze, and talk (Goodwin 1981; Levinson 1988). This has had a substantive impact on how social encounters are viewed and how online computer systems supporting social interaction may be designed. For example, in their seminal work, Jirotka et al. (1991) surface new key roles in computer-mediated communication drawing on Participation Frameworks. They show new important details of the ways that interactions start and topic changes occur within an interaction.

Inspired by Participation Frameworks, it may be worthwhile to take a look at how students in a MOOC orient towards their participation in social interaction, and the way this may evolve dynamically over the course of the interaction. Even if it is not possible to develop and verify a formal Participation Framework. I will show it's possible to make progress by exploring particular facets around how MOOC students accomplish rich and dynamic participation in social interaction.

Other work has also surfaced the fine-grained features of how participants in a social interaction accomplish seen-but-unnoticed activities. For example, research suggests participants almost automatically undertake detection and resolution of misunderstandings throughout conversation, to seemingly ensure all parties have the same shared grasp of the talk and conduct in the interaction (Schegloff 1992). Using recordings of conversation, Schegloff is able to illustrate this is accomplished practically through the sequential basis of talk. "Speakers reveal aspects of their understanding of the prior talk to which their own speech is addressed" (Sacks 1989). Participants appear to use reportable-observable features of the talk of others, including the

detailed features of speech such as shifting changes in intonation or pauses. Like Goffman's work, Schegloff's research draws heavily on practical work looking at real-world naturalistic interaction, giving an empirically-grounded explication of the reportable-observable mechanisms that allow social interaction to proceed in a productive manner. Work in areas like Computer-Supported Collaborative Work (e.g. (Heath et al. 1993; Jirotka and Wallen 2000)) and e-learning (e.g. (Kreitmayer et al. 2012)) illustrate the value of taking this kind of detailed view of interaction. In so doing, they surface the potential dangers of taking an over simplistic approach to computer system design that neglects to acknowledge and examine the kinds of rich forms of interaction that can occur around and through computer systems.

The focus on large quantities of server log data as the basis for researching interaction in the MOOC, echoes a larger trend in the study of massive online systems more broadly. While server log data-based studies remain a key area in the study of massive online systems, there is a growing corpus of literature that recognises the opportunities to surface additional important issues through the use of other kinds of data. In the next section, I will describe some of the relevant insights more broadly from other massive online platforms. These serve as examples of the wider recognition in Computer Science research of the limitations of using only large-scale online data, and show the potential of other forms of data in studying interaction online.

2.5. Social Interaction in Massive Online Platforms: The broader trend of a turn towards qualitative approaches

In this section, I will take some time to discuss a broader trend in the analysis and development of massive online systems: the shift towards more qualitative

approaches. This shows that the concern with the limitations of large amounts of quantitative data and analysis doesn't exist only around MOOCs, but is a topic of discussion and research more broadly.

Often, in studies of massive online systems through large amounts of server log data, the term 'big data' is used. However, there is no single definition of 'big data' and literature shows diverse and sometimes contradictory definitions of the term (Ward and Baker 2013). Among the most cited definition is that by the industry analysts Gartner who propose a definition encompassing "three V's": Volume, Velocity, Variety. These concepts are intended to capture the large size and rate at which data is generated, as well as the wide variety of formats and sources of data that can be synthesised (Douglas 2001; NIST Big Data 2015; Beyer and Laney 2012). The definition does not quantify the volume, velocity, or variety of data that makes it 'big data'. In contrast, Intel (2012) provides concrete numbers in their definition of 'big data', linking it to organisations "generating a median of 200 terabytes (TB) of data weekly". Microsoft's definition picks up on the requirement that 'big data' requires the application of significant computing power to perform analysis that would be impossible for a human to do. The definition creates a relationship between 'big data', 'machine learning', and 'artificial intelligence': "Big data is the term increasingly used to describe the process of applying serious computing power - the latest in machine learning and artificial intelligence - to seriously massive and often highly complex sets of information" (Microsoft 2013). Therefore, the strength of 'big data' is in surfacing new kinds of insights that would not be possible due to the substantially larger quantities than might be reasonably dealt with using manual human computation. 'Big data' tends to refer to quantitative data in a form that Machine Learning algorithms could be applied to surface insights (Lazer et al. 2014; Riff et al. 2014).

There is a growing set of literature that makes reference to the need for new approaches to studying big data. In a business context, Schroeck et al. (2012) use a survey to show the key role that analytics is going to play in making best commercial use of big data, but that innovation is still required in this area. Academic work examining the proliferation of communication on social media also unpacks some of the opportunities and challenges that lie in making good use of the large amounts of social media data that are available and the effects they have on society. Procter et al. (2013a; 2013b) study communication on Twitter, and use a combined approach to process the large amount of textual data that can be collected from the social media platform. Although natural language processing is a very active area of research, many (including Procter et al.) consider the technique not yet sufficiently reliable to replace human computation (Black et al. 2012). However, it would not be feasible to apply human computation to the entire data corpus due to its size. Therefore, Procter et al. develop a two-step methodology. In the first step, they make use of a computational tool that automatically processes large amounts of data identifying underlying structures in the corpus and a small set of potentially significant fragments. In the second step, the identified fragments are analysed using traditional manual qualitative methods. Zubiaga et al. (2016) use crowdsourcing to perform the annotation of tweets to understand the life cycle of social media rumours. Recent work has also included efforts to evolve Machine Learning approaches to the particular domain of analysing Tweets, for example Wang et al. (2016) consider emotion classification through Machine Learning, while Zubiaga et al. (2016) explore developing eight features of Tweets that may be predictive of the geographic origin of a Tweet. Webb et al. (2015; 2016a; 2016b) consider the ways malicious or false information can spread through social media, focusing on Twitter. They combine qualitative and quantitative

approaches to studying Tweets, drawing on the work of Procter et al. described above. Their qualitative approach sheds lights on issues around governance of social media, without a large-scale analysis focused on 'big data'.

In recent years, there has also been a growing network of researchers who have moved out of 'big data'-based approaches to studying online systems, towards so called 'micro-analytic' approaches. One such primary network is called MOOD ('Microanalysis of Online Data')[3]. The focus is on performing empirically-grounded studies into the details of *how* participants accomplish various activities online, and the root causes behind *why* interaction online is the way it is. This kind of 'micro-analytic' research takes this point of view towards the design of online systems: supporting interaction through design requires first delineating the way interaction is organised and the mechanics of how interaction is accomplished by the users involved. The studies by members of this network have shown new critical, detailed features of interactions (including for example, features of language) that shape the way interaction unfolds online. For example, Rintel et al. (2016) reveal interactional issues at play in the video communication application Skype Qik. They surface a flaw in the design of the technology that is causing users to drop out, rooted in issues around obligations in turn-taking during conversation. They find texting is more popular than Qik because it is a more lightweight technology in terms of turn-taking, making it easier to use. Meredith and Stokoe (2014) use a comparison of the details of interaction in conversation through Facebook Messenger with face-to-face interaction, to understand how users accomplish conversation through Facebook Messenger. They are able to identify similar mechanisms for how participants recognise and repair seeming

[3] http://moodnetwork.ruhosting.nl/

misunderstandings in communication, which are fundamental to this form of online communication through social media. The work of the MOOD network is not new, but may be rooted back in the seminal work of researchers such as Heath and Luff (1992) who have long advocated for understanding and delineation of the ways interaction in organised, as the basis for computer system design.

The types of interactions that are possible in the MOOC cannot be considered without also considering the ways that the current platform design supports certain kinds of interaction and limits others. Therefore, alongside the conceptual concern to make sure we understand interaction as much as possible, it's important to understand the impact that the existing interface design and flow has in shaping interaction. Next, I will present an analysis of the interface flow of Coursera, and present the case for the need to design a video annotation interface.

2.6. The MOOC's Interface Flow: The lack of integration between video viewing and social interaction

In this section, I will review the main types of interaction that the current Coursera MOOC platform supports, and suggest the value in developing a video annotation interface. I will undertake an analysis of the interface flow: a map of the ways the interface supports interaction with video, discussion forums, and assessments. An analysis of the interface flow is central to understanding the interaction that takes place in the MOOC, and to identify opportunities to improve the design of the interface. The design opportunity I will show through the interface flow analysis, is the dis-aggregation between viewing video by a student, and social interaction with others. I will suggest the value in developing a video annotation interface that allows a more

seamless integration between viewing the video content and interacting socially with others, including viewing others' contributions, thus supporting intra-video social interaction.

An understanding of the interface flow was developed by me over many hours of use of the platform and as an enrolled student of many MOOCs myself. From my experience, I created a taxonomy that maps out the kinds of interactions afforded by the MOOC interface, and serves as a tool to surface some interactional shortcomings and problems of the MOOC interface.

A taxonomy of the primary interactions supported by Coursera MOOCs is shown in Figure 2.1. Each table in the taxonomy represents an interface element. The top entry in each table (shaded grey) is the assigned name of the element, with the entries in rows below being the affordances provided by the element. Arrows are used to indicate the flows of movement by users permitted by the interface between different elements. Furthermore, the elements are placed into three categories that refer to the purpose of interaction: 'Interaction with Educational Content', 'Interaction with other Students', and 'Assessment of 'Learning'.

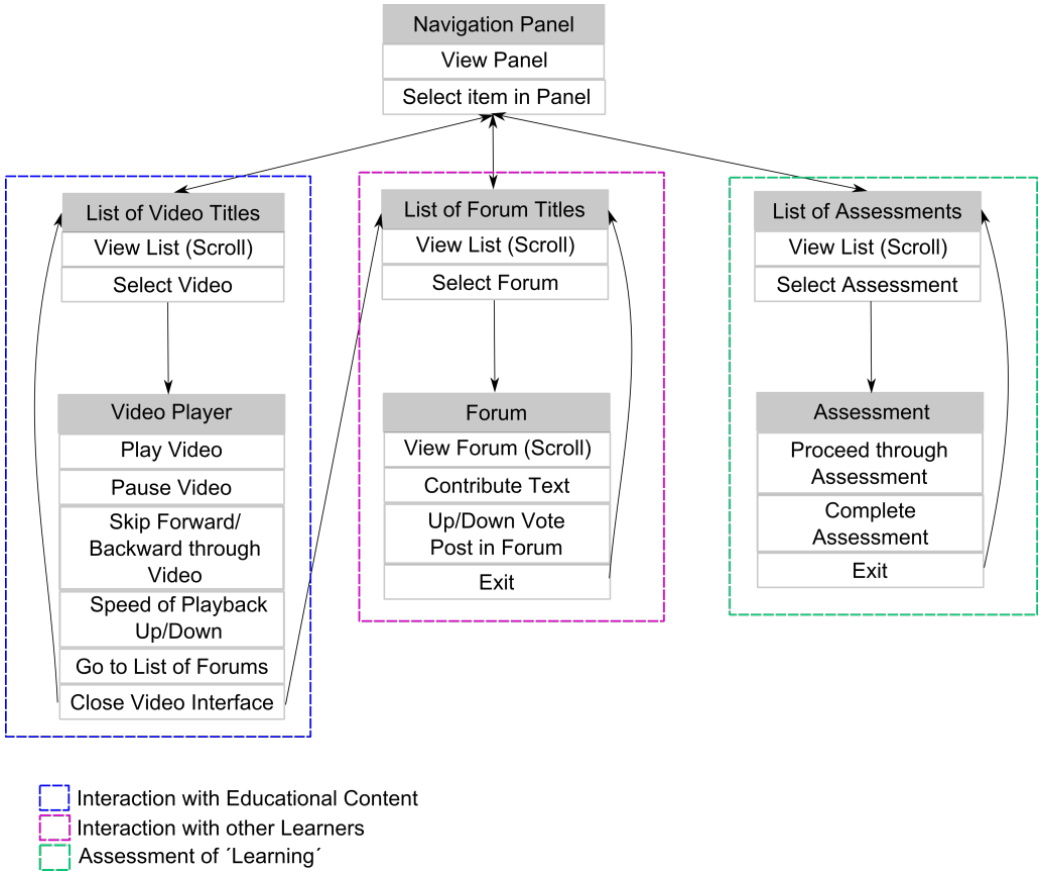


Figure 2.1. Taxonomy of interactions in Coursera MOOC

At the top of the taxonomy is the element 'Navigation Panel'. This panel is a constant feature on the left of the MOOC interface and is shown in Figure 2.2. In Figure 2.1, double arrows are used between the navigation panel element and other elements to indicate that the navigation panel is always visible.

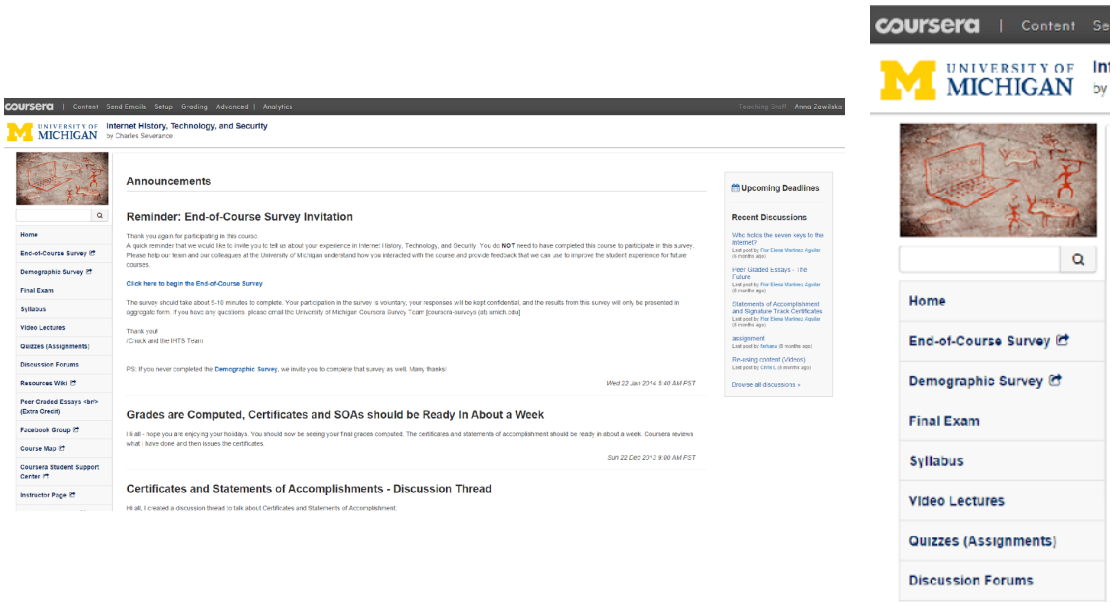


Figure 2.2. The navigation panel always on the left of the interface (left, showing the entire interface), and the panel is isolated and enlarged (right)

From the navigation panel, the user has several options for how to proceed. If they would like to interact with the educational content of the course, they select the 'Video Lectures' item in the navigation panel. This causes a list of video titles to be displayed as shown in Figure 2.3.

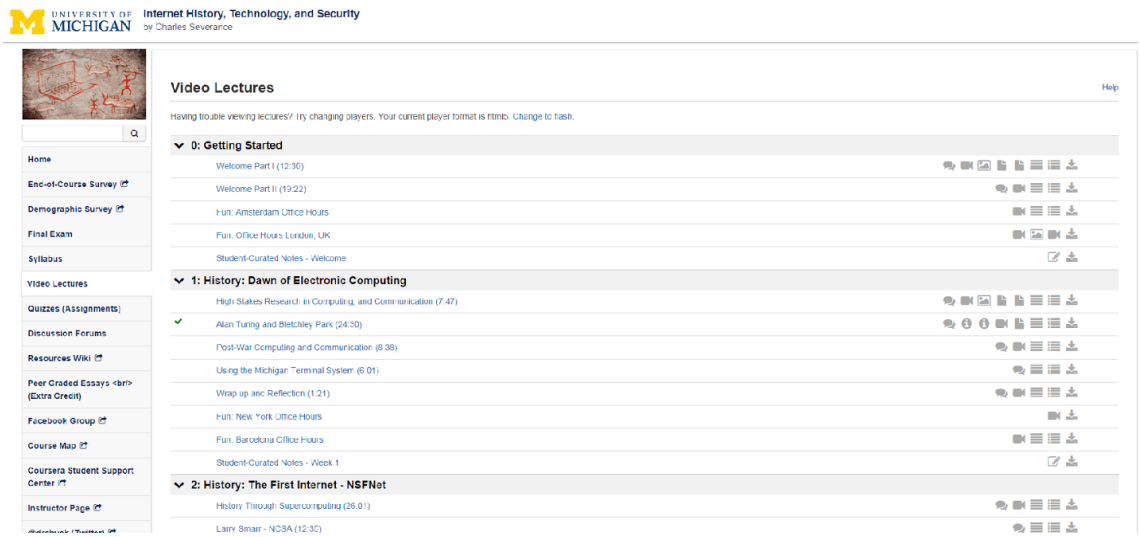


Figure 2.3. List of video titles

From this list, a student may select a particular video title that causes a video player interface to come up on the interface, shown in Figure 2.4.

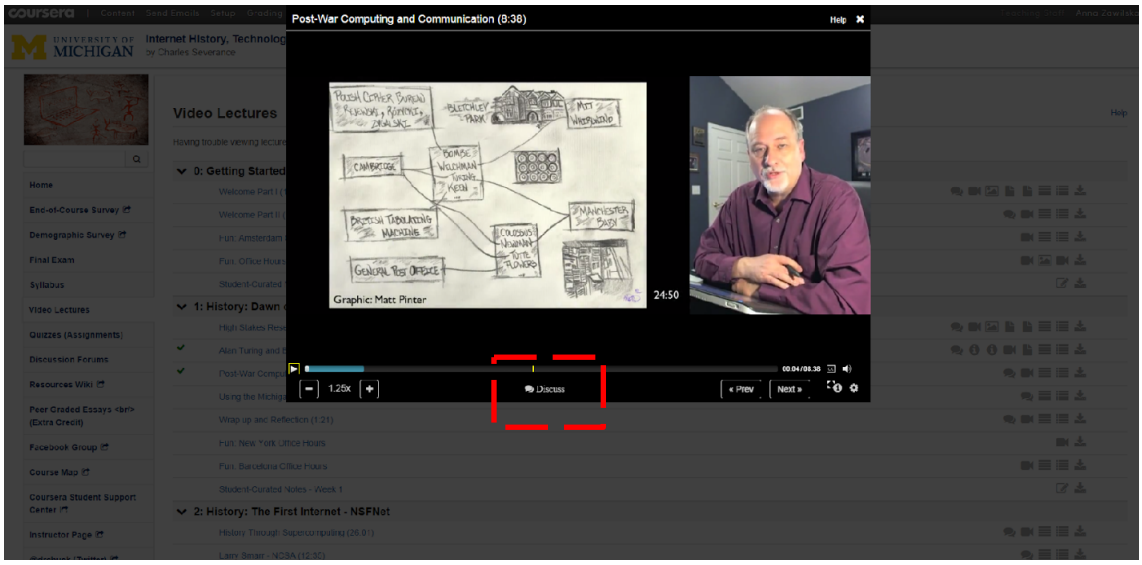


Figure 2.4. Coursera video player

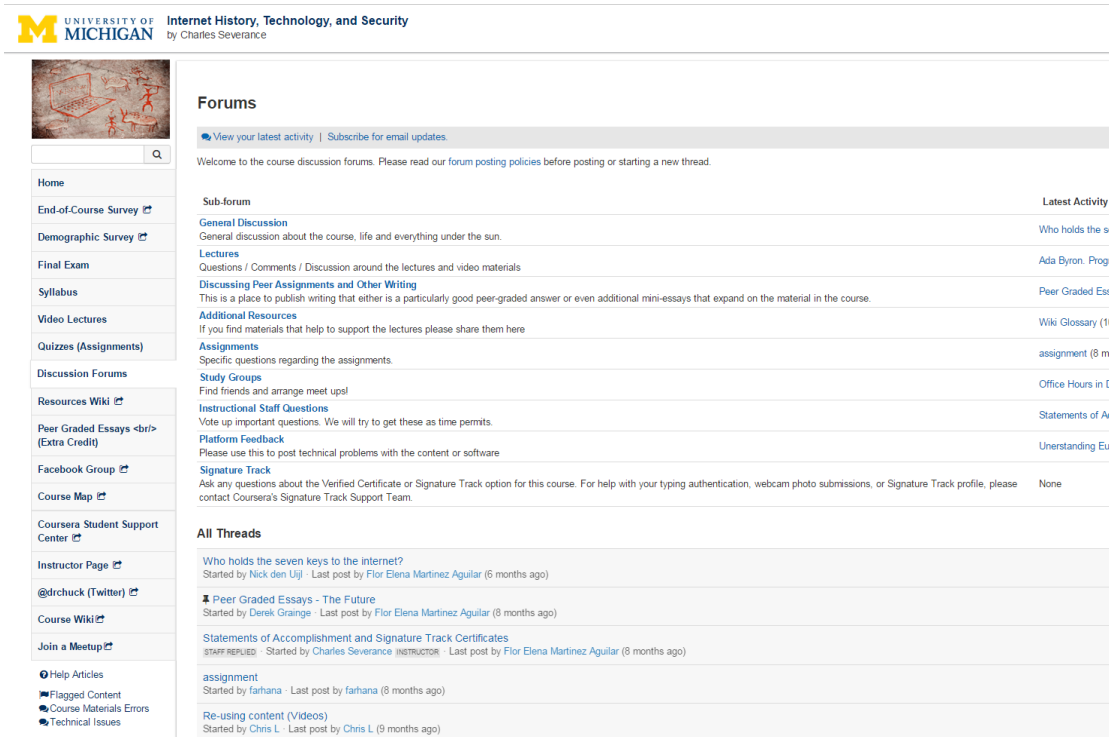
Of the three primary kinds of resources available on the MOOC (video, discussion forum, and quizzes), video is consistently the most popular resource. This may be expected given that the video provides almost all of the course content. For example, Table 2.2 shows rates for a course on Bioelectricity hosted by Duke University (Belanger and Thornton 2013).

Table 2.2. Statistics comparing popularity of video, quizzes, and forum in a MOOC (Belanger and Thornton 2013)

	Number	Percentage
Total enrolled students	12725	100%
Watched at least one video	7761	61%
Submitted at least one quiz	3658	29%
Posted in forum	800	6%

The video player interface in Figure 2.4 offers a variety of functionality as indicated in the 'Video Player' table within the 'Interaction with educational content' category of interaction in the taxonomy in Figure 2.1. This includes exiting the video interface, which returns the user to the list of video titles. A student may move directly from the video interface to the discussion forum by clicking on a 'Discuss' button at the bottom of the video player (shown in Figure 2.4 using red dotted lines), which brings up on the interface a list of discussion forum titles that may be clicked on to view the particular forum contents. Alongside each title are shown two numbers: the number of views and the number of posts for that discussion forum. A typical list of discussion forum titles is shown in Figure 2.5.

From the navigation panel, the student may also visit discussion forums by selecting 'Discussion Forums' from the navigation panel. This brings up a list of forum titles from which a student may select a particular forum. A list of forum titles and a forum are shown in Figure 2.5.



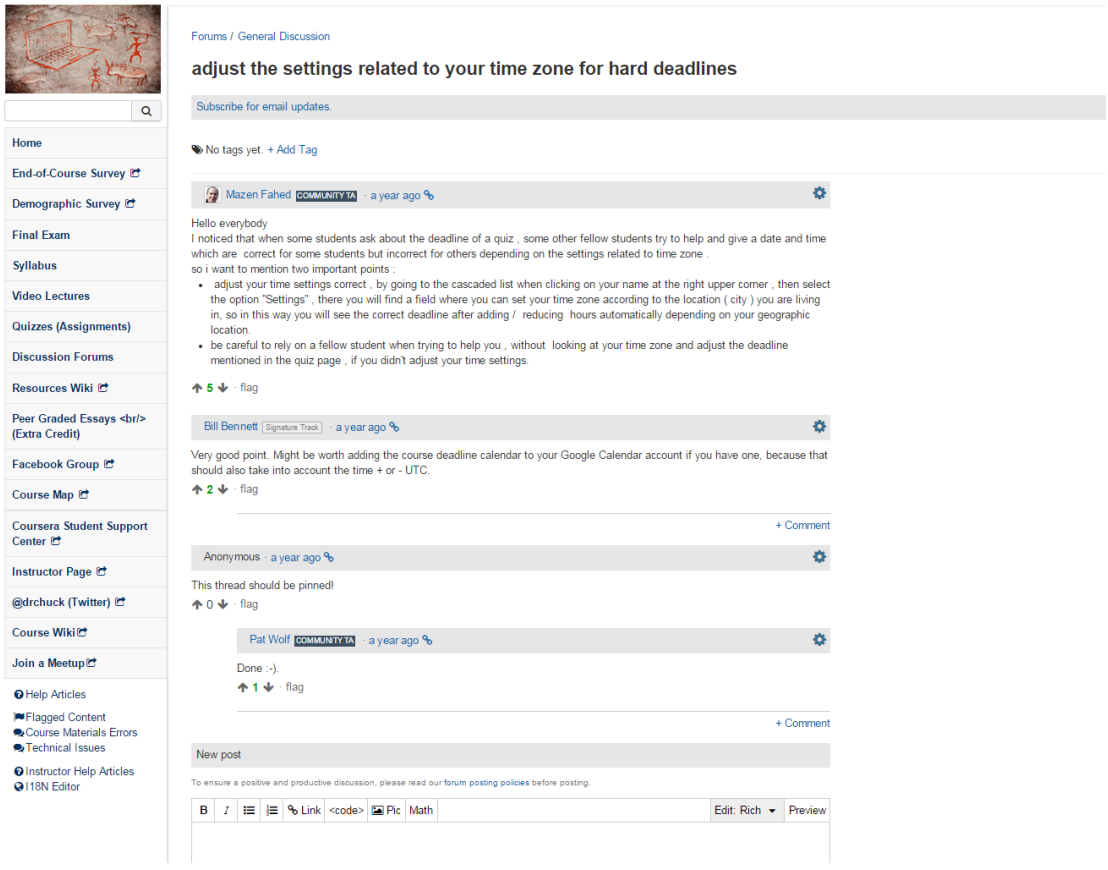


Figure 2.5. List of Coursera forum titles (top) and one forum (bottom)

From the navigation panel, the user may do assignments to earn course credit that goes towards their final score by selecting ‘Quizzes (Assignments)’ from the panel. This brings up a list of assessment titles from which a student may select an assessment to attempt. The exact nature of these assessments is determined by the course lecturer, but is normally similar to a multiple-choice quiz. In some courses, students may attempt a quiz many times before a certain deadline, to earn the best score possible. A typical list of assessments and an assessment are shown in Figure 2.6.

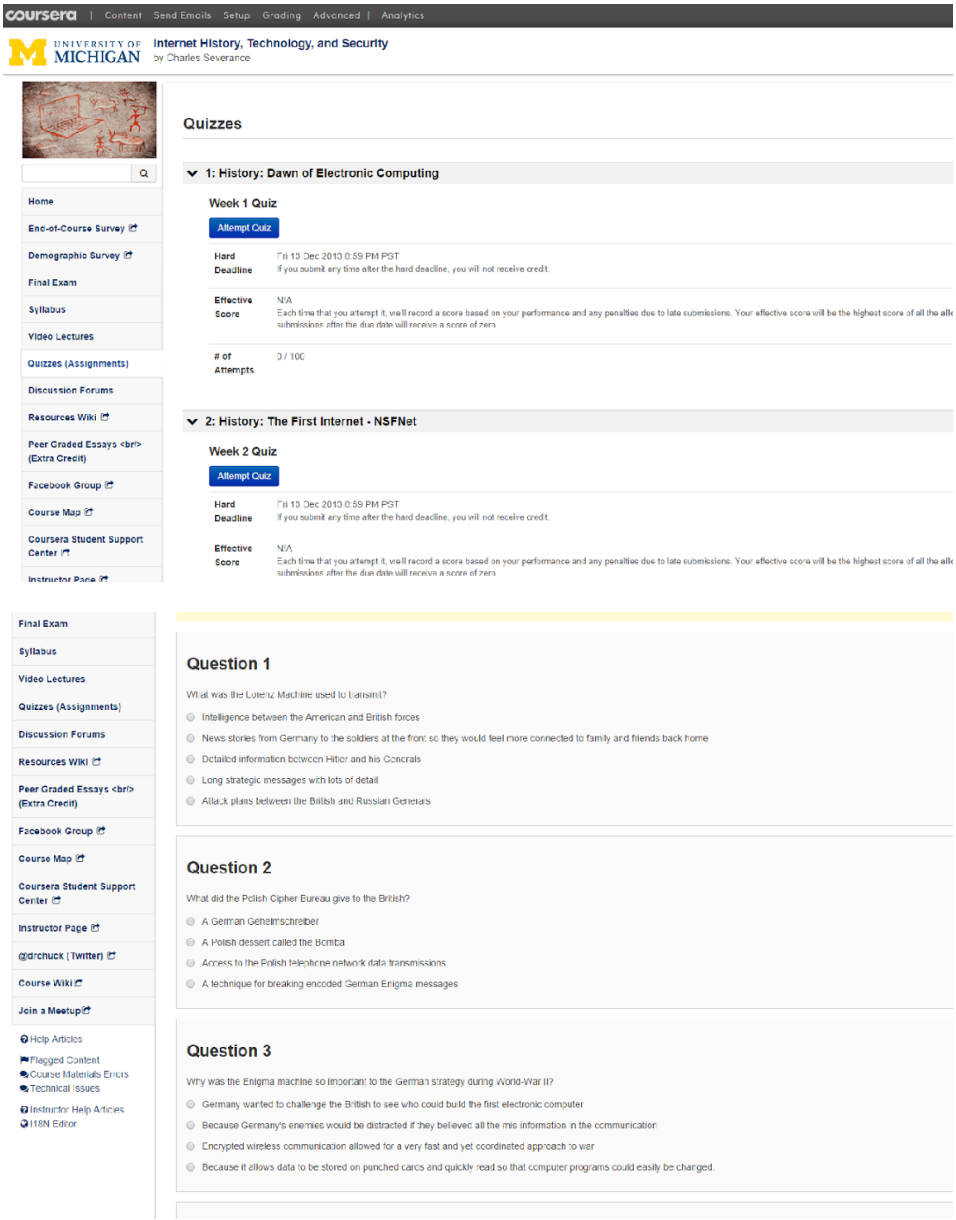


Figure 2.6. List of assessments (above) and a multiple-choice question assessment (below)

The taxonomy in Figure 2.1, along with the subsequent discussion, highlights some of the main interactional issues of the MOOC:

1. The three categories of interaction presented in the taxonomy are largely disjoint. For example, interaction with the educational video content is a primarily solitary experience for the enrolled student. While the student may move from the video player to the discussion forums using the 'Discuss' button on the video player without needing to use the navigational panel, they still have to parse and navigate the list of discussion forums, which may be large depending on the number of active students. Therefore, it is not straightforward in the interface to find a discussion in the forums that corresponds to elements in the educational video materials, which poses a challenge for creating social interaction around educational content and may deter students from doing so.

2. Students engage with the educational content (video) individually, and the video is used primarily as a broadcast medium of which students are consumers.

3. Interactionally, 'learning' is measured individually for students in standard ways based on their responses to assessment questions with singular correct responses.

As indicated by the interface flow, the extent to which the MOOC supports social interaction is substantially underdeveloped particularly around course content.

And yet, supporting social interaction in the MOOC is argued by some to be of central importance. Research suggests that contributions to the MOOC educational content by students during social interaction may be as or more important than the course content itself. The contributions generally provide a range of interpretations of the course content as well as various semantic connections by students between topics

(Pea and Lindgren 2008). Therefore, these contributions add new points of view and generate important conceptual diversity in the course content. With the existing MOOC interface flow, there is minimal opportunity for a Coursera MOOC student to experience both the original video content as well as, related to the video, the contributions of others in an organised and easy-to-view way. Given this understanding of how the current interface constrains interaction, I suggest that a worthwhile design goal for this thesis is the creation of a video annotation interface. This interface would simultaneously display the educational video content and support interaction between students, which I call 'intra-video social interaction'. This design goal will be the focus of the Computer Science work taken up from the next chapter onwards in this thesis.

2.7. Discussion

The review of literature on MOOCs in this chapter reveals that little is known about the detailed mechanics of *how* interaction (both between students and the platform, as well as social interaction) unfolds in the MOOC. Most existing research focuses on the new opportunities in data analysis that have surfaced through large quantities of server log data. These data typically consist of events such as website button clicks, discussion forums posts, and page views. Because of the sheer quantity of the data that can be collected, existing research often uses automated methods to organise and process the data. This typically results in statistically-generalisable, quantitative patterns in interaction on a broad scale: across large cohorts of students over entire courses. The focus on these kinds of studies is echoed more broadly across the study of massive online systems, where typically the term 'big data' may be used to refer to these large amounts of server log data that form the basis for analysis. Although the term 'big data' can be variably defined, it usually refers to large amounts of quantitative data

that are automatically used in Machine Learning systems and statistical analysis to surface large-scale patterns. While 'big data'-based studies still constitute a substantial proportion of existing literature into massive online systems, there is a growing research movement to expand understanding of interaction beyond 'big data' and to take a more qualitative approach. The movement focuses on the details of how interaction unfolds in these platforms and the important interactional issues at play which 'big data' may not show. For example, where 'big data'-based approaches may pick up useful patterns in when a website button click takes place, they don't help us understand at a ground level *how* or *why* that click came about in the first place. At the same time, previous Sociological work across a number of social settings, including the seminal work of Goffman (1981) and Schegloff (1992), has shown the important mechanics of interaction that are surfaced through a detailed look at interaction-in-the-wild. This work suggests the value of doing a similar kind of analysis of interaction in the MOOC. By illustrating this gap in current understanding of interaction in MOOC, this chapter serves to illuminate the conceptual territory that this thesis will aim to partially occupy and the conceptual gap it will contribute towards filling.

Expanding our understanding of how interaction unfolds in detail, will help to create implications for a set of interface requirements that better support interaction. The main design issue that this thesis will address, through an examination of the fine-grained details of interaction, is the lack of integration between the viewing of video content and social interaction on the platform. A taxonomy of the interfaces in the MOOC and the possible flows between interfaces, shows that interaction with the video content on the platform is not well integrated with social interaction. It reveals a substantial design opportunity, which can be filled through the development of a video

annotation interface that allows students to simultaneously view video as well as socially interact with other students around that video.

A study of the fine-grained aspects of social interaction in MOOCs requires methodologically moving away from 'big data' towards new forms of data. In the next chapter, I will discuss the challenges associated with this kind of study and the methodological innovation that is occurring in this area. I will then put forward and justify a particular methodological approach for the rest of this thesis.

Chapter 3.

Methodology:

A Qualitative and Ethnomethodologically-Informed Approach to Social Interaction in Massive Online Platforms

3.1. Introduction

In this chapter, I will set out the methodological stance of this thesis. In the previous chapter, I highlight the partial gap in understanding the details of how interaction unfolds in MOOCs (and in massive online platforms more broadly), where interaction has so far largely been viewed using 'big data'. Furthermore, in Chapter 2, I illustrate the ways that interactions are constrained by the current interface flow of the MOOC, and suggest a design goal for this thesis: the development of requirements for a video annotation interface that more seamlessly integrates video viewing and social interaction. In setting out the main methodological view of this thesis in this chapter, I will begin by discussing a broad methodological trend in computer system design: to take a more qualitative, iterative, and empirically-based approach to studying and designing for interaction. This is called the *turn to the social* movement. A central part of this trend is the incorporation of methodologies from Social Science into Computer Science and Human-Computer Interaction research, and I will proceed to discuss a number of these methodologies. At the heart of the methodological stance of this thesis is ethnomethodology, discussed in the final sections of this chapter, which focuses on understanding the fine-grained aspects of the sequential unfolding of interaction as it is accomplished by the participants in the setting.

3.2. A Turn Towards an Iterative, Empirically-Grounded Study of Interaction Around Computer Systems

There is a movement in Human-Computer Interaction and Computer-Supported Collaborative Work called the *turn to the social* which is associated with an increased empirical focus on understanding *how* interactions actually occur in practice through and around technology (Grudin 1990). The focus in this thesis on the fine-grained aspects of how interaction is accomplished in the MOOC fits within this broader movement and its methodological implications. In his seminal piece, Grudin (1990) describes five phases of this *turn to the social*, driven by the evolution of computer systems from solitary, to their uses in cooperative social settings with the introduction of networked computing and groupware systems.

Methodologically, the *turn to the social* is associated with a transition in the ways that interaction is studied and how understandings of interaction are used to develop requirements for computer systems. One key transition is away from the use of 'cognitive models' as a way of characterising the behaviours of users. When computers were first used in the workplace, they were designed to perform one specific simple task, operated by one individual. In these scenarios, the methodological approach centres on the use of models of human behaviour informed by cognitive theory and experimental psychology, with a focus on redesigning the interface for ergonomics (Grudin 1990). An example of such a model is GOMS (goals, operators, methods, and selection rules) (Card et al. 1983). The model incorporates fixed goals of users, the features of the system on which the users can take action, and selection rules that are the choices the user can make as they proceed through the task. The model is used to estimate how long a user will take to complete various tasks (Preece et al. p.96), as well

as how quickly users can be trained to execute specific operations on a system (John and Keiras 1996). GOMS is also used as a basis for mapping out the ways a task may be performed – in terms of a number and sequence of actions (e.g. (Shepherd 2000 p. 16)).

