Engineering serendipity in large scale learning environments

A design-based research investigation into the impact of visualising peer produced content in real-time in FutureLearn courses

Philip Tubman BA Hons, MSc

Thesis submitted for the degree of Doctor of Philosophy

Department of Educational Research and School of Computing and Communications

University of Lancaster

April 2021

Engineering serendipity in large scale learning environments

A design-based research investigation into the impact of visualising peer produced content in real-time in FutureLearn courses

Philip Tubman BA Hons MSc

This thesis results entirely from my own work and has not been submitted in substantially the same form for the award of a higher degree elsewhere.

The word length does not exceed the maximum permitted length of 80,000 words.

Signature ………………………

Luck is what happens when preparation meets opportunity

Seneca the Younger (54-15BC)

Philip Tubman BA Hons, MSc
Engineering serendipity in large scale learning environments
Thesis submitted for the degree of Doctor of Philosophy, April 2021

Abstract

This thesis investigates social learning in large scale courses, from the perspective of exploiting the benefits of massive participation. Specifically, I examine the affordances of the FutureLearn course platform, analysing their impact on learner interactions. I then create new affordances through a novel mediating artefact, the Comment Discovery Tool, and develop innovative pedagogical models which are refined through 3 phases of design-based research. The Comment Discovery Tool is an interactive visualisation of all learner commentary that allows learners to see conversations and emergent themes from the course in a non-linear fashion. In the second phase of design-based research I use formal learning design frameworks to introduce inquiry and reflection activities into the pedagogical toolkit. These are generally missing from the established model of large scale course design which values completion, progress and retention only. The third phase of design-based research continues the pedagogical innovation by encouraging learners to alter their writing style towards the development of communities of 'ambient affiliation'. This demonstrates that learning at scale requires a reconceptualisation of online courses, placing massive-ness and cooperation at the heart of the pedagogic design. This thesis is a case study into how this can be achieved by using design-based research, placing learners at the centre of the design process, and levelling up the human activity of learning to one where learners can extend the range of their own environment for the benefit of others. The research represents an original contribution because I demonstrate how real-time visualisations can encourage cooperative activity and demonstrate how pedagogical innovation can be achieved

through a rigorous user-centric analysis, starting from the materiality of the platform, and integrating theoretical frameworks. I also use a GPL open-source licence for the tool which enables others to download, remix and re-use the technology on other courses.

Table of Contents

Acknowledgements

I would like to express my gratitude to Dr Murat Oztok and Dr Phil Benachour for their supervision and support during this journey. Research is never a straight line and their encouragement, patience and collegiality has afforded me the space and time to make this project my own.

I must also extend an immeasurable thank you to 'the coders' who helped me greatly with technical developments: Michael Holland, Pierre Ciholas, Bryce Anderson, Tim Ellis, Adrian Fish, Tony Butler and Chris Lingwood. Without your support and help during this time we simply would not have a viable product, let alone a research project! Extra thanks to Tim, for the conversations throughout. Thanks also to Steve Wright, for your knowledge, support, and patience whilst I grappled with NVivo. You made the mistakes first, so I didn't have to, as with all good teachers!

Thank you to my employers Paul Harness, Andrew Meikle and Steven Powell, for your support and encouragement over the years.

Thank you to those at FutureLearn and the Lead Educators at Lancaster for allowing me to hijack your platform and courses with my pedagogical and technical interventions: At FutureLearn: Emily Pearce and Beatrix Daniel; at Lancaster: Judit Kormos, Simon Bainbridge, Sharon Ruston, Sally Bushell, Carly Stevens, Ben Dandelion, Fiona Edmonds, Sam Riches, Derek Gatherer, Alison Findlay and all the

mentors who worked on their courses. Without your support and belief none of this would be possible.

To all the members of the FutureLearn Academic Network, thank you for listening to my work in progress reports and for all your comments. To my interview subjects and those who completed the survey instruments, thank you for your time and thoughts, and to all the tens of thousands of learners who tried out the CDT, I don't know who you are, but every click contributed. I hope you discovered something serendipitous and new.

Finally, thanks to my parents for their unwavering support and encouragement throughout my entire education, and to my family, Jess and Tabitha, for their patience as I spent long hours locked away from them working on this project.

Publications derived from this research

During this project (2015-2020), I have published the following peer reviewed articles and posters to the IEEE International Conference on Advanced Learning Technologies, to the International Conference of the Learning Sciences and to the Computer Supported Collaborative Learning conference. I have also made 5 presentations to the FutureLearn Academic Network (2016-2018), 2 'work in progress' presentations internally to the Lancaster University Educational Research Department (2017), and 1 presentation to the Association for Learning Technology (ALT-C) conference (2018).

Date	Title	Publication
2016	Being Social or Social Learning: A sociocultural analysis of the FutureLearn MOOC platform	Poster; IEEE 16th International Conference on Advanced Learning Technologies, Austin, Texas
2016	Being Social or Social Learning	FutureLearn Academic Network; conference presentation
2017:1	Visualising social learning for discoverability	FutureLearn Academic Network; conference presentation

2017:2	Mass sociocultural learning: socio-technical and socio-material considerations	FutureLearn Academic Network; conference presentation
2018	Design-based research approaches towards enhancing social learning practices in MOOC platforms	Poster; 13th International Conference of the Learning Sciences, London, UK
2018	Comment Discovery Tool	FutureLearn Academic Network; conference presentation
2018	Comment Discovery and Serendipity in FutureLearn MOOCs	Association for Learning Technology (ALT-C); conference presentation
2019	Interactive visualizations to enhance social learning practices in MOOC platforms	Poster; 13th International Conference on Computer Supported Collaborative Learning, Lyon, France
2019	New Platform Affordances for Encouraging Social Interaction in MOOCs: The "Comment Discovery Tool" Interactive Visualisation Plugin	Short paper; IEEE 19th International Conference on Advanced Learning Technologies, Maceió, Brazil

2020	Can we afford it? The cybernetic determinants for pedagogical models in MOOCs	Full paper; IEEE 20th International Conference on Advanced Learning Technologies, Tartu, Estonia

These publications and presentations represent the material contribution to knowledge from the work of this thesis during 2015-2020 and the feedback and peer review process has been an enormous help in defining and refining the scope of this project.

List of Figures and Tables

Figures

Tables

Chapter 1: Introduction

1.1 Research aims

MOOCs have been a growing part of the educational landscape for just over a decade and their success in engaging millions of learners has established them as a phenomenon that is likely here to stay. At the beginning of the 2010s, they were hailed as being a major disruptive force in higher education, with Sebastian Thrun claiming in Wired magazine that in 50 years there would only be 10 universities, and his MOOC start-up Udacity would be one of them (J. Young, 2019). Others were less optimistic about MOOC learning, interpreting it as a cynical and strategic move by Silicon Valley entrepreneurs into state education (Bady, 2013), and related to this, having little pedagogical value (Baggaley, 2014).

In this thesis, I do not make binary claims about MOOCs, rather suggest that their continuity in the established educational landscape merits a close examination from a pedagogical perspective, which has been identified as an under-researched area (Bayne & Ross, 2014; Cuffe, 2015). Specifically, I examine the relationship between the affordances of the online environment, its impact on learner interactions, and its potential for innovative pedagogical approaches. There is good reason to closely examine the platform architecture, as the most common MOOC platforms are not open-source software but are designed and developed by profit seeking companies. Their relationship with academic institutions is as 'partner' rather than a 'software supplier', meaning that the tools provided on these platforms are intransient and not open for negotiation with institutions; this inevitably will restrict pedagogical approaches and lean primarily towards a profit motive. In the case of FutureLearn, it is not only the platform tools which are fixed, but also the principles of course design, such as the length of videos, the 'readability' of articles or the number of steps per week. This is explored in detail in the methods chapter where I demonstrate how these factors skew towards a progression based pedagogical approach, and profit motive.

However, attracting thousands of learners from a global audience and having a low barrier to entry creates a wide and heterogeneous cohort presenting new opportunities for pedagogical innovation. My starting point for this research is that the size and heterogeneity of the MOOC cohort can bring new pedagogical opportunities through increased possibility of serendipity, rather than being an 'overload problem' which needs to be designed around. This perspective requires recalibrating platform tools to engineer opportunities for the multitude of new serendipitous encounters which are less likely with smaller cohorts. In terms of social features, the Coursera MOOC platform uses threaded discussion forums which are common in online courses of 30-50 students but become overwhelmed quickly in a MOOC context (Brinton et al., 2014). The FutureLearn approach is to present peer produced content chronologically on each learning step, which is less susceptible to visible overload, however this approach 'hides' content by spreading it out across the course and hindering discovery. Neither approach leans towards a pedagogy which encourages genuine open participation, and both (in different ways) effectively place the learner in an isolated context, alone in the crowd.

> Currently available platforms for learning at scale are frequently impoverished on [the social] dimension, which is a problem for all, but especially under-served students who are in greater need of support. Learning in these environments is typically mostly a solitary experience. Though there are almost always discussion forums included in these environments, they are often just an appendage, and not effective in meeting the needs of learners, especially under-served learners who need more support. (Rosé & Ferschke, 2016, p. 661)

In this project, I extend the social features of the FutureLearn platform, developing the ‘Comment Discovery Tool’ (henceforth: “CDT”) which visualises all peer produced content into an interactive word cloud, filtering commentary based on user-selected terms. This new feature has the impact of representing thousands (or even hundreds of thousands) of comments by their shared semantic components and provides a means to filter and join conversations based on affinity, rather than chronology. An example of the CDT from the 2020 “Dyslexia and Foreign Language Teaching” course can be seen in Figure 1 below, with the term ‘comprehension’ selected, the cloud visualising words from comments which also have ‘comprehension’ (size represents frequency), and a list of these comments below, including a link to join in the conversation:

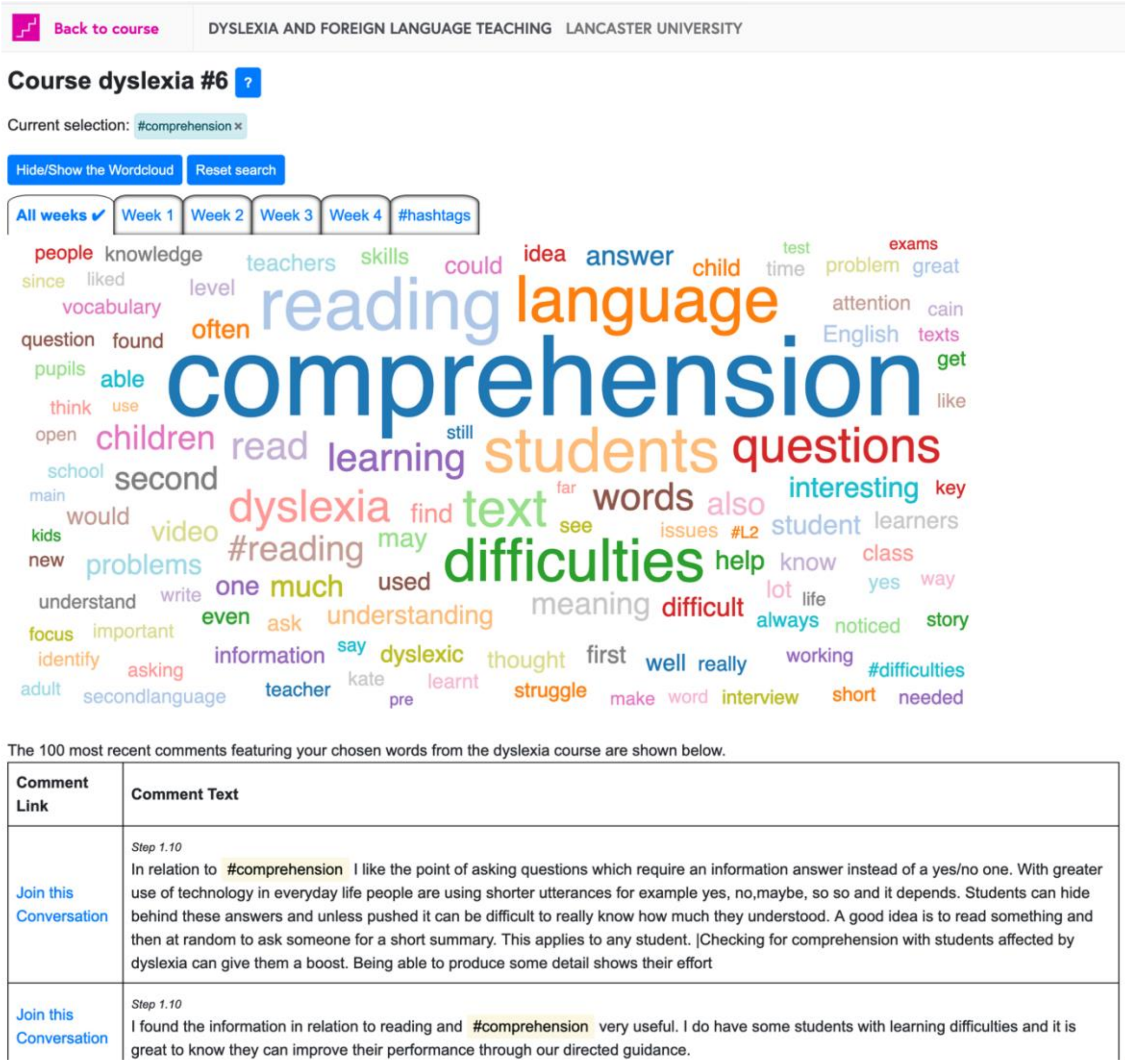


Figure 1 The Comment Discovery Tool, "Dyslexia and Foreign Language Teaching" course, 2020. There was a total of 22229 peer produced comments in this course

In the development of the CDT, I have employed 4 key theoretical frameworks to identify the gaps in platform affordances, and provide direction for design and development:

1. A cybernetic analysis of virtual learning platforms, building on the work of Britain and Liber (2004), to identify affordance gaps in the 'MOOC learning system' and where this inhibits or encourages the pedagogical model. A detailed analysis of the FutureLearn platform, and its designed affordances is explored in Chapter 4: Methodology.
2. The ABC framework (C. Young & Perović, 2016), an operationalised account of the Conversational Framework (Laurillard, 2012) which describes the patterns of learning activities within the course structure. ABC (not an acronym) is a simplified account of pedagogical patterns and I also draw from Goodyear's ACAD (Activity Centred Analysis and Design) framework (Goodyear & Carvalho, 2016) in order to take account of more complex sociomaterial elements, such as the concept of time working as an 'invisible hand', guiding the learning process. This work builds on the Cybernetic analysis of the FutureLearn platform in the Methodology chapter, by examining how the materiality of the platform affects the pedagogical opportunities. In Chapter 6, Design Based Research phase 2 (henceforth: DBR2), I examine how the new technology introduced in Chapter 5, Design Based Research phase 1 (henceforth: DBR1) may unlock pedagogical approaches that are restricted by the cybernetics of the default platform design and provide evidence through 10 semi-structured interviews of both the restrictive default platform, and the possibilities afforded by the CDT.
3. Stigmergic Design (Elliott, 2007), describing how to design collaborative systems which improve with scale. This is an important paradigm from which I pivot participation 'overload' into an advantageous state. A key component

is to break down collaborative activity into 3 constituent and co-dependent components: 1. coordination, 2. cooperation and 3. collaboration. The coordination power of the visualisation artefact is integral to its design, as described and demonstrated in DBR1 and I demonstrate in Chapter 7, Design Based Research phase 3 (henceforth: DBR3) how this can be leveraged towards a participatory pedagogical approach by allowing the users to influence the coordination affordance, creating a learning environment geared towards cooperation. The further work of this research agenda is to reinforce the ability of users to cooperate through further pedagogical innovation, and the development of a contributory culture which rewards cooperative activity.

4. Cultures of Participation (Fischer, 2011), a 5 step framework which describes the extent to which computer systems enable people to participate and to contribute actively in meaningful problems, starting from a consumption mode of interaction leading to one where users are able to design how the system should work. This focuses on the changes in the behaviour and organization of users which are brought about by technical innovations, rather than describing the new affordances. I use this to describe how the new pedagogical approaches can empower learners to take control of their own learning and learning communities, levelling up MOOC learning from an individual experience to one where learners feel they are participating in a collective enterprise. This framework is used in all the phases of DBR to measure the effect of the affordances and pedagogical interventions against the agency of the learners, as the intention is to provide learners with

increased agency to cooperate and collaborate on meaningful problems through participation in the festival of learning and sharing that is implicit to MOOCs.

This project uses a design-based research methodology (Brown, 1992) (henceforth: "DBR") to answer the overarching question of how to engineer participatory pedagogy into the social learning environment, and follows 3 iterations of the design, firstly engineering new methods of encouraging serendipitous encounters, then developing the basis for learner-led communities of 'ambient affiliation' (Zappavigna, 2011).

Each phase of DBR (chapters 5-7) focuses on different elements of the research and their research questions are emergent through the results and reflections of the previous phase. The overarching research questions for each phase are:

1. DBR1: How can a new platform tool affect levels of overt sociality, as defined by the novel conversational unit analysis, and do learners intuitively perceive the designed affordances of the intervention?
2. DBR2: Can the effects of the intervention be replicated in terms of the novel conversational unit analysis? How can the affordances and benefits of the platform tool be made more visible to learners through task proposals? How do learners qualitatively interact with the platform and the new tool?
3. DBR3: How can the pedagogical approach be enhanced through use of the intervention by affording learners the agency to extend their use of the platform?

The overall contribution that this project makes are:

1. as a case study in utilising a design-based research method to develop new tools fit for the needs of learners in MOOCs (Tubman et al., 2018);
2. the development of a heuristic model for analysing thousands of conversations based on the social materiality of FutureLearn platform, in order to measure effectiveness of an intervention in the social learning sphere (Tubman, Benachour, et al., 2019);
3. the development of new learning activities, suited for self-directed learners in MOOCs, which encourage cooperative behaviours that work in conjunction with the new affordances created by the CDT and serving as a basis for self-sustaining, peer-led community formation (Tubman et al., 2020).

The following sections of this chapter briefly introduces the main concepts and set out in more detail how they inform the project aims, then outlines the structure of the thesis and the publications which were made from this thesis, demonstrating a material contribution to knowledge.

1.2 Keystone concepts

As described above, the design and development of this research project is influenced predominantly by 4 theoretical frameworks, which are explained more fully in Chapter 3: Relevant Theories. Other concepts also influence the work, but these frameworks and the concepts from which they are made are central to the analysis of the data throughout all the phases of the research project. Cybernetic systems analysis is applied to the FutureLearn platform in Chapter 4: Methodology where the premise of the CDT artefact's design is explained, and DBR1 (chapter 5) collects evidence related to the initial technological implementation, mainly from a point of view of whether these new affordances are understood and utilised. DBR2 (chapter 6) continues the narrative to examine how these affordances could be better realised to more learners through pedagogical scaffolding, utilising learning design frameworks and the concepts within them, and DBR3 (chapter 7) uses the Stigmergic principles of coordination, cooperation and collaboration by making pedagogical and technological alterations in order to 'level up' the learners' agency by enabling them to extend the range of their environment.

1.2.1 Cybernetic systems analysis

My investigation starts with an analysis of the affordances of the FutureLearn platform, which builds on previous research by Britain and Liber (2004) that develops a framework for evaluating virtual learning environments against pedagogical models. Their framework examines the organisational structures of the environment and relates these to the Conversational Framework, which is a well-known model of effective teaching practice (Laurillard, 1993) and also the theoretical basis for the FutureLearn platform (Ferguson & Sharples, 2014). Britain and Liber's framework provides 5 criteria to measure whether the system will facilitate or inhibit the conversational framework: *resource negotiation, adaptation, self-organisation, monitoring and individualisation*. The FutureLearn platform is to some degree inflexible across all these dimensions although this is not uncommon in any Virtual Learning Environment. For example, in the case of ‘adaptation’, this dimension lies in stark contrast to the completion criteria which is so often cited as a major metric for determining the ‘success’ of a MOOC.

> It would be beneficial if it were possible for the teacher to vary the way the course or module is provided to incorporate new materials and processes, but many VLEs have to be pre-packaged and don’t support that flexibility. It may be that materials and pedagogies are not suited to the particular cohort or part of it – it should be possible to change them based on new opportunities. (Britain & Liber, 2004, p. 17)

It is contradictory to have a fixed set of learning steps, the completion of which signifies overall participation, and for materials to be changed whilst the course is in motion to suit the cohort. The design choice of fixed content/ progression metric over adaptive delivery is an example of how top-level architectural design choices affect pedagogical options. As a comparison, other courses might track completion outside the platform, through institutional record keeping, or determined solely by successful attainment in assessments rather than progression through a defined set of content; it is worth pointing out that these approaches are not scalable in the same way. Britain and Liber suggest that adaptation is intimately linked with ‘teaching practice’ and needs to be balanced in such a way that pedagogical approaches can compensate for rigid course structures. This presents a further challenge for MOOCs as the close facilitation and discussion with teachers, a requirement of the conversational framework, is also not possible at scale. It may be argued that the increase in peer participation compensates to some degree (indeed it *must*), however this again places emphasis on the material affordances of the platform as mediating artefacts, and the pedagogical scaffolding of the learning activities to encourage such cooperative practice. Too often in MOOCs learners are neither encouraged to adequately support each other, nor are the discussion features sophisticated enough to mediate effective peer interactions at scale. It is these aspects of MOOC learning that I seek to improve in this project, in terms of developing tools and pedagogical approaches which empower learners to break out of their isolation and feel part of a collective enterprise.

1.2.2 Learning Design Frameworks

In order to carefully examine the pedagogy of learning at scale, I employ the ABC learning design framework (C. Young & Perović, 2016) in order to identify pedagogical gaps in the FutureLearn platform, some of which are caused by the platform itself, and limit its collaborative potential. This framework builds on Diana Laurillard's later work: "Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology" (Laurillard, 2012), which divides teaching activities into 6 major categories: *acquisition, inquiry, collaboration, discussion, practice and production*. The ABC methodology transparently represents the proposed teaching activities of a course into one of these categories to establish a simplified pedagogical pattern applicable to any course. MOOCs are online only courses, so it is critical that they contain mediating artefacts which are sufficient for these activities.

Where MOOC platforms excel in terms of scaling access to content curated from leading academics (acquisition), and computer aided assessment for recall (practice), this comes at the expense of collaborative learning activity (inquiry, (focused) discussion, collaboration) usually associated with the sociocultural account of online and distance learning (Beetham & Sharpe, 2013; Conole, 2013; Garrison, 1997). In other words, as demonstrated above, they necessitate a more independent study model that is in stark contrast to the more accepted socio-constructivist approaches to learning (Bayne & Ross, 2014). This project examines how inquiry, discussion and collaboration can be fostered at scale in the FutureLearn platform and the second phase of DBR uses an interview method to determine how learners typically interact with the platform and the CDT plugin. It is in this second phase of DBR that I extend upon the simple ABC framework and introduce the ACAD framework (Goodyear & Carvalho, 2016) which specifically includes the environment as a key relational factor. “Time” is identified as an invisible aspect of the sociomaterial setting, and this becomes an important piece of the analysis in terms of how learners interact.