As interactions with technology become more commonplace and integrated into the everyday practices of users, models of behaviour such as GOMS, are less relevant as they don't capture the complex social and material settings in which interactions involving technology take place. The literature provides many examples of this complexity and its important impact on computer system design. For example Luff et al. (2014) consider "media spaces" which are audiovisual environments through which participants can work and interact remotely. They show how even simple referential activity, such as pointing, may be disrupted by media spaces. This is because the activity is produced collaboratively through an emergent sequence of coordinated actions, which is disrupted as the media space fragments the work setting, with the work space now distributed geographically. Dourish and Bell (2011) illustrate the important implications of Ubiquitous Computing culturally and socially, as well as technically. They present a more complex and accurate view of key topics such as infrastructure and privacy. In considering how to develop an 'intelligent' gavel for distributed auctions, Heath et al. (2009) show the complex ways that participants coordinate their activity through pace and rhythm of bidding, and the key role of various forms of visibility and witnessability in the setting.

A shift in focus away from theoretical models of user behaviour towards understanding interaction empirically also changes the flow of the design process away from *linear* towards more *iterative*. Previously, theoretical models were thought to allow for more

efficient design. The theories were sometimes said to inform the selection of which designs to explore through implementation and testing, and which to ignore (Carroll 2003 p.161). The overall design process was structured to proceed according to a *Waterfall* model (Boehm 2006). An initial design was generated which was then reviewed according to models of user activity which were seen as predicting users' behaviour, thus removing the need for contact with real users in real environments of use. Furthermore, there was no retrospective editing or modification of the system.

However, the *waterfall* approach to system design has since been critiqued around three major issues: the negligence of contexts of usage, the lack of an iterative design process, and the lack of integration between the theoretical models used in these traditional approaches and the actual experience of system users (Lewis 1990). Given how cheap and pervasive technology is becoming, there are ever-new complex ways that people use technological systems (Schmidt and Cohen 2013) – complexity which increasingly needs to be central to an iterative, empirically-grounded design process, further delegitimising traditional approaches grounded in the *waterfall* model. As there is a growth in distributed systems, and complex collectively-conducted patterns of interaction, designers have become much more sensitive to the need to relate and place *in context* the behaviour of individual users with respect to the group and organisational life within which it is located, and to iteratively study the way various versions of a system are used in practice (Bowers and Rodden 1993; Grudin 1990). In *iterative* design approaches, usually each iteration is designed to introduce a new kind of complexity or facet of interaction for examination (Larman and Basili 2003). The prototype used in each new iteration may incorporate insights from the previous iteration, and the research questions of focus for a particular iteration may be informed by the insights from the previous one.

The need to move away from traditional system design approaches is underscored by examples of computer systems that have failed in practice in part due to a lack of attention to the social settings of computer usage and work. For example, the following is a quote taken from a project attempting to build a new computer system to support the work of London ambulance dispatchers, but which was considered by many to be a failure (Page et al. 1993 p.40):

> "Management were misguided or naïve in believing that computer systems in themselves could bring about [such] changes in human practices. Experience in many different environments proves that computer systems cannot influence change in this way. They can only assist in the process and any attempt to force change through the introduction of a system with the characteristics of an operational "straight jacket" would be potentially doomed to failure."

The report suggests that part of the problem with the developed system is that it was not designed with sensitivity to the existing practice of dispatchers and the setting in which dispatchers work. This insight into *how* the dispatcher accomplish their work was missed.

The *turn to the social* also calls for a fundamental change in the goals of computer system design (Grudin 1990). Given the wider variety of tasks that computer system hardware such as mobile devices provide, a primary goal of design is no longer to create computer systems supporting a small set of simple tasks. Rather, the primary goals now encompass creating designs which are consistent with users' expectations and intentions, and which support a wide variety of user activities with recognition of the larger social and organisational settings in which the activity sits.

A further methodological consequence of the *turn to the social* is the incorporation of approaches informed by Social Science into computer system design. This includes a methodological shift in computer system design away from only quantitative models and metrics around system usage, towards qualitative research (Sommerville and

Sawyer 1998). Among the methods drawn into computer system design from Social Science, ethnomethodology is one (Jirotka and Wallen 2000; Martin and Sommerville 2004). Ethnomethodology is of particular interest in this thesis as its focus on the sequential unfolding of observable-reportable social interaction around and through an online system, aligns particularly well with the goals of this thesis to study the details of how interaction unfolds. Therefore, I will return to discuss ethnomethodology in some detail in Section 3.4.

3.3. Incorporation of Qualitative Social Science Methods into the Study and Design of Online Systems

In the previous section, I discuss the *turn to the social* which aligns with a shift in computer system design towards an empirically-based understanding of how interaction takes place through and around technology. A key aspect of this is a shift away from theoretical models of an individual's cognitive state, towards an understanding of social life, which draws on Social Science methodologies. Here, I discuss three methodologies that may be used to study interaction in the MOOC and create implications for the system design. These methodologies use qualitative data, in contrast to the traditionally-used 'big data'-based studies (Chapter 2).

3.3.1. Usability Testing: the think-aloud protocol and Constructive Interaction

In Section 3.2, I discuss that the *turn to the social* encourages the incorporation of stages into the technology design process which tests systems with real users – *usability testing* is one of the approaches that supports these kinds of tests. Usability testing takes a working version of a software system and asks a set of participants to

use it. From this, observations are made on the usability of the system and from this, design modifications may result. Studies that compare usability testing with non-empirical testing have found, among other things, that usability testing uncovers serious interface flaws, is an efficient way of finding recurring problems, and seldom produces any false problems (Jeffries and Desurvire 1991).

Usability testing is often used in combination with an iterative design approach, in contrast to the traditional *waterfall* approach (Section 3.2). Therefore, usability testing may be performed repeatedly on a series of prototypes as part of the iterative development of the system towards its final state (Eden 2011).

Two well-known and commonly-used approaches to usability testing include the *think-aloud protocol* and *Constructive Interaction*. These two approaches will be discussed in turn next.

The *think-aloud protocol* is perhaps the most well-known technique for prototype evaluation (Eden 2011 p.36). Participants are given a set of tasks to perform using a system and are asked to verbalise their thoughts to an observer as they progress through the task. The observer may need to prompt the participant to speak with statements like “what are you thinking now”, and “please keep talking as you do the task”. Think-aloud experiments may be seen to gain access to user misconceptions about the system (Nielsen 1993 p.195). The resulting data from a think-aloud experiment – the talk of the participants as they perform different actions using the interface, as well as the observer’s notes about the participant’s comments and actions – may be (and has traditionally been) associated with surfacing the user’s internal mental state (Section 3.2) through identification of key events, speech, and patterns of behaviour (Preece et al. 2002 p.381). The strength of this approach is that it typically

surfaces the most obvious or substantial flaws in a system's interface design (Douglas 1995; Jeffries and Desurvire 1991).

However, the think-aloud protocol is argued to have some substantial limitations. In particular, it may be seen to be difficult to implement correctly given the unnaturalness of asking a participant to constantly verbalise their thoughts while working through a given task (Nielsen 1993 p.196). It's common for observers to forget to prompt participants enough (Boren and Ramey 2000), and participants may find it difficult to verbalise their experiences in general (Smith and Dunckley 2002). Additionally, the experiment is awkward for the observer who cannot over-prompt the participant which makes them uncomfortable and act unnaturally (Smith and Dunckley 2002).

A collaborative think-aloud technique called *Constructive Interaction*, developed by Naomi Miyake (Draper et al. 1984), was developed to at least partially address some of the limitations of the think-aloud protocol. Constructive Interaction involves two participants working together through a shared problem or activity (Miyake 1986) with a broad objective such as together figuring out the operation of a piece of technology. After they are given their task or activity objective, the participants are encouraged to engage with each other with minimal interruption from the observer. Constructive Interaction may be more natural than the think-aloud protocol since people are accustomed to explaining their ideas and perspectives to others. The discussion between the pair of participants is usually not considered to be direct evidence of their cognitive state, rather it is evidence for what the participants understand by what they see, and their expressed rationales for proceeding in a particular way, which they make visible to their co-participant in the name of collaboration, but also incidentally to the researcher (Draper et al. 1984 p.264). This artefact of collaboration is therefore a more

naturally generated protocol (Suchman 2007 p.123). The data captured from Constructive Interaction may be video with audio of the interactions between participants as well as possibly computer-generated records such as keystrokes (Douglas 1995). The original focus of Constructive Interaction was to use it to model the iterative processes around developing understanding (Miyake 1986) through the identification of the ways new thought processes of each participant may develop through the pairs' interaction (Draper et al. 1984 p.264). Constructive Interaction has since informed the study of computer system usage in human-computer interaction research (Draper et al. 1984; Van den Haak et al. 2004; Kahler 2000; Wildman 1995).

Both the think-aloud protocol and Constructive Interaction 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 p.64).

A criticism of Constructive Interaction is that it may remain similar to think-aloud protocols if the analysis of data elicited is informed by cognitive theory, and if the results of analysis are viewed to be a reflection of information-processing events occurring inside the individual's mind (Ericsson and Simon 1980). However, without drawing on cognitive theories, Constructive Interaction has been used with relative success to identify important system usability issues and design implications in a number of research areas such as within Human-Computer Interaction (Draper et al. 1984; Suchman 1987), Ubiquitous Computing (Dourish and Bell 2011) and Artificial Intelligence (Suchman 1987).

3.3.2. Quasi-Naturalistic Experimentation

One of the challenges associated with the study of interaction around massive open online systems such as the MOOC, is establishing a site for data elicitation. This may be the reason why most data about interactions between participants through online systems are based on data retrieved from the server (Chapter 2).

A further challenge to finding a site for elicitation may be related to the system itself. For example, if I would like to explore possibilities around the ways the MOOC might better support intra-video social interaction, this is a methodological challenge as the MOOC interface currently does not support intra-video social interaction therefore there is no obvious naturalistic site for data elicitation. Furthermore, even if a new system prototype is developed, it is rare to be able to study the prototype immediately as part of naturally occurring social interaction by making it live online. For these reasons, trying to understand massive open online systems and prototype versions of the system may require a more experimental approach (Heath et al. 2010 p.144).

Quasi-naturalistic is one kind of experimental approach. In these experiments, participants are recorded undertaking simple tasks with a prototype system. The tasks are designed to seem as natural as possible, open-ended, somewhat familiar to the participants, and not require intervention from the researcher (Heath et al. 2010 p. 144). Unlike standard experimental methods, it is not important that the participants complete the task. Rather, the activity of the participants is recorded, usually by video, and is analysed with a regard to the interactional organisation through which the participants accomplish activities within the task. The analysis is usually focused on *how* the participants interactionally accomplish the task (Heath et al. 2010). This kind

of analysis may be informed by ethnomethodology which is discussed further in Section 3.4.

Quasi-naturalistic experimentation has previously proven successful particularly in revealing our limited understanding of fairly basic features of social interaction. As Garfinkel suggests, quasi-experimental data can aid "a sluggish imagination" (Garfinkel 1967). For example, in the subject area of media spaces for mediated communication in Computer-Supported Collaborative Work, quasi-naturalistic experimentation reveals pre-suppositions and resources which have largely remained unexplicated in other more conventional studies of the workplace. Many studies using quasi-naturalistic experimentation surfaces how little is known about seemingly simple and possibly everyday activities, and in particular "the competencies and skills that people bring to bear in their accomplishment" (Heath et al. 2010 p.144). I discuss the value of examining seemingly "simple activity" and possibly "everyday" again in Section 3.4.

One of the most well-known set of quasi-naturalistic experiments, which is informed by ethnomethodology, was conducted by Suchman (1987). Her work has had a substantive impact on a number of research areas including Artificial Intelligence. I will return to discuss her work in more detail in Section 3.4.

3.3.3. Online Survey and Non-Probabilistic Research

A commonly-used method for gaining a snapshot of reported views or perspectives from users of online platforms is Survey (Marshall and Rossman 2010). If free-text responses to survey questions are collected from participants, qualitative analysis can be performed to surface themes and issues that the respondent sees as important in shaping the ways they interact with a platform. There are many advantages of using an online survey, including the following three:

1. *Access.* Especially if there is an easy way to send out messages to all potential participants such as using a mass email containing a link to a survey online, it is likely relatively easy to at least provide participants with the capacity to respond to a survey with little effort to the researcher beyond developing the survey questions.

2. *Time and flexible research.* Survey responses collected online through an online survey are easily and immediately accessible online in digital format, allowing for data analysis to begin simultaneously with continuing data elicitation.

3. *Cost.* An online survey eliminates the need for paper and other costs such as manual survey delivery, printing, and converting data to digital format for analysis.

There exists a substantive corpus of research that exploits these advantages of survey to investigate online communities, online relationships, and a variety of other aspects of computer-mediated sociality (e.g. (Andrews et al. 2003; Bachmann and Elfrink 1996; Stanton 1998; Witmer et al. 1999; Yun and Trumbo 2000)).

However, an online survey also has a number of associated disadvantages, including:

1. *Inconsistencies in interpretation of questions among respondents*. There is some uncertainty with regard to the way respondents understand the survey questions. This is dependent on, for example, phrasing of and terms used in the questions. This disadvantage is more easily managed in other methodologies such as an in-person interview, where the participant has the opportunity to ask the researcher for clarification of any questions. This issue may be mitigated by piloting survey questions with an available sample of participants who may be said to be typical of the potential respondents of the survey (Marshall and Rossman 2010).

2. *Sampling*. There are significant issues around sampling and associated statistical generalisability for online surveys. The primary problem is due to so-called 'lurker' users online who do not make their presence known, therefore making it difficult to obtain an accurate sampling frame (Wright 2005). Additionally, a bias called the 'self-selection bias' is introduced into the responses (Stanton 1998; Thompson et al. 2003) because whether or not a potential participant replies to a survey question depends on their initiative. This is a form of systematic bias introduced by the tendency of some individuals to reply to an invitation for an online survey over others. These biases limit online surveys to be used for nonprobabilistic research only. This means that the findings may not be statistically generalisable.

However, there are other forms of generalisation beyond statistical, which may also be associated with deriving results from data and are useful for computer system study and development. As opposed to statistical generalisation, one may consider *logical generalisation*. Statistical generalisation is "the process by which the analyst draws conclusions about the existence of two or more characteristics in some wider population from some sample of that population" (Mitchell 1983 p.188). This is the kind of generalisation that is valued in the traditional quantitative tradition. Logical inference is independent from statistical inference and is "the process by which the analyst draws conclusions about the essential linkage between two or more characteristics in terms of explanatory schema" (Mitchell 1983 p.190). In other words, it is the consideration of the processes at play along with enabling and disabling mechanisms. This inference speaks to more general characteristics of phenomena. Although the results cannot be rigorously shown to be statistically significant, they may identify new phenomena, causal processes, counter examples or additional examples about existing theories/conceptual frameworks/phenomena. In so doing, this kind of

research may illuminate new territories for thinking about issues; particularly useful at the early exploratory stages of a study.

Online survey responses may therefore only provide material for logical inference, and thus cannot guarantee external statistical validity. However, by being explicit about the theoretical framework/conceptual framework that emerges, those who do further design research studies that are guided by the same (or similar) parameters may investigate whether the ideas described in the framework could be generalised to other settings (Marshall and Rossman 2010 p.252).

There exist many examples of nonprobabilistic research which have a major impact on Computer Science and on computer system design. One example, which will be discussed in more detail in Section 3.4, is the work of Suchman documented in her book 'Plans and Situated Actions' (Suchman 1987). Suchman uses particular instances of activity to illustrate the dependence between purposeful human action and particular context, thus discrediting methods used in traditional software engineering involving user modelling in fields including Human-Computer Interaction and Artificial Intelligence. A significant portion of the design community have taken up Suchman's arguments (Button and Dourish 1996; Duguid 2012).

3. *What respondents say versus what they do.* The researcher analysing survey responses should be sensitive to the fact that what respondents report they do may not necessarily align with their actual practice. Therefore, the results of the analysis of survey responses should be used with a sensitivity that they may provide insight into reported concepts and themes that seem important to the respondent, which may then merit further investigation or have some design implications, rather than being the objective truth about the actual practices of participants.

4. *Low response rates.* Previous literature using an online survey commonly reports low response rates. For example, email-only surveys are reported to have a rate around 20% (Kaplowitz et al. 2004), while web surveys are reported responses of as low as 2% (Petchenik and Watermolen 2011). I suggest this is not necessarily a problem in the case of massive online systems. As these systems have such a large number of participants, even if only a small fraction of them respond, these responses may still constitute a substantive corpus of data for analysis.

If free-text responses are collected from the online survey, they may be analysed using *Thematic Analysis* (Boyatzis 1998). This approach is particularly applicable for exploratory studies where no prior predictions or hypotheses are made, rather the themes and issues are left to emerge from the data. The phases of Thematic Analysis may include: (1) organisation of data, (2) immersion in data, (3) generation of categories, themes, typologies, matrices, clusters, (4) coding the data, (5) offering interpretation through note-writing, and (6) a search for alternative understandings (Marshall and Rossman 2010 p.221). In step 5, the particular interpretation of the data is expected to take on shape and validity as major modifications in the themes and issues significant in the data become rare, and data tends to begin to fall within the bounds of the established themes and issues. In step 6, the researcher may remain sensitive to alternative interpretations by questioning each interpretation and associated theme as it comes to mind. This reflexive step also reduces the risk of bias affecting the emergence of themes. The questions to ask oneself include: *is there another way to view what I have read? What are other interpretations and why does mine have particular merit among others?* (Marshall and Rossman 2010 p.221). Furthermore, if the survey data are sufficiently anonymised and if ethical clearance

permits, the elicited data may be provided for other researchers to analyse in an attempt to identify any personal bias.

As mentioned above, a limitation of surveys is that they don't necessarily shed light on the actual activities of the participant. Rather, the researcher is only exposed to the participant's report of their activity. A different methodological approach that is focused on gathering empirical data on participant's actual activity, is called *ethnomethodology*. Ethnomethodology is a central methodology drawn upon in this thesis, given my interest in understanding in detail how interaction is actually accomplished by enrolled MOOC students. Therefore, ethnomethodology will be discussed in detail next.

3.4. Ethnomethodology and Computer System Design

Ethnomethodology is an analytic orientation and methodological approach which focuses the researcher on the observable-reportable practices through which participants accomplish social interaction (Section 3.2 and 3.3). In this section, I will elaborate further on ethnomethodology. It is granted particular attention because, as I will explain, the focus of ethnomethodology as an analytic orientation aligns well with the goals of this thesis to examine the fine-grained aspects of interaction in MOOCs (Chapter 1).

3.4.1. Historical Development and Key Concepts

Ethnomethodologically-informed approaches focus the researcher on understanding moment-by-moment reportable-observable sequences of interaction as they unfold sequentially within, and are shaped by, their immediate social and material setting. Ethnomethodology has its roots in the 1950s in the work of Harold Garfinkel (Heritage

1984). Garfinkel's goals were to respecify the predominant subject matter and methodological approaches of traditional Sociology, which were predominantly structural-functionalist. Rather than using social interactions as a *resource* for inquiry that may lead to a rather structural view of interaction, ethnomethodology uses interactions as a *topic* of inquiry. For ethnomethodology, social life is produced 'from within' by members of the social setting, is practically accomplished by the settings' members, and is of a locally situated character. Most traditional sociologists thought that "the objective reality of social facts is Sociology's fundamental principle" (Durkheim 2013), and from this all Sociological reasoning and practice followed. From this view, social order was a form of concerted action performed in accordance with shared rules and norms (Parsons 1937). Garfinkel's critique of traditional Sociology included challenging its foundational 'objective reality of social facts'. He was not satisfied that social order proceeded unproblematically from these social facts, and that rather social order is *produced* in and *through* the actions and interactions of participants. The 'objective reality of social fact' is not to be assumed but studied, in the view of Garfinkel. He rejected the notion of creating abstract theories such as Actor-Network Theory (Law and Hassard 1999) and Activity Theory (Nardi and Redmiles 2002)[4].

Garfinkel used the term 'judgemental dope' to describe the decision making of society's members according to the traditional Sociological view that they act blindly in accordance with theoretically formalised rules and norms. With regard to interactions around video on massive open online platforms such as MOOCs, an ethnomethodologically-informed view may therefore inquire how the natural 'facts' of

[4] For a comprehensive review of other Sociological orientations, please see (Randall et al. 2007)

sociality around video are produced in the first place rather than having them 'in place' and theorising them.

Ethnomethodology is distinct from conventional Sociological theory because it doesn't accept any theoretical structures or factors imposed onto ordinary social activity. Rather, ethnomethodology draws attention to the ways in which participants in society *produce* their activities and how those activities are recognised by other participants as part of taken-for-granted social life (Garfinkel 1967).

Therefore, ethnomethodology is said to *respecify* Sociology with regard to what Garfinkel calls 'reflexively accountable action' (Button 1993; Heritage 1984) and in terms of order "in-and-as-of-the-workings-of-ordinary-society" (Button 1993 p.7). Ethnomethodology focuses on how people exercise these common-sense understandings by finding, within the immediate circumstances of action, the means to understand it and interpret it for practical purposes.

The value of applying an ethnomethodologically-informed view even to areas of seemingly simple or everyday activity, is emphasised by many studies which make apparent how little we know about these activities and the kinds of competencies people bring to bear on their accomplishments (Section 3.3.2). Studies of virtual reality environments such as 'Second Life' or 'World of Warcraft' have shown that contrary to initial expectations, issues around referential practice in the physical world persist in the digital domain (Brown and Bell 2004; Moore et al. 2007). These studies suggest that 'simple' and 'ordinary' activity should be taken seriously, and insights about offline social interaction may assist the study of online social interaction and development/ design.

The focus on ‘mundane’ and ‘everyday’ activity links up again with a central idea in ethnomethodology: that the production of social order is not a ‘special’ activity that is done only by certain groups at certain times, but it is part of ordinary everyday life at all times. Observable-reportable interactions form a central part of this – Garfinkel suggests everyday social interaction is not just rational social behaviour, but is dependent on being intelligible to others. One of Garfinkel’s colleagues, Sacks (1984) illustrates the way these kinds of understandings are applied in everyday conversation. He explores conversation between two friends discussing a police incident at a store and a car accident, and shows how the conversation is formulated by each participant so as to make certain parts of the conversation exceptional and others more ordinary.

With regard to data collection, ethnomethodology-informed analysis requires data that provides a resource for ethnomethodology’s interests in the "detailed and observable practices which make up the incarnate production of ordinary social facts" (Lynch et al. 1983 p.206). A researcher informed by ethnomethodology would like to look at instances of activity in detail, *within those circumstances*, to see the setting from the particular person’s perspective. Ethnomethodologically-informed approaches focus the researcher on understanding moment-by-moment reportable-observable sequences of interaction as they unfold sequentially within and are interdependent on their social and material setting.

I previously discussed that from an ethnomethodologically-informed viewpoint, accomplishing social interaction depends on the mutual intelligibility of the actions and speech of participants in a social interaction. Social interaction is not just about how individuals engage in rational social behaviour, but also how they can be *seen* by other participants to be engaging in it. From an ethnomethodologically-informed view,

everyone understands what rational social behaviour is and what it looks like, and ethnomethodology asks, *how?* In other words, ethnomethodologists capture this property of social interaction by saying it is 'reflexively accountable'. It is 'accountable' in the sense that the very observable-reportable ways it is organised provide to others the means to recognise it as what it is. It is 'reflexive' in that activity is accountable within the very fact of its production. The organisation of action within specific and innumerable circumstances of production, provides to other participants the means to recognise it and respond to it. Ethnomethodology asserts that the circumstances in which action and speech are used grant to others the means to recognise the nature of that action or speech. For example, the greeting 'Hello' is recognised as a greeting by being placed in a particular position within the course of a conversation. The 'work' this word does is dependent on how the word is used. "Hello" could equally be, for example, a request for attention or an expression of surprise – depending on how the word is used. This priority of human social activity – 'accountability' – is one the critical features of ethnomethodologically-informed studies (Dourish and Button 1998).

Garfinkel and Sacks created the term 'member' as someone who is competent in the usage of language and social action in appropriate social settings, as opposed to someone who just grasps syntax or grammar (Garfinkel and Sacks 1970). Being a member depends on the competent use of language or action appropriately for the setting, which is dependent on their understand of "what everyone knows that everyone knows" (Dourish and Button 1998 p.6) – their recognition of accountable action.

Other concepts around interaction that play a central role in an ethnomethodologically-informed approach are 'sequentiality' and 'temporality'.

Related to the previously-stated idea that the setting in which a particular word is said contributes towards making its meaning intelligible to others, an important aspect of a piece of speech or action is the sequence/s of activities in which it is embedded. Furthermore, within that sequence, the precise placing of activity or speech is important to the meanings they have as they are emergent and contingent on their just prior actions and speech in the setting. Within the flow of interaction, how an action or piece of speech relates to previous actions as well as gives way to future ones is essential to understanding the meaning of that piece of action or speech (Garfinkel 1967).

That the sense of actions or speech is identified by the circumstances in which they arise and that they in-turn contribute and advance those circumstances to form the foundation for subsequent activity – in other words, the "occasioned and occasioning character of practice action" (Heath et al. 2010 p.82) (may be referred to as the 'situated' nature of activity) draw attention to the local, practical accomplishment of social activity.

Dourish and Button summarise many aspects of ethnomethodology and its founding principles:

> "More generally, the notion of "what everyone knows that everyone knows" speaks to a form of common-sense understanding ("common" both because it is shared and because it is mundane) that is the basis of the mutual recognisability of accountable action. Ethnomethodology focuses on how people exercise these common-sense understandings by finding, within the immediate circumstances of action, the means to understand it and interpret it for practice purposes. These different elements – **accountability, membership, and common-sense understandings** – together contribute to the ethnomethodological frame of reference that places it in opposition to the view of the work espoused by … traditional sociology." (Dourish and Button 1998 p.6)

The main concepts underlying ethnomethodology inform some key aspects of the analytic commitments driving the ethnomethodologically-informed analysis of video

data of people's interactions. I will now unpack in some more detail the analytic commitments of ethnomethodology, drawing on work by Heath et al. (Heath et al. 2010 p.82):

Firstly, the sense and significance of an action or piece of speech is indicated by the immediately preceding and proceeding activities of the group participants such as responses of the other participants. An ethnomethodologically-informed analysis draws minimally on the data analyst's own interpretation of the meaning or significance of an action or piece of speech, rather drawing heavily on the observable-reportable action and speech to which it gives rise, to determine meaning and significance.

Secondly, it is important to assert there is no correspondence between a particular physical movement or piece of speech and the work it performs in a particular social interaction. As discussed previously in the case of the word "hello" – the same word or physical gesture can perform different work depending on the circumstances in which it occurs. Therefore, the meaning and significance of an action or piece of speech should be analysed in the context of the immediately preceding and proceeding actions. An important question that draws the analyst's attention to this kind of view is 'Why this now?' (Schegloff and Sacks 1973). This is in contrast to previous approaches on non-verbal behaviour such as those that attempt to create a relationship between a cognitive state and facial expression of a subject (e.g. (Ekman and Rosenberg 1997)). In contrast to this, an ethnomethodologically-informed approach is concerned with activity in terms of what it appears to accomplish within the immediate circumstances of its production and how it is treated by the participants in the social interaction.

A particular form of ethnomethodology is Conversation Analysis (CA), developed primarily by Harvey Sacks (Sacks 1992). Conversation Analysis examines fragments of naturally occurring conversation and looks within that data for the mechanisms by which conversation is accomplished by participants. Reflecting ethnomethodology's focus on observable-reportable social activity, Conversation Analysis looks at the ways different aspects of conversation (such as introducing a subject or closing off a topic of discussion) are managed by the participants in the conversation (Sacks 1992). This is similar to the previously discussed example of the word "hello" and how it may have different meanings depending on the context in which it is produced, and what comes next in the conversation.

This review of ethnomethodology reveals that its analytic orientation aligns well with the goal of this thesis to study the *details* of the how and why interaction unfolds in the MOOC (Chapter 1). As no naturalistic site for intra-video social interaction exists for the MOOC, I suggest using quasi-naturalistic experimentation (Section 3.3.2) during which the participants' interactions are video recorded. This video data may then be subject to a kind of ethnomethodological analysis in that it also looks at the local, situated, real-time organisational specifics of the social phenomena. In this case, one could call this approach *informed by ethnomethodology*. I will return to a more detailed discussion on the methodological choices taken in this thesis in Chapter 4.

3.4.2. The Use of Ethnomethodology in Design of Massive Open Online Systems

I will now consider more specifically how ethnomethodologically-informed studies of social interaction through and around technological systems, be they naturalistic or using prototypes in a quasi-naturalistic setup, can be used to inform computer system

design. There is an important conceptual point to be made about how ethnomethodology, with its focus on moment-by-moment organisation of practice as it is situated in its social and material context, may be used for the design of technology. In the face of the unavoidably transformational and interventionist nature of technology, it seems ethnomethodology becomes powerless. Dourish and Button call this the 'paradox of technomethodology' (Dourish and Button 1998). However, the key to overcoming this paradox is to see the role of ethnomethodology not to provide rules of details of interaction that must be programmed into a system, but rather to provide a way of elucidating the design specification for particular applications by showing what functional equivalents of interactional mechanisms would have to be built (Button and Sharrock 1995). The situated practice that ethnomethodology allows us to surface and study provides the grounding for the determination of important *interactional accomplishments* for which the system needs to provide functional equivalents, rather than attempting to recreate those same *details* of interaction with the final system.

Another view on overcoming the *paradox of technomethodology* is related to generalisability of social interaction in particular. Some believe there is an inherent 'generalisability' in ethnomethodologically-informed approaches to social interaction, even if only one study or experiment is conducted. This point was articulated by Sacks (1963), and is related to the previous discussion about logical generalisability (Section 3.3.3). Sacks highlights the ethnomethodologically-informed view of social order: that it is a mundane feature of everyday life and a constituent feature of the ordinary activities and common-sense reasoning that inhabits and animates it. The random encounters of individuals within a society's culture provides us with a shared resource for 'arranging' the social situations in which we find ourselves engaged, whether those

things are familiar or not. Therefore, even individual instances of social activity are generalisable because a group's participants will have *shared cultural practice* that constitutes their membership. The practice is constituted by the existence of what Sacks terms the seen-but-unnoticed "machinery of interaction" (Sacks 1984).

Assuming the participants in a group in the study or experiment draw on a similar kind of shared cultural resource as other groups performing a similar activity even if their circumstances change (such a change in the technological system), the understanding of the machinery of interaction and situated practice in the former will still be relevant to the latter group. On this note, there has been substantive work in the literature that attempts to use ethnomethodologically-informed concepts developed for co-present groups to social interaction of distributed groups. Prior literature in Computer-Supported Collaborative Work has used sociological concepts grounded in observation of in-person collaborative activity and investigated how those similar accomplishments may be seen in systems with distributed users. For example, dialogical interfaces are examined using Sacks et al.'s model of turn-taking (Button and Sharrock 1995), and computer mediated communication is examined using Goffman's Participation Frameworks (Jirotka et al. 1991). Also, as discussed, ethnomethodologically-informed studies of co-present groups have been used to inform the design of systems that support distributed groups. Examples of where ethnomethodologically-informed methods are successful in implementing technology have mainly involved work situations. There are a number of key seminal pieces in this area such as the work of air traffic controllers (Harper and Hughes 1993), city stockbrokers (Heath et al. 1994), software engineers (Button and Sharrock 1994), designers (Rogers 1993), print shop workers (Bowers et al. 1995), users of menu-driven computer systems (Heath and Luff 1993), multimedia technologies (Heath and Luff 1992), and photocopiers (Suchman

1987). This thread of work continues as Internet-based technologies increasingly permeated work settings, for example: archaeological work (de la Flor et al. 2010), the dentist office (Hindmarsh et al. 2011), auction houses (Heath and Luff 2013), and video-based media spaces (Luff et al. 2011). These studies typically collect video footage of the social setting, and then transcribe short fragments of footage to highlight the details of how interaction is accomplished. For example, de la Flor et al. (2010), examine the details in timing and delivery of an archeologist pointing out and describing a letter on an ancient tablet, as shown in Figure 3.1. The transcription of talk and activity reflect a focus on the emphasis and pauses in the participant's speech, as well as the directionality of the motion of the archaeologist's hand pointing out an ancient letter in an image of a table. Examples of this are shown in Figure 3.1.

R:	That's an 'M' certainly	*((traces 'M'))*
A:	An •'M'	
R:	'EMI' (.) an 'M' ( )	*((traces 'M' again))*
	[	
A:	°That's right°	
R:	And then an 'R'	*((traces 'R'))*
A:	Yup.	
R:	'E' 'S'	*((traces 'E' and 'S'))*

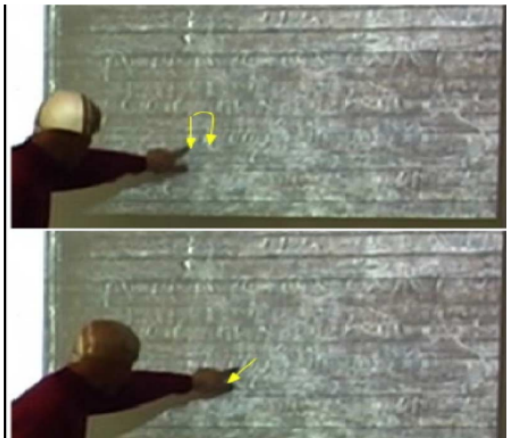


Figure 3.1. Analysis from de la Flor et al. (2010) that examines the details in timing and delivery of an archeologist's assessment of a letter on an ancient tablet

Generalisations and related abstractions from ethnomethdologically-informed studies of this kind are *generative* in that they are intended to give rise to new system requirements, which may then be tested using a prototype. Ethnomethodologically-informed insights from a particular system may be used to generate implications for new systems by: (1) looking for features of *breakdown* (Douglas 1995) which occurs when users express difficulties performing a particular activity such as when participants can't find expected objects or repeatedly select the wrong object in an interface. An ethnomethodologically-informed view will look in the participants' speech and actions to determine moments of breakdown and to understand the source of breakdown. (2) creating a running list of interactional accomplishments which appear to be a central part of the social interaction of a group in one situation, and consider how you might be able to support those kinds of interactional accomplishments in a new situation with a modified technological system that plays some role within the social interactions of the group. This links back to previous discussions on the studies that look at interaction within virtual worlds such as 'Second Life' or 'World of Warcraft' (Section 3.4.1). There, researchers find that many of the interactional difficulties of participants are a mapping from the kinds of interactional issues that co-present participants in social interaction encounter. Therefore, understanding the analogous activity in a co-present group may provide both traction and a framing of the issues a distributed group may experience. By 'framing', I mean furnishing some new concepts around which to consider the design of technology. By attempting to support the same kinds of central accomplishments of a co-present group online, it may be possible that the resulting computer system may be quite intuitive as it draws on the *shared cultural practice* that exists offline resulting in fewer instances of breakdown. The way that central issues in interaction in a co-present

group may map to social interaction in a distributed group, is an important topic for consideration in this thesis, and will be discussed again in Chapter 4. Furthermore, particularly for those users who are new to a particular online environment for social interaction, an understanding of how that kind of interaction may take place offline (and taking the ethnomethodologically-informed view that the participant would understand *shared cultural practice*), could help to inform instruction manuals or training that are intuitive.

The above-stated two points are relevant to the use of an ethnomethodologically-informed approach to studying and developing massive open online systems such as MOOCs. Particularly with regard to my goal of studying the details of interaction around and through MOOCs, it may be particularly useful to apply an ethnomethodologically-informed approach to understand at a first level usability issues, but also to be able to undertake an iterative design development. Point 2 above also opens the possibility of using studies of a co-present group performing some of the kinds of social tasks I may like to support in the MOOC with the distributed group, and to use that to develop requirements for the MOOC. However, the nature of the mapping between interaction in a co-present and distributed groups is underexplored. This, as will be explained further in Chapter 4, is one of the central issues which will be partially unpacked in this thesis.

This ethnomethodologically-informed approach to studying social interaction around and through technology, bears some similarity to emerging work focused on developing new methodologies such as co-realization (Hartswood et al. 2002) and design thinking (Kimbell 2011; Wastell 2011). “Co-realisation” is suggested, as a way to merge an ethnomethodologically-informed approach with a participatory design-

informed approach. The strength of the former is explicating the social character of work and some of its key fine-grained mechanics, the strength of the latter is promoting the value of user expertise for design. Some of the main critiques around design thinking and its evolution echoes this thesis: traditionally, design may have been driven by theories only known to the designer thus putting them in a privileged position. Some new literature suggests that rather it's more important to attend to the lived, situated, and embodied activities of the users. Both of these examples illustrate a shift methodologically in other areas to a more user-centred approach, suggesting the broader validity of the key points made in this thesis.

The work of Lucy Suchman (already mentioned in Sections 3.3.1 and 3.3.2) is an example of pioneering research illustrating how an ethnomethodologically-informed approach may be used to analyse human-computer interaction and which effects computer system design as well as other research areas such as Artificial Intelligence (Suchman 1987). Suchman's studies are *informed* by ethnomethodology rather than adhering strictly to the traditions of ethnomethodology as she used quasi-naturalistic studies. Suchman's video-based study involved pairs of users attempting to make sense of an 'intelligent' photocopier. Suchman used quasi-naturalistic experimentation with Constructive Interaction (Section 3.3.2). In partial alignment with Garfinkel in his definition of 'judgemental dope', Suchman used her studies to illustrate that even purposeful human action is situated in and organised around the particular circumstances. A key aspect of Suchman's analysis was to understand the "organisation of human-machine communication, including its troubles, in terms of constraints posed by asymmetries in the respective situation resources of human and machine" (Suchman 1987 p.126). In addition to raising questions about cognitive models around human conduct, Suchman's work informed a range of studies that

surface the importance of examining the way artefacts are situated in sequences of social interaction (Heath et al. 2010 p.136). Some argue that Suchman is responsible for introducing an ethnomethodologically-informed tradition to the research areas of Human-Computer Interaction (Dourish and Button 1998).