1.2.3 Stigmergic design paradigm

FutureLearn is the largest MOOC provider in the UK and has over 14 million users. Lancaster University has developed 21 courses in partnership with FutureLearn and has ‘run’ these courses multiple times, totaling 49 course instances. Over 200,000 learners have enrolled on these courses, generating over half a million individual comments. These raw statistics demonstrate the scale of the informational and organizational challenge.

In order to develop a tool that is able to cope with participation at this scale, I have drawn upon the concept of stigmergic design (Dron, 2006; Elliott, 2007) which builds on ideas of harvesting ‘swarm intelligence’ (Van Dyke Parunak, 2006) through the utilization of ‘signs’ left in the environment for future users. The archetypal example of this type of design in nature is in the building of large ant hills, where ants leave pheromone traces as signs to direct those who come after them (Dron, 2006). Mark Elliot develops the theoretical construct of Stigmergy in his PhD thesis by identifying 3 component parts of collaborative activity, which are dependent on each other: coordination, cooperation and collaboration (Elliott, 2007).

FutureLearn’s social features allows comments to be left on each learning step and paginates them chronologically. This allows the platform design to appear clean and encourages comments to be specific to the content of the page in question, but coordination is managed primarily by chronology and not affinity, effectively creating mini ‘time silos’ of commentary. In this sense it is feasible for learners who are on the step at the same time to cooperate, but this is a much smaller subset of the total cohort and not a reflection of the true heterogeneity of the group. Collaboration in a sociocultural sense requires that members of a group have a shared understanding of a problem and cooperate by performing roles to complete the task or problem. The heterogeneity of the cohort makes it unlikely that learners working on a learning step at the same time can work together in this manner, however if content were coordinated according to affinity, then self-directed learners can form communities of interest, or 'ambient affiliation' (Zappavigna, 2011), and a common understanding could emerge. The increased size of the cohort in MOOCs may result in further collective enterprises being developed from these serendipitous encounters.

I take an iterative, design-based research approach to developing the new plugin[1], taking account of each of the components of stigmergic design. The first iteration of the tool in DBR1 coordinates peer produced content by visualising words as an interactive word cloud, and filtering comments according to selected terms. This allows learners to discover missed conversations and provides a link for them to participate, but it is apparent from feedback in DBR1 that these affordances are only available to those who perceive this designed intent; others do not recognize this and some feel the tool is a distraction. Therefore, the subsequent phases of DBR examine the learning activity around the tool, taking learner feedback at each stage, and refining the affordances of the tool and its placement within the course to make it more useful and apparent. The final iteration of the tool in DBR3 steers the pedagogy of the MOOC towards cooperative and self-sustaining learning

[1] A 'plugin' is an extension or add-on to a computer system which adds additional functionality or extends its capabilities

communities through careful activity design which encourages peers to support each other (including those who are ‘behind them’) through their use of the tool and style of their commentary.

1.2.4 Cultures of Participation

The final ‘keystone concept’ for this thesis draws on Gerhard Fischer’s ‘Cultures of Participation’ framework (Fischer, 2011), in order to track learner behaviour against a framework for social computing which defines the extent to which users are supported to actively engage in meaningful problems. This framework speaks to all new media and relates most specifically to the ability for technical innovation to change the behaviour and organization of its users. At the base level of the framework (Level 0), Users are ‘unaware consumers’ who can only consume content developed by ‘high-tech scribes’; they have no means to become designers themselves, and Fischer claims this contributes toward “the degeneration of humans into ‘coach potatoes’, for whom a remote control is the most important instrument of their cognitive activities” (Fischer, 1998). The foundation for this framework is work by sociologists and educators Ivan Illich, Neil Postman and Malcolm McLuhan, amongst others, who bemoan how the consumer mindset permeates through society, making citizens who feel left out of political decisions and are denied the opportunity to take an active role in shaping the world.

> A consumer mindset is not limited to television. In our educational institutions learners are often treated as consumers, creating a mindset of consumerism for the rest of their lives (Fischer, 1998)

The Cultures of Participation framework grounds this project; it reminds us that developing active learners and communities is the cornerstone of participatory education and the end goal of this project. I am not developing social learning pedagogies to discover efficiencies in mass scale tutoring, rather empowering people to define their own problem spaces and meeting like-minded others with whom they can discuss, collaborate and make plans. My starting point is that the MOOC environment is a common ground, a safe space, and a place where thousands of similarly interested individuals can congregate; the tragedy is that the current platforms and pedagogies do not build on the safe space to encourage the types of spontaneous and serendipitous engagements which may spark something new. They are instead focused on keeping the space simplistic so they can shepherd thousands of users to the end, like a tour guide in an art gallery herding tourist groups towards the gift shop. My intervention (the CDT and associated learning activities) subverts this linear pathway, encourages and empowers learners to define their own affinity groupings, which includes vicarious users, and to leave visible signs which say: 'I am here, I am interested, and I want to connect'. In this respect the users are 'levelled up' to the penultimate stage of the Cultures of Participation framework: 'Level 3: Users extend the range of the environment'.

1.3 Thesis structure

1.3.1 Scoping the literature

MOOCs are a relatively new artefact of education, but the term connects the divergent discourses of e-learning (online courses) and open educational practices; there is an underlying tension in this connection which is explored in the literature chapter. I use the seminal article by Anna Sfard: On Two Metaphors for Learning and the Dangers of Choosing Just One (Sfard, 1998) as a means to highlight this fundamental tension and as framework for my own pedagogical development of MOOC learning. Sfard uses the language of metaphor to delineate between learning through acquisition and learning through participation. Formal e-learning and the types of online distance education offered by higher education institutions firmly rests on the 'learning through participation' metaphor and is grounded in sociocultural learning designs (Beetham & Sharpe, 2013; Conole, 2013); conversely, the philosophy of 'Open Education' foregrounds the free nature of resources, discusses the legalities of access and therefore prioritises 'learning as acquisition', or knowledge as a commodity.

It is common throughout the literature to divide MOOCs into 2 categories: the xMOOC (an extension of a campus course) and the cMOOC (a realisation of the 'connectivist' learning theory (Siemens, 2005) and an extension of open educational practices). I note that when making this distinction, the origin stories of learning (their 'learning metaphors') are switched – cMOOCs are an extension of open education, yet draw on a participatory metaphor for knowledge construction, and xMOOCs are an extension of campus e-learning but are often reduced to the learning through acquisition metaphorical account.

However, neither conceptualisation of MOOC explores the materiality of the learning, and both perceives technology in instrumentalist terms, as the “neutral and invisible means to achieving educational goals” (Hamilton & Friesen, 2013). Indeed, the failure to account for the first letter in the acronym, ‘M’ for “massive”, becomes a major weakness for both types of MOOC. For both conceptions of MOOC, massiveness presents a ‘problem of overload’, and it is a failure of the respective systems design to account for this adequately. The forums in xMOOCs which become overwhelmed by activity are limited by structural and material factors which reduces their utility (Gillani et al., 2014), and the distributed nature of participation in a cMOOC overwhelms learners who report difficulty ‘keeping up’ (Mackness et al., 2010). I demonstrate that neither of these findings is surprising, as discussion tools have long been a site of contention (Hewitt, 2005; Scardamalia, 2002) and navigating the intricacies of vast information networks is indeed a skill in itself (Kop, 2011).

I argue that it is through focusing on the organisation of learner behaviour that we can account for massiveness in MOOCs, and it is through leveraging learning analytics, which are common for post-hoc analysis of learning in both types of MOOC, that we can cut through these ‘overload’ problems and encourage innovative learning designs where massive participation and diversity can be embraced. The development of such a toolset draws on a line of literature from open education:

that of peer production systems and considers Caroline Haythornthwaite's model of 'light and heavyweight peer production' (Haythornthwaite, 2009), but notes the differences between systems she cites and those necessary for organising commentary in a MOOC. These differences can also be understood in terms of the 2 metaphorical accounts of learning, as peer production in the open educational sense emerges from an 'acquisition metaphor' for knowledge in terms of collaboratively building coherent knowledge bases (such as 'citizen science' initiatives); it is the function of the underlying pedagogy to develop this into a participatory learning metaphorical account in terms of the development of knowledge communities.

1.3.2 Theoretical perspectives

As discussed in the research aims section above, the broad aim of this project is to develop participatory pedagogical approaches into the MOOC educational context. Therefore, the baseline theoretical concept is 'sociocultural learning', firstly as defined by Vygotsky (1978), who later extended this into 'activity theory', and which has been applied to e-learning through the use of learning design frameworks and the concept of 'social constructivism' (Beetham & Sharpe, 2013; Conole, 2013).

> There has been a paradigm shift in the field due to new thinking around learning theories from behaviourism, through to cognitivism and finally constructivism (Mayes and De Freitas 2004). These theories led to the

> development of particular uses of technology designed to support the underpinning principles of the theories. (Conole, 2013, p. 3)

Vygotskian activity theory proposes activity can be understood through the relationship between artefacts, rules and division of labour and it is this basic insight which acts as a template for the design of the artefact (CDT) in each phase of design-based research. There is considerable overlap and alignment between this account of activity and the pedagogical premises for learning design frameworks, for example Bayne and Ross (2014) describe MOOC pedagogy as “something that emerges in complex negotiation between the platform, the teaching approaches of the academic team developing the course, and the pattern of learning interactions as the course is played out”. Goodyear’s ACAD learning design framework (Goodyear & Carvalho, 2016) talks of the interaction between environment, epistemology and social activity, and this itself aligns with Vygotskian activity theory in terms of the relationships between artefacts (learning environment/ tools), rules (suggested learning activities) and division of labour (the confluence of diverse social activity).

Although it is not possible to fully separate these elements in each phase of development, the first DBR phase focuses on the tool itself and analyses the affordances which are commonly perceived by learners, the second phase looks more closely at the pedagogical elements, such as the suggested learning activities, learners’ modes of participation and how this is encouraged by the environment. Finally, the third phase analyses the emergent social activity in terms of dividing the labour required for cooperative activity. At each stage I discuss how the findings relate to the theoretical premises of the project and which suggestions should be taken forward into the subsequent phase.

1.3.3 Methodology

The methodology chapter introduces the design-based research (DBR) methodology, which is defined as practical, and enacted in real world settings. This project seeks to engineer solutions to improve MOOCs, rather than seek to explain their phenomena, so I take a design-based approach to measuring the success and iterating the product based on feedback.

> the natural sciences are concerned with how things are … Design on the other hand is concerned with how things ought to be (Simon, 1969).

DBR is relatively new in the domain of educational research, and owes much to the pioneering work of Ann Brown, who states that DBR is engineering “interventions that not only work by recognizable standards, but are also based on theoretical descriptions that delineate why they work, and thus render them reliable and repeatable” (Brown, 1992, p. 143). DBR interventions are enacted in a natural setting to make sense of the rich systems of social relationships and technologies which make the complete picture, and designs are iterated and modified dynamically according to the situation. This does not make the design experiments less valid as they are describable by theory and may in fact contribute to developing theory rather than evaluating it.

> A design science uses and contributes to theoretical science, but it builds design principles rather than theories, and the heuristics of practice rather than explanations, although like both the sciences and the arts, it uses what has gone before as a platform or inspiration for what it creates. Teaching is more like a design science because it uses what is known about teaching to attain the goal of student learning, and uses the implementation of its designs to keep improving them. (Laurillard, 2012, p. 1)

I also use this chapter to set out a detailed analysis of the FutureLearn platform, as a starting point for the designed intervention. In this section of the methodology chapter, I examine the entire platform through the lens of Human Computer Interaction (HCI) and Usability design, comparing these methods of analysis with the Cybernetic analysis described above. Taken together, these perspectives detail the affordances of each page, and demonstrate whether they promote or inhibit the conversational framework. The lack of affinity-based discoverability limits learners to a progression-based view of learning and does not account for the wide, heterogeneous cohort of learners, which is crucial to understanding how learning happens *at scale*. This is a limitation as MOOC learners are self-directed, they need to engage with ideas that are important to their own learning goals through conversations with other learners that they couldn't encounter in other circumstances. The remainder of this chapter details the technical implementation of the CDT, in terms of the design choices I made between certain visualisation technologies and their practical implementation.

1.3.4 Findings from the phases of design-based research

There follows 3 chapters examining 3 iterations of the intervention. The first phase concerns the prototyping and details the quantitative methods for measuring effectiveness across the entire comment corpus; the second phase details how the pedagogical scaffolding works in line with the mediating artefact and examines the qualitative interactions of the learners, and the final phase makes changes to the artefact based on user feedback throughout and alters the pedagogical approach of the entire course to lean it towards a participatory pedagogy. There is a discussion section in each of these chapters following the presentation of data and findings.

1.3.4.1 DBR1

DBR1 describes how the CDT is integrated into the FutureLearn course, and measures its impact through the development of a quantitative method, which build on the work of the first pilot study (Tubman et al., 2016) and also extends on the work of Shi-Min Chua and Mike Sharples (Chua et al., 2017) who classify individual comments according to their placement in conversations. Their classification system builds on the materiality of the platform because conversations on FutureLearn are not threaded, so responses are either ‘first reply’ or ‘further reply’, and from a new person or the original commenter. I extend this schema to classify entire conversations which is the basis of a heuristic, determining whether a conversation is a quick 1-2 between 2 people, or whether it is an extensive discussion. If this is combined with the total number of unique participants in a conversation, it is possible to see which conversations are likely to have greater diversity and development of ideas within. There are a total of 7 possible categories of conversation, and the classification is scripted, meaning that it can be applied to an unlimited number of conversations. The novel classification schema and script are an original contribution to knowledge made by this thesis.

I use this quantitative analysis technique to determine whether the CDT is making a significant impact on overt sociality in courses by applying the script to conversations in non-CDT courses, and in CDT-enabled versions. I use an ANOVA comparison of means to analyse the effect of the CDT on conversation length and number of people per conversation, and also modify the heuristic to create 4 categories of ‘conversational type’, analysing the proportions of conversations in each of the categories between the 2 experimental conditions. There is also a further quantitative analysis determining correlation using a Spearman’s Rho (Weir, n.d.) on a survey instrument, to determine the correlation between learner preferences, and perceptions of the new artefact. Finally there is a qualitative (abductive) analysis of several hundred comments which are left on the CDT ‘step’, to determine how the affordance is being used and what improvements should be taken forward into the second phase of DBR.

1.3.4.2 DBR2

DBR2 describes how the CDT artefact is used by the learners through a qualitative interview method and replicates the same quantitative methods above to determine the levels of impact and correlation, drawing on a much larger dataset in both analyses. The tool is integrated into the learning design of the course by using the ABC framework to develop a pedagogical pattern and inserting where a learning activity focused on 'inquiry' will have the most impact. I place the step towards the end of each week and scaffold the activity as an opportunity for 'reflection' with the CDT artefact as a support for extensive reflection.

My interviews seek to discover how learners use the platform at large, and how they perceive the CDT supported reflection and inquiry activity. I develop 4 distinct methods of interaction with FutureLearn and establish the main method as one where learners enjoy the flexibility of learning at their preferred time, and the interactions with others on the platform as serendipitous, random, and occasionally beautiful. Most learners do not look back at previous steps unless they have a notification of a reply to their comment, and most will only read 10-20 comments on each learning step; clearly these findings are relative to each other. The CDT is perceived as a tool for extended reading as well as a tool for joining in new conversations, which suggests that the quantitative methods used above which measure overt sociality only account for part of the utility of the tool. This leads me to make further changes to the affordances of the tool and the scaffolding to develop more peer-supportive pedagogical approach, which is explored in DBR3.

1.3.4.2 DBR3

In DBR3 I make 2 changes to the way the CDT functions: firstly, by adding a 'tab' which displays terms which have been preceded with a 'hashtag' symbol ("#") and encouraging learners throughout all the steps of the course to use hashtags in their commentary, so communities of 'ambient affiliation' can emerge; secondly, by asking a simple question on the CDT page: "What words did you click on and why?". This simple question encourages learners to articulate their thought processes and I demonstrate that for learners who read between 10-20 comments (the most common), they are likely to see 4-5 comments from other learners of this type, which itself acts as an instruction for their own use of the artefact. This is an example of how tools, pedagogical approaches and stigmergic peer production principles work together. The tool provides the coordination of commentary, the pedagogical approach proposes a micro-action (sharing your thoughts) and the resultant commentary acts as a signal for the learners who come behind in terms of the interests of the learners in the cohort and how to connect with them. This novel pedagogical approach which ties all these aspects together to develop sustainable participatory learning *at scale* is another example of where this thesis contributes to knowledge. I follow up by suggesting several enhancements which could be made across the entire learning environment, but which are outside the remit of this project.

Chapter 2: Literature

2.1 Introduction

This section explores relevant themes within existing literature which inform the project's design. I first consider how learning has been historically defined in broad terms, in relation to its epistemic nature (learning as acquisition vs participation), and the wide political implications of these conceptualisations (knowledge as commodity vs community-based concepts) (Sfard, 1998). I then continue to discuss how these themes might relate to online learning (formal e-learning and informal open education) and to the emergent MOOC context.

I have also differentiated between conceptions of e-learning which have integrated learning designs grounded in sociocultural theory into formal courses, and conceptions of e-learning in the informal open education literature; this is because although they both concern online education, they foreground different aspects of learning. Where Open Education foregrounds the 'free' nature of online resources, and discusses the democratisation of content and methods of licencing and distributing content in a way which protects creators and encourages peer production (i.e. knowledge as commodity, therefore acquisition metaphor dominant), e-learning in the formal course context discusses methods of embedding participation into online environments through learning designs grounded in sociocultural theory (i.e. learning *through* participation, therefore participation metaphor dominant).

The MOOC context can broadly be divided into xMOOCs: ‘X’ for eXtension because they are supposedly extensions of campus courses in that they replicate the curriculum of a campus course online; and cMOOCs: ‘c’ for Connectivism as they emphasise “social networked learning” through a Personal Learning Environment (PLE) and are an extension of open education. Other methods of distinguishing between MOOCs have been suggested, such as ‘associative’, ‘cognitivist’ and ‘connectivist’ (Conole, 2014a, 2015) which refers to learning design frameworks.

Learning design frameworks such as 7C (Conole, 2014b) which supports Conole’s taxonomy are important and this is explored in more detail below in relation to affordances of tools and curriculum design. However, it is acknowledged here because the ‘learning activity’ C’s such as ‘communicate’ and ‘collaborate’ are necessarily restricted in current xMOOCs by structural factors such as the relationship between the affordances of the platform and the scale of participation, so from a learning platform/ environment perspective, it still makes sense to maintain the xMOOC/ cMOOC distinction at this point:

> In xMOOCs, the focus is more on the individual so communication might be mainly restricted to interaction with the tutors via email. Similarly, the Collaborate C is about fostering mechanisms to enable collaboration or

group work. Again, this is likely to be more prevalent in cMOOCs than xMOOCs. (Conole, 2015, p. 248)

On the surface, xMOOCs resemble traditional courses in terms of certain measures (curriculum, engagement, completion) and there is a body of literature which examines these factors in terms of learning analytics, or 'big data' (Ferguson & Clow, 2015; Fincham et al., 2019; Joksimović et al., 2018; Kizilcec et al., 2013). However, from an e-learning perspective, they fundamentally foreground an acquisition metaphor for learning (or an associative, cognitivist learning design framework) as a solution to the 'problem' of scale. That is to say, they necessitate a more independent study model that is in stark contrast to the more accepted socio-constructivist approaches to learning (Bayne & Ross, 2014) and this is also demonstrated by the quote from Gráinne Conole above. The elimination of important teacher-student interaction and lack of appropriately structured student-student interactions into the design of learning activities means they can only weakly represent a traditional campus course. Without these scaffolded interactions, others argue that they act as a supersized behaviourism where "increasing student numbers and reducing student support produces an educational value meal" (Baggaley, 2014, p. 160).

On the other hand, the cMOOC perspective is foregrounding a participation metaphor for learning, but without the structure of curriculum, or explicit grounding in learning design frameworks. That is to say, from some perspectives it is seen uncritically as a good thing to expose thousands of learners to each other in a learning network (Stewart, 2013), and furthermore from the related ‘rhizomatic learning’ perspective, this is encouraged because more content is produced (Cormier, 2008). However, more is not necessarily better as it creates problems of overload at the systems level (Barricelli et al., 2015), the network or community level (Bozkurt et al., 2016) and the level of the learner themselves (Kop, 2011; Mackness et al., 2010).

Neither approach adequately takes into account the nature of technological affordances which mediate the interactions or structural factors which scaffold the learning process, which are important aspects of the learning design literature (Conole, 2013). Established ideas about pedagogy and technology are reinforcing a dualistic interpretation of MOOCs; however, both these views perceive technology in instrumentalist terms, as the neutral and invisible means to achieving educational goals (Hamilton & Friesen, 2013). There is an under-researched area in taking a ‘system design’ approach to the problem and examining specific affordances of technology in relation to how learning design can respond to the different problems of overload presented when learning at scale. ‘Affordance’ is an important part of the discourse on learning design relating to earlier e-learning theory and practice (Conole et al., 2004; Conole & Dyke, 2004; Laurillard, 1993; Oliver, 2005) and should be explicitly considered, especially if the aim is to develop ‘cultures of participation’ (Fischer, 2009, 2011, 2014, 1998) or constructivist learning designs which work in the scaled learning context.

The new affordances designed in this project use visualisations of learning analytics, which are made available to learners at 'learntime' (an important distinction for how learning analytics are usually deployed to measure learning *after the event*) and use the principle of coordination to cut through the problems of participation, information and cognitive overload to enable learners to self-regulate their learning goals, whilst also embracing the large-scale eclectic and diverse participation. This ability to cut through both information and participation overload also makes room for innovative learning designs where educators can embrace massive participation and diversity to satisfy learners' personal goals.

2.2 Historic conceptualisations of learning

Rather than review the entire literature on the multitude of learning theories and definitions of learning, this section reviews a seminal paper by Anna Sfard on the debate between educational researchers on the question 'what is this thing called learning?' This article does not critique or support one theory or another, but reduces theoretical standpoints to their metaphorical origins:

> [Metaphors' …] special power stems from the fact that they often cross the borders between the spontaneous and the scientific, between the intuitive and the formal. Conveyed through language from one domain to another, they enable conceptual osmosis between everyday and scientific

> discourses, letting our primary intuition shape scientific ideas and the formal conceptions feed back into the intuition. (Sfard, 1998, p. 4)

I use this as a starting point for a discussion on learning theory because it is possible to unpack these broad conceptualisations in terms of both their epistemological standpoints and relation to the process of learning, and the resultant political implications in terms of the value of educational provision in society. These are the important factors when considering their use in MOOC pedagogy, which is a novel and emergent learning context.