3.5. Discussion

This chapter provides a methodological stance for this thesis, which aligns with my goals to understand details of *how* and *why* interaction unfolds the way it does in the MOOC. The *turn to the social* (Section 3.2) refers to the movement within Human-Computer Interaction that emphasises methodologically taking an empirically-grounded, qualitative, and iterative approach to understanding interaction and designing computer systems. Part of this shift involves incorporating Social Science methodologies, into computer system design. I review some Social Science-based methodological approaches and related issues around examining and understanding interaction in massive open online systems such as MOOCs (Section 3.3). This includes usability testing, the use of quasi-naturalistic experimentation, as well as online surveys with qualitative analysis of responses.

In Section 3.4, particular focus is given to the methodological approach called *ethnomethodology*, where I discuss some of its underlying principles in terms of social interaction as well as the way it may inform studies of interaction and requirements for technology. Ethnomethodology focuses on the details of *how* interaction is accomplished in terms of its observable-reportable features, usually drawing on detailed transcriptions of short fragments of video data capturing interaction. The strength of this methodological stance is that an empirically-based understanding is

surfaced of exactly how interaction unfolds, including what issues and resources shape that unfolding. This gives an insight into foundational issues around why interaction is the way it is, rather than only looking at decontextualised events as part of that interaction, such as clicks or page views, which is the case for 'big data'-based approaches.

There are several limitations of an ethnomethodologically-informed methodological stance. One key limitation is the lack of immediate statistical generalisability of the gleaned insights. However, there is an inherent logical and conceptual generalisability based on the fact that we all participate in society in some kind of an ordered way in our day-to-day lives (Section 3.4.2). Existing literature suggests that understanding this detail in interaction can expand our conceptualisations of interaction and provide important design implications, that 'big data'-based approaches may gloss over. Therefore, this methodological view may serve to complement other approaches, rather than replace them. In Chapter 4, I will take up the methodological view from this chapter, and walk through a particular set of methods in this thesis to progress towards answering my research goals (Chapter 1).

Chapter 4.

Method:

A Three Stage Iterative Study of Social Interaction, and Design of Proof-of-Concept Video Annotation Interface

4.1. Introduction

This chapter will take up the methodological stance from Chapter 3, and lead into a more in-depth discussion of the methods used in this thesis. The goals of this thesis are to understand the details of how interaction unfolds in the MOOC, in accordance with an ethnomethodologically-informed view, and to apply that to the design of a proof-of-concept video annotation interface (Chapters 1 and 2). The shift towards an empirically-focused, iterative approach to understanding and designing for interaction, is part of the wider 'turn to the social' movement described in Chapter 3. In this chapter, I will begin by outlining the broad aspects of my approach, including a three-phase iterative approach and the decisions made around data collection and analysis. I will then proceed to present some of the details of each phase, and reflect on the strengths and limitations of the selected methods.

4.2. Broad Overview and General Issues

This section will cover the broad features of the method used, and some general practical issues around method that will play a role throughout the experimental work in this thesis. At the most broad level, the method I put forward in this thesis consists

of three iterations, following the *iterative design process* for a video annotation interface. The use of an iterative approach is cognisant of the complex and unpredictable ways that prototypes may be used (see Section 3.2). As is a norm in iterative design approaches (see Section 3.2), each new iteration in this thesis serves to examine a new facet around social interaction with each new iteration incorporating some new element of complexity and a new prototype version of a video annotation interface. The prototype in each new iteration incorporates insights from the previous iteration, and addresses a set of research questions that emerge from the previous iteration. An overview of each phase will be given next, with more detail given in Sections 4.3 to 4.5.

Phase 1 - Broad and Narrow Context of Social Interactions in MOOCs

This phase starts at a general level by exploring the contextual issues around interaction in the MOOC. Most existing literature on the MOOC using a 'big data'-based approach gives a sense of the distribution over time of events like clicks or page views, which are the artefacts of interaction, but with minimal insight into the root causes of why interaction unfolded in that way and the interactional context into which those events fit (Chapter 2). This phase uses a pilot survey study followed by a series of quasi-naturalistic experiments with the current MOOC video interface, to explore the various contextual issues that come to bear on the way interaction unfolds.

Phase 2 - The Individual and the Collaborative - Intra-Video Social Interaction of a Co-Present Group

This phase takes up the understanding from phase 1 around interaction viewed in context, to create a first prototype of a video annotation interface: a low-fidelity prototype using pen and paper, and a tablet computer. The focus of this phase narrows

to understanding how individual and collective activities of a co-present group are orchestrated to accomplish intra-video social interaction. This phase centres on a set of quasi-naturalistic experiments, where co-present groups interact with the prototype. Their activity is video recorded and analysed from an ethnomethodologically-informed perspective. This enriches our understanding of intra-video social interaction, and results in new requirements for the proof-of-concept video annotation interface used in phase 3.

Phase 3 - Exploring Interactional Issues at the Root of Intra-Video Social Interaction for Distributed Groups

In this final phase, the implications for a video annotation interface for the MOOC are taken up to create a proof-of-concept video annotation interface. This phase focuses on studying the impact of the video annotation interface on the detailed mechanics of how a distributed group accomplishes intra-video social interaction. The experimental setup involves collecting both video data as well as server log data on the activity of a distributed group in a quasi-naturalistic experiment. By collecting data about the group's activity both remotely and through video collected at the site of each participant, I explore the relative merits of both types of data and create a rich picture of interaction. This enhances our understanding of intra-video social interaction and is used to created a final iteration of requirements for a video annotation interface for the MOOC.

As part of executing these phases, there are a number of practical issues and considerations that shape the experimental method, including the availability of participants, possible venues for quasi-naturalistic experiments, and hardware setup requirements. These issues are discussed in turn, next.

4.2.1. Recruitment of Participants and Ethical Approval

The recruitment and involvement of participants in the studies of this thesis were carefully managed and considered. In aligning with the broad trend that most enrolled students in MOOCs have at least a tertiary level of education (see Section 2.2), I decided to source participants by advertising on the University of Oxford mailing lists. In particular, I had access to two lists: those for the Department of Computer Science and those for the Saïd Business School. I don't consider it a substantial limitation that the participants in my experiments were affiliated with only these two University of Oxford departments. While these participants may have been biased towards framing educational content in Computer Science- or business-related issues as opposed to, for example, Sociological or Anthropological issues, I was primarily interesting in *how* they accomplish interaction rather than the particular nature of their discussion. No effort was made to randomise participants with respect to race or gender (or any other characteristic that traditional Sociological theory states may impact social interaction). In aligning with an ethnomethodologically-informed approach, it is assumed that any significant race or gender issues would be reflected in the video data recording the interactions.

For each experiment, I filtered participants to ensure they didn't know the particulars of each others' prior knowledge and experiences. This reflects partially the fact that enrolled MOOC students generally have access to the general demographic information about the broader enrolled student group, but not the details of the prior experiences and knowledge of individual students (Jordan 2013). Upon meeting, just before the study was carried out, the participants were asked not to talk to any extent about their prior experience or knowledge. In order to ensure that the participants

were as active as possible during the studies, I also filtered participants to those that stated they had at least a moderate interest in the subject matter of the educational content shown to them as part of the quasi-naturalistic experiments, and made sure none of the participants had previously seen the educational content.

4.2.2. Data collection and analysis: Video Data of Interaction and Post-Experiment Interviews

Across all three phases of this thesis, I performed quasi-naturalistic experiments where I video-recorded the activity of groups of participants interacting with different versions of a video annotation interface, in aligning with an ethnomethodologically-informed approach. In each study where a group's activities were video recorded, at least two cameras were used to capture their activity from various angles and minimise the chance of missing an important interaction. After each experiment, a post-experiment interview of a relatively unstructured nature was conducted, to provide participants with opportunities to provide any insights not reflected in the video data. Where these interviews reveal important insights, they will be discussed in Chapters 5 to 7.

For analysis, the video data was broken down into short fragments. These fragments are represented in this thesis using transcripts and video data screenshots, as shown in Figures 4.1 and 4.2. In aligning with the ethical clearance[5] received for this research, the participants' identities are anonymised in the transcripts and video screenshots in

[5] This DPhil has been granted full ethical clearance by the Social Sciences and Humanities Interdivisional Research Ethics Committee, University of Oxford, allocated the code C1 13-025.

this thesis. The transcription approach used draws on the system developed by Atkinson and Heritage (Atkinson and Heritage 1984) and is explained in detail in Appendix 1.

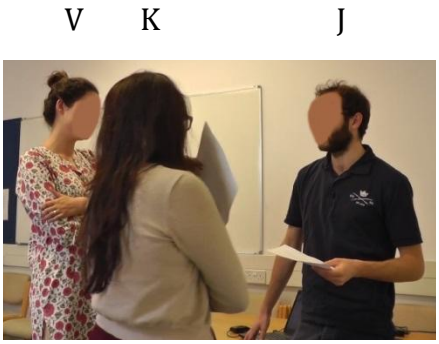


Figure 4.1. An example screenshot of video collected collected in Phase 2 with participants V, K, and .

Sample Fragment

1	K:	Based on patterns that they see they (0.2) could be really wrong (0.5) sometimes
2	J:	=Actually that's a good point
		[
3	V:	Surely a GP a lot of it is (0.1) I mean the reason we don't have just a robo:t (0.2)
4		deciding
		[
5	K:	Ya
6		true

Figure 4.2. Example of fragment transcript from data analysis in this thesis

The transcripts and video data screenshots are written to highlight details in interaction such as, pauses in speech, emphasis in speech, relative timing of speech and gestures of the participants, and shifts in the physical orientation of participants towards each other and various aspects of their environment.

Given the level of detail at which interactions are examined, there is a potentially large number of aspects of interaction that could be analysed. However, given the

constraints of keeping this thesis a reasonable length, a process of selection of phenomena and fragments was undertaken as follows. First, I performed extensive and multiple reviews of the video data, initially noting many aspects of interaction that could be of interest. I then began to iteratively, with further consultation of the video data, refine the list to those aspects of interaction that appeared to be most substantive and fundamental to the accomplishment of the group's activities, and so that various combinations of the related issues and themes weaved a coherent narrative for this thesis. During this process, I also presented my work and refined the transcripts and data sets at data sessions with my research group.

The tasks given to participants during the quasi-naturalistic experiments were developed carefully to be open-ended, to minimise the ways the task might shape the behaviour of the group in a specific direction (as explained in Section 3.3.2). An important balance that has to be maintained is formulating experimental tasks that are open-ended enough to allow for participants to use the systems 'naturally', without making the task too open to be confusing. The tasks are centred around asking participants to use the computer system provided to view educational videos, and consider engaging with the videos and each other. I asked participants to consider contributing liberally.

There was no requirement for participants to accomplish anything in particular during their quasi-naturalistic experiment. The way I drew on the co-present group's activities was to look at interactionally *how* the participants collaborate, for example: the way the participants may accomplish shifts between viewing the video and discussing; the way potentially conflicting ideas are presented and perhaps negotiated by the group; the way participants may appear to come to an understanding of each other's views;

and the way the participants may make reference to the video contents during discussion. My focus is on delineating fundamental aspects of social interaction and aspects of the individual's participation in it, such as the ways that an individual's activities might be coordinated with the activities of others.

Training was also carefully planned at each phase so that participants were comfortable using the computer systems before participating in the quasi-naturalistic study. A challenge around training is ensuring it doesn't take a lot of participants' time, which makes it harder to recruit participants. To manage this, all training was piloted with colleagues from the Department of Computer Science at the University of Oxford, to trade-off the thoroughness of training with time required.

4.2.3. Venues and Hardware Requirements

The venues selected for the quasi-naturalistic experimentation were in the Department of Computer Science and Saïd Business School, at the University of Oxford. They were selected to be large enough so that there was a comfortable amount of room for the participants to move around. Where the participants required pen-and-paper, for example in phase 2, an abundance of these resources were provided so as not to restrict the activity of participants.

It was a challenge methodologically to source the cameras and laptops for all the quasi-naturalistic experiments. For example, as shall be explained later in this chapter, the phase 3 experiment required six cameras, six tripods, and three laptops. These were sourced from friends, from my supervisor, and from the IT support team at the Department of Computer Science. The hardware requirements served to limit the size of the group used in phase 3. All the participants also had to be trained on how to use

the laptops and prototypes in the experiments. Through providing brief training at the beginning of every study, I attempted to ensure participants were as comfortable as possible using the laptops and other computer systems, so that it did not hinder their activity unnecessarily.

4.2.4. Educational Content for Prototypes

For the experiments, I required primarily two types of educational materials to populate the prototypes and systems used by participants: educational videos and forum posts from a live educational platform. These materials were sourced through two key partnerships with existing educational platforms: from the Coursera MOOC 'Internet History, Technology, and Security'[6]. and from the online educational platform, GOTO[7], at the Saïd Business School at the University of Oxford. In order to at least partially control for various kinds of educational videos used in each experimental phase, different videos were used across the phases. The videos were also chosen to be introductory in nature so that they are broad enough to be able to engage a wider scope of experiment participants. In phase 1, I used a video on the introduction to security, in phases 2 and 3 I used different parts of videos on 'security' and 'big data'. Participants were also screened to ensure that they had at least a moderate to high interest in the subject matter of the videos for that experiment.

---

[6] http://www.coursera.org/learn/internet-history/home/info I am grateful to Prof. Chuck Severance for giving me access to their materials.

[7] GOTO (Global Opportunities and Threats: Oxford) is a platform developed at the Saïd Business School for present and past MBA students which aims to present information regarding, and encourage discussion. http://goto.sbs.ox.ac.uk/. I am grateful to the GOTO team for giving me access to their materials.

Having discussed the overarching decisions made around method for this thesis, in the following sections I will discuss in more detail the methods associated with each phase of this thesis's empirical work.

4.3. Phase 1 - Broad and Narrow Context of Social Interactions in MOOCs

As discussed in Chapter 2, 'big data'-based approaches to interaction in the MOOC so far provide a relatively interactionally-decontextualised view of interaction. That is, the existing literature gives us an understanding of the distribution of different events (such as clicks or page views), but not necessarily access to the sequence of interaction through which those events emerged, and which gives rise to those events. This phase is broad, and focuses on attempting to reveal important contextual issues around interaction in the platform; placing events such as clicks, lecture video viewed, and forum views, back into the context of the sequences of interaction into which they fit. In addition to gaining a new view of interaction in the MOOC, I draw out some preliminary qualitative requirements for the proof-of-concept video annotation interface to more seamlessly connect video viewing and social interaction. This phase uses one survey, and two sets of quasi-naturalistic studies, which will be described next.

Stage 1: Survey

In the first part of this phase, I undertook an exploratory pilot study through the administration of an online survey to the enrolled students in the live Coursera MOOC 'Internet History, Technology, and Security'. I gained access to the students of this

MOOC through a collaboration with the course Professor, with whom I worked as a Teaching Assistant for the course over the period May-June 2013.

The survey question is intended to probe at so-called *broad contextual* issues: the larger elements of a person's context such as some main features in their personal or professional lives, which are not likely to change substantially within one session of usage of the platform. Survey questions were piloted with three or four colleagues from the University of Oxford (see Section 3.3.3 for a discussion on piloting surveys). The piloting allowed me to, firstly, check the interpretation of the question by respondents, to try to reduce the likelihood of the question not being understood as intended, and, secondly, test which questions appear to probe at contextual issues around interaction. It is important that the survey question be open-ended and elicit free-text response from students. The reason for this is to attempt to preserve the integrity of the data by minimising the influence of the survey on the participant's response, allowing them to shape the free text how they feel best and use the vocabulary of their choice; all of which is of value to the analysis.

As a result of the piloting process, one compound question, consisting of two parts, was chosen. The piloting process illustrated that responses to this question surfaces important contextual issues, but without explicitly or directly asking about context and thus minimising any bias. The questions asked the MOOC students to reflect on the targets and goals of their participation in the MOOC as well as their interests and understandings. The question is: *Where in your everyday life do you hope to apply what you learn in this course? How do you hope to apply what you have learnt?* I collected free-text response to this question which I then analysed using Thematic

Analysis (Boyatzis 1998) (as explained in Section 3.3.3). The result of the analysis of these responses will be explained in Chapter 5.

Overall, 670 responses were collected. I undertook analysis of this on my own, regularly presenting key themes (with a sample of responses) at data sessions with my research group. Researchers looking to reproduce this kind of analysis at this scale may expect analysis can take around 2 months of focused work.

Stage 2: Usability Testing

Following on from the survey, I focus on gaining a view of the details around how user-platform interaction progresses and the important contextual issues that shape interaction as it unfolds. To do this, I draw on the tradition of *usability testing* (see Section 3.3.1), using a method informed by Constructive Interaction (see Section 3.3.1), and conduct two experiments. Screenshots of the video data collected from the experiments are shown in Figure 4.3. In an attempt to surface the more fundamental aspects of usability and related contextual issues, I look in particular for aspects of interaction that come up in the data from both experiments. On the left, the participants are labelled B and M, with B closer to the camera. On the right, the participants are labelled K and L, with K closer to the camera.

Figure 4.3. Screenshots of video collected during phase 1, with two pairs of participants - with participant B near the camera and M away from the camera (left) and participant K near the camera and L away from the camera (right)

An important part of a quasi-naturalistic experimental usability study, is to not provide a strict task (Section 3.3.2). Therefore, the task given to participants is open-ended: to explore the subject matter of the Coursera MOOC *Internet History, Technology, and Security*; and to consider the extent to which and ways in which the course may be of interest to them.

Each session requires 2 hours: 1 hour with the participants, and an additional 1 hour before and after to set up the venue. The resource requirements for this experiment are: 2 cameras, 2 tripods, 1 desk, and 1 laptop.

By capturing the unfolding of interaction in real-time, I am able to see how those interactional events like clicks or page views, captured by 'big data'-based approaches, fit into the context of the sequence of interaction into which they sit. This surfaces some of the contextual issues and factors that shape and drive forward interaction, which will be discussed further in Chapter 5.

Stage 3: Intra-Video Social Interaction with Video and Social Interaction Separated Out

In this final stage of phase 1, I take a first exploratory look at intra-video social interaction in a co-present group. Importantly, I modify the computer system the participants used in order to free the participants from the current Coursera interface flow, and to allow for seamless access to both educational video content, as well as social interaction. I am interested in how the participants orient towards each other, video content, discussion forums, and other online resources. Therefore, the modified setup consists of each resource being presented on a separate device as shown in Figure 4.4.

S G

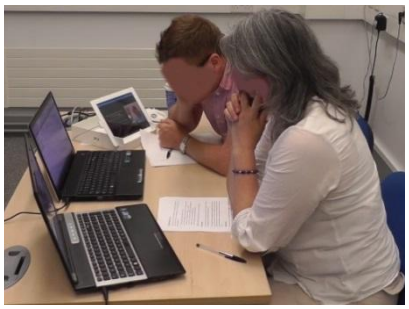


Figure 4.4. Screenshot of video collected during phase 1, with one pair of participants - participant G near the camera and S away from the camera

From the participants' perspective, the left laptop shows a Google search engine page giving them access to any online resources, the middle laptop shows the discussion forums pertinent to the video provided, and the right tablet plays the video (where participants can control the playing/pausing/fast forward/rewinding of the video). The discussion forums are real world ones taken from the Coursera MOOC 'Internet History, Technology, and Security' and thus contain real discussion generated by real enrolled MOOC students. As each resource is associated with a different device, the participants need to undertake substantive physical movements to turn their attention to different resources, thus surfacing in a more noticeable manner the ways the pairs interact with and around the resources. I was interested in how the participants collaboratively view the video and draw on the available resources, using observable-reportable speech and gesture to accomplish this.

This experimental setup points to the ways a video annotation interface may operate since it allows simultaneous and seamless access to viewing a video as well as social interaction (both synchronously within the pair of participants as well as asynchronously with the posts in the discussion forum). This experimental setup therefore breaks up the existing interface flow as presented in Section 2.2. Currently, in

terms of interface flow, the viewing of video is disjoint and separate from any interaction with discussion forums or other online resources.

Although both participants are able to reach and exercise control over all three devices in front of them, I acknowledge that the experimental setup has a limitation as the participants don't have equal access to the devices. For example, it is easier for S to reach the tablet and it was easier for G to reach the left laptop. I remained sensitive to this as I performed data analysis, as will be discussed in Chapter 5.

The participants are asked to perform an open-ended task: to explore the subject matter through a set of provided videos from the Coursera MOOC 'Internet History, Technology, and Security', to consider the ways in which the content is of interest to them, and to consider engaging with the discussion forum or online resources as they view videos. The participants are given freedom to skip, play, pause, fast forward, or rewind the videos as they wish. The participants report having no prior experience with Coursera MOOCs, and said they are at least moderately interested in the subject matter of Internet history, technology, and security.

This session requires 2 hours of time: 1 hour with the participants, and an additional 1 hour before and after to set up the venue. The resource requirements for this experiment are: 2 cameras, 2 tripods, 1 desk, 3 devices that can connect to the internet.

4.4. Phase 2 - The Individual and the Collaborative - Intra-Video Social Interaction of a Co-Present Group

This phase builds on the understanding of interaction from phase 1 by focusing more in-depth on the ways that larger groups of participants accomplish intra-video social interaction. My goal is to examine and delineate some fundamental aspects of the social interaction of the groups, and in particular the constituent features of social interaction and aspects of the individual participants' involvement in it. The data shows some interesting ways that individual and collaborative work in the group comes together as part of the progression of intra-video social interaction. The details of the analysis will be presented in Chapter 6.

This phases uses two quasi-naturalistic studies of co-present groups. Snapshots of the video data from the experiments are shown in Figure 4.5. I chose to use two groups of variable size, to capture dynamics associated with group size, but keeping the group size relatively low to remain within the range of the resources in terms of physical space, sourcing participants, and time to analyse data. I selected one group of 3 and one of 5.

Data snapshots showing the two participant groups in this chapter are given below in Figure 4.5. Both groups watch the same video: an introduction to big data.

Figure 4.5. Screenshots of video data in phase 2 with two groups of participants – V, K, and J (left), and A, B, C, D, and E (right)

For each experiment, at least two cameras are set up in the experiment venue to capture the interaction from different angles, and a wireless microphone was worn by one of the participants to increase the likelihood of capturing all the facets of the speech between the participants.

The prototype constructed for this phase is a paper-and-pen prototype. The features of this prototype will be justified in-depth in Chapter 6. The requirements for the prototype are that it provide the group with a way to simultaneously have access to a video that they could play/pause/fast forward/rewind as they wish, and that there is a clear way to associate an annotation to a moment in the video. The prototype consists of a tablet mounted onto a white board as shown in Figure 4.6. The participants could associate an annotation with a moment in the video by placing on the board a magnetic strip starting below the moment in the video of interest as indicated in the horizontal scroll bar displayed on the tablet. At the end of the magnetic strip, the participants could place a paper annotation, held to the white board by the magnetic strip.

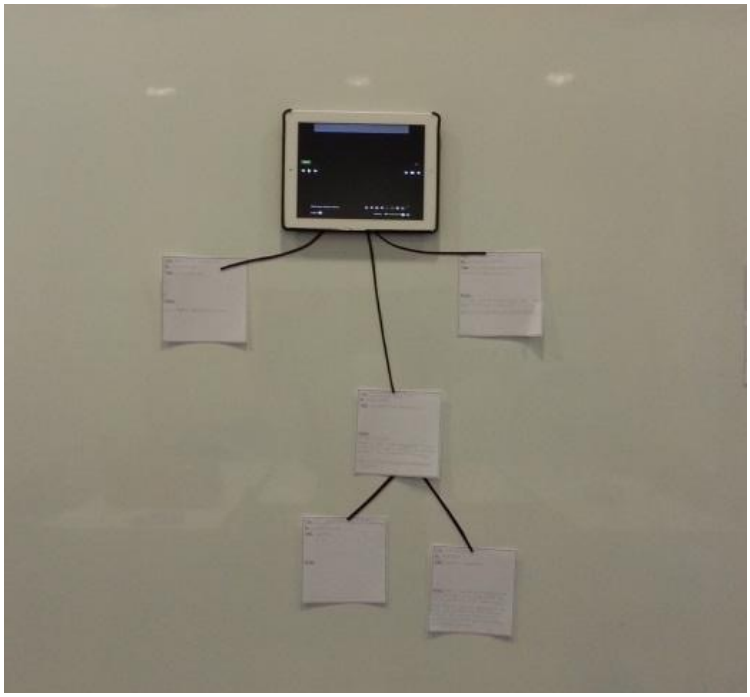


Figure 4.6. Low-fidelity prototype mounted on whiteboard for phase 2

The video used is provided by the team at the GOTO learning platform at Saïd Business School[8]. In order to try to make the prototype seem as realistic as possible, it is pre-populated with video annotations to surface the way participants orient towards them. The pre-populated annotations are taken from an educational platform (from the forums within GOTO), and annotated to the appropriate times in the video. At the same time, a laptop with a Google search page is provided on a table by the white board to allow for any searching or viewing of content online.

As explained in Section 3.3.2, tasks for quasi-naturalistic experiments should try not to shape the behaviour of the group in a specific direction. Therefore, the task set for the each group is open-ended: to collaboratively explore the topic of big data by watching a video, to consider how and to what extent the video is interesting to them, and to consider documenting this using annotations.

[8] www.goto.sbs.ox.ac.uk

Researchers looking to reproduce this kind of study need to arrange a space available for roughly 2 hours. The session with the participants requires around 1 hour, with another hour for preparation.

In addition to providing new conceptual insights into the details of how interaction is accomplished, these experiments give rise to a further set of requirements for a proof-of-concepts video annotation interface. These requirements take into account an understanding of the mechanics of how individual and collaborative interaction takes place in the MOOC. The insights from this chapter are used to create system requirements as well as to generate further guiding questions for the final iteration in Chapter 7.

The use of an analysis of the interaction of a co-present group to create implications for an online system supporting a distributed group, is a central issue for exploration in this thesis (presented in the research goals in Chapter 1). One way to gain traction in the investigation into the fine-grained aspects of the accomplishment of interaction online is to begin the study with an examination of the ways that co-present participants accomplish a similar kind of social interaction, and to generalise the interactional accomplishments of the co-present group to create implications for a system that supports similar interactional accomplishments for a distributed group (Section 3.4.2). A more detailed discussion of this will be presented in Chapter 6.

4.5. Phase 3 - Exploring Interactional Issues at the Root of Intra-Video Social Interaction for Distributed Groups

This final phase puts into practice a proof-of-concept video annotation interface whose features are informed by the studies in phases 1 and 2. I use a quasi-naturalistic

experiment to study how a distributed group of participants accomplishes intra-video social interaction. My goal is to delineate fundamental aspects of social interaction, including aspects of individuals' participation in social interaction and their seeming reasoning as they interact with the system.

An important aspect of the method in this phase is the simultaneous collection of data both remotely (as a traditional 'big data'-type study would do) as well as local reasoning by participants around the platform. Local reasoning was collected through video data that captures the conversation of the distributed participants across different sites, as shown in Figure 4.7. In aligning with the Constructive Interaction methodology[9], at each site, there is a pair of participants. I use three pairs of participants, as this is a manageable group size given the resources available for this thesis. Each pair uses the system as one user - each of the pairs was assigned a neutral and random system user name, one of: Participant A, B, or C.

Figure 4.7. Video data screenshots from phase 3 with three pairs of participants – A and S (left), C and J (centre) and B and M (right)

[9] I also ran this experiment with a version using individuals in a think-aloud setup. This experiment is considered relatively unsuccessful and therefore the resulting data is not presented in this thesis. A primary difficulty in running the experiment was limited resources, which meant I was not able to arrange a researcher to sit locally with each individual to prompt them to 'think-aloud'. Therefore, the participants often under-explained their actions and reasoning, leading to minimal insight being drawn from the experimental data.

The features of the proof-of-concepts interface will be explained in detail in Chapter 7, as they depend on the insights gleaned from phases 1 and 2 of this thesis. To give the reader a sense of the interface, one screenshot is shown below in Figure 4.8. The interface allows simultaneous access to video and social interaction through discussion threads.

Figure 4.8. Screen capture of prototype used in phase 3

The three pairs of participants use the system simultaneously - by this it is meant they are all asked, starting at the same time, to view the same video and, if they like, interact with each others' contributions as well as provide their own contributions.

In order to be able to compare how interaction unfolds for the proof-of-concept video annotation interface and for the current MOOC interface, this study is designed to have two parts:

In the first part of the study, the participants are asked to, at the same time, engage with the same video and associated contributions using the prototype. In the prototype, they view an introductory video to big data provided by the GOTO learning platform from the Saïd Business School. The interface is pre-populated with

contributions from real-world students on the GOTO platform to simulate other users having used the system previously at a different time.

In the second part of the study, the participants switch to, at the same time, engaging with a video and associated discussion forums in Coursera. The participants are asked to view an introductory video to data security from the Coursera MOOC *Internet, History, Technology, and Security*.

The purpose of running the experiment with both the prototype and Coursera interfaces is to elucidate some differences in social interaction supported by the Coursera platform and by the prototype, to consider the extent to which the prototype supports the kinds of social interaction it was designed for, and the ways the prototype features support them. The purpose of this comparison is not to make a definitive statement about which interface is 'better', but rather contrast the kinds of interactional accomplishments seen for each system.

As I am interested in capturing the expressed reasoning of the participants, a key element of this phase is to record both the data about their interaction that could conceivably be collected remotely online (such as their 'clicks' or the comments they submit online) as well as their reasoning as their interaction unfolds and emerges. In order to capture fully the activity of the participants, the following setup was performed:

- There were two cameras in each room – one from the rear and one from the side, in order to capture as many interactions as possible including the laptop from the participants' perspective. This captured the so-called 'local reasoning' by each pair of participants.

- A screen capture, with audio, from each device was made, for two purposes: first to capture the exact actions the participants performed on the interface, and second because the laptops had in-built microphones and so in addition to the cameras this captured the speech of the participants.

Methodologically, this phase requires six cameras with tripods, and three laptops. The experiment requires the simultaneous use of three rooms (one room for each pair), so that the pairs are isolated from each other physically and can only interact with each other through the platform. Each room requires an Internet connection and the rooms need to be close enough for me to move between them in case of any extraordinary system issues needing debugging. Therefore, in terms of method, this experimental setup requires a high number of resources which limit the number of experiments run to one, and the number of participants to six. More details about the methodological challenges of running this experiment will be described in Chapters 7 and 8.

Around 2.5 hours are required in total to run this type of experiment. In my experience, I needed around 1 hour to set up the equipment, and then 1.5 hours to train participants, allocate and seat them in the three rooms, and have participants perform the task.

The screen captures, video data, and remote data together give me a rich view of the fine-grained aspects of how interaction unfolded through the prototype and through Coursera. As will be presented in Chapter 7, I was able to compare the insights from each, to explore the extent to which it is possible to support the kinds of social interactional accomplishments seen in the co-present group in the distributed group. I am also then able to consider any mappings between the practice seen in the co-present groups and those in the distributed group, providing some insights into the

ways social interaction may map or be transformed, and what some of the effects of this may be on social interaction.

4.6. Discussion

In this chapter, I draw on the goals of the thesis presented in Chapter 2 as well as the discussion of methodologies in Chapter 3, to explain and justify the methods used. In Section 4.2, I present an overview of the main, broad parts of the methodological approach including the selection of a three-phase iterative design process which culminates in the study of the details of intra-video social interaction by a distributed group using a proof-of-concept video annotation interface. In Sections 4.3, 4.4, and 4.5, I present more particular information about each of the three phases respectively, with details about each prototype left to be explained in Chapters 5, 6, and 7 as they are related to the insights gleaned from each study.

The experimental setups used in this thesis are only one variant of a number of different methods that could be used, and therefore have limitations. In this thesis, I chose to begin with a broad exploration of contextual issues around interaction, before focusing on intra-video social interaction of co-present and then distributed groups. Any one of these phases could have been selected as the sole focus in this thesis, however performing all three phases is important in this thesis in order to meet the substantive design goal of completing an iterative design process to create a proof-of-concept video annotation interface, as stated in the research goals in Chapter 1. Furthermore, the process of selecting the most 'significant' aspects of interaction to analyse (as described in Section 4.2.2) is relative to my view on the data, and it is possible that a different researcher would have picked up on different interactional

aspects. Therefore, the insights gleaned in this thesis do not necessarily cover all interesting or important aspects of interaction. Finally, several decisions about method, including the sourcing of participants and size of the participant groups for quasi-naturalistic experimentation, were made on the basis of practical issues around sourcing venues and hardware (camera equipment and laptops). However, as the data analysis in Chapters 5, 6, and 7 will show, these experiments provide important analytic insights. In Chapter 8, I will draw together together these insights to meet the substantive, conceptual, and methodological goals stated in Chapter 1.

Chapter 5.

Phase 1 -

Broad and Narrow Context of Social Interactions in MOOCs

5.1. Introduction

In this chapter, I will present the results of analysis for the first phase of the iterative approach towards developing a proof-of-concept video annotation interface. I draw on the findings of this thesis so far: in particular that most of our current understanding of social interaction and video viewing in MOOCs is rooted in data collected remotely, like server log data. This data gives us a largely interactionally decontextualised view of interaction.

I start with an exploratory pilot study that explores themes of context within free-text survey responses from students in a live MOOC. What emerges is a wider view of context, looking at the 'broad context' into which interactions in the MOOC are embedded. 'Broad context' is defined here to mean the larger elements of a person's context such as some main features in their personal or professional lives, which are not likely to change substantially within one session of usage of the platform. I use survey and qualitative thematic analysis to explore broad contextual issues and surface a set of themes and issues. The survey elicits free text responses of enrolled students

in the live Coursera MOOC called 'Internet, History, Technology, and Security'[10]. As I'll explain, many of these issues may drive the way students use the interface.

Over a set of subsequent studies, I will drill down into the data moving towards an ethnomethodologically-informed approach that most directly deals with my research questions (Section 1.2). I will then turn to a more granular view of interaction, focusing on 'narrow context' and intra-video social interaction in the current MOOC platform. 'Narrow context' is defined here generally as: the nature of the activity as it occurs in the social and material surroundings of the user within a time frame close to before, after, or during the activity. I will use Usability Testing to understand some important details of the mechanics of interaction with video on the platform.

Finally, continuing my focus on narrow context, I will take the first step in exploring intra-video social interaction, using a quasi-naturalistic experiment with a prototype system that liberates participants interactionally from the interface flow of the current MOOC platform. The existing interface flow separates interaction with video content from social interaction (Section 2.5). In this prototype, participants have simultaneous access to videos, discussion forums, and each other, thus seamlessly integrating video viewing and social interaction. This provides insights into how participants ordinarily accomplish intra-video social interaction. The insights gleaned in this chapter inform a set of issues and implications for the design of a prototype to be used in the next phase in Chapter 6.

The study of interaction in context brings to light a number of interactional issues and characteristics. It reveals interaction as quite complex and delicate, relying upon a sequence of contingent, emergent, and situated observable-reportable activities.

[10] http://www.coursera.org/learn/internet-history

Issues that will be discussed next include: those associated with delicate timing of activities, the capacity for participants to be a part of different activities at once, and the role of public availability of activity.

5.2. Survey - A Pilot Study Exploration of Broad Contextual Factors Potentially Impacting Interaction[11]

Several studies illustrate that enrolled MOOC students participate in different ways. Existing research has most commonly used large quantities of quantitative data on MOOC students such as the number of video views, number of forum views, number of forum posts, time spent logged into the MOOC, number of videos at least partially viewed (Section 2.3). This data is often used to create categories of participation (e.g. (Kizilcec et al. 2013)). The lens with which we derive student groups is usually in terms of these artefacts of interaction.

In this study, I take an alternative approach, viewing and categorising student participation and interaction qualitatively in terms of the broad contextual factors that shape them: the students' motivations and goals, and students' reported areas of application of the course content. I use qualitative data in the form of free-text responses to an open-ended survey question administered to enrolled students in a live Coursera MOOC. From these responses, some themes emerge which may constitute a set of qualitative categories of participation rooted in broad contextual issues. As I will show, these themes are useful even if they provide only a snapshot of

[11] An alternative version of an analysis of the data in this section may be found at: Zawilska, A., M. Jirotka, and M. Hartswood. "Surfacing Collective Intelligence with Implications for Interface Design in Massive Open Online Courses." Social Collective Intelligence. Springer International Publishing, 2014. 187-202.

issues at one time, and don't reveal how and why they may dynamically change over time.

Since I was the single coder in analysing the data, it is possible there are biases in the categories of themes surfaced. Validation of themes was performed partially by discussing a random sample of responses from each category at research group sessions so that at least, at a broad level, the validity of the main categories could be challenged. However, the goal here is not to come up with a set of categories that are broadly valid, since that is likely to change over time and between groups (this is only one single snapshot). Rather, it's to show the diversity in the various expressed goals of the students.

This study has some similarity to existing work that examines student's motivations in taking MOOCs (e.g. (Belanger and Thornton 2013; Wilkowski et al. 2014; Kizilcec 2015)). However, a key difference is the method with which these motivations are revealed. Many existing survey studies take a quantitative-based approach including using predefined categories of motivations. They take these quantitative motivation metrics and may build models to use them to predict course completion.

As opposed to using a pre-defined set of categories of broad contextual issues, this study allows categories to surface more organically through the survey responses, by keeping the survey questions open and eliciting free-text responses. This arrangement minimises the impact of the survey question on responses, forcing the respondents to independently select the particular terminology and subject of their responses.

The compound question administered to the students is: *Where in your everyday life do you hope to apply what you learn in this course? How do you hope to apply what you have learnt?* This question was asked to enrolled students in the Coursera MOOC

Internet History, Technology, and Security, in June 2013. The response rate (rr) to the survey question is 19%, calculated as shown in equation 1.

$$rr = no.\,unique\ submissions \,/\, no.\,unique\ students\ who\ viewed\ survey\ invitation \quad (1)$$
$$= 670/3596$$
$$= 19\%$$

This response rate is reasonable when compared to documented response rates from other online surveys. Email-only surveys are reported to have a rate of around 20% (Kaplowitz et al. 2004). Web surveys have had reported responses of as low as 2% (Petchenik and Watermolen 2011).