Sfard (1998) summarises the discussion on theoretical positions of learning by speaking of the ‘acquisition’ metaphor for learning and the ‘participation’ metaphor. Whilst the acquisition metaphor has been the dominant language of learning, it carries with it a paradox which was first exposed in Plato’s Meno:

> But in what way will you look for it, Socrates, this thing that you don’t know at all what it is? What sort of thing, among the things you don’t know, will you propose to look for? Or even if you should meet right up against it, how will you know that this is the thing you didn’t know? (Plato & Grube, 1976)

By this definition all learning is impossible; however, while the debate has continued in philosophical circles, education has successfully deployed the language of acquisition to explain what concepts are, how they relate to each other hierarchically, and how they can be internalized and transferred from one person to another. The acquisition metaphor for learning is bound in social and political practice because this is an ontological definition of knowledge, defined as a commodity in a way which equates to property, power and social structure. That is to say, when knowledge is seen as currency and defined into an epistemological hierarchy of conceptual units, access to that currency, social mobility and hoarding of that currency defines social practices in a much more far-reaching way: this is long known, and Francis Bacon first used the phrase "knowledge is power" in 1597. The political debates around distribution of 'knowledge as currency' is also reflected in the discussion below on online open education, although these discourses focus on free access to knowledge or knowledge as a 'public commodity'.

> If knowledge is conceived of as a commodity, it is only natural that attitudes toward learning reflect the way the given society thinks about material wealth. (Sfard, 1998, p. 8)

The participation metaphor for learning posits that learning is an activity which is necessarily situated, whose goal is 'learning to become part of a community', perhaps through a kind of 'apprenticeship' model, as posited by Lave and Wenger (1991), or through participation in an activity system (Engestrom, 2014).

> The theory of situated learning (Brown, Collins,& Duguid, 1989; Lave, 1988; Lave & Wenger, 1991), the discursive paradigm (Edwards & Potter, 1992; Foucault, 1972; Harre & Gillet, 1995), and the theory of distributed cognition (Salomon, 1993) are probably the best developed among them. All of these are theories of a new kind, differing from the old doctrines not only in their vision of learning but also, and perhaps most importantly, in their epistemological foundations and in the underlying assumption about the mission of research on learning. (Sfard, 1998, p. 12)

Indeed, Lave argues against the division of knowledge into transferable units, suggesting that the conception of 'transfer' is "seriously misconceived" (Lave, 1988, p. 39). The participation metaphor for learning constitutes learning as a community activity rather than a commodity, so replaces all talk of private possessions with discourse about shared *activities*.

> The promise of the participation metaphor seems, indeed, quite substantial. The vocabulary of participation brings the message of togetherness, solidarity, and collaboration. (Sfard, 1998, p. 8)

E-learning which supports a constructivist approach draws from these situated and participative learning theories and these theoretical standpoints have been foregrounded when using learning design frameworks (such as the 7C framework referred to above) to integrate learning technologies.

However, Sfard argues that whilst this removes the paradox at the centre of the acquisition model of learning by replacing units of knowledge with an activity (of becoming), it may be wrought with its own problems as it could be viewed as niche by the actual learners themselves, who are brought up on the language of acquisition. This has proven generally problematic in the realisation of the participation metaphor for learning in e-learning courses by students who ‘just want things to be simply presented and easily acquired’.

Sfard concludes her review of the 2 metaphors for learning by acknowledging that in practice it is best not to be dogmatic about pedagogy, and that pragmatism is the best course of action in a teaching situation:

> Educational practices have an overpowering propensity for extreme, one-for-all practical recipes. A trendy mixture of constructivist, social-interactionist, and situationist approaches-which has much to do with the participation metaphor-is often translated into a total banishment of "teaching by telling," an imperative to make "cooperative learning" mandatory to all, and a complete delegitimatization of instruction that is not "problem based" or not situated in a real-life context. But this means

> putting too much of a good thing into one pot. Because no two students have the same needs and no two teachers arrive at their best performance in the same way, theoretical exclusivity and didactic single-mindedness can be trusted to make even the best of educational ideas fail. (Sfard, 1998, p. 10)

Knox (2014) argues that designing for learning in the MOOC context is especially critical because of its global reach, and the danger of uncritically accepting an 'acquisition' design due to its easy-scalability is that it also may have the effect of reducing innovative teaching practice:

> Given the international scope of MOOCs, continued research may need to explore the extent to which massive enrolments might reduce the diversity of instruction in particular disciplines. (Knox, 2014, p. 171)

2.2.1 Conceptualising learning in e-learning

E-learning is resistant to clear definition because it encompasses a range of approaches, from those grounded in communication such as 'computer supported collaborative learning', 'networked learning', 'technology enhanced learning', collective 'knowledge construction' through to more behaviourist 'teaching machine' approaches such as 'computer aided assessment' which favour economies and efficiencies of practice. However, from the point of view of Learning Design and its integration into formal educational practice, such as university level programmes of study, it is generally associated with the uses of technology as a mediator for communication and participation, rather than content delivery, and I explore this definition in this section.

I will firstly review the conceptions of learning from the point of view of how e-learning has changed the formal educational structures (e.g. programmes of study and systems of quality assurance), then onto the overlapping concern of how e-learning has affected the domain of ‘open education’, which is not exclusive to e-learning but developments in web 2.0 technology have cemented its place as an important piece of the literature. It is from this backdrop that a critique of learning in MOOCs can begin. MOOCs occupy the space in between formal and non-formal education (Walji et al., 2016), which brings about tensions between how learning is evidenced as outcomes in formal education vs how learning is conceptualised in non-formal settings. In negotiating these tensions, and designing for scale, many MOOCs do not build on the transformative aspects of e-learning, rather focusing on content delivery and computer aided assessment. However, by focusing on design of the technical affordances of MOOC platforms and learning designs at the level of student activity, we can start to alter the balance and retrieve some of the transformative impacts brought about by e-learning and associated technologies.

Haythornthwaite and Andrews describe e-learning as:

> a transformative movement in learning, not just the transfer of learning to the online stage…. not bounded by institutional structures of course, programs or degrees, but instead as embracing the way learning flows across physical, geographical and disciplinary borders… as perpetual, sustained over a lifetime and enacted in multiple, daily occurrences as we search for information to satisfy our learning needs and contribute content that promotes our and others' understanding. (Haythornthwaite & Andrews, 2011, p. 2)

This foregrounds the participatory and cooperative nature of this type of learning experience ("contribute content that promotes our and others' understanding") and so learning designs which use technology in formal settings tend to be premised on situated learning theories, grounded in Vygotsky's idea that higher level knowledge is necessarily socially created (Vygotsky, 1978) or models of participatory learning experiences which value "Legitimate Peripheral Participation in Communities of Practice" (Lave & Wenger, 1991), both of which follow the 'participation metaphor' and focus on the 'activity' of learning (rather than the acquisition of concepts). These are both examples of increasingly dominant theoretical positions adopted by educational researchers over the past 30 years as the behaviourist pedagogies are supplanted by sociocultural theories and constructivist approaches to learning.

Gráinne Conole defines this shift:

> There has been a paradigm shift in the field due to new thinking around learning theories from behaviourism, through to cognitivism and finally constructivism (Mayes and De Freitas 2004). These theories led to the development of particular uses of technology designed to support the underpinning principles of the theories. (Conole, 2013, p. 3)

2.3 e-learning in the formal educational setting

Technology, combined with this 'constructivist turn' in pedagogical theory, has brought about new approaches for teaching and learning at university. Garrison (1997) called this "the post-industrial age of distance education". He contrasts the industrial approach of teaching which has clearly defined divisions of labour: the educator as a content creator (for postal or didactic pedagogical methods) and the student a receiver without a choice. The post-industrial approach is typified by the educator taking a facilitation role with the student at the centre of the learning process, collaborating, constructing meaning and confirming understanding. In terms of distance education, computer conferencing has driven this step change.

However, the step change isn't limited to distance education, and forms of constructivist pedagogical approaches, such as inquiry-based learning, problem-based learning or case-based learning are accelerated into formal educational settings by the introduction of digital technologies, in the form of 'blended' approaches, often termed 'active learning'. These approaches place the student in the centre of the learning activity and often use the affordances of communication technologies such as discussion groups, wikis and mobile tools to facilitate the interactions. When the student/ learner is in the centre of the activity of learning, learning happens *through* participation, which relies heavily on the 'participation metaphor' discussed earlier.

An example of an inquiry based learning activity concerns the product 'Knowledge Forum', developed iteratively by Marlene Scardamalia and Carl Bereiter, which enables students to work together to construct knowledge through collaborative research (Scardamalia, 2002). Learners research an authentic problem and post their findings in the form of links or 'notes' in the database, and then make pedagogically based connections or scaffolds between these nodes, such as 'rise above' (for synthesis), 'our theory', 'did you know', 'our evidence'. Their work is situated in middle school classrooms and is an example of how a self-directed, technology supported pedagogy can teach children research skills, and help shift their thinking from simply recording information to making evidence based interpretive accounts.

This approach is similar to problem-based learning (PBL) in that it is based in constructivist principles of authentic problems and collaborative effort. PBL originates in the earlier days of medical education in the mid 1950's when the concept wasn't necessarily technology supported: groups of 8-10 students were presented with a ‘patient with symptoms’ and had to generate an evidenced diagnosis. Students must discuss the problem, generate hypotheses, conduct self-directed, inquiry based research and then reconvene to evaluate the group’s efforts and collaborate towards an answer. The ‘learning’ in this activity necessitates participation, and the students will construct meaning through bringing together their conversations and research. Therefore, they generate their own learning objectives based on their analysis of the problem. As a constructivist teaching method, it has been exported to other subjects, with problems ranging from “Do asteroids in space pose a problem, and what should we be doing about it?” to “What caused the flooding in the Mid-West in 1993 and what should be done to prevent it happening again?” (Wilson, 1996). These constructivist pedagogical methods often work best when supported by technology like that used in the ‘knowledge forum’ case above.

Other examples of ‘active learning’ might be use of a discussion forum for groupwork, reflective learning logs, use of blogging tools, crowdsourcing information and any idea which places the student at the centre of the learning, constructing his/ her own meaning. Whilst technology is never sufficient for this type of activity (it requires a matching constructivist approach to learning), it often plays a necessary and sometimes crucial mediating role: perhaps to record conversations on a discussion board, aggregate blog posts or pins on a map, to create mind maps, to host links to further research or many other activities. The role of the designer is to work out which tools and resources to make available in advance, considering their in-built affordances and suitability for the task, and allowing for flexibility if required to respond to the needs of the students. This is often done with design-based methodologies, using toolkits and frameworks from the emergent field of learning design.

2.3.1 Learning Design

Emerging in the early 2000s, Learning Design (LD) is a still relatively new research field. It is related to, but distinct from, the more established field of Instructional Design (ID). ID focuses on the specifics of designing learning materials that meet a given set of learning objectives. Koper and Olivier (2004) argue that ID models consider learning with a limited pedagogical perspective, namely that 'in order to learn, a single learner has to work through a sequence of learning objects' which suggest that learning is a process of consuming content. LD, on the other hand, emphasises the pedagogic intent and high-level design principles which underpin a piece of curriculum. LD is firmly positioned within a tradition of sociocultural educational research, and references the work of Vygotsky and others who develop activity based theories for learning (Engestrom, 2014; Vygotsky, 1978). Put rather simplistically, ID focuses on the teacher and what the teacher does, while LD focuses on the students and what they do.

Gráinne Conole defines Learning Design as:

> a methodology for enabling teachers/ designers to make more informed decisions in how they go about designing learning activities and interventions, which is pedagogically informed and makes effective use of appropriate resources and technologies. This includes the design of resources and individual learning activities right up to curriculum-level design. (Conole, 2013, p. 8)

She makes the case that when the teacher is facilitator of learning in this rapidly shifting paradigm brought about by e-learning and technological innovation, they cannot be expert in using all the new technologies or be at the forefront of innovation in terms of tools and resources, therefore having a common language for sharing methodologies is important (the simpler the better!). This systematic methodology of designing for learning requires a shift from designing learning exercises in an individual and intuitive (craft based) way to something which is explicit and follows a process to capture requirements about learning/ cognitive activity and can explore the affordances of tools. As stated above, the theoretical basis for this work is sociocultural, and learning designs are supported via a range of mediating artefacts (tools).

There is little literature discussing how to embed LD into MOOC design. Indeed, Veletsianos and Shepherdson (2016) conduct a literature review from 2013-2015 and conclude that there is a paucity of research examining instructor-related topics on MOOCs. The little that does exist includes Conole (2014a, 2015) who has discussed the classification of entire MOOCs across different dimensions of context and learning and Donald et al. (2017) build on this to describe the design of a specific course which uses this classification system. Glasgow University has also utilised UCL's ABC (Arena, Blended, Connected) framework (C. Young & Perović, 2016) which builds on the types of teaching activities as defined by Diana Laurillard (2012). This work is a useful foundation to the conversation on LD for MOOCs but does not envision how new tools or affordances could enhance the overall design. The Comment Discovery Tool (CDT) as a novel mediating artefact builds on these Learning Design methodologies and adds new opportunities for designing new interactions into the LD toolkit. This locates the focus of my research directly on the available toolset in a MOOC platform and what it affords in terms of possible learning activity design, something which Donald et al. (2017) grapple with in their design of a MOOC using FutureLearn's limited toolset, and which is a limitation of the ABC curriculum design techniques when applied to MOOC learning.

2.3.2 Learning Design Frameworks

Specific Learning Design Frameworks and toolkits are a method of communicating types of constructivist pedagogy between educators. This is neither a theoretical framework, nor a methodology, rather a series of toolkits and frameworks to embed constructivist and active learning principles into a blended course, a fully online course, or an individual learning activity.

Laurillard (2012) uses the term ‘science’ to explore this type of design in her book “Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology” and defends the use of ‘design science’ by aligning teaching practice with computer science, engineering or architecture. These sciences aim to improve the world rather than describe or explain it, as the natural sciences. This is not to say that it does not contribute to theory, rather it does so with different methodologies:

> A design science uses and contributes to theoretical science, but it builds design principles rather than theories, and the heuristics of practice rather than explanations, although like both the sciences and the arts, it uses what has gone before as a platform or inspiration for what it creates. Teaching is more like a design science because it uses what is known about teaching to attain the goal of student learning, and uses the implementation of its designs to keep improving them. (Laurillard, 2012, p. 1)

Conole et al. (2004) suggest that although these constructivist approaches are oft-quoted in e-learning literature, much of what happens could actually be described using the language of behaviourism and didactic teaching. Their toolkit for learning design allows theory to be adopted flexibly into practice across different teaching situations. The toolkit creates 3 theoretically grounded dimensions operating on a continuum (individual > social; non-reflection > reflection; information > experience). These dimensions are useful when selecting tools (or mediating artefacts) for learning as they allow for the learning designer to break down all student activity across these dimensions, and this will inform either the development or choice of mediating artefact.

Laurillard (2012) takes an historic angle, and suggests that a language for describing and formalising pedagogical patterns is necessary to heuristically share teaching practice to a wide audience. As such she has suggested 6 activities which make up the vast majority (if not all) of teaching situations. These are learning through: *Acquisition, Inquiry, Discussion, Practice, Collaboration, Production*. This framework is different because it gives learning designers a theoretical grounding from which to start designing pedagogical patterns at the level of the course and from there it is possible to choose mediating artefacts.

In terms of toolsets, Laurillard suggests that educationalists have rarely designed their own tools but that they have absorbed them from other contexts; for example, slide presentation tools were primarily a business invention. The rapid growth of digital toolsets over the past few decades combined with changes to types of educational contexts has been a shock for which educationalists need to adapt. Laurillard introduces this challenge:

> Precisely because of their potential to change education unbidden, it is imperative that teachers and lecturers place themselves in a position where they are able to master the use of digital technologies, to harness their power, and put them to the proper service of education. This is a theme that will repeat throughout the following chapters – education must now begin to drive its use of technology. (Laurillard, 2012, p. 2)

Laurillard’s framework has been operationalised by colleagues at UCL (C. Young & Perović, 2016) who have linked activities with tools for each learning type, for example ‘web search’ for inquiry or ‘video lectures’ for acquisition. They have started work mapping FutureLearn-specific tools, for example ‘video’ with acquisition and ‘quiz’ with practice. This mapping informs the learning design process on FutureLearn, but as noted above by Conole (2015) the collaborative and communicative aspects are restricted by the toolset and scale of participation.

Goodyear’s ‘ACAD’ (Activity Centred Analysis and Design) framework (Goodyear & Carvalho, 2016) has been developed as an approach to analyzing complex learning networks, so is related to the MOOC context. It acknowledges that learning networks are co-configured by the participants and designers, and this is explicitly valued and encouraged. In this respect it is a framework for designing for scale.

> In describing the changes in scale and complexity in networked learning witnessed over the last 20-30 years, Goodyear et al. (2015) talk about a shift from the 'virtual classroom' to the 'learning city'. So we have to ask: are the tools, ideas and methods that were appropriate for furnishing the virtual classroom adequate to the challenges of city planning or urban design? (Goodyear & Carvalho, 2016, p. 219)

As with the other LD frameworks, activity is placed at the centre of ACAD, and since human activity is best understood in the Vygotskian manner as both physically and socially situated, the framework needs to be able to identify and represent such things as material artefacts, digital tools, social structures, divisions of labour and other organisational arrangements that shape and are shaped by the human activity. However, activity cannot be 'designed'; we can only 'propose tasks' which can influence subsequent activity, meaning that there is an indirect relationship between tasks and learning outcomes, mediated by tools/ physically situated objects, and divisions of labour/ socially situated factors. In terms of designing tools for learning, as with this project, it is useful to highlight that this framework shares the activity-centric qualities of Laurillard and Conole, but explicitly includes the physically situated artefacts as crucial mediators of the activity which is missing from the previous adaptations of the 7C and ABC frameworks to MOOC learning. In other words, activity is physically, epistemically and socially situated and mediates outcomes. This holistic focus supports the design-based research methodology that is taken by this project.

There are several other Learning Design frameworks such as ICAP framework (Chi & Wylie, 2014) and constructive alignment (Biggs, 2003) which align activity with learning outcomes and assessment, SOLE toolkit (Conole & Fill, 2005) which all feature in Learning Design for 'traditional' constructivist pedagogies in a formal university setting. My work with FutureLearn MOOCs uses the ABC framework, based in Laurillard's 'Teaching as a Design Science' for mapping affordances to activity, and Goodyear's ACAD framework, because of its focus on holistic analysis of physical and social factors in learning networks, which inform future designs.

2.3.3 "Affordances" of technology

> The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. (Gibson, 1979a, p. 115)

Given the above discussion regarding learning designs and their relationship with the potential of technology, it is useful to discuss what is meant by the term affordance, which is commonly used to describe the properties which are designed into an environment. 'Affordance' is a contested term in educational research, but it is an important concept for this design project as it describes the relationship between an environment and a user. The original use of the term was to describe an 'action possibility' of the environment which is independent of the user (Gibson, 1979b) although it is also argued that affordances can include the *perceived* properties of the environment too (Gaver, 1991). Other perspectives regarding HCI and systems design make the case that there is a distinction between information which signals the affordance and the affordance itself, so the signals can incorporate affordances (perceived and actual), conceptual models, and constraints (physical, logical and cultural) (McGrenere & Ho, 2000; Norman, 1988, 1998). Norman and then Neilsen take this idea of ease of perception of an affordance as a crucial part of their work into 'usability'. This is particularly pertinent when considering the affordance of a learning tool, because the 'signal' is at the same time the instructions for engagement in a learning activity (which has social and cultural implications), and the 'design' or 'usability' of the tool itself. According to Norman's model, non-technical (cultural and physical) conventions may make it difficult for the educator to change practice in the necessary manner that the technology can be incorporated into the teaching, or that the conceptual model about the learning process (learner or educator) may prevent them perceiving an affordance at all.

Jones (2015) identifies affordance as one of the key constructs which are fundamental to networked learning which takes the Norman/ Neilsen conception of 'perceived affordance' and applies this to human activity:

> [affordances…] are best understood as relational – they are what the tool offers a particular user, or class of users. Pushing this further, analysing the entanglement of tools (etc) in human activity, with a view to understanding how to (re)design or select new tools that are better for

> the job, requires an understanding of how one tool can be said to be better than another, for a class of users and for a class of task. (Goodyear & Carvalho, 2016, p. 222)

Oliver (2005, 2011, 2013) critiques the whole concept of affordance, suggesting that it is too vague to have any value, preferring to theorise 'technology' using a 'social' traditions such as Science and Technology studies, or Actor-Network theory. He suggests that Gibson's account is too materialistic and does not account for perception in accounting for an affordance, whilst an account which builds an ontological argument on 'possible perceptions' is not useful either because any list of potential interactions will always be *impossibly* incomplete. A potential position on 'perceived' affordance is to take a normative perspective and describe 'typical' or 'expected' interactions, which is the position taken by Norman. For Oliver, this is not good enough in terms of a precise definition, and he builds an argument where technology affordance is cast stylistically as 'genres' of cultural activity, or 'cultural text', which can be discussed drawing on the work on Foucault, Science and Technology studies and Actor-Network theory, applied to the tool in question and related to wider societal or social norms at the same time.

I take on board Oliver's critique of the 'traditional' definition of affordance but argue that that unit of analysis for Oliver is too wide. In this learning design project, I am not analysing 'the technology of MOOCs' as a broad cultural movement in opposition to other possible educational provision (as cultural texts); this is the kind of work of

Knox (2015), who is relating the concept of the Massive Open Online Course to the theoretical tradition of posthumanism. I am rather examining users' interactions within a single system and attempting to create a platform or foundation for designing learning which emphasises social interaction rather than individual consumption (participation rather than acquisition), and Norman's model does allow for the use of affordance as a lens to examine other cultural, physical and conceptual factors. I am not making claims about transient elements of online environment, such as the representation of 'the rug' in Oliver (2013) which is the visual cue for 'conversation' in the cited Second Life case study and I am not claiming that the word cloud visualisation 'causes' learning (which would be a very blunt Gibsonian interpretation!); moreover that it creates the interactional opportunities which are important yet missing parts of the basic platform, and can be exploited by a good learning design which also takes into consideration the social, conceptual and contextual factors referred to in Goodyear's ACAD framework. In terms of this goal, it is more appropriate from a practical perspective to align with a definition of affordance such as McGrenere and Ho (2000), Norman and Neilsen (1998), Goodyear and Carvalho (2016) or Jones (2015) (i.e. a "soft" cause in which technology causes change but is not the sole determinant) as this works alongside the analysis from the ACAD framework for understanding emergent learning activity. Oliver (2005) himself states in a section named "I have no affordances, yet I must design":

> [researchers and designers…] persist in using the term affordance because of the desire to speak about technology in a general enough way that it can be theorised in its own right (Oliver, 2005, p. 410)

2.3.4 Summary

I have demonstrated above that in the formal university setting, a popular ‘flavour’ of e-learning has emerged alongside a turn towards constructivist pedagogies, some of which are mediated through technology which follows a ‘participation metaphor’ account of learning, as defined by Sfard (1998). In order to communicate methods of combining technology with the constructivist pedagogical approaches, a new field of study called ‘Learning Design’ has emerged, and several toolkits and frameworks have been developed to aid teachers and designers. These frameworks map learning activities onto affordances of technology. I identify this focus on affordance as a major gap in the literature specifically relating to the design of MOOCs and I have identified that there is a paucity of research regarding the general use of learning design frameworks in MOOCs, albeit to a lesser extent than the gaps in analysing MOOCs from a systems/ affordances perspective. Therefore, I propose new tools to be developed and integrated into the MOOC experience, which can be analysed in terms of the resultant emergent activity.