Through thematic analysis (Section 3.3.3), themes capturing a set of different goals of participation are surfaced, as summarised in Table 5.1.

Table 5.1. Summary of surfaced themes

Category: Qualitative theme	Description of students within this category
General interest	The student as the curious explorer, looking to expand current knowledge with non-professional interests in mind
Current profession	The student as seeker of knowledge for current professional work
Future profession	The student as seeker of knowledge for future professional path
Focus on a sub-topic	The student as seeker of information from a sub-topic of the course, in this case online privacy/security
Everyday usage	The student as seeker of knowledge relating to everyday use
Historical trends	The student as the curious seeker of knowledge relating to historical trends

Each theme surfaced from the data will now be discussed in turn using illustrative snippets of survey responses. Each group of snippets were selected to be a representative group, judged qualitatively by me. Each snippet is coded to a particular respondent by a numerical label placed after the quote. Each theme will be reviewed before the overall implications for this thesis are discussed. Many of the individual student responses are found to fit into several different themes.

Theme 1. General Interest - the student as a curious explorer, looking to expand current knowledge with non-professional interests in mind.

Number of responses within this theme: 202

> *"At 77 years of age, I'm just expanding my knowledge of the electronic world and my fascination with it." - 10*
>
> *"I'm glad to have an opportunity to learn more about this history, but I don't see any particular application. Maybe some of the interviews will inspire me to pursue bright ideas. I think that, in general, the study of history is apt to be a useful pursuit, even when the specific applications to current affairs cannot be specified." - 512*
>
> *"Mainly I will use it like a general knowledge. It can be useful during interactions with my friends and [colleagues], especially with those working on computer sciences as I have a lot of them." - 333*
>
> *"I hope to use it with my sons, who have computer degrees, in order to contribute to conversations with them." - 61*
>
> *"So much of my life involves internet technology that I am sure it will be consistently useful in work, home and education. I want to be able to explain this to my children before they get their first devices, and I want to be able to help explain security and other user issues to my not-so-tech-savvy parents." - 24*

Responses within this theme point to students who participate in the MOOC without a professional goal or sometimes without any readily definable focus at all beyond 'general interest'. These students appear to be motivated by how their learning can contribute broadly to their general knowledge, or by a non-focally motivated interest in the subject area. This does not mean such participation is necessarily purposeless, for

instance, some students report taking the course to gain tools for communicating knowledge about the Internet to family members and friends. In this case, students may be seen to convert course content into social capital by acquiring forms of knowledge valued in their local social networks. MOOCs are thus revealed to be part of a broader learning network that goes beyond a dyadic relationship with the enrolled student. This illustrates the rich ways external social situations may come to bear on how a person relates to content. It highlights the importance of the course content as a resource to be leveraged and used within social settings. It speaks to the importance of bringing in social interaction into the learning of the content. In practice, the course participants might like to practice talking about the content with others.

Theme 2. Current profession – the student as seeker of knowledge for current professional work.

Number of responses within this theme: 147

> *"In my work with developing countries - I want to help them understand the origins of the technology and how they can use it to their advantage" - 22*
>
> *"Some of my staff are responsible for our company's website, but I have no IT background at all. I am hoping this course will give me a better understanding of today's Internet and some of the issues our firm needs to prepare[d] to deal with." - 662*
>
> *"I am doing research on hacktivist culture from a sociological, psychological perspective in my master degree at the moment. Furthermore I will use it for my PhD." - 660*

Responses within this theme point to students who wish to apply the course content to their current professional contexts. One profession that appears very regularly is that of teaching. The data show two primary ways in which teachers respond that they may benefit from taking the course. Two illustrative quotes are shown next, each from a different participant.

"I teach a course about cyberpsychology at the university I work in, and [this] is precious information for contextualising that material." - 23

"I teach Vocational English lessons at an ICT high school. I am an English Language Teacher who is interested in technology. So, I think, by learning a lot in this course, I can impress my pupils and set an example for them to dig deeper for their learning." - 79

In the first quote above, there is evidence for teachers wishing to use the MOOC knowledge to enhance the courses they teach. In the second quote, the teacher reports hoping to use their enrolment and participation in the MOOC as evidence to their students of the value of education. The value of the MOOC as a resource for educators has, since this study, been picked up as a pertinent direction for development in new research (Hicks 2015) although this topic is not a focus in this thesis since this deals with issues around professional competencies and skills of educators, which is not a group of users I'm focused on.

In practice, therefore educators may have general genres of information they search for as they navigate through the course. Also, a current professional setting is something that participants in that setting know best about, so teachers may benefit from socially interacting with others

Theme 3. Future profession – the student as seeker of knowledge for future professional path.

Number of responses within this theme: 132

"I hope to apply [what I learn] from this course to my work with my own internet startup" - 49

"I plan to [self-learn] AP Computer Science and I hope to gain insight on this rather interesting aspect of a subject I plan to delve deeper into next year. I hope to keep this course as general knowledge and to maybe, one day, be able to contribute or start a conversation in my academic future with the unique material in this course." - 501

"I hope to use this information in giving me a new direction in my career in the military. So far the lectures have been helpful in my understanding of the INTERNET and how it has become what it is now and how it will grow." - 398

"I plan to get back into the programming field after taking 20 years off to raise my family (and a few other things). What better place to start than with the history of what happened since I've been out of touch." - 301

The responses grouped under this theme illustrate that the course content may be contrasted by students to practices in particular industries with attention paid to how the content may illuminate areas in these industries for future professional opportunities. Because these opportunities are not conceived separately from each student's own personal requirements, what counts as an opportunity will depend on the personal history and interests of the student, and what the student predicts will hold value for them in the future. Although there is a strong sense of individuality present here, I am struck by how valuable it may be for students to share their understandings of how content may relate to future professional opportunities: it may broaden others' appreciation for what opportunities there are or pique their interest in a previously unconsidered career path. The role of social interaction could be to build on the course content; to share stories of how this knowledge has been put into practice in different professional settings.

Theme 4. Focus on a sub-topic – the student as seeker of information from a sub-topic of the course, in this case online privacy/security.

Number of responses within this theme: 87

"I hope to use what I have learnt in my classes about computing Internet and privacy as [an] introduction. Most of the time I stress about privacy and security of Internet" - 270

"I hope to apply what I learn from this course in my family's business, which involves home security and networking." - 43

"When I check into a hotel, I want to no longer worry about the security of the wifi because I will know exactly how it works." - 450

"Better understand the current privacy/terrorism prevention/wikileaks and snowden turmoils" - 212

The responses here are characterised by a focus by the student on the part of the course concerned with online security and privacy. The data provide evidence for the student reasoning through these topics with regard to personal security practices and available software. To a lesser degree, responses also indicate an interest in understanding aspects of privacy/security online in order to understand related news stories popular at the time of the survey such as Wikileaks (Fildes 2010) and Edward Snowden (BBC 2013). The course content is framed in wider discourse and controversy which may be seen as involvement in processes of norm formation and the linking of the academic and the public realms by the student. If students could share their insights through social interaction on the MOOC, this may help others understand and value personal security practices as well as assist the collective in scoping and collecting perspectives on current issues important to the public.

Theme 5. Everyday usage – the participant as seeker of knowledge relating to everyday use.

Number of responses within this theme: 166

"Given that computers are and will continue to be a part of our everyday life, this course is giving me the historical background to understand the people and events that effect how I use computers today. Much like taking an American History course to understand how our country came to be, this course helps me understand how my computer usage came to be. Now, when I hear terms, I am not as bewildered as I was before taking this course. It is essential to understanding why I exist in this computer tech world." - 181

"I hope I can learn more about who is behind things like the browser I use, the programing of applications I games I use." - 370

The responses in this theme are characterised by students expressing that the course content has inherent value because of the everyday necessity of Internet technologies in modern society and the widespread use of particular technologies such as the Internet browser. In these responses, students appear to link the practical everyday software with wider patterns of meaning and an understanding of the forces that shape the technological landscape. It is the narrative behind a particular piece of technology which may assist the student to make sense of their participation in it. This may be linked to broader spiritual or political discourse. Shared perspectives of the everyday value of technology may inform others about the value of particular software, and may spark counter debates or discussions, and it may spark conversation about the value of software, its primary purpose, and shortfalls.

Theme 6. Historical trends – the student as the curious seeker of knowledge relating to historical trends.

Number of responses within this theme: 79

> *"I believe that history goes in circles and sometimes in order to answer complicated questions of future we may need to look back into past. There is a chance that such a problem or similar was already solved." -69*
>
> *"I hope to use the knowledge of the course to the analysis of technology and its impact in society. In history you can learn tendencies. I'm part of the Free Software Movement, and we're trying to construct the concept of popular technology in my city." - 60*
>
> *"One benefit of the course is sharpening ability to teach others about the power of independent, innovative thinking. Another benefit, based in particular on Week Four lectures, is practical and emphatic reinforcement of personal values and principles. Noble civic motivations do not always spring to mind in analysing the march of technology. Yet, a frequent theme embedded in much of the course dialogue so far is just that. How to make society better[?] The applications of this concept in daily life are endless." - 46*

The responses in this theme point to students who express interest in what course content may indicate about the historical narratives of the development of Internet technology and what it may indicate about future trends both in technology but also more broadly in society. For example, the third quote illustrates a student using the course content to reinforce narratives with a moral/ethical dimension ("personal values and principles" - 46) in technological development. Students sharing these perspectives may assist others in contextualising current events as the effect of technology on society increases, and create semantic arcs that may speak to discussion about the future of technology and society. The idea that students may prefer a different semantic ordering and arc of the course content than established by the course lecturer is mentioned in the discussion of the everyday usage theme, but is perhaps more clear for this theme since the data speak more directly to the issue of narratives. In this case, the student may be interested in particular changes over time and wish to focus on the characteristics of the transformation; for example, the circumstances around a particular change; the drivers, and the results. The students may be interested in drawing out several arcs of history, which is done better by humans than by machines. The participants may be less targeted, as reflected in the General Interest and Everyday Usage themes and therefore their particular interests may emerge with the course content. They may come across it serendipitously.

The themes surfaced from the survey responses cut a cross-section across student participation and interaction that is different to most previous studies of participation. Previously, participation and interaction at a broad scale has been given a primarily 'big data' treatment - by examining patterns in the artefacts of interaction over time, and finding groups within that, for example using Machine Learning (Section 2.4). In

contrast, these categories reflect the broad contextual issues that may come to bear on how students participate and interact in the MOOC.

These themes together highlight three key issues around social interaction in MOOCs:

Firstly, this suggests we can add a level of richness to events like a 'click' on a webpage button, or the view of a video on the MOOC. In some literature, these events tend to be aggregated. Viewed through the lens of the revealed themes here, an interactional event like a webpage button 'click' could have very different meanings because it could be embedded within different broader contexts. For example, the act of stopping viewing a video (click of a 'stop'/'close' button on a video frame), may occur because a student doesn't see the applicability to their real-world working environment (related to the 'current profession' theme) or it may be because they don't see the practical value of the video for protecting themselves online (related to the 'everyday usage' theme). And yet, 'big data'-based approaches to interaction tend to group all webpage button 'clicks' together, thus abstracting away these differences.

Secondly, these themes shed light on how one version of a piece of educational content cannot completely serve all the different ways that students might want to put that content into practice. For example, a student may see a video on 'data security on the Internet' as valuable because of their broad interest in the design of computer system, while another student may be interested in the historical and sociological issues that have made data security a pertinent issue. These two students, though viewing the same video content, may have different ways they'd like to build on and apply that knowledge. Many of the goals appear to necessitate some additional human input and social interaction around the content. For example, if a student may like to know how some piece of course content relates to the everyday practices associated

with a profession, this understanding is best provided by interacting with a person who has experience with the profession rather than by a machine. This suggests how valuable it may be to develop the MOOC video interface so that students can contribute to the video content based on their own experiences. This provides further empirical support to the substantive goal of this thesis to develop a proof-of-concept video annotation interface.

This issue also has relevance to the debate around measuring student 'success' in MOOCs (Section 2.2). The themes here suggest that completion rate is not an adequate indicator of success, if success is defined to be the extent to which a course helps a student accomplish their goals. Only portions of a course may be relevant with respect to how a participant wants to put the educational content into practice.

The broad contextual issues here open up new questions around the unfolding of interaction in the MOOC with respect to narrow context. As a result of the varied broad contextual issues at play, the details of interaction may be different between students and the reasons *why* may be different. For instance, depending on how targeted their goals are, different students may be inclined to be more or less active in creating and leading conversations in forums. A student who has a very particular question regarding the everyday usage of an Internet technology like the browser may be more likely to instigate a discussion on this topic than a student who is driven more by a general interest who may not have particular questions in mind but may rather just wish to explore the content and discussions others are having. The details of how interaction unfolds in these two cases may be different – one participant may be eager to elicit particular responses to a question, the other may undertake more viewing of others' discussion with occasional questions which emerge gradually. As another

example, students who are attempting to learn how to use the content as social capital may wish to view others' discussions in order to get to grips with pertinent vocabulary or a sense of current topics of interest and relevance, than instigate a new conversation themselves immediately. As another example, if a student pauses a particular video at a certain moment in time, these themes open up the question – why? It may be because the video is not of interest as it relates to their broad context, either because they already know it or because it's not of relevance, they may be considering remaining within the course but skipping ahead in the course, or they may be attempting to create a new discussion based on something in the video which is very relevant to their broad context. In this way, unpacking some of the implications of these broad contextual themes, bring us back to some of the research questions at the heart of this thesis, directs us towards studying interaction in its narrow context, and invites us to explore interaction further through an ethnomethodologically-informed lens.

In the next section, I will take a closer look at the narrow context into which interaction in the MOOC in embedded, taking a deeper dive into understand the fine-grained details of how interaction with video unfolds in particular.

5.3. Usability Testing - More Than Just 'Viewing a Video'

Although video is by far the most popular feature of the MOOC, our current understanding of what drives student-video interaction is relatively simplistic (Section 2.4).

Usually, existing studies classify video-viewing in terms of a binary categorisation of either 'having been viewed' or not (e.g. (Kizilcec et al. 2013; Weller 2014)). A video is

classified as having been 'viewed' if the student at least partially played the video (even if they did not watch the entire video). Other interactions with video picked up in 'big data'-based studies include events, or artefacts, of interaction such as, clicks on the play/pause button of the video. By examining instances of interaction that reveal the narrow context into which these events may be embedded, I will show that these events are not all equal in nature but may be driven by dynamic and emergent interactional issues. The data shed light onto two substantial facets around interaction that emerge as substantial in both usability studies:

(1) If a participant has started watching the video, they don't unproblematically and unquestionably view all MOOC videos or, once they've started watching a MOOC video, don't necessarily complete watching it. Rather, they continually appear to assess the value of the content, considering whether to fast forward or stop watching.

(2) Students do not passively 'consume' the video data, but appear to *actively* attempt to create semantic links between the content and their prior history and knowledge, resulting in various framings of the content by different students. This serves to highlight the level of activity that a student undertakes while viewing a video, which is not captured in 'big data'-based approaches.

An example is illustrated in Fragment 5.1, shown in Figure 5.1. This fragment reflects an aspect of student-video interaction that surfaces as central across the pair of usability studies conducted (see Chapter 4 for rationale about how fragments are chosen).

Just before Fragment 5.1, the participants, B and M, have paused the video. In Line 1, B expresses a wish to fast forward the video ("What I would li::ke to do now is…fast fo:::rward", Line 1) but with a condition ("where you can", Line 1) which she appears to

elaborate in Line 3 (“where you can see”): expressing a wish to “see” (Line 3) ahead in the video.

Fragment 5.1

1	B:	What I would li::ke to do now is to have a fast fo:::rward ((laugh)) where you can
2	M:	((laugh)) But you can see (0.1) just can::t
		[
3	B:	Ya but where you can see (0.1) what is (0.1) what he (0.1) because he
4		might say some more impo::rtant things

Figure 5.1. Fragment 5.1 transcript

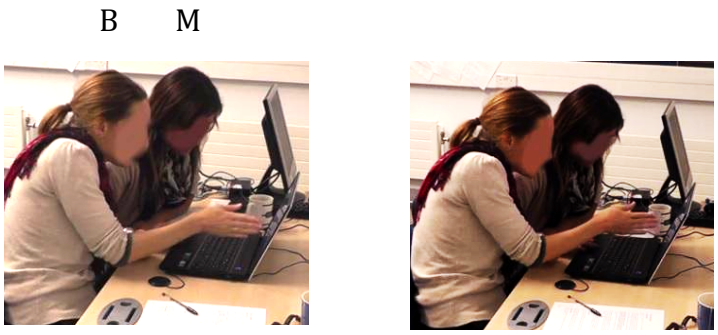


Figure 5.2. Illustrative data snapshots showing B’s hand movement during Line 3 of transcript in Figure 5.1

During her speech in Line 3 (“(0.1) what is (0.1) what he”), B moves her right hand horizontally left and right quickly near to the scroll bar of the video shown on the laptop screen (illustrated in Figure 5.2). This appears to draw M’s attention (indicated by the direction of M’s gaze) along the video scroll bar, who then clicks ahead on the scroll bar while talking in Line 2 (“you can see…just can::t”). B expresses an interest to get some information about what the rest of the video holds without necessarily needing to view the rest of the video. M responds to this with a statement that the interface doesn’t support this, displaying an understanding of B’s expressed interest.

In talk by B after the fragment, she reveals that she has noticed the remaining time in the video scroll bar, and would like to know more about what the video holds without necessarily having to invest time in watching the rest of the video. And so, B is implicitly and actively applying a kind of criterion to her video-viewing; in this case time constraints.

Data from the other participant pair, L and K, display similar instances of activity, at different stages of video-viewing. There are fragments of activity that illustrate how they proceed with the selection of a video from a list of video titles shown in the interface. The data suggest they make relevant a set of emergent criteria when making value judgements, including their expressed mood or feeling at the time. In one case, participants L and K chose a particular video from a list of video titles because they expressed being in the mood for something "fun", and perceived a video title 'Fun times in Amsterdam office hours' to be suitable to this mood. At no point in the experiment do the participants select a video title without expressing some criteria and value judgement to the titles.

The data echo a concept which underpins various traditional approaches to studying information foraging: *information scent*. This is defined as the cues users perceive in their local environment (proximal cues), while exploring and searching for information, to judge distal sources of information and to navigate towards them (Carroll 2003 p. 165). As a concept, information scent also has wider relevance to the Internet as web pages may contain a lot of information and hyperlinks to other web pages and information. For example, in a typical webpage, hyperlinks in the form of underlined link text would be a form of proximal information. Any theories around information scent are not of primary interest in this thesis because I am interested in shedding light

onto the empirical details of how interaction is seen to be unfolding in the MOOC (Section 1.1) which most studies on information scent deal with using cognitive models and theories (e.g. (Ponceleon and Dieberger 2011)). However, these fragments give us some insight into the practice of discerning distal sources of information around video using immediate information. The data here suggest that the process of following a kind of information scent, which is usually viewed as a cognitive process, may be made visible through observation of the practices and interaction within this social setting (Lynch 2006).

Part of this 'information scent' is judging the content in front of you (proximal cues) - and the data show this is not objective. What we see in the data is that seemingly value judgements on the content are made commonly by linking the content to prior experience and knowledge. For example, in one scene, participant B pauses the video and connects a piece of content on the speed of Internet to high frequency trading, stating that this is a topic in which she has previously been interested. This contrasts to M's framing of the video content, who expresses that she does not see the value of the video content in the same way. The participants' subsequent discussion on high frequency trading uses the video and B's prior experience as a seed for conversation. In another instance from the participants L and K, they are seen to select particular videos and show interest in content because of more personal reasons. For example, in one instance, L states that she selects a video because its title contains the name of a historical character whose picture is hung in her son's bedroom. Participants therefore may variably orient to a particular piece of video content by seemingly creating semantic links between the content, and prior experience and knowledge. This could have relevance to the wider discourse around MOOCs on whether course completion is a good measure of the 'success' of a MOOC student (Section 2.2). The data suggest

that a participant may make personally meaningful selections of only parts of the course based on prior knowledge and experience that may be heterogeneous across the student group, therefore suggesting course completion may not be a good success measure for the MOOC.

The fragments show that the user's behaviour is contingent on local factors that are not pre-determined: the participants express apparent reasoning for their actions by bringing in prior knowledge and experience, which they don't verbally set up in advance of the experiment but only near to the moment of the video content to which it relates.

The value judgement therefore appears to co-emerge with the video content, even if broad contextual factors such as prior knowledge and experience don't change substantively over one session of system usage. This echoes some of the findings around *situated plans and actions* by Suchman (Section 3.4) which refers to the emergent nature of interactions with systems that draws on narrow contextual issues such as the perceived system affordances at that time and, perhaps, a person's mood. This work shows that the same quality of interaction exists for interaction with video content in the MOOC.

The studies of interaction so far use the current MOOC system, and therefore interaction is shaped and scoped by the existing interface flow (Section 2.5). In working towards a proof-of-concept video annotation interface, I need to gradually step outside of this interface flow and explore new kinds of interaction on the platform. I will take this first step in the next section, where I will create an experimental system that allows for the seamless integration of video viewing and interaction in a way that the current MOOC interface does not.

5.4. Constructive Interaction - New Interactional Patterns With An Alternate Interface Flow

In the studies above, I do not modify the MOOC interface, and therefore the interactional possibilities in those settings are influenced by the existing MOOC interface flow (Section 2.5). In progressing towards the design of a proof-of-concept video annotation interface, I am interested in exploring new interactional possibilities, which requires I develop prototypes that liberate participants from the restricted interface flow of the existing MOOC platform. The existing MOOC interface flow tends to dis-aggregate video viewing from social interaction, limiting possibilities for intra-video social interaction (Section 2.5).

By opening up the video viewing experience so that it is no longer isolated from interactions with social interaction, I begin to examine *how* users might accomplish social interaction around the video content in the MOOC, *if* the interface flow were modified so users have the freedom to interact with each other and other resources while viewing video content. This understanding can help re-think the conceptions that underpin the development of the video interface - and help generate a set of issues for a proof-of-concept video annotation interface such as, the resources the participants use, and some of the mechanics of the intra-video social interaction.

This part of the method uses two participants collaboratively using a prototype that consists of three devices (Section 4.3). The first device is a tablet to their right, showing MOOC video that can be played, paused, fast forward, or rewound. The second device is a laptop to the left of the tablet, presenting the course discussion forums. Finally, a laptop to their left provides an Internet portal in case participants wish to look

something up online. The videos and discussion forums are live ones from the Coursera MOOC 'Internet History, Technology, and Security'.

In the following sections, I describe three aspects of the participants' social interaction that came through in the data substantively, according to my analysis. First, in the next section, I will describe the ways that participants appear to collaboratively establish a shared view of the video content. I will then discuss two types of transitional activities that the participants undertake: a 'divergence' and 'convergence' in activity, as well as transitions between video-viewing and social interaction.

5.4.1. Working Towards A Shared Grasp of Video Content

The data show a regular pattern where the participants appear to collaboratively construct and align around one point of view on a video. As their social interaction unfolds, participants appear to make a purposeful effort to check their apparent grasp of the video content with each other, through discussion where each verbalises their understanding of the course content, and then clarify and resolve any discrepancies in their expressed understandings. Thus, they perform interactional work to create a sense of shared grasp of the video content. The accomplishment of a kind of 'shared grasp' is defined to be, in observable-reportable terms, when the participants appear to express, in their talk and gestures, agreement and alignment around one expressed point of view.

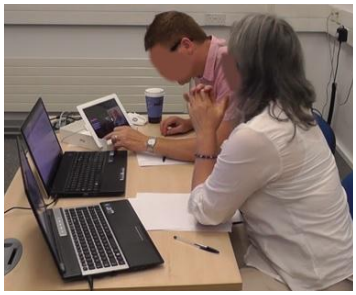
Fragment 5.2 is a particularly illustrative example of this. At the beginning of the fragment, the participants, G and S, are viewing a video, transcribed under 'Vid' in Figure 5.3 below. At the end of Line 1, S pauses the video as shown in Figure 5.4(i).

Fragment 5.2

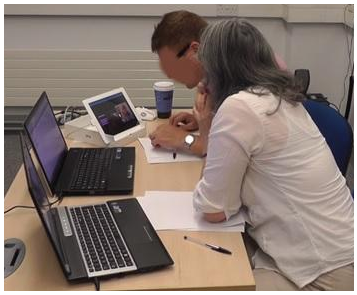
1	Vid:	*So my favourite analogy to talk about how packets work is*
2	S:	(2.0) Uh:: ((laugh)) (2.0) oh OK so there are bits of packets aren't they
		[[
3	G:	Here is it (0.5) code Ah::
4		I See (1.5) All right

Figure 5.3. Fragment 5.2 transcript

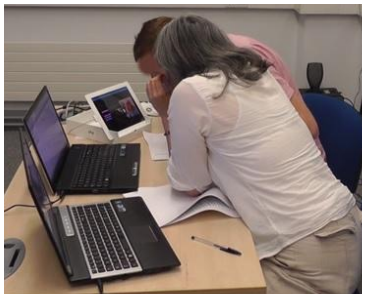
S G



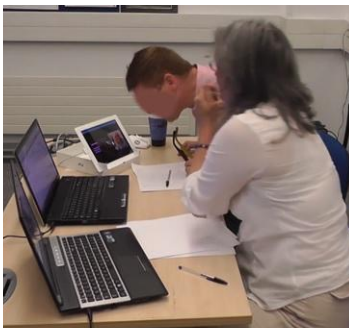
(i) Line 1



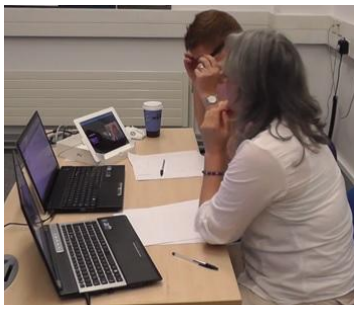
(ii) Line 2 – towards beginning



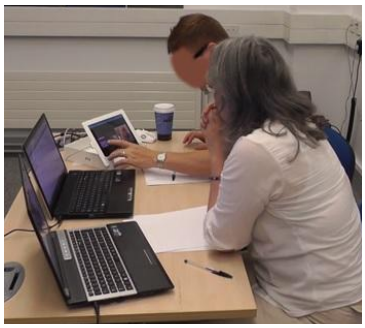
(iii) Line 2 – towards end



(iv) Line 4



(v) Line 4



(vi) End of Line 4 – S restarts video

Figure 5.4. Illustrate data snapshot for Fragment 5.2

In Line 2, until the end of the second (2.0) pause, both participants lean forward towards the tablet, and direct their gaze towards it, as shown in Figure 5.4(ii) to (iii). The portion of the fragment to which I draw the reader's attention, is the contents of the speech in Lines 2 to 4. The tablet is displaying a schematic image, the contents of which have not yet been explained by the lecturer in the video. In Line 2, S appears to speculate on what the image represents ("so there are bits of packets", Line 2), and

invites G to express her opinion (“aren’t they”, Line 2). At the same time, in Line 3, G offers an alternative interpretation of the image, that it displays code (“Here is…code”, Line 3) but later appears to agree with S’s interpretation (“Ah:: I see…all right”, Line 3 to 4).

The speech of the participants suggests they are attempting to come to some shared expressed understanding of the image displayed on the tablet, as they both verbalise suggestions in Lines 2 and 3, with G ultimately expressing agreement with S’s suggestion in Lines 3 to 4 (“I see…All right”, Lines 3-4). The interweaving talk serves to allow the participants to align around a view of what’s being displayed, resolving any potential misalignments without planning or explicitly stating that they’re doing so. In this fragment, as in many others, the participants usually undertake a discussion on their understanding soon after the related moment or segment of the video, and appear to persist in this discussion, explaining and modifying their views until a shared orientation appears to be established verbally.

The seeming negotiation by the participants towards a shared grasp occurs gradually. The participants accomplish a shift into a period of negotiation in a delicate and complex way, without either participant explicitly calling out that this is what they’re trying to achieve. Fragment 5.2 illustrates how the participants’ gestures and speech appear to play a central role in this accomplishment. The extent to which the participants lean in and out, and the use of G and S’s glasses appear to shape the movement of the participants towards discussion. As shown in Figure 5.4(i), when S initially presses the pause button on the tablet, the participants are somewhat leant away from the tablet. In Line 2, G puts on her glasses and begins to lean forward, and soon after S leans in too, as shown in Figure 5.4(ii) and (iii). Then, during the (1.5)

pause in Line 4, G leans back again, followed by S as shown in Figure 5.4(iv) and (v). Without explicitly agreeing on this, S restarts the video at the end of Line 4 as shown in Figure 5.4(vi); an act that is made visible over time by his arm movement, shifting his finger towards the 'play' button. Each act is made sense of in light of the previous one, and in turn gives way to the next act which gradually shifts the pair away from discussion towards returning to video viewing.

If the video data snapshots in Figure 5.4(i) and Figure 5.4(vi) are compared, it appears that at the end of the fragment the participants have completed a series of movements where they have returned to the position from which they departed at the beginning of the fragment. This kind of movement aligns with existing literature on co-present interaction on the 'home position' (Schegloff and Sacks 2002). It is defined to be the same place, and regularly the same position, where a large number of observed moves and sequences begin and end. It serves to illustrate, in seen-but-unnoticed ways, a bounded episode of movement.

In addition to verbalising their understandings of the video content, the participants also appear to actively incorporate into their conversation related prior knowledge. They do this unproblematically, without explicitly stating that this is the case. In the shown fragment, the participants pause the video and begin to interpret the graphic shown on the video, before it is explained by the professor later in the video, and so they must be drawing on prior knowledge. In other instances, G and S are seen to build on the materials presented by the course professor, using the video content as a seed for further discussion. For example, there is an instance where a concept around communication on the Internet (called 'packet switching') is being presented. S jumps in to say that packet switching is a simple concept but more difficult in terms of

implementation, which was not mentioned in the video. In the post-experiment interview, it turned out that S was aware of this because of a previous professional job.

This particular aspect of interaction is concerned with the focused co-creation of a seeming common expressed point of view on a piece of video content, by the participants. However, further analysis of the data shows that participants do not always maintain a common orientation to each other or an artefact in the environment. Rather, their attention appears to diverge and converge dynamically, as will be explained in the next section.

5.4.2. The Divergence and Convergence of Activity

Even though the pair of participants proceed through the task together, they do not always appear to both focus exclusively on their discussion or the same artefact in their environment (which I term to be when the participants' activity is 'converged'). Rather, the data suggest that participants may spend periods of time splitting their attention between more than one artefact, or between discussion and artefacts (which I term to be when the participants' activities are 'diverged'). This appears to occur in emergent and seen-but-unnoticed ways, without seeming to be preplanned or explicitly stated. Rather, the participants shift dynamically between periods of converged and diverged activity. An examination of the narrow context illustrates some aspects of how the participants appear to seamlessly accomplish this dynamic divergence/convergence interactionally.

In the following fragment, 5.3, shown in Figure 5.5, there is an illustrative example of a case where G and S's activity has diverged and where at one point G's attention is split between discussion with S and reviewing discussion forums on the central laptop. Just

before the fragment, G and S are both leaning forward toward the central laptop. Previously, they finished watching a video on the topic of the Internet protocol of 'Packet Switching'. S is talking about an aspect of packet switching which he says confuses him. At the beginning of the fragment, S is still explaining this aspect ("within the IP protocol…", Line 1).

Fragment 5.3

1 2	S:	within the IP protocol you've got (0.2) different kinds of (0.5) underlying (0.5) protocol
3	G:	Mhm::
4 5 6 7	S:	So you can send data you know in like (0.5) uhm in slightly different ways or using the IP (0.1) protocol which itself is sent physical over the wires using other (0.5) protocols and I never totally understand how that all (1.0) you know works together
		[
8 9	G:	Mm:: ri::ght kind of the und- underlying infrastructure (2.0) and protocols
10	S:	Mhm::

Figure 5.5. Fragment 5.3 transcript

S G

(i) Line 1

(ii) Line 8-9 "kind of the und-underlying Infrastructure"

(iii) Line 9 ("and protocols")

Figure 5.6. Illustrative data snapshot from Fragment 5.3

While S is speaking in Line 1, he is leant back away from any device and the direction of his gaze is no longer directed at the central laptop. Rather, his gaze becomes directed at G as shown in Figure 5.6(i). Also during his speech, S makes several hand gestures, as shown in Figure 5.6(i). Over the course of Lines 1 to 9, G does not, however, shift back away from the middle laptop nor does her gaze seem to substantially move from the central laptop. Rather, she remains leant forward and as S starts talking in Line 1, moves her hand forward to use the trackpad of the laptop, to scroll the page down, giving the impression that she is reading the text on the screen while S speaks. And yet, while G appears to be engaging with the discussion forum in the central laptop, she appears to also maintain a form of engagement with S as she expresses agreement with him in Line 3 ("Mhm::") as well as attempts to recapitulate what S has said in Lines 8 to 9 ("ri::ght kind of the…underlying infrastructure…and protocols"). From Line 7, S shifts the direction of his body orientation and gaze back toward the central laptop as shown in Figure 5.6(ii) to (iii).

The two participants have accomplished a divergence and subsequent convergence of activity. While S was speaking in Lines 4 to 7, G's attention appears to have been split between listening to S and participating in the discussion as well as reading the contents of the central laptop. Although G was not physically turned towards S, S appeared to have understood that G was still listening to him as he continued speaking. This appears to have been accomplished by S creating seeming utterances of agreement ("Mhm::") in Lines 3 and 8. G's listening to S is confirmed later by her speech in Lines 8 and 9 where she picks up directly from what S was saying previously. The divergence in the attention of the participants appears to have been facilitated in part by G's ability to use both her sense of sight and hearing to give attention to the central laptop screen as well as S's speech.

The public availability of each other's gestures and utterances appears to play a central role in the way both G and S express their points, as well as how G was able to pick up on when S was nearing the end of his point (as he turned toward the central laptop) leading to G's speech in Line 8. G's statement like "mhm", while not contributing substantively to S's point, served to indicate her attention to his point, even though physically she was oriented towards the laptop. G's arm position, with her hand operating the track pad on the laptop, served to illustrate to S that her attention was partially directed towards the Forums at the same time. All of this was accomplished without G and S explicitly agreeing to do so.

The fragment illustrates how, although engaged in social interaction with each other, the participants do not simply and explicitly always focus their joint attention on a singular artefact or their discussion. Rather, it may be more complex and rich, constituted by at times dynamic divergence and convergence of activity. Although each participant is undertaking their own sequence of activities, their interactional accomplishments are not distinct. G's speech ("mhm") is made sense of in light of the foregoing talk and gestures of S, just as S understands G's utterance in light of these activities and G's conduct more generally. The activities are continually being understood in the light of the ongoing interactional activities of their co-participant. The context is being continually reshaped and renewed by the contributions of the participants in the setting. Because of the way that the activity and context is being reshaped by the activities of the participants, and how tightly they are intertwined, any boundary between the 'individual' and 'collaborative' activities of the participants is blurred.

The data in this section illustrate the complexity and delicacy of the ways that transitions between diverged and converged activity are accomplished. These transitions appear seamless and are co-produced using the visible actions and utterances of the participants. In the next section, I will discuss another type of transition - between video-viewing and social interaction - illustrating some similar qualities between these types of transitions.

5.4.3. The Situated Interplay Between Discussion and Playing/Pausing the Video

In this section, I will continue to examine the way participants accomplish a kind of transition as part of their social interaction. In particular, I will discuss aspects of how participants accomplish the shift between video viewing and discussion, revealing it to be complex, delicate, and co-produced in dynamic seen-but-unnoticed ways. A pivotal act within this transition is the pressing of the play/pause button for the video on the tablet, by one of the participants. However, while the play/pause button plays a role, it usually doesn't neatly, simply, or uniformly result in a transition between video-viewing and discussion. Rather, the ways in which the act is orchestrated with the other activities taking place in the setting, is critical to the transition, as illustrated in the following fragment.

In Fragment 5.4, we see a scene where G makes a comment that is taken up for further discussion. At the beginning of the fragment, the video is playing as transcribed under 'Vid' in Line 1 as shown in Figure 5.7.

Fragment 5.4

1	Vid:	*Don't worry about it it just comes out (0.5) so anyway we have the routers*
		[
2	G:	Ah that's (0.5) so::: what's the word
3		don't worry about it it just comes out (0.5) like I thought we were supposed to
4		be learning something
		[
5	S:	Don't
6		You think (0.2) if I'd signed up for this course I'd be I'd be like that point that's
7		why I'm doing this course ((laugh))
		[
8	G:	Oh my
9		go:::d ((laugh))

Figure 5.7. Fragment 5.4 transcript

S G

(i) Line 1

(ii) Line 2 ("so::: what's the word")

(iii) Line 2-3 ("don't worry about it it just comes out")

(iv) Line 5-6 ("Don't you think (0.2) if I'd signed up")

(v) Line 6 ("I'd be I'd be like that point")

Figure 5.8. Illustrative data snapshot from Fragment 5.4

In Line 2, G begins to make a critique ("Ah that's…so::: what's the word", Line 2) of the video segment they've just viewed in part by echoing the words heard in the segment ("don't worry about it it just comes out", Lines 2-3). During her speech in Lines 2 to 4, G remains in the same position as at the beginning of the fragment, leaning back in the chair as shown in Figure 5.8(ii) and (iii). G goes on to further explain her critique in Lines 3 and 4 by seemingly implying the statement doesn't support any kind of learning as she perceives it ("I thought we were supposed to be learning something", Lines 3-4).

S pauses the video at the moment where the video sound cuts out at the end of Line 1 after G has started speaking as in Figure 5.8(ii). After pausing the video, S turns the direction of his body orientation and gaze towards G in Line 2 as shown in Figure 5.8(iv). S initially stays quiet as G talks in Line 2, before beginning to speak as indicated in Line 5. S seems to take up G's point and appears to express agreement by saying that if he'd have signed up for the course, he would want the lecturer to also unpack and teach more about the topic rather than saying "don't worry about it it just comes out" as G pointed out ("if I'd signed up for this course…I'd be like…that's why I'm doing this course", Lines 6-7). While saying "like" in Line 6, S is tapping his pointed finger on the table as if in a demanding-like gesture, as shown in Figure 5.8(v). S seems to express

further dissatisfaction in Lines 7 and 8 with “Oh my go:::d”, and at the end of the fragment both participants laugh.

The data suggest that *timing* of the discussion of the participants relative to the video, is an important aspect of how the interaction unfolds. The participants tend to discuss a segment of video very soon after that segment takes place. In Fragment 5.4, G begins speaking over the video initially (Line 2), echoing the words of the video (“don’t worry about it it just comes out”, Lines 2-3). Through this echo, G is able to establish a link between her talk and that portion of the video, and S can understand G’s utterance in light of that video. The discussion is made intelligible by their shared experience of viewing the video, and the association of the conversation to that moment by S echoing the words in the video.

Timing in the delivery of talk also appears to be important. The (0.5) pause and the extension of “so:::” provides an opportunity for S to stop the playing of the video, to re-orient his attention towards G, to pick up her comment, and to co-participate in developing it further into conversation. The act by S of pausing the video at that point, appears to then be understood by the participants as drawing both parties into a discussion that picks up the point raised by G.

The shift between video-viewing and discussion is a gradual transition, accomplished in part by the gesture of pressing the play/pause button but determined by the participants’ activities leading up to and after the act. In this setting, the physical orientation of the participants’ heads and bodies also appears to be a central mutual resource to collaboratively accomplish this transition. Whether or not discussion develops further appears to depend on the ways the participants visibly shift the direction in which their gaze and bodies are oriented. A play/pause of the button

doesn't in itself guarantee that the participants transition between video-viewing and discussion. Rather, it is the ways in which that act is orchestrated with other activity in the setting, that shapes the nature of the transition. The data show that the following activities in the setting have an impact: which participant instigated the transition (like G's speech in Line 2 of Fragment 5.4), and which participant replied and *how* (like S's pressing of the pause button and shift in orientation towards G starting in Line 2 of Fragment 5.4). In contrast to Fragment 5.4, there are other periods of activity where even though G makes a statement about the video (as in Line 2 of Fragment 5.4), S doesn't pause the video and discussion doesn't develop from there. Rather, S at times responds with a short snort or chuckle and by maintaining his body and head orientation towards the tablet. In these cases, this appears to be treated by G as a sign of agreement (or lack of interest by S in discussion) who doesn't continue her talk, and the pair continues to view the video rather than discussing further. At other times, G leans forward to pause the video before making a comment about the video, but discussion isn't taken up, again in part because S doesn't substantially shift his head and body orientation, and only chuckles or lightly snorts. So, no one participant or act serves to determine the nature of their transition between video-viewing and discussion, but it is how these things co-emerge, delicately and sequentially, over time that comes to determine how interaction unfolds.

The public availability of the participants' activities, and the physicality of the setting - the need to physically move to operate the devices - appears to be a central resource for the participants as they shift between video-viewing and discussion. The fine-grained aspects of how the participants move their hands and arms around the act of pressing the play/pause button appears to be important in setting up whether or not they transition between video-viewing and discussion. In the presented Fragment 5.4,

S is seen to make a relatively quick gesture to reach forward, press the pause button on the tablet, and to then retract his hand from the button (Figure 5.8(ii)). That act in part serves to secure the further talk from G in Line 3 onwards in Figure 5.8, and cause the pair to shift away from video viewing into substantial discussion. This hand and arm movement by S contrasts with movement seen in other activity fragments where substantial discussion does not proceed a participant pausing the video. For example, there are fragments in the data where G presses the pause button but holds her finger over the button while she talks. This appears to give the impression of a more tentative statement by G who dismisses her own point, and presses the play button for the video on the tablet almost immediately after she's completed her talk, not giving any substantial space of time for S to talk. This again illustrates how the activity of the participants around a press of the pause/play button on the tablet shapes the significance of the act in complex ways. The public availability of their physical poses and orientations towards the devices appears to play into how the participants accomplish the bounding of an episode of discussion. The participants appear to settle in a pose during video viewing - as soon in Figure 5.8(i) - with S leaned slightly forward and G leaned slightly back in their chairs. They appear to use this at least partially as a kind of 'home position' (Schegloff and Sacks 2002) - beginning in this pose at the beginning of a transition from video-viewing to discussion, and returning to a similar pose as the participants round up a period of discussion and return to video-viewing.

And so, the data suggest that no one particular activity of the participant, or an act/ event like pausing/playing the video, determines the shift between video viewing and discussion. Rather, it is how the activities of the participants sequentially give rise to each other, and how the act of playing/pausing of the video is embedded within them, that serves to gradually pivot and transition activity between video-viewing and

discussion. It is the way that these are all coordinated together in an emergent way, collaboratively, that shapes and facilitates transitions between video viewing and discussion. The exact role of the play/pause in the unfolding sequence of interaction appears to depend on who speaks around the act, the details of that speech act (for example, if it's a laugh or a more substantive utterance), the physical activity of the participants around the act (such as shifts in body orientation or gaze direction, and hand gestures). This is all underpinned by the public availability of the participants' talk and gestures that are a resource to the other as part of the sequential unfolding of activity. All of these complexities in the fine-grained details of unfolding interaction are navigated seemingly unproblematically and ordinarily by the participants in a contingent fashion.

Having detailed the most prominent central aspects of interaction that surface from the data, in the next section I will synthesise the findings to present some conclusions and set up the next phase of studies in Chapter 6.

5.5. Conclusion

Where a large portion of prior work into social interaction and video viewing in MOOCs has used interactionally decontextualised metrics collected remotely, such as server log data, in this chapter I have undertaken an exploration of interaction around and through the MOOC that considers the broad and narrow contextual issues that may shape and be shaped by interaction.

The studies in this chapter reveal that student-video interaction and intra-video social interaction can be quite complex and delicate, relying upon a sequence of contingent, emergent, and situated observable-reportable activity. The usability studies (Section

5.3) show how the metrics used to characterise student-video interaction, like the ways students are categorised in a binary fashion into having 'viewed a video' if they start viewing a video, gloss over rich interaction like the possibly ongoing following of an 'information scent' and individual seeming value judgments of course content. The studies examining social interaction with a modified prototype (Section 5.4), illustrate the unfolding of the social interaction was not simply deterministically accomplished by one act or one participant alone, but is co-produced in an emergent fashion. This co-production appears to be constituted by a delicate and complex sequence of unfolding gestures and utterances.

The goal of this chapter is not to suggest that the MOOC should be designed to attempt to reproduce or replicate these complexities in social interaction for a distributed group in a MOOC. It may not be possible to develop the MOOC to support or allow for the richness and complexity of this interaction. More importantly, the work in this chapter may be used to reconsider the types of mechanics of social interaction that the MOOC supports, and to rethink the conceptions of interaction that underpin the development of the MOOC to support collaboration and interaction. These studies surface that holding onto a too simplistic conception of interaction in the MOOC may unduly constrain the potential of the platform.

The interactional issues surfaced here should be seen as provisional, part of the iterative design process undertaken in this thesis, needing further testing and development. Nevertheless, it is useful to outline the key interactional issues that are apparent from the studies in this chapter:

- **Timing:** The data suggest that the interleaving of discussion with video viewing, rather than discussion taking place only before or after video-viewing, is a central

aspect of the participants' social interaction. Discussing a video fragment soon after it occurred, including pausing the video to develop discussion, helped the participants make sense of each other's talk and appears to have been accepted unproblematically as an ordinary way of referencing and developing a video fragment through discussion.

The current MOOC doesn't support this kind of interleaving of discussion and video viewing. Rather, the discussion forums in which social interaction is contained, are not designed to easily be linked to a moment of video, and cannot be viewed by the participant as they are watching the video (Section 2.5).

- **Public availability of activity:** The seen-but-unnoticed and delicate ways that participants accomplished their intra-video social interaction, appears to centrally depend on the participants being able to use the detailed features of each other's talk and gestures as a resource. For example, in working towards a shared grasp of video content and in diverging/converging activity, the participants' head and body orientation, and the details of their arm movements playing/pausing the video, appeared critical in making the activity intelligible to both participants who go on to accomplish their collaboration seemingly unproblematically. Furthermore, no single gesture or piece of talk appears to operate alone in determining the ways in which social interaction unfolds, but rather it is constituted by a sequence in which gestures and utterances co-emerge in a coordinated and co-dependent manner.

  The video interface in the current MOOC, and the interface of discussion forums, don't provide any visibility of the synchronous activities of others. Therefore, the

participants are interactionally impaired, unable to draw on the activities of others to coordinate their activity in a meaningful way.

- **'Video viewing' is not a passive and simple activity:** In these studies, we can see the ways in which the artefacts of interaction (that are typically the focus of 'big data'-based studies) such as 'video views' and 'click's may be a gloss over important interactional complexity and richness in video viewing. For example, students may actively engage in considering the value of the video, and seem to use the video and interface as a resource for attempting to judge whether or not to continue viewing the video. In most 'big data'-based studies, an act like a 'click' would be undistinguished between users because there is limited visibility on the activity preceeding it, during it, and proceeding it, which gives it meaning. Furthermore, video content may not be homogeneously and unquestionably 'consumed' by the student. Rather, students may undertake an active framing of the video content - a framing that is shaped by the narrow and broad contextual issues at play, and informed by the student's prior knowledge and experiences.

  By dis-aggregating the video content from discussion forums, the current MOOC interface presents the single point of view of the contents of an educational video. It doesn't show the variety of perspectives and framings of the content that could be contributed by the student group.

- **Engagement in Multiple Activities:** Even though the participants in these studies are engaged in joint activity, the analyses show that participants can at the same time orient towards various resources in the environment, such as different devices or to each other. And yet, each participant's activities are continually being understood in light of the ongoing activities of the other participant in the setting.

The narrow context is continuously available to be reshaped by the contributions of the participants. At times, broadly, participants appear to together focus on accomplishing transitions between video viewing and discussion, and continuously collaboratively work towards aligning around a common view of the video. However, within these and other activities, they may dynamically shift into and out of more diverged activity - apparently able to accomplish simultaneously orienting towards the activities of the other participant, while sometimes orienting towards a different device.

The existing interface flow of the MOOC tends to constrain students within either video viewing, or social interaction, or assessments (Section 2.5). It doesn't seem to allow for an inter-connection between these activities which would support participants in easily shifting between these various activities.

These fine-grained aspects of interaction that appear to be central to the ways participants ordinarily accomplish intra-video social interaction, would be difficult to surface through a 'big data'-based methodological approach because they are detailed and seen-but-unnoticed by the participants. Only by looking at the context through which those 'big data' events take place, can we understand some of the issues that shape and give meaning to those events. By examining interaction qualitatively in quasi-naturalistic settings, we can glean insights about the nuanced mechanics of how interaction is accomplished. This suggests the relevance of the kinds of ethnomethodologically-informed usability studies and methodologies pioneered by those such as Suchman (Section 3.4) in studying MOOCs. One trade-off of this qualitative approach is that it shows a deeper and richer cut of interaction, but doesn't have the immediate statistical generalisable quality of 'big data'-based approaches.

However, it can help to fill in gaps in understanding of some of the nuance of interaction, and therefore is a good complement to 'big data'-based approaches - a topic which will be taken up and developed in subsequent chapters.

This first phase of the iterative approach to developing a proof-of-concept video annotation interface, reveals some of the broad and narrow contextual issues that shape interaction in the MOOC more generally, and interaction with and around video more particularly. In the next phase, I'm interested in continuing to understand the details of how intra-video social interaction is accomplished, but by larger groups and through a new prototype whose features draw on the issues surfaced here. The general guiding questions for enquiry that emerge from this phase, and which drive the next phase of research in Chapter 6 are:

1. Keeping with the provision of simultaneous access to resources while viewing a video, how do larger groups interact when there is an annotation system that associates contributions with moments in a video? How do groups also interact with existing interactions tied to moments in the video?

2. How do larger groups of participants interact as part of collaborative video annotation? Is there any mapping between the interaction seen in this phase and those seen in larger groups? In what ways are new kinds of practice seen to emerge for larger groups? What are some fundamental aspects of their social interaction and the individual participant's involvement in it?

Chapter 6.

Phase 2 - The Individual and the Collaborative: Intra-Video Social Interaction of a Co-Present Group

6.1. Introduction

Although 'collaboration' is a concept of interest in several existing studies of the MOOC, there is minimal literature that unpacks and delineates the mechanics that constitute it. The idea of collaboration is usually taken up in literature on MOOCs to refer to a particular kind of activity that requires the action and contribution of many (e.g. (Al-Atabi and DeBoer 2014; McAuley et al. 2010; Burke and Mahoney 2013; Staubitz et al. 2014)). This is not only a sociological concern, but may also have profound implications for the success of the MOOC designed to support this kind of interaction. Strong evidence for this lies in the area of Computer Supported Collaborative Work (CSCW). CSCW contains several seminal works that illustrate how challenges to developing systems to support and enhance participation in large-scale multi-party events, depends foundationally on delineating the organisation of collaboration within the ecology of naturalistic or quasi-naturalistic setups (Galegher and Kraut 1990; Grudin 1988; Orlikowski 1992; Luff et al. 2011; Luff et al. 2014; Button 2012). This is in contrast to traditional views that focus on the kinds and capabilities of technologies that could be incorporated into online systems.

There isn't a clear reason for why research into collaboration in the MOOC hasn't drawn deeply on work done in CSCW. This may be because there is no naturally-

occurring site in which to perform the kinds of studies typical in CSCW (Chapter 4), and because this approach resists the kind of large-scale statistically-generalisable quantitative results that appear to dominate MOOC research to date.

In this chapter, I will explore a few aspects of the organisation of collaboration by two groups of participants undertaking intra-video social interaction, around MOOC video content. Snapshots from video data of the groups are shown in Figure 6.1. They are performing an open-ended task of collaboratively viewing and discussing an educational video.

Figure 6.1. Screenshots of video data in phase 2 with two groups of participants – V, K, and J (left), and A, B, C, D, and E (right)

In the following sections, I use example fragments to bring out three patterns of social interaction seen across both groups. These patterns appeared most central to the groups accomplishing their social interaction. By examining the ways the participants coordinate their actions with each other, I provide a sense of the delicacy and complexity of their cooperative work. I show how even seeming individual tasks are systematically coordinated with the activities of co-participants in seen-but-unnoticed ways (Garfinkel 1967). As I will show, an understanding of this, can help develop the ideas that underpin the development of the MOOC video annotation interface to support social interaction.

6.2. The Act of Playing/Pausing Video as a Resource for Interaction

A prominent pattern of activity seen in the data, is the pressing of the playing/pausing button on the tablet. This is particularly central during transitions by the group between video-viewing and discussion. Superficially, the act of playing/pausing the video may be seen to be an individual one: it is performed by one person pressing a small area of the tablet with their finger. However, an analysis of the fine-grained details of interaction leading up to, during, and directly after, the act, reveals it to be a more complex act of a collaborative nature.

Fragment 6.1 shows an example of how the group transitions from video-viewing to discussion, seemingly triggered by the action of participant E. Just before this fragment, the participants are watching the video in relative silence, as indicated by the speech transcribed under the name 'Vid' (for Video) in Line 1 of the transcript below. The group members are all directing their gaze to the tablet as shown in Figure 6.3(i).

Fragment 6.1

1	Vid:	*powerful axis (1.0) the vision of the future is actually to*
2	E:	So probably there is a point that I want to make is (1.0) with the retail
3		that option will be more quicker↑ (1.0) than the private (0.5) sector

Figure 6.2. Fragment 6.1 transcript

D B C E A

(i) Just before Fragment 6.1

(ii) Line 1 (“powerful axis”)

(iii) Line 1 (“the vision of the”)

(iv) Line 1 (“future is actually to”)

Figure 6.3. Illustrative data snapshots from Fragment 6.1

As the video is playing at the beginning of Line 1, E shifts his body orientation towards the group as shown in Figure 6.3(ii). A, C, and D react by turning their gaze toward E in Line 1 during “the vision of the” as shown in Figure 6.3(iii). At the same time, C moves his hand away from his face and forward as if in a possibly questioning pose. Thereafter, in Line 3, during “future is actually to”, A moves forward to pause the video as shown in Figure 6.3(iv).

In Line 2, the group members turn their body orientation and gaze direction toward E. E begins to explain a point about the retail and private sectors that relates to the video content they were just viewing (“so probably there is a point”…, Line 2).

The public availability of the acts and talk of the participants are a critical resource to the group, in accomplishing the transition from video-viewing to discussion. For

example, the act of pausing the video by A is a substantial aspect of the transition. The physicality and public availability of this act – the gradual extension of her arm towards the tablet – provides the rest of the group with a resource to notice the way the social interaction is progressing and to influence it. In this fragment, the other participants don't appear to express an observable-noticeable objection to the pausing of the video. All the group members ultimately display an understanding of the purpose of this act by A, and follow her in re-orienting towards E, which sets the stage for E to further express his point (Lines 2-3). At the same time, E aligns with the actions of A, since he pauses his speech until A completes the pressing of the pause button.

While the pressing of the pause button was a major physical gesture that shaped the interaction, the transition is not accomplished by that act alone. Rather, it is orchestrated through a whole sequence of fine-grained gestures. Although A is the individual participant who paused the video, the act of pausing the video is both influenced by the activity of the rest of the group, and goes on to influence the activity of others. A's pausing of the video followed, and is made sense of in light of, both the activity of E moving forward into the conversational circle (Line 1), and the gesture made by C (– changing the direction of his gaze towards E and making what may be interpreted to be a questioning gesture with his hand). The subsequent activity - E's talk and the re-orientation of the group towards E - in turn can be understood with respect to A's activity. In seen-but-unnoticed ways, without explicitly declaring it in advance, all the group members together co-produce the transition, underpinned by the public availability of their talk and gestures.

The play/pause button is a key artefact in the fragment of interaction, playing a complex and delicate role. In one way, the button serves the more obvious purpose of

allowing the video to be paused, giving E the opportunity to take centre stage in the social interaction and express his point. In a more complex way, the button is also a resource for all the participants to make sense of the talk and visual conduct of each other, based on how various participants orient towards it. A, while using it to pause the video, at the same time uses it to make visible the intention to shift the nature of their social interaction away from video-viewing, towards discussion. In this way, the button serves to make A's act a resource for the rest of the group. The button is temporarily and dynamically brought into the interaction, and its relevance is shaped by, and shapes, the ongoing talk and gestures within which it is embedded.

Over the course of the co-production of the transition in Fragment 6.1, the participants dynamically inhabit a number of asymmetric participatory modes. The pausing of the video, although performed individually by A, is part of a sequence of activity in which all the participants are involved. The pausing of the video by A is performed in response to the movement of E, and appears to serve the purpose of supporting E interactionally in further elaborating his comment by re-orienting the rest of the group towards him. The pausing of the video by A thus shapes the involvement of the other participants, as it precedes their re-orientation away from the video towards E. Therefore, the transition is constituted by a coordinated sequence of acts with the participants playing various kinds of participatory roles.

No two participants act the same way in terms of their activity's fine-grained details, and yet together they all orchestrate the transition, without explicitly stating that this is their goal. A's act of pausing the video is shaped by the activity of another group member, E, and serves to shape and transition the activity of the entire group. E is perhaps an instigator, but the pressing of the pause button by A, and the gradual re-

orientation of the group (thus in a way appearing to accept the direction in which interaction was unfolding), all together accomplish the transition seamlessly. A's pausing of the video is made sense of in light of the foregoing activity of E. In the same way, the other participants unproblematically align with A's act, and give meaning to it with their subsequent talk and conduct. None of the various roles and activities of a participant can be disentangled from the activities of the others around that time.

While pausing the video is a central pivot around which the activity of the group transitioned from video viewing to discussion, the transition cannot be attributed to the single instant in which the pause button was pressed. This act alone does not change the course of the social interaction. Rather, the wider sequence into which the act of pausing the video is embedded, provides the key to understanding how the transition is accomplished by the group. The transition is constituted by a sequence of unfolding activity that is co-produced over time, which the data illustrate to be complex and delicate.

Furthermore, there is no evidence in the video data to suggest that the transition is pre-planned, nor is the strategy for accomplishing the transition explicitly stated. Rather, the transition has an emergent quality. At each moment, the activities of a participant are being understood in light of, and influenced by, the ongoing activities of their co-participants. Therefore, the direction in which the social interaction unfolds is continually being reshaped and renewed by the contributions of the participants in the setting.

In the current MOOC interface, there is no way of understanding the synchronous activity of others while logged in. The data suggest this may impair the extent to which coordinated activity and social interaction can develop among students. In particular,

providing some visibility of transitions into and out of video-viewing, can help orchestrate social interaction seeded on fragments of educational video content.

The analysis so far shows the complex role an individual's activity can play in the unfolding of social interaction, and the delicate ways individuals' activity come together as part of the sequential unfolding of social interaction. The group members participate in a variety of different and complex ways. In next section, I will continue to consider the heterogeneous ways people may participate in groups, highlighting the various degrees of participation people can undertake. Even when participants appear at first glance to be on the periphery of a group, their activities can still substantially influence the unfolding social interaction of the group.

6.3. The Complexity and Dynamicity of Peripheral Participation

In the previous section, the analysis suggests that even an individual action such as pausing a video, is organised in relation to the emergent conduct of other group members, with each member participating in a different way. It is the relative orchestration of these different modes of participation together, that accomplish the further unfolding of social interaction. In this section, I explore this concept of heterogeneous participation further, by looking at how participants can accomplish peripheral participation.

Peripheral modes of participation feature substantively in CSCW literature across a number of different settings, and is referred to through the concepts of peripheral awareness, engagement, and/or participation. In existing literature, the idea of 'peripheral awareness' emerged mainly through workplace studies as a concept to call out participants' abilities to stay sensitive to the activity of others while engaged in

distinct tasks (Heath et al. 2002). 'Peripheral participation' has been defined in slightly different ways across literature, usually referring to a person who maintains some kind of monitoring of tasks of others without being directly involved (Heath et al. 2002; Monk and Watts 2000). At other times, the idea of 'peripheral monitoring' is used as one piece of how peripheral participation might be accomplished. A deeper unpacking of these related concepts exists across a number of different workplace settings including the City Dealing Trading Room (Heath et al. 1995), Air Traffic Control (Harper and Hughes 1993), Medical Emergency settings (Sarcevic at al. 2012), and the collaborative settings of cybersecurity analysts (Ahrend et al. 2016). Drawing on similar kinds of analysis to this existing work, I will show that in intra-video social interaction, the participants appear to employ the ability to maintain a form of peripheral engagement in the activities of others which can serve to shape their and the group's activities. These data build on the literature by illustrating examples of peripheral monitoring during intra-video social interaction around educational videos.

I will show how peripheral participants do not necessarily exercise any less influence over the activities of a group than the core group members. Furthermore, I'll show how transitions by participants into and out of these peripheral participatory modes are unproblematically accomplished dynamically and in seen-but-unnoticed ways.

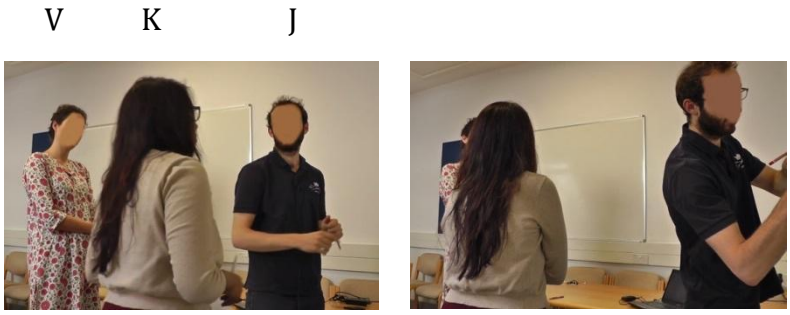
Next, I will present two fragments: the first shows the transition of a participant from a more central to a more 'peripheral' participatory mode, and the second shows their transition back into a more active mode.

Just before Fragment 6.2 shown in Figure 6.4, the participants are speaking to each other, with the video paused. In Line 1, K speaks as shown in the transcript below ("Based on patterns…", Line 1).

Fragment 6.2

1	K:	Based on patterns that they see they (0.2) could be really wrong (0.5) sometimes
2	J:	=Actually that's a good point
		[
3	V:	Surely a GP a lot of it is (0.1) I mean the reason we don't have just a robo:t (0.2)
4		deciding
		[
5	K:	Ya
6		true

Figure 6.4. Fragment 6.2 transcript



(i) Line 2 (ii) Line 3-6

Figure 6.5. Illustrative data snapshots from Fragment 6.2

Initially, the participants are physically arranged into a kind of conversational circle, defined to be when the participants stand in a semi-circular configuration around the tablet with their bodies oriented towards the centre of the circle and where the participants are mostly visible to each other.

In Line 2, as J speaks, he moves to get a piece of paper, magnetic strip, and pencil before turning back to K as shown in Figure 6.5(i). In Line 3 to 6, V and K speak, discussing the work of a doctor and how it is more than only providing information and data to a patient ("…we don't just have a robot…", Line 3). At the beginning of Line 3, J moves to the whiteboard, shown in Figure 6.5(ii). After Line 3, J has his back turned to

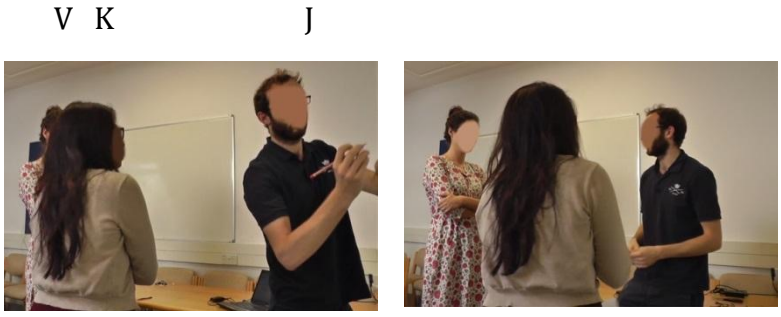
V and K who now have their body orientation and gaze directed towards each other, thus J is temporarily excluded from the conversational circle as shown in Figure 6.5(ii). J is now participating peripherally with respect to the rest of the group because of his relative position to the conversational circle of V and K.

Fragment 6.3 in Figure 6.6 illustrates how, a short time later, J moves back into the conversational circle and therefore back into a more central participatory mode with respect to V and K. In Line 1 shown in Fragment 6.3, J moves away from writing the annotation, and shifts his body orientation and gaze back towards V and K as shown in Figure 6.7(ii), who in turn shift their body orientation and gaze to J.

Fragment 6.3

1 2	V:	Just a value judgement (0.5) You look at someone and you (0.5) you recognise something that's intangible
3	K:	=Ya:: (0.5) that's true
4 5	J:	(2.0) Ya (2.0) that's a good point because ofte::n (0.1) when you go to a doctor it starts with how you feel it's like very (1.0) ya
6	K:	=Ya
7	J:	(1.0) Ah:: (1.5) not measurable in a sense

Figure 6.6. Fragment 6.3 transcript



(i) Line 1 (ii) Line 4-7

Figure 6.7. Illustrative data snapshots from Fragment 6.3

In Lines 1 and 2, V and K are still mostly in their own conversational circle, however in Line 3, they begin to orient towards J and recreate the larger conversational circle including J. By Line 4, J is re-included into their conversational circle as shown in Figure 6.7(ii). J at the same time begins to speak in Line 4.

An important aspect of Fragment 6.3 is that J's speech in Line 4 leads directly on in topic from the conversation V and K were having while he was not a part of their conversational circle ("Ya…that's a good point because", Line 4). J continues to add to the points V and K made by talking about how the way you feel can be subjective ("when you go the doctor it starts with how you feel it's…not measurable in a sense", Lines 4, 5, and 7).

The continuation of the single thread of discussion as J shifts into and out of the peripheral mode of participation suggests he was aware of V and K's discussion throughout. The data show how he uses different physical senses to monitor their conversation. At times, while turned away, since J couldn't see V and K, it seems like he used his sense of hearing to continue to monitor their conversation. At the same time, V and K unproblematically align with J's activity throughout, appearing able to make sense of J's activity, even though they appear at one point to close the social circle around themselves.

The pencils and whiteboard in the setting play a central role in making J's activities intelligible to V and K. While inhabiting a more peripheral role, turned away from the social circle, J is turned towards the whiteboard, and is using the pencil and paper provided to undertake the addition of a note to the whiteboard. While this is an individual activity, it serves to establish J's participatory mode with respect to the wider group, and is a resource for the other participants in navigating J's varying role in the

group. Therefore, J's access to these artefacts, and his use of them, is critical to the accomplishment of the activity seen in Fragments 6.2 and 6.3.

Although J's actions stand out, since they are distinct from V and K's activity, his transitions into and out of a peripheral role are not individual but rather accomplished through a collaborative effort. In coordination with J's physical movement, the conversational circle (constituted by the way V and K arranged themselves in the space) appears to have been dynamically shifted and adjusted by V and K as J moves into and out of his more distinct and peripheral role. The public availability of V and K's discussion while J is out of the conversational circle also appears to be a resource to J in being able to retake his position in the conversational circle, and may be a deciding factor in the timing of when J rejoins the circle. V and K appear to be aware that J may be overhearing their discussion while he is turned away from them, as they unproblematically adjust the conversational circle to accept J and listen to J's speech that leads on directly from theirs that took place while J is turned away from them.

There is no evidence in the data that J's peripheral participation is pre-planned in any way. V and K do not seem to request an explicit explanation of what J is doing when he leaves and rejoins the social circle over the course of Fragment 6.2 and 6.3. Rather, V and K appear to unproblematically accept J's movement out of the social circle, and his overhearing of their conversation - which J then uses as a resource when rejoining the social circle. This echoes many of the features of so-called 'outlouds' in work settings (e.g. (Heath et al. 1995 p.9)) where this appears to be a central part of how particular work tasks are accomplished.

J's peripheral participation is not associated with a single moment, but accomplished over time cooperatively. J instigates the transition into a peripheral role by his

movement in Line 1 of Fragment 6.2, and is perhaps complete in Line 4, but it cannot be attributed to any moment or act in particular. There is a distinct emergent and gradual nature of the transitions between peripheral and more central forms of participation.

This data can contribute towards the ideas that underpin how we could develop the video interface of the MOOC. It suggests the value of creating some kind of virtual space that allows individual participants in a group to act independently, contributing if they like, while at the same time making some facet of their independent activity visual to others who can coordinate themselves accordingly. The data suggest the importance of taking into account and supporting various kinds of peripheral participation into the design of the MOOC. This is beneficial for all the members of group.

The analysis in this section further shows the complex and heterogeneous participatory modes that together constitute social interaction and drive collaboration forward. No one participant alone accomplishes the social interaction, and each new activity of an individual participant is made meaningful by the previous activities of the other participants, serving in turn as the basis for meaning of the subsequent participant's activity. It is this continuous orchestration of activity by all the participants together that appears to move the social interaction forward. In the next section, I will continue to illustrate this kind of orchestration characteristic of the group's activities, by examining the ways in which the participants appear to, through speech, work towards establishing a shared sense of each other's talk and activities, in seen-but-unnoticed ways.

6.4. Safeguarding a Shared Sense of Each Other's Talk and Conduct

The analysis so far illustrates the ways in which the continuous orchestration of activity by participants is a central part of the unfolding of intra-video social interaction by the group. In this section, I will show how the participants appear to use each other's gestures and utterances to detect misalignments in understanding, as expressed through talk and action. In cases where a misalignment seems to be detected by a participant, the data illustrate that they undertake a period of social interaction through which they attempt to resolve any seeming misunderstanding to secure what I have called the 'shared grasp' of each other's talk and conduct. A shared grasp appears to be secured when there is agreement and understanding expressed in the speech and gestures of the participants, such as saying 'yes' or through nodding of the head. A participant seemingly detecting a potential misunderstanding, for example, might produce a speech or gestures that appear to bring something up for consideration or clarification, using an instrument such as a clarifying question.

At the beginning of Fragment 6.4, K appears to be speaking about data handling and management and how people may feel about their medical record storage ("how people will feel about…medical records", Line 1). K's statement may be seen as rather tentative, as indicated by K's termination of her turn in Line 2 with quiet speech (so quiet that her exact words could not be transcribed with confidence in Line 2) as well as "I don't know" at the end of Line 2.

Fragment 6.4

1 2	K:	how people will feel about it (1.0) especially with medical records and stuff like that (( )) (0.2) I don't know
3	J:	So the ethical issues↑
4	K:	=Ya
5	V:	=M::

Figure 6.8. Fragment 6.4 transcript

(i) Line 1: "about it (1.0) especially"

(ii) Line 4

Figure 6.9. Illustrative data snapshots from Fragment 6.4

In Line 3, J responds by saying "So the ethical issues↑". It seems, through the use of an emphasised "so" in Line 3, that J is offering a term ("ethical issues") to capture the theme of K's previous statement, to seemingly clarify and confirm it, since K answers with approval ("Ya", Line 4). J's statement also appears to be a request because of the rising shift in intonation at the end of Line 3. Therefore, J's statement could be a form of recapitulation of the statement which K has just made. In so doing, J appears to make an expressed form of his understanding of J's statement in Lines 1 to 2 publicly available to the other participants. Although so far only J and K are speaking to each other, J's response appears to have an influence on V, too, who expresses an agreement and understanding in Line 5 ("M:::") almost at the same time as K expresses her agreement in Line 4 ("=Ya"). Therefore, a sense of a shared grasp of the talk of K in

Line 1 appears to have been achieved through both J's recapitulation of K's speech and the expressions of agreement and understanding by K and V. J's recapitulation, although an individual act, was pre- and pro-ceeded by the activities by other participants and appears to have an important direct impact on the shared grasp of the talk by the whole group, including V.

The strategy employed by J to clarify K's speech, appears to be sensitive to the details of the talk and activity of K. J's statement in Line 3 (his somewhat individual act) appears to have been coordinated with the speech of K, as it seems that J waited for K to complete her sentence, and offered his recapitulation in response to what seems like her uncertainty in the point she makes in Lines 1 to 2. The (0.2) pause before the 'I don't know' (Line 2) represents some uncertainty, and a potential winding down of her comment in Lines 1 and 2. J appears to be sensitive to this and times his comment relative to it. The end of K's speech appears to also be connected to the gestures she's making: she appears to be moving her arm substantially as she talks (seen in Figure 6.9(i)), but her arm comes to rest in a vertical orientation with her hand on her lips as she completes her talk in Line 3 (shown in Figure 6.9(ii)). The orientation of K with respect to the whiteboard and tablet appears to influence when J begins to talk, since his talk starts as K shifts away from the whiteboard. In leading directly on from where K has wound up her point, J's statement in Line 3, is made sense of with respect to K's immediately preceding talk, and so the timing of the talk is central to V and K seemingly making sense of J's talk.

Over the course of the group establishing a seeming sense of shared grasp of K's talk, V and J's activity are shaped in different ways by K's activity at the beginning of the fragment. J appears most vocal in responding in Line 3. Although silent for most of the

fragment, V appears to use the interaction between K and J as a resource to establish a seeming shared grasp because she appears to express a form of agreement in Line 5.

While at face-value, one may consider this fragment as a case of J simply, individually, requesting clarification on K's statement; by examining the fine-grained aspects of the interaction to show the work that the various participants' talk and gestures do, the analysis shows that this activity has a distinct collaborative nature. Furthermore, the participants appear to have accomplished this shared grasp without explicitly setting up and stating it as a goal. It seems to be unproblematically undertaken in seen-but-unnoticed ways, drawing on each other's gestures and talk as a resource. It is accomplished through a strategy that emerges gradually and collaboratively. The mechanics of how a group might come to a seeming shared sense of each other's talk and conduct is not substantially incorporated into the way the MOOC platform is designed. These data begin to paint a picture of how important this can be in the unfolding of social interaction, and suggest the value of bringing this into the wider conceptual framework used to design and assess MOOCs in the future.

Over the course of the last three sections, I have shown the complex ways in which social interaction is underpinned by an interactional orchestration of talk and gesture, continually, by all the participants in the setting. This may be true, even in cases where at first glance it appears that one individual is acting somewhat alone in the setting (like pressing the pause button, moving into and out of a peripheral participation mode, or asking for clarification on another participant's talk). In the next section, I will synthesise the insights discussed so far, to surface some new considerations for intra-video social interaction in the MOOC, and set up for the next phase of analysis in Chapter 7.

6.5. Discussion

The concept of 'collaboration' has been referred to in studies on MOOCs, without explicitly delineating the mechanics that might make up and drive forward social interaction. In this chapter, I have undertaken an exploration of some of the fine-grained observable-reportable aspects of how participants might accomplish intra-video social interaction, and reveal some of the important nuances in the mechanics that underpin collaboration.

The data illustrate how the 'individual' and 'collaborative' may be very tightly interlinked and interdependent, making it difficult (if not impossible) to separate the two concepts. In the fragments, activities which may at first glance be considered 'individual', appear to be systematically coordinated with the actions of other participants. All of the fragments illustrate how activity by the group is usually not dependent on any one person, but appears to be co-produced by all the participants together in a complex and delicate sequence of interaction.

The data illustrate the rich kinds of participatory modes of the participants and how they continuously act in a contingent manner with respect to the emergent conduct of others. Individuals are seen to co-produce the collaboration through an orchestration of heterogeneous and dynamic participatory roles. For example, I find that a transition from video viewing to discussion may be constituted by an individual's utterance or action to indicate the desire to speak, followed by the seen-but-unnoticed and emergent re-orientation of the rest of the group away from the video towards the individual. All these complex mechanics must work together to accomplish the delicate transition. They are underpinned by the public availability of the gestures and utterances of the participants that become resources drawn on by others.