2.4 Learning within an Open Education context

This section will examine learning in an open education context, which in many ways stands in contrast to the work detailed above. Open Education does not refer to structured courses and the movement is grounded in a 'philosophy' of openness, rather than a theory of learning, so the formal Learning Design toolkits seen in participation-based e-learning above are not central to the discourse. Learning is necessarily informal and emphases self-directed individuals designing custom pathways and collecting resources in a Personal Learning Environment (PLE) or portfolio (Attwell, 2007). In terms of production, Open Education can be cooperative, in the case of collaborating on a defined project, such as Wikipedia or other peer production systems, or individual, in the case of the licensing of any learning resource. This is again related to the philosophy of 'free' and 'open' rather than any constructs around learning theory or the process of successful learning.

Open Education is concerned with the *production and protection of resources*. Whilst 'openness' suggests ideas of self-directed learning, which are constructivist in nature due to the implicit student-centred approach, much of the literature concerns *access* to 'knowledge resources' (Hylén, 2006), which is based in the language of the 'acquisition' metaphor for learning: knowledge is viewed through the lens of commodity which the open education proponents wish to be in the public domain. This is not entirely without cause, as patent and copyright law protecting intellectual property acts as a control of knowledge, so it is important to develop parallel systems of protection for content creators to state how a resource may be used, re-used or remixed (Bissell, 2009).

The lack of defined curriculum which is common to informal learning using open educational resources is both an opportunity and a challenge. It is challenging because by itself it does not offer much insight into the cognitive *processes* of learning towards outcomes, but it is an opportunity in that learning can be self-directed at the level of personal interest, rather than structured into a course setting with pre-defined and designed-for knowledge outcomes. In theory, learners can set their own learning goals and decide when they are fulfilled, but in practice the overload of resources available on the web can also restrict critical learning experiences and create untrustworthy echo chambers, such as the discourse on 'flat earth', conspiracy theories and other 'fake news'.

An interesting development in terms of technologies to support open education within a constructivist frame is the Personal Learning Environment, which is a platform to aggregate content from the web and social media, so can be used as a reflective, sharing and production system for inquiry-based learning. The concept of the cMOOC is an extension of this type of open education and attempts to link students together into a supportive communication network so they can help each other construct knowledge through discussion and discovery. My critique of this approach is that it isn't transparent and many give up because they find this type of learning disorientating (Kop, 2011; Mackness et al., 2010). This seems to align with the quote above from Sfard about a propensity to develop pedagogies at the extremes:

> […] no two students have the same needs and no two teachers arrive at their best performance in the same way, theoretical exclusivity and didactic single-mindedness can be trusted to make even the best of educational ideas fail. (Sfard, 1998, p. 11)

Another strand of literature on Open Education is in the sense of open peer production systems or collective intelligence such as Wikipedia (although there are many others). This idea is relevant to this project because the designed intervention in this project surfaces the 'collective intelligence' which emerges in a scaled learning context through a visualisation which is designed to encourage social interaction and discovery. There are differences in the methods used for peer production systems in terms of collectively building a coherent knowledgebase, as in Wikipedia, and utilising the collective intelligence in order to help self-directed learners in a MOOC, and these issues are explored in a later section.

2.4.1 Open Educational Licences

As discussed above, open education has primarily been defined as a philosophy about access to knowledge. Several public copyright licences have been defined which enable the free distribution of content. An example of this is the Creative

Commons[2] which has several derivatives depending on how the content or resource can be shared or remixed. For example, content can be shared with or without attribution, for commercial or non-commercial, and with or without the rights to create derivative works (Bissell, 2009). There are similar developments in the area of open source software publishing, and the most commonly attributed licences for open source software (which can be conceptualised as a specialised type of OER with the specific intention of remixing and improving) is the GNU General Public Licence (GPL), and this has been used for well-known large-scale collaborative projects such as WordPress, Git, MySQL and many more. It is also the licence type for the 'Comment Discovery Tool' application developed for this project.

Unsurprisingly, the technicalities of these licences are extremely complex, as they involve legal status, rights in contract law, and terms of use when integrated in non-GPL licenced products. It is not for this thesis to explore all these details, rather to demonstrate that much of the discourse on Open Education concerns the status of knowledge objects, rather than being a theory on the process of learning. This means that the discourse on 'Open Education' relies on the language of the acquisition metaphor for learning. In contrast, Lave (1988, p. 39), drawing on a participation metaphor, suggests that the conception of 'transfer' (of concepts) is "seriously misconceived" and that learning occurs through the activity of becoming part of a

[2] https://creativecommons.org/licenses/ details the different licence types available in the creative commons framework

community. That is not to say that OERs cannot have value within learning designs that are rich in participation, but that they are not sufficient to explain the cognitive processes involved in learning.

2.4.2 Open Education and Peer Production Systems

Open Education can also refer to the collaborative development of knowledge; this is best seen in the development of Open Source Software (OSS), where individuals have been able to download source code and contribute new code in order to improve or add bespoke features. Large scale projects, such as the Linux operating system, are examples of software which has been developed in this way. It can also refer to the construction of learning resources, such as Wikipedia, where any user can edit any page to improve the resource. There is great potential in this type of crowdsourcing, especially in learning designs for MOOCs, but there is also a note of caution in terms of the tasks and the division of labour required to scaffold these activities. Recall Goodyear's claims above about designing for learning and the indirect link between task and outcome, with physical, epistemic and social factors acting as mediating factors. All 3 factors need to be considered holistically!

Stackoverflow.com is an example of an online discussion forum for answering technical questions and has become a massive repository of information on a range of subjects; it has developed a reputation score based on 'upvoting' by the community. However, stackoverflow.com values a very structured kind of knowledge such that questions and answers are concisely articulated, and posts are aggressively edited to avoid duplication. Wikipedia also employs strict rules for participation to build a coherent knowledge base. The rule-based systems on both these platforms suppress free flowing conversation in favour of structured, well-defined contributions, and contributors accept or rapidly learn what these baseline standards are. It is certainly an option for MOOC designers to develop features which mimic these systems, but there are dangers in that learners may cease to contribute when they see their comments being edited, deleted, and downvoted. This approach would also seem to go against the interpretive and knowledge building nature of sociocultural learning which values knowledge construction through synthesis of ideas. Indeed, an upvoting, rule-based method ironically reinforces a canonical epistemology and positivist pedagogy, albeit not from a single teacher to a lecture theatre, but from groups of experts. There is perhaps value in this approach, in certain types of courses, to organise contributions, but it is not the approach taken in this project.

Other peer production systems, such as iSpot Nature, 'peer teacher', and some citizen science initiatives also have a well-defined problem space where new contributions add to an overarching (database) structure. Haythornthwaite (2009) has developed theory to explain models of participation, utilising concepts of 'light and heavyweight peer production', but there are key differences in the forms of participation she cites and the context of actually designing learning at scale. Lightweight peer production involves making a micro-contribution to a structured dataset, like the structured systems cited above. In these cases, contributions are rewarded with points, which acts as a motivator for the community, and gamification type systems can be developed. The example of heavyweight peer production cited is the production of peer reviewed academic work, where the community is much smaller, the contributions require more effort, and the rewards are reputation in journals and by peers. User generated data in MOOCs is less structured and requires interpretation, as heavyweight peer production, but the contributions are small in term of meaning value within a single comment and are available in a structured format in the form of a relational dataset, as with lightweight peer production. It is large scale, so systems for coping with overload need to be designed for coordination. One option is that individual contributions could be scored and prioritised according to their score, like in the reward models of lightweight production, but this would create a new problem in the tension between the development of scoring schemas or algorithms and a sociocultural pedagogy. Neither technology nor the social context can be neutral in a sociomaterial sense; for example, if league tables were developed based upon number of posts, number of 'liked' posts, or number of posts which stem a conversation, each would presuppose a fixed idea of value, which in itself would change the nature of the participation, restrict diversity and suppress niche themes or serendipitous encounters.

2.4.4 Summary of Online Open Education

Open Education is a philosophy about access to knowledge rather than a theory of learning. It is used to describe the process of informal learning in the internet age, but it is steeped in the language of acquisition because it concerns OERs, and a large body of work is directed at defining its legal status such that everything from small learning objects or individual assets to whole software applications can be protected from patent and copyright law.

Another formulation of Open Education is in the construct of peer production systems, in which large databases of knowledge such as Wikipedia, iSpot nature or computer software are collaboratively produced. This is an important idea for this research project, as ideas of crowdsourcing are key to exploiting the massive participation that comes with MOOCs towards collaborative activity. The challenge is to design mediating artefacts and learning designs which can make use of large scale and diverse participation without also creating divisions of labour or software that imposes preset structures on the data, and therefore restrictive pedagogical methods, as is seen with many 'lightweight' peer production systems.

2.5 Learning in MOOCs

The above discourses on e-learning and open education form the backdrop for understanding how learning might be constructed in a MOOC environment. In formal education, e-learning has developed discourses around Learning Designs grounded in a participation metaphor for learning, and affordances of Web 2.0 tools have supported learner centric pedagogical approaches such as inquiry or problem based learning. In the informal learning environment, this is developed further into the practice of cMOOCs. However, other Open Education discourses have centred around a philosophy of knowledge as public domain resources or objects, through development of open licences. This strongly favours the language of acquisition, although learning as a process or theory is not explicitly part of these discourses. Peer production systems also fit within the realm of open education, but they often speak about the development of structured datasets rather than community practice. This study looks at peer production from the point of view of building multiple, overlapping communities within a MOOC and providing a means for learners to navigate, develop and participate in those emergent communities.

'MOOC' stands for Massive Open Online Course, so this means a merger between the divergent discourses on *Open* Education as well as e-learning (*"online course"*), which as explained above is not necessarily a comfortable combination in terms of understanding the process of learning. MOOCs also occupy the space in between formal and non-formal education (Walji et al., 2016), which adds complexity to any analysis on learning because of the tensions explained above between acquisition (of knowledge) in terms of Open Education and participation in activity or community in terms of sociocultural learning designs. Within these constructs, only the cMOOC conceptualisation of participation in open networks fits neatly onto both sides but it is difficult to conceive of this as a 'course'; however the term 'MOOC' is now used loosely to describe an evolving ecosystem of open online learning environments, encompassing a spectrum of course designs ranging from networks of distributed online resources (cMOOCs) to structured learning pathways centralized on proprietary or open source platforms (xMOOCs) (Rodriguez, 2012).

xMOOC type courses deploy an acquisition based approach to understanding learning and it is in this arena that DeBoer et al. (2014) suggest that traditional educational 'course' variables (curriculum, engagement, completion) need to be re-conceptualised in order to understand the MOOC context in terms of diverse and personal goals. cMOOC advocates suggest that students learn by making connections and cyclically navigating networks but this idea overlooks research findings that many MOOC students are often overwhelmed by the peer-to-peer approach and disorientated by fragmented nature of the course materials (Andersen & Ponti, 2014; Knox, 2014; Mackness et al., 2010).

> It becomes clear that if people are learning on these vast, disparate information networks, they need the ability to understand the intricacies of the networks in order to negotiate their structures. (Kop, 2011)

A further problematic aspect for both analyses is how they cope with the other term: '*Massive'*. xMOOCs have dealt with this by simplifying their approach to pedagogical behaviourism, or cognitivism, where cMOOCs have complicated their product to something which only the most digitally literate people are able to use (which ironically deprecates their status as 'open' because it is not presented in an accessible format for most). The starting point for this design project is the idea that new affordances/ mediating artefacts can enrich learning designs such that participation based activities can be scaffolded into massive courses in a way that is not overwhelming or disorientating for learners. This is practically achieved by a mediating artefact which reveals learning analytics to students in the form of an interactive visualisation, affording the ability to navigate the massive participation and the formation of communities through use of keywords.