The opportunity for asynchronous social interaction was offered by adding onto the whiteboard contributions from students taken from the online course. These contributions were a springboard for group conversation twice in the first group and once in the second. In these cases, the group appeared to notice and briefly reflect upon the contributions, however quickly the group's focus appeared to shift to their own discussion. What we can see then is that engagement in synchronous discussion tends to draw the participants in, more so than asynchronous discussion, however this could be because the participants in the room may be obliged to pay attention to each other more than contributions on the whiteboard. Future work should experiment with providing more comments pre-populated in the prototype to create increased opportunities for asynchronous communication. Alternatively, it's possible that adapting the task given to participants, so that it explicitly asks them to participate in asynchronous communication, would provide more insight into how this communication unfolds.

It may not be possible to reproduce these complexities and this nuance of social interaction for a large distributed group using the MOOC video interface. However, more importantly, the analyses in this chapter bring under reconsideration the types of mechanics that we aim to support for intra-video social interaction. These new conceptual considerations can then help in more effective consideration of the ideas that underpin the development of a MOOC video annotation interface to support social interaction. Holding onto a too simplistic view of collaboration, may unnecessarily constrain the potential of the MOOC video interface.

The interactional insights surfaced here are provisional, and will be built up further in subsequent chapters. Nevertheless, it is useful to outline the key interactional issues that are surfaced through the analyses in this chapter:

- **Body movement and orientation:** Over the course of unfolding interaction, participants show a competency for making sense of a wide variety of body movements (including delicate ones), and drawing on them for the production of their subsequent activity. For example, participants orient to subtle gestures that reflect one participant's request to verbally express a viewpoint, that gives way to the subsequent body movement of another participant, to pause the video. These two activities can be made sense of in light of each other, and the participants use each other's body movement and orientation to accomplish the activity, without explicitly stating that they are doing so. Body movement and orientation are also used to indicate varying degrees of active participation, and may roughly demarcate a social circle. However, the social circle can be dynamically and delicately negotiated in seen-but-unnoticed ways (Garfinkel 1967). Therefore, body orientation and movement are instruments through which to create varying kinds of participatory modes. It also allows for the appropriate timing of talk and other acts as part of attempting to seemingly establish a sense of shared view around one of the participant's talk or gestures.

  The video interface in the current MOOC and the discussion forums, don't provide any visibility of the synchronous activity of other participants. Therefore, the participants aren't able to draw on these activities, as they appear to do in the data presented here. This may compromise interactionally

the participants' abilities to negotiate the social interaction and coordinate their activities with others.

- **Use of artefacts:** Certain artefacts within the physical setting of the co-present group, appear to play a critical role in the unfolding of social interaction. Artefacts, and features of artefacts, are a key resource used in interaction. Furthermore, the talk and gestures of the participants in some cases could only be made sense of with respect to the state of these artefacts. For example, the play/pause button on the tablet (Section 6.2), when acted upon by a participant, served as a core aspect of a transition between video viewing and discussion, but in itself did not accomplish that transition. Rather, it was the sequence of social interaction within which it was embedded; the orchestrated interaction of all the participants leading up to, during, and after the act that gives the act meaning and drives social interaction forward. The use of the pencil, and the movement of participants with respect to the whiteboard on which the tablet was mounted, also serves as a resource to assist the participants in the group to navigate the dynamic transition by some participants into and out of peripheral participatory roles (Section 6.3). More generally, the body orientation and gaze direction of the participants towards and away from the tablet on the whiteboard, serves to establish what they were referring to through talk, and is used to navigate conversation. In establishing shared grasp of talk, this is used to establish a kind of boundary of one participant's speech, giving way to another's (Section 6.4). The strategies of usage of the artefacts also don't appear to be explicitly pre-established by the participants. Artefacts are used dynamically and emergently, at times used only momentarily, with their sense and relevance made clear with respect to the

unfolding talk and activity of the participants. Therefore, having regular access to the ways that co-participants orient towards artefacts and their features, is critical to the accomplishment of social interaction.

By only showing the final submitted discussion forum posts, and by separating out the resources in the MOOC into disjoint interactional streams, the current MOOC interface gives minimal visibility into the ways that participants are orienting towards and engaging with multiple artefacts provided within the MOOC.

- **Heterogeneous and complex participatory modes:** The ways different group members participate over the course of the unfolding of social interaction, may be varied, dynamic and complex. Their activities also appear to be delicate and navigated in seen-but-unnoticed ways. For example, the data illustrate a transition between viewing a video and discussion may be accomplished by at least 3 different kinds of participant: one who instigates the transition, another who presses the 'pause' button on the video, and a third who aligns with the activities of the others in orienting away from the video towards the conversational circle. All of these different constructed modes of participation co-emerge over time seemingly unproblematically. None of these different participation modes alone could accomplish the transition so it's the coordinated effort of all participants that drives social interaction forward. Another example shown in this chapter is the case of establishing a seeming shared sense of the actions and talk of one participant. Again, it's the complex and delicate orchestration of the activity across participants that culminates in

seeming shared sense of alignment. None of the participants act in the same way, and yet they all contribute to the overall collaboration.

The MOOC interface doesn't support participants in taking up a number of different participatory modes. For example, by dis-aggregating discussion forum materials from video content, participants are not able to take up peripheral participatory modes - simultaneously viewing video while engaging peripherally in discussion. A sense of the activities of others, which are a resource for students to construct different modes of participation, is also not available in the current MOOC interface.

- **Timing:** As in the previous chapters, we find that discussion is generally interleaved with video-viewing. The two activities are tightly coupled and help give each other meaning. The meaning of the conversation by the participants, is provided by the video that plays near to the time that the discussion starts. Furthermore, the participants navigate and accomplish appropriate timing with respect to each other's activity by drawing on the details of each other's gestures and talk that are available in the co-present setting. The emergent nature of the activity of the group also serves to highlight that no one collaborative activity is contained within a single instant, rather any particular act may take place over an extended period of time, dependent on a complex and delicate sequence of interaction before it, during it, and after it.

  The current MOOC video interface doesn't reflect a sensitivity to the timing of discussion with respect to the video content. Rather, discussion and video contents exist in two separate interactional streams (Chapter 3). This restricts interactionally the ways in which discussion around video can emerge.

The data suggest several differences and similarities related to group size. One key difference is the complexity of how the group coordinates their actions and talks in accomplishing social interaction. For example, we saw in the studies that playing/ pausing a video for the larger group contains many more modes of participation, with all groups members collaboratively accomplishing transitions in activity. In the smaller group, fewer members had to co-accomplish collaboration through coordinated action and talk, so each member had to take into account fewer actions from others in the setting. However, there are many similarities between the groups. The themes surfaced above were seen across both studies, suggesting they may be agnostic to group size. That being said, it's possible other issues would arise should the group scale to much larger sizes. This is discussed further in Section 8.6 when areas of work for the future are outlined.

Aligning with an iterative approach to system design (Chapter 4), I will use the insights in this chapter along with previous ones to generate requirements for the system in Chapter 7. The following implications for the system supporting the distributed group in the third phase are proposed:

1. Although it may not be possible to give distributed participants online the kind of public availability of each others' speech and gestures that seems to central for the co-present group, I suggest it may be possible to give online participants a sense of awareness of the presence and activity of other. The limitation of the online environment is that it's difficult to reproduce the ability to be sensitive to all the rich physical cues of others that we get in a physical setting, in a virtual one. One can give some awareness by firstly allowing participating simultaneous visibility of the video, annotation writing, and conversation by others. Secondly, the interface may be

designed to include some indicator that keeps participants aware of new changes in discussion threads.

2. The data illustrate how conversation may develop around one instant or segment of video, therefore an implication for the prototype is to attempt to preserve the capacity to have a thread of conversation associated with a moment or time segment in the video.

3. While the prototype in this phase supports annotation only to an instant of the video, the data illustrate that video segments may commonly form the basis for discussion. Therefore, an implication for the prototype is to allow for annotation not only to an instant in a video but also to a time segment.

In Chapter 7, a prototype incorporating the implications just outlined will be used in an experiment by a distributed group of participants. A general guiding question for Chapter 7 derived from this phase is: Given a prototype which attempts to support some of the kinds of interactional accomplishments seen in this chapter, what insights can I gain about the nuance of the practice of social interaction of a distributed group? How do participants in a distributed group accomplish their social interaction, what evidence can be gleaned about how they orient towards the social interaction and accomplish it by looking at the reasoning that participants undertake? How do they orient toward social interaction when confronted with many different conversation threads?

Chapter 7.

Phase 3 - Exploring Interactional Issues at the Root of Intra-Video Social Interaction for Distributed Groups

7.1. Introduction

In this Chapter, I will undertake the final stage of the 3-part iterative development of a proof-of-concept video annotation interface. In this stage, I will explore the effects that a new proof-of-concept interface has on intra-video social interaction of a distributed group, and surface a set of issues that can be taken up further in future work. The features of the interface draw on the insights about intra-video social interaction revealed in previous chapters.

Methodologically, I will take a 'blended' approach that uses both a 'big data'-based view on interaction, as well as an ethnomethodologically-informed one. In previous chapters, I've made regular reference to the way 'big data'-based approaches to studying the MOOC can contrast and complement a more qualitative ethnomethodologically-informed approach. In this chapter, I will explore one way for how the two approaches can be used together to reveal a fuller picture of intra-video social interaction.

The methodological setup involves a small group of distributed students as shown in Figure 7.1. The goal of the setup is to simulate a distributed group using the platform synchronously. A pair of participants is used at each site in aligning with the Constructive Interaction tradition (see Chapter 4 for details about the rationale behind

this approach). The participants' activities are collected through two channels: the first is using the video footage, and the second is by collecting the ultimate contributions and actions of the participants on the platform which could be collected remotely. The latter mode of data collection mirrors that of a 'big data'-based study.

Figure 7.1. Three pairs of participants – A and S (left), C and J (centre) and B and M (right)

In the study, the participants are asked to view an educational video at the same time, to set up the possibility of synchronous intra-video social interaction. The participants are given the open-ended task to watch the video and consider talking to each other if they want to. The interface they all use is also pre-populated with existing discussions taken from the real-world educational platforms.

The study has two parts: in the first part, the participants perform the task using the Coursera interface, while in the second part they use the new proof-of-concept interface. This is to provide for a comparative study of how each system shapes the details of the mechanics of interaction.

There is overlap of this with existing research into computer-mediated communication (CMC). Methods for CMC research continually evolve as the technologies that support it do. Different facets of CMC can be taken up for study, with different analytic approaches (Romiszowki and Mason 2004). My interest in details of interaction around intra-video social interaction is perhaps most closely related to two genres of CMC

research. The first takes a conversation analysis-based view on text-based online discussions like in forums across a number of domains. For example, Jirotka et al. (1991) unpack the conversational issues in a series of discussion thread posts among distributed participants playing a game online. As another example, González-Lloret (2011) examines the way conversation tends to support learning of a new language. That there are interactional problems in forums has long been recognised in the literature. Herring (2006) unpacks some of the problems with text-based discussion forums in terms of interactional coherence, and the reasons why, despite the limitations of forums, they remain popular. Most of this first genre of CMC has focused on the traces of social interaction that can be collected remotely. The second main genre of CMC that most directly informs some of the work in the chapter looks at synchronous computer-mediated communication in action, focused on capturing video data of a participant actively undertaking CMC. Work on this has been done across a number of different domains. In his seminal works, Whalen unpacks the way interaction unfolds in a number of workspaces including call centres (e.g. Whalen et al. 2004) and public safety communication centres (e.g. Whalen 1995; Kameo and Whalen 2015). Rintel is a pioneer of examining the impact that voice- and video-over-IP technologies have in interactions, taking a closer look at how sent video messages are constructed through interaction and how they are then received (e.g. Rintel 2010; Rintel 2015) surfacing interesting interactional issues at the heart of this kind of communication. The CMC supported through more experimental technologies like media spaces are a subject for more experimentation (e.g. (Luff et al. 2011)).

For my experimental setup, I am interested in capitalising on all the forms of data that I can gather: the video data of each pair of participants as they interact with each other and the platform, as well as the forum posts that I can collect remotely. In that sense,

the data that I'm looking at and some of the analytic approaches I bring to this study draw a bit on both genres of CMC work. The result is that I can track end-to-end the interactions through which a forum post is created, ultimately what text is submitted online in the post, and interactionally how participants receive that post. This gives a rich view of CMC that hasn't substantially been captured in existing literature on the MOOC. An alternative approach to this study would be to just look at the data collected online, or just focus on the interactions captured in the video. However, this would not give as full a picture of the development and construction of a forum post therefore I leave this alternative for potential future work.

I will show how through my chosen approach I am able to see aspects of interaction which could be captured remotely, as well as the detailed aspects of interaction captured locally at each site in the video data. The video data reveal how the participants orient towards the video content and discussion forums, including interactionally what drives and shapes the creation and submission of their own contributions to the platform. Since the study includes the participants using both the Coursera and new proof-of-concept video interfaces, it illustrates some of the ways the new interface transforms and better supports intra-video social interaction.

Over the course of this chapter, I will cover three patterns in intra-video social interaction that are chosen because they occur very regularly, and appear central to the way the participants accomplished the task in this study (Section 4.5). The three aspects are: how video viewing is interrelated to interaction with discussion forums, the interactional underpinnings of how misalignment around discussion forums can develop and persist, and the stepwise progression by participants into developing and submitting a contribution to the platform. Before discussing each of these aspects, in

the next section, I will review the main features of the proof-of-concept interface, and how they draw on the insights from previous chapters.

7.2. Key Features of the Proof-of-Concept Video Annotation Interface

Building on the studies from previous chapters, the key features of the proof-of-concept video annotation interface draw on the surfaced understandings around intra-video social interaction. The previous analyses reveal important nuance of the mechanics of intra-video social interaction, and expand aspects of existing conceptualisations of the interactional resources that support intra-video social interaction, including:

- The key role of awareness of others' activities

  The data have shown the varying ways that participants emergently co-produce social interaction, by orchestrating their talk and gestures. Visibility and public availability of each other's activity plays a central enabling role in this.

- The important link between a participant's contribution and the associated video fragment

  As social interaction unfolds, analysis in previous chapters shows the important relationship between a contribution (such as a spoken statement in a co-present group) and a fragment of video to which it related. Typically, the analysis shows that the contribution is linked to the fragment by the contributor, either by explicitly referring to the relevant fragment in talk, or through the relative timing of the contribution to the video fragment (the participant speaking soon after the relevant video fragment). In the co-present group, the participants' experience of both the relevant fragment and the

contribution, serve to be a resource for the group in seemingly relatively unproblematically find alignment around a contribution.

- The seeding of a discussion around a video fragment

  The data show the ways in which discussion can develop, based on one person's contribution around a video fragment.

There are four key features that are implemented in the new interface, and whose effects will be explored in the analysis. Each feature is described in Table 7.1, alongside the insight on which they draw. The features will be explained in detail below the table.

Table 7.1. Insight on Intra-Video Social Interaction, with corresponding feature in proof-of-concept interface

Insight	Feature in Proof-of-Concept Interface
- The important link between a participant's contribution and the associated video fragment	Simultaneous visibility of the video and the discussion forums.
	A widget that is always visible and which indicated synchronously when others have made a contribution, acting as a synchronous indicator of the activities of others. If you click the widget, it displays the recently added contributions of others in the forums.
- The important link between a participant's contribution and the associated video fragment - The seeding of a discussion around a video fragment	Ability to tag a discussion forum thread to a moment or fragment of video Direct time-tagging of a discussion forum post (and subsequent discussion) to a moment in the video so that it can be reviewed/viewed alongside that forum post
	The ability to view the video moment or fragment associated with a discussion forum.

The proof-of-concept interface is developed by adapting an existing system (Pea et al. 2004)[12]. As shown in Figure 7.2, the interface provides simultaneous visibility of the

[12] I wish to thank Pea and his collaborators for assisting me with this system.

video and the discussion forums: a view of the video on the left, and discussion threads on the right. Each contribution by a participant has an associated time stamp, linking it to a moment or fragment of the video.

Figure 7.2. Screen capture of prototype

An example discussion thread is shown in Figure 7.3. After the first post in the discussion thread, the other contributions are labelled as comments. At the top of the first post is a time-stamp or thumbnail which can be selected, and which will then replay that moment or fragment of the video so that it can be viewed alongside the discussion forum.

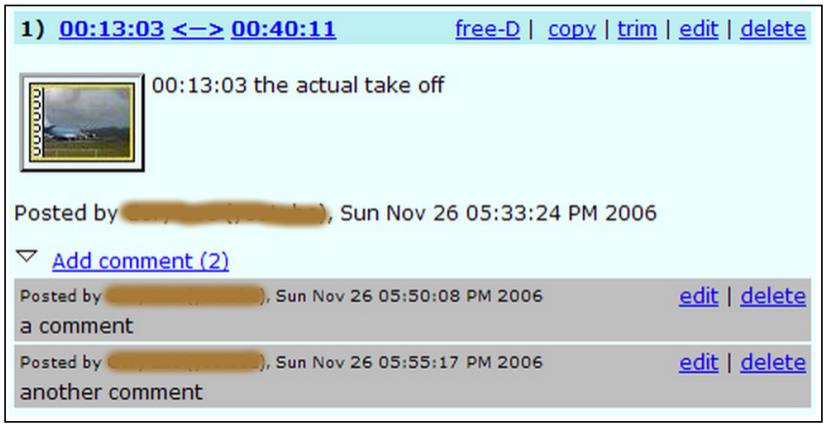


Figure 7.3. Screenshot of prototype discussion thread with contained comments

Furthermore, the system provides an update widget that sits above the video as shown in Figure 7.4. The update widget notifies a user in real-time of new contributions by

other users by turning red as well as indicating the number of updates (Figure 7.4, right). Users may then click 'update' as shown in Figure 7.4 to view the new contributions.

Figure 7.4. Screenshot of the update widget showing no updates (left) and one new update (right)

Having reviewed the key features of the new proof-of-concept interface, I will now proceed to the data analysis that reveals some of the main ways that these features appear to shape the details of interaction around educational video content.

7.3. Interactional Integration of Video Viewing and Discussion

In this section, I will explore the ways in which the two interfaces shape the mechanics of intra-video social interaction. In a sense, this is at the heart of this thesis, as it speaks directly to the disintegration of the social interaction from video viewing, that was identified through an analysis of the Coursera interface flow in Chapter 3. In this section, I perform a comparative analysis by dividing the study into two parts: in the first part, all participants use the existing Coursera interface, and in the second part, the participants transfer to using the new proof-of-concept interface. This allows me to capture data on interactions with both systems in order to compare and contrast them. This section will show that, although both systems have a video interface and have discussion forums, they support two fundamentally different modes of intra-video social interaction.

I start by considering a typical transition between video-viewing and engagement with discussion forums, for the existing Coursera interface. As I will show, the analysis suggests that this transition appears to depend on participant's expressed recollection of particular words and terms, previously seeing in discussion forum titles. An example of this is shown in Fragment 7.1 shown in Figure 7.5.

In Fragment 7.1, just as it starts, the participants are viewing a video in the Coursera interface (indicated by a transcript of the video 'Vid' in Line 1). In Line 2, J makes a verbal comment as the video keeps playing, that's related to the contents of the video ("Apart from the government", Line 2).

Fragment 7.1

1	Vid:	*No one cares but it's good to assume that in the future it may be*
		[
2	J:	Apart from the government
3	C:	Well ye:::s (1.0) I think someone posted about (0.2) you know
		[
4	J:	I think actually there was:: (0.2) something along those
5		lines so::: let's (0.2) let's see what they (0.2) let's go back here
		[
6	C:	Yea::::h
7		There was something about GCHQ

Figure 7.5. Fragment 7.1 transcript

In response to this, at the beginning of Line 3, C pauses the video to say that she recalls someone posting about it in the discussion threads ("I think someone posted about… you know", Line 3). J appears to almost simultaneously also recall the post ("I think actually there was::..something along those lines", Line 4) and suggests they go view it

("let's see…let's go back there", Line 5). C recalls a particular term used in a discussion thread title, 'GCHQ' ("there was something about GCHQ", Line 7).

After this fragment, the participants turn their attention to the discussion forums; appearing to find the forum that has the term 'GCHQ' in it. After finding it, they read its contents, and appear to come to the conclusion that the forum does not cover the facet of data security they were expecting. The discussion forum they found covers a news story about a political facet of GCHQ, which the participants express they don't find interesting.

It is informative to consider the ways in which the interface offers resources and affordances to connect a discussion forum to a portion of educational video content. The data suggest that the video interface offers minimal resources - and rather the participants appear to draw on prior experience viewing the discussion forums. In the sequential unfolding of interaction in the fragment, the shift in attention towards viewing the discussion thread seen above, seems to depend primarily on the the participants' apparent capacity to recall a word ('GCHQ') in a previously seen discussion thread.

Fragment 7.1 shows the potentially problematic nature of the seeming dependency on recalling words in a discussion forum title, to seemingly link video content and discussion forums. First, the linking depends on a participant having witnessed that discussion forum in the first place, which may make it difficult for a student just starting the course who hasn't spent much time reviewing the discussion forums. Secondly, it also depends on being able to recognise and understand those terms in the right way. Other participants may not have recognised the term 'GCHQ', and even, as in this case, when participants recognise the term, the forum discusses an aspect of the

topic which seems to be different to the particular facet of the topic in which the participants express interest. Interactionally, the participants don't appear to draw on any resources in the interface to facilitate the interweaving of discussions with video-viewing. As seen in the fragment, this may present substantial challenges to the student being able to contextualise discussion forum posts within educational video content. Since the transition appears to be primarily triggered and driven by the participants themselves rather than any aspect of the interface, the transition is of a 'proactive' nature.

On the other hand, the details of the mechanics of the interweaving of discussion forum engagement with video viewing, appear different for the new proof-of-concept interface. The first feature that appears to play a central role is the update widget. The data show that participants regularly orient towards and use the widget: discussion within the pairs shows that they typically alter their activity in response to the widget reflecting a new contribution by others. An example of this is presented in Fragment 7.2.

Just before Fragment 7.2, the participants are viewing a video, and the update widget changes status to show a new contribution has been added by other participants. As the video continues to play, in Line 1, S appears to invite A to check the new contribution using a question as indicated by the rising shift in intonation ("shall we check the new response↑", Line 1).

Fragment 7.2

1	S:	Shall we check the new response↑
2	A:	M:: (1.0) shall we do that↑
3	S:	((Yea:::h it looks interesting))
		[[
4	A:	Yeah Shall we pause the video

Figure 7.6. Fragment 7.2 transcript

A expresses agreement in Line 2 (“M::”, Line 2) and Line 4 (“Yeah”, Line 4). There appears to be some reasoning around whether or not to pause the video and view the update (“Shall we check the new response↑”, Line 1; “shall we”, Line 2), but then the participants pause their viewing of the video in response to the update widget.

The widget thus appears to be an instrumental resource in creating this shift. As opposed to Fragment 7.1 where the transition is ‘proactive’, here the transition has a more ‘reactive’ nature since the transition is instrumented through an interface feature rather than requiring the participant to independently instigate it.

Furthermore, the concurrent availability of the discussion forums alongside the video in the new interface, and the fact that a portion of the discussion is shown in the interface (rather than a title) appears to open up new possibilities for discovering forum content. This is in contrast to the Coursera interface which requires explicitly changing interfaces to search for a discussion forum. An illustrative example of this is shown in Fragment 7.3.

Just before Fragment 7.3, B and M are reviewing others’ contributions in the prototype. In Line 1, B directs M’s attention to a particular contribution made on a video segment using the mouse on the screen. B expresses interest in Line 1 (“this is

interesting", Line 1) to which M appears to mostly agree in Line 2 ("yeah", Line 2). The discussion thread they are viewing is shown in Figure 7.8.

Fragment 7.3

1	B:	Uh::m (0.2) I mean this is interesting how they say this
2	M:	What (0.2) what yeah
3	B:	Play
4	M:	(8.0) ((laugh))
		[
5	B:	((laugh))

Figure 7.7. Fragment 7.3 transcript

Transcript:

1 2 3 4	Love the way they've been talking about profit and how much business can get from big data, then suddenly there's a couple of shots of children as if to make this more "human". If this is about profit, SAY it's about profit. Don't try and pretend it's about 'think of the kids' :)

Figure 7.8. Screen capture and transcript of discussion thread viewed by participants in Fragment 7.3

As shown in Figure 7.8, the original contribution has a joking quality to it, making fun of how the video displays an image of children among what is perceived to be a more serious discussion about big data in business. The emoticon at the end of the contribution (":)") in Figure 7.8 reflects a playful nature. When B says "Play" in Line 3, he clicks on the contribution to view the video segment to which the contribution is associated. During the (8.0) pause in Line 4, they are viewing the video segment. As the video segment reaches the part showing the video frame with the children, both M

and B laugh as shown in the transcript in Figure 7.7 in Lines 4 and 5. Just after the fragment, B and M add a comment to the thread, about how funny they found the video segment.

The fragment illustrates how participants are being drawn into discussion forums, and how this is facilitated by the visual availability of the forum in the interface. The participants appear to unproblematically rapidly engage with the many forums displayed in the interface - since the video data reveal they direct their gaze towards the forums, and tend to call out and discuss only particular ones in which they express interest. An analysis of the data collected shows that generally the participants spend less time establishing a sense of the nature of the forums, compared to the Coursera platform where forums are presented as a long list of forum titles. This reduced barrier to finding threads of interest in the new interface, may increase the chances of the user making their own contribution. At the same time, it's worth calling out that a list of forum headings is a sensible first step in curating forums for a large group in a scalable way. This data just highlight some of the dangers of this, that can be used as part of the ideas that underpin the design and assessment of the MOOC as a platform. There are no substantial tried-and-tested interface alternatives to heading-based forums for large-scale platforms, and new research into visualisations would be required. I will come back to this when I draw out my overall contributions and future work, in Chapter 8.

Fragment 7.3 also illustrates the important role played by being able to explicitly associate a forum post with a video fragment, in supporting social interaction in the new interface. Since those viewing a forum post can view the same portion of video that the forum creators did around the time they created the post, this feature

supports a kind of shared experience of the video. In this case, an analysis of the video data show how both the creators and readers of the forum post laugh at a very similar moment in the video fragment: the moment when the frame with children is shown. Thus, the capacity for any participants to see this associated video fragment creates a kind of alignment in their reactions. As part of the sequential unfolding of the interactions of participants B and M, this shared experience appears to have the effect of drawing the participants into the discussion since the pair proceed to submit a comment to the thread.

And so, the data suggest that although both systems support access to, and transitions between videos and discussion forums, the mechanics of the arc of the transition is quite different. In the case of Coursera, the transition has a more *proactive* quality - with a seeming dependency on prior experience with the platform. In the case of the new interface, the interface appears to be much more of a resource to allow participants to be drawn into discussion threads and thus has a more *reactive* nature.

This section shows two different patterns of interaction for transitioning between video-viewing and discussion forums. Once a conversation begins in the discussion forums, there are a number of potential interactional challenges and pitfalls. One example of how this can happen, with the nuanced mechanics of interaction unpacked, will be explained in the next section.

7.4. Potential Persistence of Misalignments Around Contributions

In itself, it's not a new insight that conversation through discussion forums is associated with interactional challenges and difficulties, particularly when compared to in-person conversation (e.g (Eastman and Swift 2002; Thomas 2002; Vonderwell 2003; Morrow

2006)). However, the methodological approach taken in this section provides a way to see interactionally what might underpin some of these issues, which has not been done substantially before. I illustrate these issues by studying the ways in which the participants engage with one forum post. Using the video data, I am able to capture the interactions of the pair C and J when they created a post, as well as the interaction of the other two pairs who view and react to it.

Fragment 7.4 shows a part of the creation of the forum post of interest in this section. The contribution is made by pair C and K while they are watching a video on 'big data'. The particular frame of the video they reference is shown below in Figure 7.9.

Figure 7.9. Screenshot showing the logo discussed by participant in Fragment 7.6

In the data, J appears to spot the logo to the left of the video speaker's name in the bottom left-hand corner of the frame, as shown in Figure 7.9. J calls out the logo and points to it. He mentions to C that he believes the speaker is from the Saïd Business School. C encourages him to add a comment about it but, as shown in Fragment 7.4 (Figure 7.9) in Line 1, C asks for clarification about *how* J knows this to be the case ("How do you know↑", Line 1).

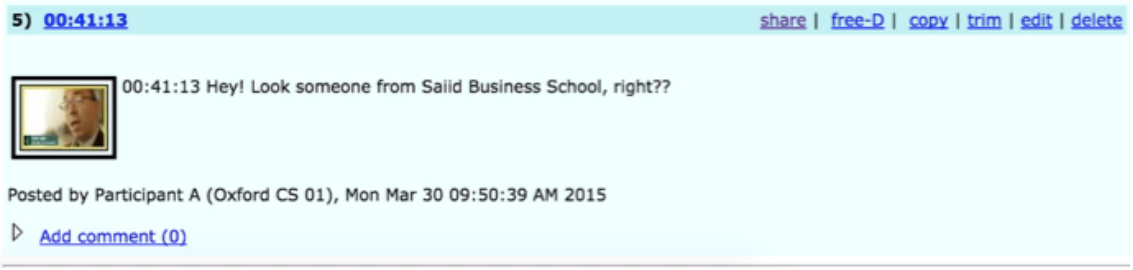
Fragment 7.4

1	C:	How do you know↑ s-
2 3	J:	Schoo::l (0.2) I think I think the logo is (0.2) maybe I'm wrong maybe it's the logo of this format I d-

Figure 7.9. Fragment 7.4 transcript

To C's question, J responds in Lines 2 to 3 with a clarification that he believes the logo is of the Saïd Business School ("Schoo::l...the logo", Line 2). However, J does not express this confidently, adding that he thinks he may be wrong ("maybe I'm wrong", Lines 2 to 3). After the fragment, C encourages J to create a contribution regardless of his uncertainty.

The final text contribution by C and J on this topic is shown in Figure 7.10. The contribution may be said to reflect J's uncertainty in whether the logo is that of Saïd Business School because the contribution ends with "right??" which represents a question suggesting he seeks a response from others.



Transcript:

1	Hey! Look someone from Saiid Business School, right?

Figure 7.10. Screenshot and transcript of the discussion thread added by participants in Fragment 7.4

An analysis of the submitted contribution shows no mention of the logo, which we understand to be the reason behind the post based on the video data collected. Since

there doesn't seem to be an explicit agreement between C and J to leave out mention of the logo in the contribution, this seems to occur without the participants openly planning to do so.

As I will show next, this is an important consideration when considering the ways the other participants seem to orient towards the contribution. I will now consider the response by each pair to this post in turn.

Firstly, B and M appear to recognise that the discussion forum creator may have seen and recognised the Said Business School logo. However, they don't appear to reach a strong agreement nor do they draw on the forum itself to come to this conclusion.

A portion of B and M's reasoning while viewing the text contribution is shown in Fragment 7.5. Before the fragment, M expresses some uncertainty in how the contributors know that the speaker is from Saïd Business School. B then states, as illustrated in Line 1 of the fragment, that B recognises the Saïd Business School logo because he knows the logo as prior knowledge ("I know the logo…check…this is the Saïd business school logo", Lines 1-2).

Fragment 7.5

1	B:	I- It's I know the logo so OK let's check this (1.0) will it play this part↑
2		(0.2) yeah (0.2) this is the saïd business school logo
3	M:	Oh may- maybe he did that before seeing the whole video (0.5) because
4		at some point it says like business school in the video

Figure 7.11. Fragment 7.5 transcript

During "this (1.0) will it play this part ↑" (Line 1), B clicks the thumbnail associated with the contribution which brings up the video frame from Figure 7.9. In Line 2, B moves the mouse over the logo to show M when he says "this is" in Line 2. Even with

B correctly (though unbeknown to them) identifying how the contribution creators, C and J, knew the speaker was from the Saïd Business School, M still appears unconvinced. M suggests that he believes he recalls the video mentioning the Saïd Business School at another time ("at some point it says like business school in the video", Lines 3-4). It turns out, therefore, that each participant seemed to think C and J's comment was obvious, but for different reasons. B knows the logo, where M recalls seeing the words 'saïd business school' at a different time during the video.

In the case of B and M, the video frame with which the contribution is associated, is a resource for the pair to offer a seeming explanation for the original contribution.

This may be contrasted to the response by the other pair - A and S. The interactions of this pair illustrate how they appear to come to a shared interpretation of the discussion thread, but don't refer to the logo at all.

Just before Fragment 7.6, A and S have clicked the thumbnail associated with the contribution in the prototype and are viewing the video frame (shown previously in Figure 7.9) with which C and J's contribution is associated.

Fragment 7.6

1	A:	Oh OK so they added some extra information about where he's where he's from
2		(1.0)
3	S:	Or maybe where he might be from

Figure 7.12. Fragment 7.6 transcript

In their speech, there is no evidence that A and S recognise the logo or consider it to be the reason why C and J added the discussion thread. Rather, they appear to assume C and J have some extra information ("OK so they added some extra information", Line 1). S seems to pick up on the tentative nature of the statement ("Or maybe where he

might be from", Line 3), likely due to the question marks at the end of the contribution ("Hey! Look someone from Saiid Business School, right??"). After the fragment, the participants leave reviewing this contribution and return to watching the video. So, although, like participants B and M, they also review the fragment associated with the contributions and appear to search for clues in it, they don't express an orientation towards the logo.

These two fragments illustrate the varying ways that students can orient to one discussion forum post. Both participants appear to express assumptions about the nature of the post. While we might expect the participants to primarily draw on the contents of the contribution in order to seemingly gain a sense of it, it appears that it is only one of a number of resources they may draw on. Despite the fact that there is no explicit confirmation of their expressed interpretation of the post, both pairs of participants appear to come to a sense of meaning of the contribution in a way that neither appear to express as problematic. Neither pair also suggest asking for clarification within that forum thread.

The data analysis in this section is a particularly good example of the potential advantages of using both the video data and forum post data into analysis. By contrasting the various ways that a message is constructed on one hand, and then ultimately received by others, we can draw out interactional issues that couldn't be picked up if only the forum text alone were analysed.

These fragments illustrate the rich ways that participants can orient towards a post in the discussion forum, that would not be apparent if only the 'big data' about the interaction were used. Big data would support only the collection of the final post and perhaps the fact that the other participants had spent some time viewing it. These

potential communication pitfalls and some of the interactional origins behind them seen in the video data, couldn't be picked up. This study shows how a view towards the details of interaction through an ethnomethodologically-informed approach, can surface important interactional complexities in how participants orient towards a contribution and its implications for a shared orientation by the distributed group. On a more individual level, there are also important complexities associated with the creation and posting of a contribution by a participant in the platform, which will be discussed in the next section.

7.5. Stepwise Progression Through a Contribution to the Platform

Having considered some of the ways social interaction has been shaped by the new interface, in this section I will shift focus to consider what the the study reveals about the interactional details of how a potential contribution to the platform is developed and submitted to the platform. In particular, the data show that both in the case of the old and new interfaces, there appears to be a stepwise progression into contribution.

From a 'big data' perspective, we typically have access to the acts of the participants contributing to a forum, like clicking the 'submit' button, and the final text. The analysis here shows that this process is emergent. It illustrates some of the key resources that the participants draw on, as part of this complex and delicate unfolding sequence of interaction. The analysis shows that the development of this contribution has a stepwise progressive quality with a number of subsequent stages. Understanding the issues at play at each stage, as they are surfaced through the interactions of the pairs, reveals new opportunities to design the interface to support contribution.

The notion of stepwise progression into and out of activity echoes a substantial amount of seminal work in Conversation Analysis across a number of different settings (e.g. (Jefferson 1984; Heath et al. 1994)). At its core, the notion reflects recognition that any one act is part of a larger sequence, closely related to what comes before and after it, developed over time. That same conceptualisation of activity is unpacked here for the creation of a contribution.

Activity towards a contribution typically starts with participants expressing interest in a video fragment, near to the time of viewing it. This video fragment then forms a seed for a discussion between the participants and the contribution to the platform. If the participants agree to create a contribution, this begins a process that involves a series of criteria through which the participants shape their contribution. Rather than a contribution being an echo of what is discussed by the participants, the contribution appears to be *produced* through the discussion and only reflects a part of it. A comparison of the ultimate forum posts, with the interaction data captured in video, show the extent to which the post only captures a portion of the local discussion of the pair.

For example, Figure 7.13 shows an example of a contribution submitted by pair C & J. This is the typical artefact of interaction that could be collected remotely, en masse, through 'big data'. The comment is a response to a forum post about a story where the chain store Target predicted a teenager was pregnant based on her purchase and online browsing history.

1 2 3 4	Yes, I've also heard about that story. It might make people more aware of privacy concerns about big data. However the average consumer might not know how they unknowingly share data with retailers and even if they knew, they might not be willing to adapt their behaviour to change this.

Figure 7.13. Transcription of comment contributed by participants

Since we have video data showing the interactional circumstances through which this response is constructed, I can see that there are aspects of the discussion between the two participants that is not reflected in the contribution. An example is shown in Fragment 7.7.

Fragment 7.7 contains a substantive point, made by S over Lines 1 to 2; that retailers care about most people not those doing research in privacy or related topics. A expresses agreement with this in Line 3 ("That's right yes", Line 3).

Fragment 7.7

1 2	S:	Or at least most people (1.0) and as retailer you care about most people you don't care about people doing research in privacy or something
3	A:	=That's right yes

Figure 7.14. Fragment 7.7 transcript

The point made in Fragment 7.7, related to the concerns of retailers with respect to people's opinions on data security, is not mentioned directly in the contribution.

The data suggest the production of a contribution, including how certain aspects of discussion are filtered out of the contribution, is a complex and emergent process. The participants appear to undertake this without explicitly stating so, although an analysis of their interactions provide some insight into the kind of criterion they apply in shaping the contribution and deciding which aspects of their local discussion goes into the contribution.

I'll now walk through the main steps in the progression, illustrating some of it using data fragments.

Firstly, there is a seeming quality assessment stage: where participants discuss whether a potential contribution fits within the standard and flavour of existing posts, and how it compares to other posts. These are typically subjective assessments of what's been seen by the participants in the forums. Once the participants express a decision to continue creating the contribution, the participants continue to shape the contribution based on a new set of criteria. For example, in the post-experiment interview, the participants express that they felt obliged to keep their text contributions relatively short, and in developing the text contributions, aimed to remove any aspects of the contribution they felt were not central to their point, or were 'obvious' or natural extensions of the points they already made. These criteria are not pre-established or articulated by the participants in the study.

The discussion by the participants also reflects a seeming awareness of an 'audience' to the contribution: the other participants that will be reading the post. An example of this is shown in Fragment 7.8. Just before this fragment, the participants complete typing a contribution and are both directing their gaze towards the text. A suggests a modification to the text, adding a further text description: "maybe we should sa::y… more details are found here or something like that" (Lines 1-2).

Fragment 7.8

1	A:	Well (0.2) maybe we should sa::y (1.0) uh (0.5) this this gives more details are
2		found here or something like that
		[
3	S:	Yeah but people might figure it out
		[
4	A:	OK they'll figure it out yes
5	S:	=With XKCD

Figure 7.15. Fragment 7.8 transcript

A's response to this in Line 3 appears to reference the other MOOC students that may be audience to the contribution, using "people" to refer to them in "Yeah but people might figure it out" (Line 3), referring to the particular term in the contribution "XKCD" in Line 5 that S appears to believe gives the audience enough information ("people might figure it out…with XKCD", Lines 3 and 5). A appears to agree with this in Line 4, again seemingly making reference to an audience using "they" in Line 4 ("they'll figure it out yes", Line 4). The "they" in this case is a reference to the other students who may be audience to this post.

A final facet of the interactions of the participants caught in the video data, which is illustrative of the production by the participants of the final contribution, is the act of typing. The video data show that the typing of the contribution by participants is interweaved with the discussion through which it emerges. As one important aspect, the pace of typing appears to coincide with an expressed sense of confidence in what they're typing. Furthermore, the way the text is edited reflects a changing sense of meaning of the contribution as it is being shaped during production by the participants. For example, initially participant A types:

1 2	However the average consumer might not know how they unknowingly share data with retailers so th

Figure 7.16. Transcript of initial piece of text typed by participant

Before deleting the letters 'so th' (Line 2 above) rather typing:

1 2 3	However the average consumer might not know how they unknowingly share data with retailers and even if they knew, they might not be willing to adapt their behaviour to change this.

Figure 7.17. Transcript of edited piece of text typed by participant

The addition of "even if they knew" (Line 2) changes the meaning of the contribution because it highlights a facet of the relationships between consumers and privacy issues. This is an emergent quality of the production of the contribution that is revealed only by considering both the ultimate contribution as well as the details of how interaction unfolds, which an ethnomethodologically-informed approach provides.

The issue of pace and editing while typing is related to the notion that a contribution is developed over time, and that there is important insight to be gleaned about a contribution by examining this process, as much as it is to analyse the final text contribution. This shines light on the value of this approach which examines both the final contributed text and the details of the interaction through which it emerges.

Exposing the stepwise production process associated with a contribution, and unpacking some of the ways that a contribution is produced through interaction (rather than being an echo of the interaction of the pair), can expand our conceptualisation of how students move into more active participation and may opens up new design possibilities for encouraging engagement. If we understand the criteria participants apply as part of contribution, we can design the interface to provide this information,

reassure participants, and support them towards contribution. I will return to this design issue again later in this thesis.

7.6. Discussion

The 'blended' methodological approach taken in this chapter serves to surface some of the interactional origins and issues around those artefacts of interaction that we typically see through 'big data'-based approaches. The understanding that social interaction by a distributed group can be difficult, particularly relative to a co-present group, in itself is not a new insight. However, these difficulties as they are manifest in the details of interaction, have not been substantially explored before until now. The use of an experimental setup that draws on Constructive Interaction (Chapter 4) provides, among others, a view of the interactional resources that the participants draw on, and the criteria they express as salient. By viewing data that can be collected remotely, alongside video data, it's possible to surface details of how intra-video social interaction shapes and is shaped by the features of the interfaces - both the existing Coursera interface and the new proof-of-concept interface.

The key interactional issues surfaced here should be seen as provisional, and can be taken up for further testing and development in future work (described in Chapter 8). Nevertheless it is useful to outline the key interactional issues that are apparent from the studies in this chapter:

- **Use of Artefacts in the Interface for Considering Discussion forums**

  The data show how the proof-of-concept interface better supports the alignment of distributed group members around discussion forums, by enriching a text discussion forum post with an explicit tag to a video fragment.

This can help give participants a clear sense of the meaning of the forum post, alleviating some of the dependencies seen in Coursera on seemingly remembering key words in discussion forum titles. However, the enriching of a forum post in this way, doesn't always create alignment, and the data show examples of how seeming misalignments can occur and persist, particularly where participants don't explicitly question or clarify assumptions they seem to make.

- **Use of Artefacts in the Interface for Shifting between Discussion Forums and Videos**

  The ways in which the interface make the activities of others easily accessible particularly when viewing video, appears to play an important role in shaping the mechanics of how and why a shift in activity occurs between video-viewing and engagement in discussion forums. In the case of the Coursera interface, the shift appears to be *proactive* on the part of the participant: dependent on their seeming recollection of discussion forum titles and prior experience reviewing the forums. In the new interface, the transition has a more *reactive* quality where the interface itself plays a central role in instigating the shift, re-orienting the participants away from the video and towards the discussion forum.

- **Engagement in Multiple Activities**

  The participants appear to unproblematically, and without an explicitly-stated strategy, navigate the simultaneous access to the video player, discussion forums, and update widget in the new interface. These aspects of the interface appear to be used as a resource for the accomplishment of quickly reviewing discussions, and in the discovery of forum posts of interest. The data suggest

these activities have the effect of making understanding posts easier, and of seeming to drawing participants into more active engagement with discussion forums.

- **Timing**

The data show the various ways that the interactions around and through the platform by participants are emergent; shaped and changed over time. This makes *time* a significant dimension to consider in studying interaction in the MOOC. For example, the analysis shows how contribution to the platform is a process that emerges over time, in a stepwise fashion, and therefore understanding this progression in sequence is an important part of designing for intra-video social interaction. As part of this sequence of interaction, participants may employ emergent criteria to develop a contribution. The criteria can also depend on how much time the participants have spent viewing the discussion forums, since they may refer to a sense of 'audience' when shaping the contribution. Also the way a contribution emerges over time - the ways it is typed and edited - reflects an important nuance in the process of developing a contribution.

Both the current MOOC and proof-of-concept systems are relatively static in their display over time. That is, they don't take into account that participants may require different kinds of information at different points in their journeys in the MOOC over time. For example, understanding that a contribution to the platform is an emergent process, one that may be foreshadowed by an assessment of the quality of a potential contribution by the student, opens up new possibilities for encouraging participants to contribute. At the moment the

participant is creating a contribution, it may be useful to alter the interface to show summary statistics of the enrolled participants and the nature of the existing discussion forums, to support the process of constructing that contribution. This can ease the barrier to contribution, and may help bring students into more active modes of participation. All these insights are grounded in a sensitivity to the changing needs of a participant over time.

From a design perspective, we can see how the key features in the proof-of-concept interface appear to make certain aspects of intra-video social interaction easier, particularly in reducing the barriers in meaningfully linking and contextualising discussion within educational video content. Also, a kind of shared experience of the video can be created, which is much more difficult to do for a distributed group than a co-present one. In other ways, several of the interactional challenges from the Coursera platform appear to persist. For example, several challenges around preventing misalignments around forum posts, are still to be addressed. These issues will be picked up in final chapter of this thesis, and folded into substantive design requirements.

In these studies in particular, it turned out the most significant examples of social interaction, in terms of moving the activity forward, was the interaction the distributed pairs of participants had with each other through the platform, synchronously. Therefore the focus of analysis here is the ways a group of synchronous users might interact with each other with only a relatively small set of asynchronous messages populating the forums. In future work, it would be interested to modify this methodology to populate the screen with more pre-existing forum posts (creating more opportunity for asynchronous communication). Alongside this, more work would

be required on how to help users visualise a large corpus of asynchronous messages (see Section 8.6 where this is outlined for future work).

In this Chapter, I've explored using a kind of 'blended' methodological approach to examine intra-video social interaction of a distributed group. The video collected gives a sense of the nuanced interactions through which the contributions and activities on the platform emerge. This gives an initial sense of how two different approaches - one looking at data collected remotely from the platform, and another looking at qualitative video data - could be used together to surface new insights. In the next chapter, I will pick up these insights and along with findings from previous chapters, and synthesise them into the overall contributions of this thesis.

Chapter 8. Discussion

8.1. Introduction

In this chapter, I will draw out the implications of my findings from previous chapters to present the main conceptual, methodological, and substantive contributions.

In Sections 8.2, 8.3, and 8.4, I will discuss the conceptual, substantive, and methodological, contributions of this thesis respectively. The conceptual contributions speak to the new ideas and issues surfaced about intra-video social interaction. The substantive implications are key considerations and issues, based on what we know about social interaction as a result of the studies, that should inform the requirements of these systems going forward. The methodological contributions bring to light the advantages and challenges associated with the ethnomethodologically-informed approach taken in this thesis, learned through experience.

In Section 8.5, I draw out how my work contributes to the wider discourse on the MOOC as a learning device. Finally, in Section 8.6, I reflect upon possibly fruitful future avenues of research based on this thesis.

8.2. Conceptual Contribution: Insights into Interaction

The studies in this thesis bring to the fore key fine-grained aspects of how participants accomplish intra-video social interaction, which serve to extend and enrich our understanding of interaction in the MOOC beyond most existing studies. In this section, I will chose three key themes that appear most significant and central to the ways participants accomplish intra-video social interaction.These conceptual themes are: the

manifold implications of a ‘click’, the mechanics of how participants transcend boundaries in space and time, and participation.

8.2.1. The Manifold Implications of a ‘Click’

Many existing studies into student-platform interaction perform analysis on ‘click’ data collected remotely. These may be clicks made on a number of different elements in the interface. Two of the most common types of clicks analysed in existing studies are: clicks on buttons to submit contributions to discussion forums (e.g. (Kizilcec et al. 2013; Whitmer et al. 2014)) and clicks on the play/pause button in the video interface (e.g. (Guo et al. 2014; Kim et al. 2014)). In most existing studies, the ‘click’ is a central unit of measurement, and subject to statistical analysis such as aggregation over time.

Typically in ‘big data’-based studies, there is an assumed one-to-one mapping between a type of click and its purpose, that allows the clicks to be aggregated and grouped together. An analysis of the statistical characteristics of a particular type of click is then taken as a proxy for different facets of interaction. For example, some studies use clicks on buttons that submit entered text to a discussion forum, as a proxy for engagement (e.g. (Kizilcec et al. 2013)). Other studies examine clicks on video control buttons (e.g. the play/pause button and rewind/fast forward buttons) to identify ‘problematic’ portions of a video (e.g. (Kim et al. 2014)). The rationale is that if students tend to pause and review certain parts of a video, the video fragments should be made easier to understand. This information is then a resource to the course professor who can improve the educational video content.

The analyses in this thesis show the ways in which an act like a click can have a one-to-many mapping to its effect or purpose in interaction. This holds for student-MOOC interaction and for intra-video social interaction.

At a more individual level, the reason behind a click to pause a video may be, for example: an attempt to reflect on whether a video is worthwhile and the following of an information scent to try find educational content that is perceived as more relevant to a student (Section 5.3), it can be a pause as part of a shift towards discussion forums (Section 6.2), or it can be a shift to other videos (Section 5.4). The pausing of a video is not necessarily an indicator of a problem, and all the clicks on the video controls can't necessarily be aggregated into a set of simple and succinct categories of activity.

In a group, the act of clicking, if made visible to others, can play a key role in helping everyone coordinating their activity, enabling everyone to accomplish their ongoing and emergent social interaction. For example, the analyses in this thesis showed how the progression into and out of a click of a button was a resource for others in the co-present groups in seen-but-unnoticed ways (Sections 6.2, 5.4). That similar kind of principle held true for the distributed group (Section 7.3), where the widget was used as an interactional resource, helping participants dynamically shift between video content and discussion in a way that they appear to have found useful. In those cases, a click that played the purpose of submitting a contribution to the discussion forum, triggered the widget and was therefore also a resource to other participants who used it as part of their whole experience of interacting with and through the platform.

I propose that a useful analogy can be drawn here to the concept of a 'manifold' from fluid dynamics[13]. In machinery, a manifold can be a pipe that branches into several openings. It captures the idea that one bulk of fluid originating from one source down a single pipe, can then be split and channeled into different destinations. Put differently, a physical phenomenon originating from one source, can be split up so that it goes on to have multiple and varied influence elsewhere: a one-to-many mapping between one act and its ultimate purpose and effect. A similar concept can apply to an act like a 'click' in the MOOC: the physical act itself doesn't necessarily simply map to one effect or purpose. It can play a heterogeneous and varied role in the context of a social interaction.

8.2.2. Transcending Spatial and Temporal Boundaries

In conducting studies that examine intra-video social interaction for both co-present and distributed groups, this thesis provides insight into the ways the accomplishment of social interaction by co-present participants maps to the distributed group who are now trying to transcend the temporal and spatial boundaries they face.

Conceptually, using the ways participants accomplish social interaction offline as a kind of ground truth for the study for distributed groups, can be illustrative. This has been shown to be the case in much existing literature on workplace collaboration in CSCW (de la Flor et al. 2010; Jirotka et al. 2005; Jirotka et al. 2013). I argue that similar types of parallels can apply to MOOCs, and be a meaningful way of constructing the conceptual picture of intra-video social interaction.

[13]"a pipe or chamber branching into several openings", https://www.google.co.uk/#q=define+manifold

Attempting to create and unpack how offline and online interaction echo each other, can be a key resource in understanding why social interaction in the MOOC is the way it is. This can help in the endeavour to design systems that better support social interaction to help participants transcend temporal and spatial boundaries. Drawing on my analyses, I see three key concepts around social interaction that meaningfully apply to both the co-present and distributed groups:

- Public availability of each other's activity helps drive social interaction forward

The capacity of group members to have an awareness of each other's activities, is a key part of how participants in the co-present group coordinate their activities and accomplish social interaction. For example, it's central to how they attended to multiple different resources at the same time (Sections 5.4.1, 5.4.2, 6.3), and transitioned between video viewing and discussion (Sections 5.4.2, 6.2). This awareness of each other's activities persists in the distributed group, where the participants used the widget (that reflects the activities of others) as a resource to move between video viewing and discussion, and regularly refer to other distributed users when formulating and submitting their contributions to the platform (Section 7.5).

- Immediate context of an act helps give the act a sense of meaning

Whether co-present or distributed, insight into the immediate context of an act by a participant, like the submission of a contribution to the discussion forums, appeared to be an important resource in the seeming attempt to understand the purpose of that act within the group's activities. In the co-present groups of Chapters 5 and 6, the coordination of everyone's activity in the moments leading up to and after an act like the pressing of a play/pause button on the video, shows how the progression of the

social interaction through that act was orchestrated through the activity leading up to and just after the act. It does not depend on the act alone. In the distributed group in Chapter 7, the participants' ability to review the video segment associated with a forum post, appears to be a central resource to the participants in seeming to make sense of that contribution. Being able to see the video that inspired the contribution appeared to give the participants a better sense of the meaning of that contribution. And so, giving visibility into the narrow context around a contribution in the interface, is a key resource for the participants as part of social interaction.

- Timing is an important characteristics of intra-video social interaction

The way activities of the group participants are coordinated over time, is a core interactional resource for the participants in the co-present groups in Chapters 5 and 6, where the analysis shows the sequential unfolding of activity to be delicate and emergent. A similar feature of activity was seen in the distributed group in Chapter 7, where the interface capability of anchoring a discussion forum post to a moment in time in a video, appears to be an important part of how participants seemed to make sense of that post.

The widget feature in the proof-of-concept interface (that provided participants with a synchronous sense of the activities of others), is taken up and used as a resource over the course of the study. The incorporation of this feature that supports synchronous interaction with others, appears to be a central part of how participants accomplished intra-video social interaction in a way that they described as easier than in the traditional MOOC interface. Participants also reported that this ability to see and respond to others synchronously is meaningful and useful to them as part of how they might ordinarily wish to interact with others around the video in the MOOC.

8.2.3. A More Complex View of Participation

The subject of participation is central in a substantial portion of existing research into MOOCs. A lot of research explores and unpacks the various ways that student participation in a MOOC contradicts traditional models around participation and learning (Section 2.2). In this thesis, the studies highlight some new insights around the way participatory categories are defined, and the inferences made from them.

Typically, the participation categories are grounded in big data metrics of interaction, like clicks or video views, and studies tend to take for granted that students generally remain within these same categories or move within categories only up to a couple of times within one course (e.g. (Kizilcec et al. 2013; Clow 2013; Santos et al. 2014; Grünewald et al. 2013)).

I have surfaced some new aspects of participation in the MOOC, because of the focus on the granular aspects of interaction in the MOOC. Firstly, the data show how modes of participation can change in a dynamic and emergent fashion, which goes against the view of the broad largely static categories of participation in previous studies.

The data illustrate how the way a person interacts with the platform depends on the alignment they seem to see between the content and their prior experiences (Sections 5.3, 5.4), and participants can be drawn temporarily into more active modes of participation depending on the exact contents of the interface. At other times, participants may temporarily shift into and out of more peripheral roles with respect to the rest of the group (Section 6.3).

In Chapter 7, we see for the distributed group how the extent to which a person engages with and participates in a discussion forum could emerge moment-by-

moment as part of a stepwise progression. In fact, a lurker and an active participant could start interactionally with very similar behaviours (e.g. both may start by looking at a particular discussion forum) but their sequence of interactions may diverge because, for example, a lurker may feel like they don't have a substantive enough contribution to make (Section 7.5). Therefore, these two categories of participation may be quite related and not distinct, which goes against much existing literature.

The studies also reveal other nuances around whether participation is *proactive* or *reactive* (Section 7.3). Participants seem to quickly move temporarily into reactive modes of participation in response to the widget when it synchronously changed status to reflect the activities of others. Participants appear to temporarily move into proactive modes of participation when they seem to recall having seen something in the forums which they interpreted as relevant to the video content they were viewing. And so, the nature of participation from these studies appears much more dynamic and emergent than the picture of participation provided by most prior work. This emergent nature of participation opens up new territories for exploring how we might both support the participation better, but also how we might be able to influence or alter the mode of participation of the student.

Most existing research also tends to group students into one type of participatory category, where I showed how one student can simultaneously accomplish different kinds of participation towards different artefacts and other group members. A student may be actively contributing to a forum, while maintaining more of a peripheral participatory mode towards video content (5.4, 6.3). The different ways of participating can be in flux and can overlap in complex ways.

8.3. Substantive Contribution: Implications of new conceptualisations of interaction on system design

My analyses provide a more granular view of how participants accomplish intra-video social interaction. This view illustrates some key mechanics of interaction that can inform important concepts underpinning requirements for the video interface of the MOOC. In this section, I will draw out a set of issues to inform requirements for a new type of MOOC interface which personalises the participant's experience on the site so that their interactions are supported dynamically. The items discussed in this section are not concrete requirements for designs. Rather, they can serve as part of a foundation for requirements going forward, and be built upon by other work in this domain (in important areas such as asynchronous interaction; see Section 8.6 for an outline on future work).

8.3.1. A Dynamic Interface that Varies the Information It Displays Over Time

The interactional resources and ways a participant accomplishes social interaction can vary dynamically, on a moment-by-moment basis. The interface (in both the current MOOC and the proof-of-concept one) were relatively static in what they display over time; the resources they provided to a participant over time. A key issue that should inform design of the MOOC in the future is a sensitivity in the changing needs of the participant over time.

A system that is informed by this sensitivity opens new possibilities to shape the participation of students. In Chapter 7, we see how there is a pattern of a stepwise progression into and out of contributing (or not contributing) to a discussion forum.

There are particular aspects of this progression that tend to move a participant towards contributing, like a sense that their contribution is valid among the others that exist (Section 7.5). So an example feature could be one where a participant is reassured and led through this process. This might help shift students who might otherwise be lurkers, into more active modes of participation.

Chapter 5 illustrates how participants appear to consider the relevance of content and navigated towards content that aligned with their expressed prior interest and experience. Understanding how to support this kind of 'information scent' particularly in the early stages of a course and over time, may be an important part of better supporting interaction in the MOOC.

At an even more detailed level, we see how the pace and rhythm of typing by a participant tends to correlate with an expressed sense of confidence or hesitation in a contribution, and this may inform a system that serves to reassure and lead students through these periods of interaction. This may assist students in becoming more active participants (Section 7.5), and shifting participants out of, for example, 'lurker' modes of participation.

There are key privacy and ethical issues surrounding tracking details of a person's actions online (Martin 2016). For example, there are important considerations around how to acquire user's permission to track (e.g. Chauhan and Rathore 2015), such as whether to make tracking opt-in or opt-out. I would like to flag this as an important topic that merits further study in the future to unpack associated considerations and how they might impact the MOOC's design.

8.3.2. Mutual Visibility of Activity is a Key Interactional Resource

The availability of each other's activities appears to be a central interaction resource for both the co-present and distributed participants. Chapter 6 shows how 'individual' and 'collaborative' activity are tightly interweaved and linked together. In that sense, it's impossible to support an individual without giving them the resources to navigate the activities of others, and have their activities be seen by others.

When designing a system so that it scales to support large groups of students, creative design work will need to be done in the interface to give a view of the activity that is easily understandable and intuitive to the student. A basic example of this kind of interface element is the widget in the study in Chapter 7. The widget system appears to be meaningful in helping people gain a sense of the synchronous activities of others, but may not be scalable to a wider group and so may need more consideration in future research. I suggest that some features that may be grounded in this issue and which could be useful are:

- An eye tracker system that automatically annotates a contribution with the visual activity of the creator, to help give more information about the context under which the annotation was created and thus reducing the chances of any seeming misunderstandings of the nature seen in Chapter 7.
- Some kind of a dashboard that shows simultaneously who is viewing various discussion forums of interest, and who is typing a contribution to various discussion forums.
- Providing a larger set of possible responses to a discussion forum such as the ability to 'like' a post, or perhaps be able to request some clarification with a click of one button. Right now, it's only possible to submit a typed response,

which doesn't reflect the wide variety and scope of responses that are available to participants in a co-present setting.

In a traditional physical classroom, it may be inappropriate for a student to split their attention between a teacher, and other students or resources. However, in contrast, these studies suggest that by shifting the classroom online, this kind of split of attention is possible and may be beneficial to the student. It's another example of how a social space, and some of the norms for social interaction within it, may be fundamentally transformed with the shift to online digital spaces, like seen in other domains such as game playing (e.g. Brown and Bell 2004; Moore et al. 2007).

8.3.3. Providing Simultaneous Visibility of Resources to Create One Cohesive Experience

Where the current MOOC interface tends to present only one type of resource (for example, forum, or video) at one time, the studies here showed the value of capitalising on participants' capacity to navigate and meaningfully use together a number of different resources at once.

The capacity for participants to ordinarily engage simultaneously with both video content and each other, was shown in several instances across the studies of the co-present groups in Chapters 5 and 6. There, the participants are able to dynamically interweave and overlay periods of interacting with content presented to them through a laptop (Chapter 5) or tablet (Chapter 6), and each other. This was done in seen-but-unnoticed ways and appears to be central to the ways they accomplish intra-video social interaction.

The same competencies around engagement in multiple activities at once, were seen in the distributed group in Chapter 7, where the proof-of-concept interface presented them with a simultaneous view of the video content, discussion forums, and a widget that synchronously reflected the activities of others. The participants are able to navigate this interface and appear to move between the various activities, seemingly drawing on various facets of the interface, unproblematically. The analysis shows this navigation was central to how they accomplish intra-video social interaction.

The participants in all the settings are proactive in orchestrating for themselves the way they use these various resources over time, and they appear to coordinate these resources as one cohesive experience. Therefore, it's important to provide the participants with the right resources to help them quickly and easily access and manage a number of different resources in the course. Doing this to a certain extent in the proof-of-concept interface in Chapter 7 creates a more intuitive experience for the participants, and participants report feeling they had more of the right kinds of interactional resources at hand to perform the kinds of activities they wanted to. This is in contrast to the existing MOOC interface where participants reported feeling 'blind' about where to find various resources in the course. Therefore, I propose that it's important to give participants simultaneous access to a set of resources in the interface, but designed thoughtfully so that they are meaningful with respect to the activities the participant is trying to orchestrate and accomplish as part of intra-video social interaction. In this respect, it may be interesting to do future work on visualisation of resources. A more advanced interface that builds on the proof-of-concept system in Chapter 7, could be transformative.

8.4. Methodological Contribution: reflections on my ethnomethodologically-informed approach

The methodological approach of this thesis is rooted in the interest in fine-grained aspects of social interaction in the MOOC. The insights here build on existing work that has contributed to bridging the gap between micro-analysis of qualitative data, and large amounts of data processed en-masse usually with quantitative-based methods. This thesis is one of the first pieces of research to consider the use of ethnomethodologically-informed approaches to studying social interaction of co-present participants, alongside big data on social interaction like 'clicks'. At the time of writing this, most existing studies combine detailed textual analysis with big data-based approaches (e.g. Procter et al. 2013; Tolmie et al. 2017), so this thesis contributes towards expanding the scope of the micro-analytic approaches considered alongside big data-based ones.

As I'll discuss next, the approach provided opportunities to surface these new aspects of interaction, but at the same time presented key challenges that should be considered by anyone in the future who wishes to take up a similar approach.

8.4.1. Advantages: focus on empirical mechanics of interaction that others may gloss over

The use of video data and the ethnomethodologically-informed analysis of the captured activity is tremendously important and valuable with regard to the goal of surfacing and analysing fine-grained features of social interaction. This echoes and affirms a lot of existing research that discusses the value of this approach across a number of domains including workplace studies and computer-supported collaborative

work (Heath et al. 2010; Jirotka and Luff 2006). The power of this approach is in bringing to the fore those seen-but-unnoticed strategies for accomplishing social interaction, which are highly unlikely to be collected with methods where participants self-report their activity, like survey. Many existing studies examining interaction illustrate how some of the ways participants accomplish social interaction is 'invisible' to them because they implement these seen-but-unnoticed competencies and processes (Goffman 1974; Kendon 1992).

Furthermore, an ethnomethodologically-informed approach to social interaction gives a very rich and in-depth understanding because of the practice of developing and evolving transcripts of activities in group data sessions. As part of these sessions, the transcripts and associated video become objects for analysis and reflection by a group of participants who all bring their ideas and experiences to bear on the data and the transcript. This development of transcripts can't be codified and performed automatically (at least for now), illustrating how an ethnomethodologically-informed approach can pick out key aspects of interaction that other automated or 'big data'-based methods could not.

8.4.2. Challenges: performing the method can be quite difficult and time-consuming

The main challenges around an ethnomethodologically-informed approach are the amount of focused effort required to learn and perform the analysis of data, and the effort to set up and run studies.

Performing ethnomethodologically-informed data analysis requires the abandonment of preconceptions around social interaction, which I found challenging. The approach

requires that the *how* of seeming ordinary activity be unpacked, and that all analysis be rooted in the empirical. I found it most challenging to abandon my habit of jumping to conclusions about the cognitive state of the participants in the experiment, and what the purpose of the interactions are. Drawing on ethnomethodology requires abandoning those assumptions to focus rather on the observable-reportable features of how participants in the setting appear to actively construct their interactions over time. The development and evolution of a transcript is very time-consuming since many detailed aspects of activity need to be captured accurately. This requires patience in reviewing snippets of activity many times to be very familiar with their details and nuances. Therefore, the pace at which the analysis is done is much slower than other approaches where a piece of data is analysed only briefly.

Another challenge is in formulating the task for the quasi-naturalistic experiments. The tasks are supposed to be designed to emulate 'natural' activity so that, as far as possible, it resonates with naturalistic activity. To this end, often the tasks are supposed to be open-ended (see Section 3.3.2). However, my experience is that creating too open-ended tasks could also be confusing to participants who reported being uncertain about what to do, and that this is distracting to them while doing the task. I had to carefully devise the tasks to allow them to have enough guidance that they felt confident in their activity and were not distracted by feeling uncertain. It was also useful for me to provide illustrative scenarios and examples to the participants as part of their training. To this end, more explicitly explaining the exploratory nature of the research, reassuring the participants that there is no 'right' or 'wrong' way to perform the task, and reassuring the participants of their rights as stated in the ethical clearance to be anonymised (thus possibly reducing their self-consciousness) helps in making the participants more comfortable with an open-ended task.

I also found it challenging to train participants in the systems they used during the experiment. Given time constraints, training could not be too long but it's desirable to have participants as comfortable as possible with a system. I developed an 'instructional sheet' for each experiment and a short session where participants could use the system and get familiar with it before I collected data. It took practice for me to learn how to create an information sheet that had the right amount of information in it, which I accomplished by piloting several versions of it with colleagues before using it in a study. This approach helped me create short and effective sheets and session plans.

The use of quasi-naturalistic experimentation is a challenge from the perspective that it also limits the number of and demographic of the participants. I had to draw from those participants who live near to the site of the experiment. The number of participants was limited to the number that could fit into available venues at the University of Oxford. From an analysis perspective, I expect it would also be a challenge to run a quasi-naturalistic experiment with a very large group as the interactions may get very complex therefore making it difficult to delineate organisational aspects of social interaction.

The experimental setup for the final phase 3 study is exceptionally complex in terms of arranging venues and hardware. The setup involves sourcing and using six cameras, arranging venues with plug points for the cameras, finding three venues with minimal outside noise, and arranging an Internet connection in all three venues. As I couldn't be in more than one room at any time, it was a challenge to prepare for the possibility that some of the equipment (cameras or laptops) might have a technical problem. As there are many pieces of equipment in the phase 3 study, the chances of these kinds of

technical problems are relatively high. To try to pick up any technical problems, over the course of the study I moved regularly and quietly between the rooms to check that all the equipment appeared to be running correctly. I made sure to position the participants in each room away from the door, so that I could look into each room through the door to check that everything was still running smoothly, without distracting the participants from their activities.

Furthermore, the phase 3 study is particularly challenging from a data management and analysis perspective as a lot of data was captured in a variety of forms. I collected 6 streams of camera footage (2 cameras in each room), 3 screen capture videos from each laptop in each room, and remotely collected the activity of each participant on the platform. This creates conditions for quite complex data management, organisation, and analysis.

8.5. Implications for the MOOC as a Learning Device

The insights around interaction surfaced speak to some of the current discourse in the literature and mainstream media about the success of the MOOC as a learning device. In particular, I'd like to call out two general points.

Firstly, this thesis adds empirical evidence towards the ongoing debate within the educational technology community about the success of the MOOC as an educational platform (Chapter 2). In traditional educational settings, completion rate is often taken as a critical success metric for a course. Teachers are incentivised to help move their students through all the content and assessments to the end of the course. More recently, there has been push-back against using this metric for MOOCs. It has largely

hinged on the idea that course completion is no longer adequate, given the evolving needs of students.

My research tends to support this push-back and shines a light on a more complex picture of what a 'successful' MOOC might look like. Completion rate only captures the extent to which all student watched all videos and completed all assessments along a single schedule. It doesn't substantially capture anything about social interaction on the platform. It also doesn't take into account the rich and varied ways students might interact with video content. My data contains examples throughout showing how even short fragments of MOOC video content can be seeds for rich social interaction and conversation. Furthermore, new social trends like diversification of career paths and flexible working hours, means that even one MOOC can serve a myriad of purposes to different students (e.g. (McLoughlin and Lee 2007; Klamma et al. 2007). My exploratory study of broad contextual issues (Chapter 5) gives further evidence to support this. Students seem to have a number of different goals for the course, that makes only portion of it relevant to them. In light of these changing needs of students, a one-size-fits-all approach to education is losing relevance (e.g. (Punie et al. 2006; Bryant 2006)).

Secondly, my findings can contribute to making MOOCs a richer learning experience. The deep understanding of social interaction that comes from an examination of the fine-grained details of how its accomplished is a critical stepping stone towards better supporting it in the MOOC. While I can’t discuss in-depth the pedagogical implications of my work with respect to existing learning theories (this would constitute another entire DPhil), I can highlight some of the important ways social interaction plays into a lot of contemporary ideas about learning. Social interaction around educational

content is often seen as a critical part of effective learning. Where traditional approaches to learning paint the student as largely passive consumer of content, modern theories emphasise the need for the student to put what they've learned into action through social interaction. Social interaction is thought to help create a deeper level of learning (Garrison and Anderson 2003; Johnson and Johnson 1987; Kreijns et al. 2003; Savery and Duffy 1996). By studying interaction in detail, this thesis provides implications on the design of the MOOC that can help it develop into a more effective learning platform.

The next frontier in online learning is largely a software-centric one. The proliferation of the internet and the MOOC so far has proven we have the computer hardware to scale the provision of educational resources to larger groups of students online than ever before. The next challenge is to develop the software to curate the learning experience online so it is more social and personalised than ever before. The goal of the ethnomethodologically-informed approach taken here is to contribute concretely to the ideas and design issues that help us conquer this frontier and make good quality education accessible to more than ever before.

8.6. Future Work

I see several areas for future work building off this thesis. In this section, I outline each theme and the potential contributions that could come from each.

8.6.1. Ethical issues around increased visibility of each individual's activities on the platform

Part of this thesis advocates for potentially collecting more granular data on the activity of individuals - like tracking students' movement between the video interface and the

forum interface. Important ethical questions might include those around allowing participants to opt in and out, and how to adequately gain consent, and what the main personal privacy risks might be. A contribution from this could be an ethical framework and set of issues around this.

8.6.2. Development of new workflows and visualisation to represent the activity of others

The widget presented in the proof-of-concept interface of Chapter 7 is a basic interface feature that gives one participant a sense of the synchronous activities of others. It has its merits in the small group within the study in Chapter 7, but may not scale to a large group of students as is typical for MOOCs. I can see the basic widget feature in the proof-of-concept interface being developed, for example, into a dashboard that provides a sense of the activity and presence of others. This kind of dashboard may also incorporate more novel features like generating audio or various visualisations to allow participants an easy sense of the presence and activity of others. This may be particularly challenging given the large student group numbers. Artificial Intelligence algorithms and approaches to personalising the way information is displayed in the dashboard, may also be an interesting avenue to explore, subject to an adequate study of the possible ethical issues and consequences. I can also imagine that at scale, interactional issues not surfaced in this thesis could become important. The participants sourced for my studies are all accomplished at social interaction within the same kind of setting since they all live in Oxford. It would be beneficial to consider issues that may arise when international students from different geographic regions interact online. An outcome of this kind of future work could be new creative ways of

representing and visualising key aspects of activity of the MOOC student group, and a scoped out set of associated issues.

8.6.3. Methodological innovation to integrate big data-based and ethnomethodologically-informed approaches

Future work could build on this thesis by exploring ways to integrate ethnomethodologically-informed studies like this one, with 'big data'-based approaches. I don't offer any substantive recommendations for how the two kinds of approaches might be usefully combined, and what the associated issues and considerations may be. Future work could focus further on the methodological possibilities for fully connecting the approaches and validating the extent to which 'small data'-gleaned insights are witnessed in the real-world big data systems. A contribution of this could be a new kind of methodological approach and principles for studying MOOCs.

8.6.4. Expanding on the step-wise progression into and out of more active forms of participation

I have revealed how participants shift into and out of more active forms of participation, calling out its stepwise structure. This would be useful to researchers investigating ways to encourage more active forms of participation (e.g. Nawrot and Doucet 2014). This work suggests the value of providing different resources for each step, perhaps in the form of different prompts in the interface. There would be a variety of usability issues (you don't want to annoy the participants with too many pop ups or prompts!) and ethical issues (for example, with regard to consent) that would need to be unpacked. Experimentation with the way the interface could be designed to

shape the stepwise progression into, through, and out of possible contribution. An outcome of this future work could be a set of best practice principles and considerations for how to design MOOCs to increase likelihood of active participation.

Appendix 1. Transcription Notation

The transcription notation used is intended to capture those properties of the embodied conduct which match the priorities and interests of analysis. The orthography is based on the system developed by (Atkinson and Heritage 1984)

1. Simultaneous utterances

Two utterances which start simultaneously are symbolised by a single left-hand bracket between two lines, with the bracket positioned at the moment the two utterances begin to co-exist.

A: I think he's having difficulty because he seems to be uncertain

[

B: Oh yes

2. Contiguous utterances

When there is no interval between adjacent utterances and utterances do not overlap the two utterances are linked with an equal sign.

A: I wouldn't say that it's particularly difficult=

B: =Yes definitely

3. Intervals between utterances

When there are intervals in a stream of talk, they are timed in tenths of a second and inserted in parenthesis. These may be a pause within an utterance:

A: That was when I was (0.5) maybe three or four years old

or between utterances.

A: Do you think we should go for it?

(1.0)

B: For which one?

(0.3)

A: The first option

4. Characteristic of speech delivery – extension of sound or syllable

An extension of a sound or syllable is indicated using a colon after the syllable, with each colon corresponding roughly to one tenth of a second.

A: I re::ally think that is not the best option

5. Transcription doubt

When the words being said in the utterance are unclear in the data, double parentheses mark the phrase that's unclear and contain the best estimate of what is being said.

A: Look here this is what you said ((the system has not been changed))

B: ((laughs)) Yes so it is I'm quite surprised about that

Appendix 2. List of Publications

The list below is of all publications associated with this thesis.

- “Interface Requirements for Harnessing Collective Intelligence in Massive Open Online Courses”, The Grace Hopper Celebration of Women in Computing, October 2014

- “Surfacing Collective Intelligence with Implications for Interface Design in Massive Open Online Courses”, Collective Intelligence Conference, June 2014

- Zawilska, A., M. Jirotka, and M. Hartswood. "Surfacing Collective Intelligence with Implications for Interface Design in Massive Open Online Courses." Social Collective Intelligence. Springer International Publishing, 2014. 187-202.

- Zawilska, A., M. Jirotka, and M Hartswood. “Surfacing Collective Intelligence with Implications for Interface Design in Massive Open Online Courses” - poster at the Collective Intelligence Conference, June 2014 (Abstract available at: humancomputation.com/ci2014/papers/Active%20Papers%5CPaper%2050.pdf)

- Zawilska, A., S. Albury. "An Ethnomethodologically-Informed Approach to Designing Online Platform Interfaces to Support Social Interaction" - a presentation at the 3rd International Symposium: Micro-Analysis of Online Data (MOOD-Z) in July 2015.

- Zawilska, Anna, and Steven Albury. "An Ethnomethodologically-Informed Approach to Interface Design for Social Interactions around Video Online."Ninth International AAAI Conference on Web and Social Media. 2015.

- Zawilska, Anna, and Steven Albury. "An Ethnomethodologically-Informed Approach to Interface Design to Support Collective Web Practice Around Video." Proceedings of the ACM Web Science Conference. ACM, 2015.

References

Ahrend, J. M., Jirotka, M., & Jones, K. (2016). On the collaborative practices of cyber threat intelligence analysts to develop and utilize tacit Threat and Defence Knowledge. In 2016 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (CyberSA) (pp. 1–10). IEEE. http://doi.org/10.1109/CyberSA.2016.7503279

Al-Atabi, M., & DeBoer, J. (2014). Teaching entrepreneurship using Massive Open Online Course (MOOC). Technovation, 34(4), 261–264. http://doi.org/10.1016/j.technovation.2014.01.006

Andrews, D., Nonnecke, B., & Preece, J. (2003). Electronic survey methodology: A case study in reaching hard-to-involve Internet users. International Journal of Human-Computer Interaction, 16(2), 185–210.

Atkinson, M. J., & Heritage, J. (1984). Structure of social action: Studies in conversation analysis. Cambridge University Press.

Bachmann, D., & Elfrink, J. (1996). Tracking the progress of e-mail versus snail-mail. MarketingResearch, 8(2), 31–35.

Balch, T. (2013). MOOC Student Demographics (Spring 2013). Retrieved August 7, 2013, from http://augmentedtrader.wordpress.com/2013/01/27/mooc-student-demographics/

Bali, M. (2014). MOOC pedagogy: gleaning good practice from existing MOOCs. Ournal of Online Learning and Teaching, 10(1), 44.

BBC. (2013). Profile: Edward Snowden. Retrieved August 21, 2013, from http://www.bbc.co.uk/news/world-us-canada-22837100

Belanger, Y., & Thornton, J. (2013). Bioelectricity: A Quantitative Approach. Retrieved September 24, 2016, from http://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/6216/Duke_Bioelectricity_MOOC_Fall2012.pdf

Beyer, M. A., & Laney, D. (2012). The Importance of “Big Data”: A Definition. Retrieved September 16, 2016, from https://www.gartner.com/doc/2057415/importance-big-data-definition

Black, W., Procter, R., Gray, S., & Ananiadou, S. (2012). A data and analysis resource for an experiment in text mining a collection of micro-blogs on a political topic. In LREC (pp. 2083–2088).

Boehm, B. (2006). A view of 20th and 21st century software engineering. In Proc. of 28th international conference on Software Engineering (pp. 12–29).

Boren, T., & Ramey, J. (2000). Thinking Aloud: Reconciling Theory and Practice. IEEE Transactions on Professional Communication, 43(3), 261–278.

Bowers, J., Button, G., & Sharrock, W. (1995). Workflow from Within and Without. In Proceedings of European Conference on Computer-Supported Cooperative Work ECSCW’95 (pp. 51–66). Stockholm, Sweden.

Bowers, J., & Rodden, T. (1993). Exploding the interface: Experiences of a CSCW network. In Proc. of Human Factors in Computing Systems (pp. 255–262).

Boyatzis, R. E. (1998). Transforming qualitative information: Thematic analysis and code development. Sage.

Brend, K., & Naylor, S. (1996). Teaching and learning in science: a new perspective. In British Educational Research Association Conference. Retrieved from http://www.leeds.ac.uk/educol/documents/000000115.htm

Brown, B., & Bell, M. (2004). CSCW at play: “There” as a collaborative virtual environment. In Proc. of Computer-Supported Collaborative Work (pp. 250–259). ACM Press.

Bryant, T. (2006). Social Software in Academia. EDUCAUSE Quarterly, 29(2), 61–64.

Burke, K. R., & Mahoney, J. (2013). Meaningful collaboration: revitalizing small colleges with MOOC hybrids. Hybrid Pedagogy.

Button, G. (1993). Ethnomethodology and the Human Sciences. Newcastle Upon Tyne: Cambridge University Press.

Button, G., & Sharrock, W. (1994). Occasioned Practices in the Work of Software Engineers. In Requirements Engineering: Social and Technical Issues by Jirotka, M. and Goguen, J.A. (pp. 217–240).

Button, G., & Sharrock, W. (1995). On Simulacrums of Conversation. In The Social and Interactional Dimensions of Human-Computer Interfaces. Cambridge University Press.

Button, G. (2012). What does “work” mean in “ethnomethodological studies of work?”: Its ubiquitous relevance for systems design to support action and interaction. Design Studies, 33(6), 673–684. http://doi.org/10.1016/j.destud.2012.06.003

Card, S., Moran, T., & Newell, A. (1983). The Psychology of Human-Computer Interaction. Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Carroll, J. M. (Ed.). (2003). HCI Models, Theories, and Frameworks - Toward a Multidisciplinary Science. Morgan Kauffman.

Chauhan, S., & Rathore, S. (2015). Ethics in Behavioural Targeting: Mapping Consumers Perceptions. Business Law and Ethics: Concepts, Methodologies, Tools, and Applications, 303–319.

Chengjie, Y. U. (2015). Challenges and changes of MOOC to traditional classroom teaching mode. Canadian Social Science, 11(1), 135.

Clow, D. (2013). MOOCs and the funnel of participation. In Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13 (p. 185). New York, New York, USA: ACM Press. http://doi.org/10.1145/2460296.2460332

de la Flor, G., Luff, P., Jirotka, 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 Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI) (pp. 473–482).

Douglas, L. (2001). 3d data management: Controlling data volume, velocity and variety. Gartner.Retrieved, 6, 1.

Douglas, S. A. (1995). Conversation analysis and human-computer interaction design. In The Social and Interactional Dimensions of Human-Computer Interfaces (pp. 184–203). Cambridge, UK: Cornell University Press.

Dourish, P., & Bell, G. (2011). Divining a digital future: mess and mythology in ubiquitous computing. MIT Press.

Dourish, P., & Button, G. (1998). On “Technomethodology”: Foundational Relationships between Ethnomethodology and System Design. Retrieved from http://www.madepublic.com/corti/tp/docs/Dourish-ethnomethodology.pdf

Draper, S. W., Riley, M. S., & O’Malley, C. E. (1984). Constructive interaction: a method for studying human-computer interaction. In IFIP Conference on Human-Computer Interaction. Retrieved from http://libra.msra.cn/Publication/1926927/constructive-interaction-a-method-for-studying-human-computer-interaction

Duffy, T., & Jonassen, D. (1992). Constructivism and the technology of instruction: A conversation. Hillsdale, NJ, USA: Lawrence Erlbaum Associates

Duguid, P. (2012). On rereading. Suchman and Situated Action. Le Libellio Vol. 8, 3–9. Retrieved from http://lelibellio.com/wp-content/uploads/2013/01/Libellio26.pdf#page=3

Durkheim, E. (2013). Durkeim: The Rules of Sociological Method: And Selected Texts on Sociology and Its Method. Palgrave Macmillan.

Eastman, J. K., & Swift, C. O. (2002). Enhancing Collaborative Learning: Discussion Boards and Chat Rooms as Project Communication Tools. Business Communication Quarterly, 65(3), 29–41. http://doi.org/10.1177/108056990206500304

ECampus. (2013). Top 10 reasons for low MOOC completion rates. Retrieved July 25, 2013, from http://www.ecampusnews.com/top-news/top-10-reasons-for-low-mooc-completion-rates/?

Eden, G. (2011). The contextual evaluation framework: a prototype evaluation technique for e-Research. University of Oxford. Retrieved from http://ora.ox.ac.uk/objects/uuid:e816cf1a-3514-400c-9f2f-e022b56710d9

Ekman, P., & Rosenberg, E. L. (1997). What the face reveals: Basic and applied studies of spontaneous expression using the Facial Coding System (FACS). Oxford University Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal protocol as data. Psychological Review, 87(3), 215–251.

Eynon, R. (2014). Conceptualising interaction and learning in MOOCs. Retrieved from http://www.moocresearch.com/wp-content/uploads/2014/06/C9146_EYNON_MOOC-Research-InitiativeEynon9146Final.pdf

Fildes, J. (2010). What is Wikileaks? Retrieved August 21, 2013, from http://www.bbc.co.uk/news/technology-10757263

Gabbert, E. (2010). What's a good click-through rate for ... Retrieved May 31, 2013, from http://www.wordstream.com/blog/ws/2010/11/22/what-is-a-good-click-through-rate-for

Galegher, J., & Kraut, R. E. (1990). Technology for intellectual teamwork: Perspectives on research and design. Intellectual Teamwork: Social and Technological Foundations of Cooperative Work, 1–20.

Gannes, L. (2012). Stanford Professors Launch Coursera With $16M From Kleiner Perkins and NEA. Retrieved February 1, 2017, from http://allthingsd.com/20120418/stanford-professors-launch-coursera-with-16m-from-kleiner-perkins-and-nea/

Garfinkel, H. (1967). Studies in Ethnomethodology. New Jersey, NY, USA: Prentice-Hall.

Garfinkel, H., & Sacks, H. (1970). Formal Structures of Practical Action. In J.C. McKinney and E.A. Tiryakian (Eds), Theoretical Sociology. New York: Appleton Century Crofts.

Garrison, D. R., & Anderson, T. (2003). A framework for research and practice. In E-Learning in the 21st Century (pp. 32–47).

Gee, S. (2012). MITx - the Fallout Rate. Retrieved February 1, 2013, from http://www.i-programmer.info/news/150-training-a-education/4372-mitx-the-fallout-rate.html

Goffman, E. (1972). Encounters. Penguin.

Goffman, E. (1981). Footing. In Forms of Talk (pp. 124–159). Basil Blackwell Publisher.

Goffman, E. (1974). Frame Analysis. Cambridge, MA, USA: Harvard University Press.

González-Lloret, M. (2011). Conversation Analysis of Computer-Mediated Communication. CALICO Journal, 28(2), 308–325.

Goodwin, C. (1981). Conversational Organization: Interaction Between Speakers and Hearers. Academic Press.

Grudin, J. (1988). Why CSCW applications fail: problems in the design and evaluationof organizational interfaces. In Proceedings of the 1988 ACM conference on Computer-supported cooperative work - CSCW ’88 (pp. 85–93). New York, New York, USA: ACM Press. http://doi.org/10.1145/62266.62273

Grudin, J. (1990). The computer reaches out: the historical continuity in interface design. In Proceedings SIGCHI conference on Human factors in computing systems: Empowering people (pp. 261–268).

Grünewald, F., Meinel, C., Totschnig, M., & Willems, C. (2013). Designing MOOCs for the Support of Multiple Learning Styles (pp. 371–382). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-40814-4_29

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement. In Proceedings of the first ACM conference on Learning @ scale conference - L@S '14 (pp. 41–50). New York, New York, USA: ACM Press. http://doi.org/10.1145/2556325.2566239

Harper, R., & Hughes, J. (1993). What A F-ing System! Send 'em all to the same place and then expects us to step 'em hitting: Making technology work in air traffic control. In Technology in Working order by Button, G. London, UK: Routledge.

Hartswood, M., Procter, R., Slack, R., Voß, A., Büscher, M., Rouchy, P., & Rouncefield, M. (2002). Co-realisation and participatory design. *Scandinavian Journal of Information Systems*, *14*(2), 9–30.

Hartswood, M., Procter, R., Slack, R., Voß, A., Buscher, M., Rouncefield, M., & Rouchy, P. (2008). Co-realization: Toward a principled synthesis of ethnomethodology and participatory design. In Resources, co-evolution and artifacts (pp. 59–94).

Heath, C., Hindmarsh, J., & Luff, P. (2010). Video in Qualitative Research: Analysing Social Interaction in Everyday Life. SAGE Publications.

Heath, C., Jirotka, M., Luff, P., & Hindmarch, J. (1993). Unpacking Collaboration: The Interactional Organisation of Trading in a City Dealing Room. In Proceedings of 3rd European Conference on Computer Supported Collaborative Work ECSCW '93. Milano, Italy.

Heath, C., Jirotka, M., Luff, P., & Hindmarsh, P. (1995). The Individual and the Collaborative: the Interactional Organisation of Trading in a City Dealing Room. Journal of Computer Supported Collaborative Work, 3(1), 147–165.

Heath, C., & Luff, P. (1993). System Use and Social Organisation. In Technology in Working order by Button, G. London, UK: Routledge.

Heath, C., & Luff, P. (1992). Media Space and Communicative Asymmetries: Preliminary Observations of Video-Mediated Interaction. Human-Computer Interaction, 7(3), 315–346.

Heath, C., Luff, P., Vom Lehn, D., Hindmarsh, J., & Cleverly, J. (2002). Crafting participation: designing ecologies, configuring experience. Crafting Participation: Designing Ecologies, Configuring Experience, 1(1), 9–33.

Heath, C., Jirotka, M., Luff, P., & Hindmarsh, J. (1994). Unpacking collaboration: the interactional organisation of trading in a city dealing room. Computer Supported Cooperative Work (CSCW), 3(2), 147–165. http://doi.org/10.1007/BF00773445

Heath, C., & Luff, P. (2013). Embodied action and organisational interaction: Establishing contract on the strike of a hammer. Journal of Pragmatics, 46(1), 24–38. http://doi.org/10.1016/j.pragma.2012.01.002

Heath, C., Luff, P., Vom Lehn, D., Yamashita, J., & Kuzuoka, H. (2009). Enhancing remote participation in live auctions. In Proceedings of the 27th international conference on Human factors in computing systems - CHI 09 (p. 1427). New York, New York, USA: ACM Press. http://doi.org/10.1145/1518701.1518918

Heritage, J. (1984). Garfinkel and ethnomethodology. Cambridge: Polity Press.

Herring, S. (2006). Interactional Coherence in CMC. Journal of Computer-Mediated Communication, 4(4). http://doi.org/10.1111/j.1083-6101.1999.tb00106.x

Hicks, K. (2015, March). Why MOOCs Are Great for Teacher Development. Edudemic Connection Education & Technology. Retrieved from http://www.edudemic.com/5-moocs-educators-should-take-as-students/

Hindmarsh, J., Reynolds, P., & Dunne, S. (2011). Exhibiting understanding: The body in apprenticeship. Journal of Pragmatics, 43(2), 489–503. http://doi.org/10.1016/j.pragma.2009.09.008

Jefferson, G. (1984). On stepwise transition from talk about a trouble to inappropriately next-positioned matters. Structures of Social Action: Studies in Conversation Analysis, 191–222.

Jeffries, R., & Desurvire, H. (1991). User interface evaluation in the real world: A comparison of four techniques. In Proc. of Human Factors in Computing Systems (pp. 119–124).

Jirotka, M., Lee, C. P., & Olson, G. M. (2013). Supporting Scientific Collaboration: Methods, Tools and Concepts. Computer Supported Cooperative Work (CSCW), 22(4–6), 667–715. http://doi.org/10.1007/s10606-012-9184-0

Jirotka, M., & Luff, P. (2006). Supporting requirements with video-based analysis. IEEE Software 23.3, 42–44.

Jirotka, M., Procter, R., Hartswood, M., Slack, R., Simpson, A., Coopmans, C., … Voss, A. (2005). Collaboration and Trust in Healthcare Innovation: The eDiaMoND Case Study.

Computer Supported Cooperative Work (CSCW), 14(4), 369–398. http://doi.org/10.1007/s10606-005-9001-0

Jirotka, M., & Wallen, L. (2000). Analysing the workplace and user requirements: challenges for the development of methods for requirements engineering. In Workplace Studies: Recovering Work Practice and Informing System Design edited by Luff, P., Hindmarsh, J., Heath, C. (pp. 242–251).

Jirotka, M., Luff, P., & Gilbert, N. (1991). Participation frameworks for computer mediated communication. In Proceedings of the Second European Conference on Computer-Supported Cooperative Work ECSCW '91 (pp. 279–291). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-011-3506-1_21

John, B., & Keiras, D. (1996). Using GOMS for User Interface Design and Evaluation: Which Technique? Transactions on Computer-Human Interaction, 3, 287–319.

Johnson, D. W., & Johnson, R. T. (1987). Learning together and alone: Cooperative, competitive, and individualistic learning, 2nd ed. Prentice-Hall, Inc.

Jordan, K. (2013). MOOC Completion Rates: The Data. Retrieved July 25, 2013, from http://www.katyjordan.com/MOOCproject.html

Kahler, H. (2000). Constructive interaction and collaborative work: introducing a method for testing collaborative systems. Interactions, 7(3), 27–34.

Kameo, N., & Whalen, J. (2015). Organizing Documents: Standard Forms, Person Production and Organizational Action. Qualitative Sociology, 38(2), 205–229. http://doi.org/10.1007/s11133-015-9302-7

Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. Public Opinion Quarterly 68.1, 94–101. Retrieved from http://www.uwyo.edu/studentaff/_files/docs/survey_calendar/kaplovitz_hadlock_levine_a_comparison_of_web_and_mail_survey_reponse_rates.pdf

Kendon, A. (1992). The negotiation of context in face-to-face interaction. Rethinking Context: Language as an Interactive Phenomenon, 11, 323.

Khalil, H., & Ebner, M. (2013). How satisfied are you with your MOOC?-A Research Study on Interaction in Huge Online Courses. Proceedings of World Conference on Educational. Retrieved from http://www.academia.edu/download/41645720/568f5ba535d73.pdf

Kim, B. (2001). Social constructivism. Emerging Perspectives on Learning, Teaching, and Technology, 1(1), 16–25.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks inonline lecture videos. In Proceedings of the first ACM conference on Learning @ scale conference - L@S '14 (pp. 31–40). New York, New York, USA: ACM Press. http://doi.org/10.1145/2556325.2566237

Kimbell, L. (2011). Rethinking Design Thinking: Part I, Design and Culture. The Journal of the Design Studies Forum, 3(2), 285–306.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In Proceedings of 3rd Conference on Learning Analytics and Knowledge. Leuven, Belgium.

Kizilcec, R. F., & Schneider, E. (2015). Motivation as a Lens to Understand Online Learners. ACM Transactions on Computer-Human Interaction, 22(2), 1–24. http://doi.org/10.1145/2699735

Klamma, R., Chatti, M. A., Duval, E., Hummel, H., Hvannberg, E. H., Kravcik, M., … Scott, P. (2007). Social software for life-long learning. Journal of Educational Technology and Society, 10(3), 72–83.

Kolowich, S. (2013). Why Some Colleges Are Saying No to MOOC Deals, at Least for Now. Retrieved July 25, 2013, from http://chronicle.com/article/article-content/138863/

Konnikova, M. (2014). Will MOOCs be Flukes? Retrieved from www.newyorker.com/science/maria-konnikova/moocs-failure

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. Computers in Human Behavior, 19(3), 335–353. http://doi.org/10.1016/S0747-5632(02)00057-2

Kreitmayer, S., Rogers, Y., Laney, R., & Peake, S. (2012). From participatory to contributory simulations: changing the game in the classroom. In ACM SIGCHI Conference on Human Factors in Computing Systems 2012 CHI'12 (pp. 49–58). Austin, Texas, USA. Retrieved from http://oro.open.ac.uk/33544/1/kreitmayer_CHI2012_contributory.pdf

Larman, C., & Basili, V. R. (2003). Iterative and incremental development: A brief history. Computer, 36(6), 47–56.

Law, J., & Hassard, J. (1999). Actor network theory and after. Oxford: Blackwell.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: traps in big data analysis. Science, 343(14).

Levine, A. (2012). Owning your massive numbers. Retrieved February 1, 2013, from http://cogdogblog.com/2012/11/27/owning-massive/

Levinson, S. C. (1988). Putting Linguistics on a Proper Footing: Explorations in Goffman's Concepts of Participation. In Erving Goffman: Exploring the Interaction Order (pp. 161–227). Polity Press.

Lublin, J. (2003). Deep, surface and strategic approaches to learning. Centre for Teaching and Learning. UCD Dublin, nd, 806-825.

Lucena. (2013). About MOOC Completion Rates: The Importance of Student Investment. Retrieved July 25, 2013, from https://lucenaresearch.com/about-mooc-completion-rates-the-importance-of-student-investment/

Luff, P., Patel, M., Kuzuoka, H., & Heath, C. (2014). Assembling Collaboration: Informing the Design of Interaction Spaces. Research on Language and Social Interaction, 47(3), 317–329. http://doi.org/10.1080/08351813.2014.925680

Luff, P., Yamashita, N., Kuzuoka, H., & Heath, C. (2011). Hands on hitchcock. In Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11 (p. 43). New York, New York, USA: ACM Press. http://doi.org/10.1145/1978942.1978951

Lynch, M. (2006). Cognitive activities without cognition? ethnomethodological investigations of selected "cognitive" topics. Discourse Studies, 8(1), 95–104.

Lynch, M., Livingstone, E., & Garfinkel, H. (1983). Temporary Order in Laboratory Work. In Knoll-Centina and Mulkay (Eds), Science Observed. London: Sage.

Marshall, C., & Rossman, G. B. (2010). Designing qualitative research. Sage.

Martin, D., & Sommerville, I. (2004). Patterns of Cooperative Interaction: Linking Ethnomethodology and Design. ACM Transactions on Human-Computer Interaction, 11(1), 59–89.

Martin, K. (2016). Understanding privacy online: Development of a social contract approach to privacy. Journal of Business Ethics, 137(3), 551–569.

McAuley, A., Stewart, B., Cormier, D., & Siemens, G. (2010). In the Open: The MOOC model for digital practice. SSHRC Application, Knowledge Synthesis for the Digital Economy.

Mcauley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). The MOOC model for digital practice. Retrieved from http://www.flickr.com/photos/ecstaticist/3570660643/

Mcloughlin, C., & Lee, M. J. W. (2007). Social software and participatory learning: Pedagogical choices with technology affordances in the Web 2.0 era. ICT: Providing Choices for Learners and Learning, 664–675.

McMahon, M. (1997). Social Constructivism and the World Wide Web - A Paradigm for Learning. In ASCILITE conference. Perth, Australia.

Meredith, J., & Stokoe, E. (2014). Repair: Comparing Facebook “chat” with spoken interaction. Discourse & Communication, 8(2), 181–207. http://doi.org/10.1177/1750481313510815

Mitchell, J. C. (1983). Case and Situation Analysis. Sociological Review, 31(2), 187–211.

Miyake, N. (1986). Constructive interaction and the iterative process of understanding. Cognitive Science, 10(2), 151–177.

Monk, A., & Watts, L. (2000). Peripheral participation in video-mediated communication. International Journal of Human-Computer Studies, 52(5), 933–958. http://doi.org/10.1006/ijhc.1999.0359

Moore, R. J., Ducheneaut, N., & Nickell, E. (2007). Doing Virtually Nothing: Awareness and accountability in massively multiplayer online worlds. Computer Supported Collaborative Work, 16, 265–305.

Morrow, P. R. (2006). Telling about problems and giving advice in an Internet discussion forum: some discourse features. Discourse Studies, 8(4), 531–548. http://doi.org/10.1177/1461445606061876

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.

Nawrot, I., & Doucet, A. (2014). Building engagement for MOOC students: introducing support for time management on online learning platforms. In Proceedings of the companion publication of the 23rd international conference on World wide web companion (pp. 1077–1082).

Nielsen, J. (1993). Usability Engineering. Academic Press Inc.

Orlikowski, W. . (1992). Learning from Notes. In Proceedings of the 1992 ACM conference on Computer-supported cooperative work - CSCW '92 (pp. 362–369). New York, New York, USA: ACM Press. http://doi.org/10.1145/143457.143549

Page, D., Williams, P., & Boyd, D. (1993). Report of the Public Inquiry into the London Ambulance Service. Retrieved from http://www0.cs.ucl.ac.uk/staff/A.Finkelstein/las/lascase0.9.pdf

Pappano, L. (2012). The Year of the MOOC. Retrieved February 1, 2013, from http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=all&_r=1&

Parsons, T. (1937). The Structure of Social Action. New York: McGraw-Hill.

Pea, R., & Lindgren, R. (2008). Video collaboratories for research and education: an analysis of collaboration design patterns. IEEE Transactions on Learning Technologies, 1(4), 235–247.

Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., & Hoffert, E. (2004). The diver project: Interactive digital video repurposing. MultiMedia, IEEE, 11(1), 54–61.

Petchenik, J., & Watermolen, D. J. (2011). A cautionary note on using the Internet to survey recent hunter education graduates. Human Dimensions of Wildlife, 16(3), 216–218.

Ponceleon, D., & Dieberger, A. (2011). Hierarchical Brushing in a Collection of Vido Data. In Proc. of 34th Hawaii International Conference on System Sciences (pp. 1–8).

Preece, J., Rogers, Y., & Sharp, H. (2002). Interaction design: Beyond human-computer interaction. The Craft of Network Computing, 11(4), 34.

Procter, R., Crump, J., Karstedt, S., Voss, A., & Cantijoch, M. (2013). Reading the riots: what were the police doing on Twitter? Policing and Society, 23(4), 413–436. http://doi.org/10.1080/10439463.2013.780223

Procter, R., Vis, F., & Voss, A. (2013). Reading the riots on Twitter: methodological innovation for the analysis of big data. International Journal of Social Research Methodology, 16(3), 197–214. http://doi.org/10.1080/13645579.2013.774172

Punie, Y., Cabrera, M., Bogdanowicz, M. Zinnbauer, D., & Navajas, E. (2006). The future of ICT and learning in the knowledge society.

Randall, D., Harper, R., & Rouncefield, M. (2007). Fieldwork for Design. London, UK: Springer-Verlag.

Reigeluth, C. M. (2013). Instructional design theories and models: An overview of their current status. Routledge.

Riff, D., Lacy, S., & Fico, F. (2014). Analyzing media messages: Using quantitative content analysis in research. Routledge.

Rintel, S. (2010). Conversational management of network trouble perturbations in personal videoconferencing. In Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction - OZCHI '10 (p. 304). New York, New York, USA: ACM Press. http://doi.org/10.1145/1952222.1952288

Rintel, S., Harper, R., & O'Hara, K. (2016). The Tyranny of the Everyday in Mobile Video Messaging. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 4781–4792). http://doi.org/10.1145/2858036.2858042

Rintel, S. (2015). Omnirelevance in Technologised Interaction: Couples Coping with Video Calling Distortions.

Rogers, Y. (1993). Coordinating Computer-Mediated Work. Computer Supported Collaborative Work, 1(4), 295–315.

Romiszowski, A., & Mason, R. (2004). Computer-Mediated Communication. In Handbook of research on educational communications and technology by Jonassen, D. H (pp. 397–431).

Ross, J., Sinclair, C., Knox, J., & Macleod, H. (2014). Teacher experiences and academic identity: The missing components of MOOC pedagogy. Journal of Online Learning and Teaching, 10(1), 57–63.

Sacks, H. (1984). Notes on methodology. In J. M. Maxwell & H. J. (Eds.), Structures of Social Action: Studies in Conversation Analysis (pp. 21–27).

Sacks, H. (1992). Lectures in conversation: Volumes I and II. Oxford, UK: Blackwell.

Sacks, H. (1989). Lectures 1964-1965. Human Studies, 12(3–4), 211–393.

Sacks, H. (1963). Sociological description. Berkeley Journal of Sociology, 8, 1–16.

Santos, J. L., Klerk, J., Duval, E., Gago, D., & Rodríguez, L. (2014). Success, activity and drop-outs in MOOCs an exploratory study on the UNED COMA courses. In Proceedings of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14 (pp. 98–102). New York, New York, USA: ACM Press. http://doi.org/10.1145/2567574.2567627

Sarcevic, A., Marsic, I., & Burd, R. (2012). Teamwork errors in trauma resuscitation. ACM Transactions on Computer-Human. Retrieved from http://dl.acm.org/citation.cfm?id=2240161

Sarcevic, A., Marsic, I., & Burd, R. S. (2012). Teamwork Errors in Trauma Resuscitation. ACM Transactions on Computer-Human Interaction, 19(2), 1–30. http://doi.org/10.1145/2240156.2240161

Savery, J. R., & Duffy, T. M. (1996). Problem-based learning: An instructional model and its constructivist framework. In Constructivist learning environments: Case studies in instructional design by B.G. Wilson (pp. 135–148). Eaglewood Cliffs, NJ, USA: Educational Technology Publications.

Savery, J. R., Savery, J. R., & Duffy, T. M. (1994). Problem Based Learning: An instructional model and its constructivist framework. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.4403

Schegloff, E. A. (1992). Repair after Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation. American Journal of Sociology, 97(5), 1295–1345.

Schegloff, E. A., & Sacks, H. (1973). Opening up Closings. Semiotica, 7, 289–327.

Schegloff, E. A., & Sacks, H. (2002). Home position. Gesture, 2(2), 133–146.

Schmidt, E., & Cohen, J. (2013). The New Digital Age: Reshaping the Future of People, Nations and Business. John Murray Publishers Ltd.

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). Analytics: The real-world use of big data. IBM Global Business Services, 1–20.

Selwyn, N., & Bulfin, S. (2014). The discursive construction of MOOCs as educational opportunity and educational threat. Retrieved from http://www.moocresearch.com/wp-content/uploads/2014/06/C9130_Selwyn-Bulfin-MRI-final-report-publication-report.pdf

Shepherd, A. (2000). HTA as a framework for task analysis. In Task Analysis (pp. 9–24).

Siemens, G. (2014). Connectivism: A learning theory for the digital age. Retrieved February 1, 2017, from http://er.dut.ac.za/handle/123456789/69

Sinha, T., Li, N., Jermann, P., & Dillenbourg, P. (2014). Capturing "attrition intensifying" structural traits from didactic interaction sequences of MOOC learners. arXiv Preprint arXiv, 1409(5887). Retrieved from http://arxiv.org/abs/1409.5887

Smith, A., & Dunckley, L. (2002). Prototype Evaluation and Redesign: Structuring the Design Space Through Contextual Techniques. Interacting with Computers, 14(6), 36–40.

Sommerville, I., & Sawyer, P. (1998). Requirements Engineering: A good practice guide. West Sussex, UK: John Wiley & Sons, Ltd.

Spaniel, W. (2013). Can We Stop Caring about MOOC Completion Rates? Retrieved July 25, 2013, from http://wjspaniel.wordpress.com/2013/05/27/can-we-stop-caring-about-mooc-completion-rates/

Stanton, J. M. (1998). An empirical assessment of data collection using the Internet. Personnel Psychology, 51(3), 709–725.

Staubitz, T., Renz, J., Willems, C., & Meinel, C. (2014). Supporting Social Interaction and Collaboration on An XMOOC Platform. In Proc. EDULEARN14 (pp. 6667–6677).

Suchman, L. (2007). Human-machine reconfigurations: plans and situated actions (2nd editio). Cambridge University Press.

Suchman, L. A. (1987). Plans and Situated Action: The problem of human-machine communication. Cambridge, UK: Cambridge University Press.

Takayasu, M., Sato, K., Sano, Y., Yamada, K., Miura, W., & Takayasu, H. (2015). Rumor diffusion and convergence during the 3.11 earthquake: a twitter case study. PloS One, 10(4), e0121443. http://doi.org/10.1371/journal.pone.0121443

Thomas, M. J. W. (2002). Learning within incoherent structures: the space of online discussion forums. Journal of Computer Assisted Learning, 18(3), 351–366. http://doi.org/10.1046/j.0266-4909.2002.03800.x

Thompson, L. F., Surface, E. A., Martin, D. L., & Sanders, M. G. (2003). From paper to pixels: Moving personnel surveys to the Web. Personnel Psychology, 56(1), 197–227.

Tolmie, P., Procter, R., Rouncefield, M., Liakata, M., & Zubiaga, A. (2017). Microblog Analysis as a Programme of Work. ACM Transactions on Social Computing, to appear.

Van den Haak, M. J., de Jong, M. D. T., & Schellens, P. J. (2004). Employing think-aloud protocols and constructive interaction to test the usability of online library catalogues: a methodological comparison. Interacting with Computers, 16, 1153–1170. http://doi.org/10.1016/j.intcom.2004.07.007

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: a case study. The Internet and Higher Education, 6(1), 77–90. http://doi.org/10.1016/S1096-7516(02)00164-1

Wang, B., Liakata, M., Zubiaga, A., Procter, R., & Jensen, E. (2016). SMILE : Twitter emotion classification using domain adaptation. In CEUR Workshop Proceedings (Vol. 1619 (pp. 15–21). Sun SITE Central Europe.

Ward, J. S., & Barker, A. (2013). Undefined By Data: A Survey of Big Data Definitions. arXiv Preprint arXiv, 1309(5821). Retrieved from http://arxiv.org/abs/1309.5821

Wastell, D. (2011). Managers as designers in the public services: Beyond technomagic. Triarchy Press Limited.

Webb, H., Burnap, P., Procter, R., Rana, O., Stahl, B. C., Williams, M., … Jirotka, M. (2016). Digital Wildfires. ACM Transactions on Information Systems, 34(3), 1–23. http://doi.org/10.1145/2893478

Webb, H., Jirotka, M., Stahl, B. C., Housley, W., Edwards, A., Williams, M., … Burnap, P. (2016). Digital wildfires. ACM SIGCAS Computers and Society, 45(3), 193–201. http://doi.org/10.1145/2874239.2874267

Webb, H., Jirotka, M., Stahl, B. C., Housley, W., Edwards, A., Williams, M., … Burnap, P. (2015). "Digital Wildfires." In Proceedings of the ACM Web Science Conference on ZZZ - WebSci '15 (pp. 1–2). New York, New York, USA: ACM Press. http://doi.org/10.1145/2786451.2786929

Weller, M. (2014). Characteristics and completion rates of distributed and centralised MOOCs. Retrieved from http://www.moocresearch.com/wp-content/uploads/2014/06/C9131_WELLER_MOOC-Research-Initiative-OU.pdf

Wen, M., Yang, D., & Rose, C. (2014). Sentiment Analysis in MOOC Discussion Forums: What does it tell us? Educational Data Mining.

Whalen, J. (1995). A technology of order production: computer-aided dispatch in public safety communications. Situated Order: Studies in the Social Organization of Talk and Embodied Activities, 187–230.

Whalen, M., Whalen, J., Moore, R., Raymond, G., Szymanski, M., & Vinkhuyzen, E. (2004). Studying workscapes. Multimodal Discourse Analysis, 208–229.

Whitmer, J., Schiorring, E., & James, P. (2014). Patterns of Persistence: What Engages Students in a Remedial English Writing MOOC? In Proc. of 4th International Conference on Learning Analytics and Knowledge (pp. 279–280).

Wikipedia. (2016). Manifold (fluid mechanics). Retrieved July 23, 2016, from https://en.wikipedia.org/wiki/Manifold_(fluid_mechanics)

Wildman, D. (1995). Getting the most from paired-user testing. Interactions, 2(3), 21–27.

Wilkowski, J., Deutsch, A., & Russell, D. M. (2014). Student skill and goal achievement in the mapping with google MOOC. In Proceedings of the first ACM conference on Learning @ scale conference - L@S '14 (pp. 3–10). New York, New York, USA: ACM Press. http://doi.org/10.1145/2556325.2566240

Witmer, D. F., Colman, R. W., & Katzman, S. L. (1999). From paper-and-pencil to screen-and-keyboard: Toward a methodology for survey research on the Internet. In Doing Internet Research: Critical Issues and Methods for Examining the Net by S. Jones (pp. 145–161).

Wright, K. B. (2005). Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. Journal of Computer-Mediated Communication, 10(3). Retrieved from http://jcmc.indiana.edu/vol10/issue3/wright.html

Yang, D., Sinha, T., Adamson, D., & Rosé, C. P. (2013). Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In Proceedings of the 2013 NIPS Data-driven education workshop (p. 14).

Yun, G. W., & Trumbo, C. W. (2000). Comparative response to a survey executed by post, email, and web form. Journal of Computer-Mediated Communication, 6(1). Retrieved from http://jcmc.indiana.edu/vol6/issue1/yun.html

Zubiaga, A., Voss, A., Procter, R., Liakata, M., Wang, B., & Tsakalidis, A. (2016). Towards Real-Time, Country-Level Location Classification of Worldwide Tweets. arXiv Preprint arXiv, 1604(7236). Retrieved from http://arxiv.org/abs/1604.07236

NIST Big Data Interoperability Framework: Volume 1, Definitions. (2015, October). http://doi.org/10.6028/NIST.SP.1500-1

The Big Bang: How the Big Data Explosion Is Changing the World. (2012). Retrieved September 24, 2016, from https://news.microsoft.com/2013/02/11/the-big-bang-how-the-big-data-explosion-is-changing-the-world/#sm.0000ugsllarkmfma11m5bf2twxx0z