The following sections explore how learning is conceptualised in the 2 main forms of MOOC context and explains how massiveness creates deficiencies in each approach and how new mediating artefacts and pedagogical approaches can begin to resolve these problems.

2.5.1 Learning in xMOOCs: 'Learning' as analytics

DeBoer et al. (2014) suggest that educational 'course' variables need to be re-conceptualised for MOOCs, and Fincham et al. (2019) attempt this by grounding their analytics in Christenson and Reschly's theory of student engagement (Christenson et al., 2012). However, Veletsianos et al. (2015) argue that it is important to look at more than just log files and online interactions in order to fully understand participation in MOOCs, and I have found no account of learning analytics which attempts to reveal analytics to learners whilst they are learning, which is the starting point for this research project and a key recommendation of the work on Social Learning Analytics (Buckingham Shum & Ferguson, 2012).

All work on learning analytics starts from a point of data collection, so examines clickstream or trace data. This obviously excludes data points which are not visible, such as reading a comment, or working outside the platform (even offline!) in order to 'learn', practices which are demonstrated by the qualitative research of Veletsianos et al. (2015). It is also important that what *is* included is predicated on an acquisition metaphor for learning and equates completion or repetition of knowledge (in terms of completing quiz type assessments) with 'learning'. These data points can only ever represent 'remember', 'describe' and sometimes 'apply' levels in Bloom's famous taxonomy (Bloom, 1956), and this stands in stark contrast to the sociocultural learning designs previously employed in formal courses where learning was designed as an activity of participation, and assessed through reflective blogs or groupwork portfolios, activities which demonstrate learning to the degrees of 'analyse', 'evaluate' and 'create' in the taxonomy.

Siemens (2012) famously said "cMOOCs focus on knowledge creation and generation whereas xMOOCs focus on knowledge duplication" which outlines the behaviourist nature of xMOOC design.

Reich (2015) encourages MOOC researchers to raise the bar in terms of sharing datasets and connecting engagement analytics to theoretical positions on learning:

> Few MOOC studies make robust claims about student learning, and fewer claim that particular instructional moves caused improved learning. We have terabytes of data about what students clicked and very little understanding of what changed in their heads. (Reich, 2015, p. 34)

Joksimović et al. (2018) and Fincham et al. (2019) respond to this challenge by setting out how the data points can work together within a theoretical frame. The work of Joksimović et al. (2018) begins with a systematic literature review of MOOCs from 2012-2015 to determine what the most common approaches are to operationally defining and measuring learning outcomes. This study reveals that learning in MOOCs is typically studied through the analysis of the trace data combined with discussion or survey data, generally derived from a single course. Their suggested data points attempt to operationalise Christenson and Reschly's engagement construct and work on dropout prevention, which they claim can allow learning analytics to move out of the single platform towards generalisable constructs that function outside of the literature on MOOCs.

> Such a conceptualization would also allow for moving beyond observing student "click data" and exploring how quantity and quality of

> interactions with the course content, peers, and teaching staff could predict course outcome and persistence. (Joksimović et al., 2018, p. 70)

This development towards sophistication in the learning analytics literature begs the question of what actionable insight is being produced, and for whom? Analytics that allow us to 'predict dropout' could however be conceptualised as 'teaching analytics' or 'academic analytics' rather than being 'learning analytics' in that they provide MOOC teachers with insights as to who may or may not be able to complete a course and may also be useful for institutions to make strategic decisions. Buckingham Shum and Ferguson (2012) might call these 'academic analytics' or 'business intelligence'. That is to say, in and of themselves, they do not actually help learners.

The operationalisation of engagement factors to use as proxies for learning also has several problems. For example, behavioural analytics are steeped in an acquisition metaphor for learning (i.e. engagement with content):

> For MOOCs, this [behavioural] form of engagement can still be defined through participation in discussion forums, viewing lectures, following course activities, or number of times student accessed course wiki pages (e.g., Li et al., 2015; Santos et al., 2014; Sinha & Cassell, 2015). (Joksimović et al., 2018, p. 67)

It is also not clear that there are known methodologies for operationalising more complex constructs such as the cognitive aspects of learning, and whilst ‘quality of discourse’ (below) has been used in previous e-learning as a proxy for cognitive engagement (Garrison, 1991; Gunawardena et al., 1997; Henri, 1992) there is little to suggest how ‘quality’ can be scaled for the MOOC context:

> The rationale behind [cognitive] engagement is grounded in the premise that learning and understanding in computer-mediated learning are primarily expressed through the artefacts students generate in the learning process (Goodyear, 2002; Jones, 2008). Thus, studying learning in MOOCs should account for the quality of discourse, as a proxy for students’ cognitive engagement. (Joksimović et al., 2018, p. 68)

Furthermore, the contribution from Veletsianos et al. (2015) which digs deeper into students’ activities when they learn in MOOCs reveals practices related to learning which cannot be accounted for even with a sophisticated and multi-faceted analysis of available engagement data, such as working off platform, taking notes and also combining these with notes of related MOOCs. This highlights the fundamental problem with defining learning in MOOCs through engagement analytics: the open online context sits somewhere in between formal and informal learning, so analyses which attempt to predict dropout or operationalise course outcomes do not account for the informal nature of open education, and this is summed up by Stracke (2017) who claims “high drop-out rates are preferable in MOOCs as they show the diversity of personal objectives by the MOOC learners". Stracke continues to argue that MOOCs should be designed in a more personalised way, to offer multiple pathways through content, which attempts to straddle the same dilemma of formal education with fixed course outcomes (definable through analytics) and informal education with only personal outcomes (undefinable, subjective, and diverse). Whilst Stracke's approach diversifies the MOOC into several pathways, it does not speak to the axiomatic principles of informal education that self-directed learners have an infinite diversity of personal needs and can freely choose their engagement with multiple OERs towards those needs.

The above analysis poses 2 problems regarding learning as analytics in MOOCs. Firstly, by operationalising constructs related to learning in terms of analytics it necessarily prioritises a certain way of engaging with courses, which is distinct from how engagement with OER is conceived in Open Education and presents its own ethical challenges regarding how it may be implemented. Secondly, whilst types of analytics can be sophisticated in terms of how they relate to certain learning theories, they are only currently scalable when applied to metrics which are steeped in the acquisition metaphor for learning (i.e., how many times have you watched this video? what score did you get in the test? how often are you logging in?) which again speak only to the lower orders of critical thinking in Bloom's taxonomy. When analytics are required to answer questions on more complex constructs like cognitive engagement, they are much harder to scale as they refer to 'quality of discourse' where 'quality' is a more complex construct to operationalise. Simply put, constructs around learning like this refer to the higher levels of Bloom's taxonomy such as analyse, synthesise and create and these are almost always subject to human interpretation.

2.5.2 Social Learning Analytics

Buckingham Shum and Ferguson (2012) present a brief history of 'learning analytics', from their origins in business intelligence, to their employment as 'academic analytics' in terms of predicting dropout, measuring engagement and making strategic institutional decisions which is broadly discussed above in relation to common approaches to understanding learning through analytics in xMOOCs.

They highlight the challenge in implementing analytics which have pedagogical and ethical integrity, in the context where power and control of data is extremely important. They draw on large scale surveys such as the 1997 'World Values Survey' and propose that social learning and participatory pedagogies are a growing trend in global learning preferences[3]. They also draw on the work of Hagel et al. (2010) to propose that innovation requires social and interdisciplinary knowledge. Following this summary, they propose a series of 5 'social learning analytics' (social [*network,*

[3] https://en.wikipedia.org/wiki/World_Values_Survey supports 'human development theory' by suggesting that the largest increase in individual agency occurs in the *transition from industrial to knowledge societies*.

discourse, content, dispositions, context] analysis) and describe tools and systems which have been used to support these approaches in online social learning. It is not important at this stage to describe each analytic in detail or the examples of tools which have been used to support learning using this approach, but it is important to highlight that specific toolsets, affordances and learning designs are required to support these social learning analytics.

Furthermore, regarding the nature of power in terms of utilising analytics, they conclude that the data needs to be shared with learners themselves to set a new context for thinking about learning:

> While [social learning analytics] may be quite legitimately deployed as institutional tools to yield insight for educators and administrators, equally, they should be seen as tools to be placed in the hands of the very subjects being analysed – the learners. A key lesson from the social web paradigm, and a long-held aspiration of end-user customizability researchers, is that when empowered with appropriately flexible tools, an ecosystem grows in which new roles are created for different kinds of user to customise their tools (MacLean et al. 1990). Users are best placed to refine and combine them in new, more effective configurations which stakeholders designing or observing from 'outside' could not accomplish. It would indeed be ironic if the ways in which Social Learning Analytics tools were deployed did not honour and promote the open, critical

> dynamics that underpin much of the participatory, social web philosophy — dynamics which such tools seek to make visible. (Buckingham Shum & Ferguson, 2012, p. 21)

The quote above highlights some key aspects of this research project: firstly, tools need to be developed along the lines of the social learning analytics, secondly that these analytics need to be presented to learners themselves, and thirdly that when the users have control of the analytics they will be able to make use of them in diverse and possibly more effective ways than can be designed for by those 'outside' the actual learning experience.

2.5.3 Summary of learning in xMOOCs and Learning Analytics

xMOOCs necessitate a more independent study model that is in stark contrast to the more accepted socio-constructivist approaches to learning (Bayne & Ross, 2014) because the toolsets included in them scaffold an acquisition or transfer approach to learning. These toolsets also support the operationalization of certain constructs about learning such that they can be analysed as 'learning analytics'.

Learning analytics in themselves are not fundamentally flawed, but the limited manner in which they are deployed is ethically and practically problematic in terms of who controls the data, and their tendency to be used post-course for actionable insights which are only of value to teachers. Analytics themselves can be tied to existing learning theories (Fincham et al., 2019; Joksimović et al., 2018), but are currently limited in terms of what data can be collected and how they are related to other data points. New toolsets are required to expand the range of data collected into the realm of 'social learning analytics' and these data should be available to learners in order to support participative learning designs and innovation in terms of how analytics are employed (Buckingham Shum & Ferguson, 2012).

Conole (2014a, 2015) attempts to break apart the xMOOC/ cMOOCs division by relating them more closely with the 7C's learning design framework but acknowledges that the toolset and the massive nature of MOOCs means that several of the interactive (or social) 'C's', such as communication and collaboration are not well supported for xMOOC platforms, which introduces the delicate relationship between context, learning tools, and pedagogic approach and supports the notion that new toolsets should be produced to enable these types of activities in MOOC learning designs.

This project aims to introduce a new learning tool into an xMOOC dominant platform in order to allow more flexibility in terms of learning design and learner behaviour, which is the contribution of this project to the literature.

2.6 Learning in cMOOCs

cMOOCs are a web 2.0 or ‘networked’ pedagogical model and were in fact the first conceptualisation of MOOCs. In 2008, George Siemens and Stephen Downes experimented with public engagement in their campus class, “Connectivism and Connective Knowledge”, which became known as CCK08. This course was the first to incorporate open learning and personal learning environments with distributed content and was a practical experiment with Siemen’s earlier theoretical work defining Connectivism as a ‘learning theory for the digital age’ (Siemens, 2005). The term “MOOC” was indeed coined by Dave Cormier from these early experiments in Connectivism.

Connectivism is a theory of learning which emphasises “social networked learning” through a learner-centric and distributed platform. Individual contributions in a social network create ‘nodes’, and learning journeys can be created through learners wandering between these nodes; typical connectivist MOOCs (cMOOCs) value contributions from peers as equivalent to contributions from tutors (Siemens, 2005).

Kop and Hill use the following conceptualisation of learning in a connectivist system:

> The learning process is cyclical, in that learners will connect to a network to share and find new information, will modify their beliefs on the basis of

> new learning, and will then connect to a network to share these realizations and find new information once more. (Kop & Hill, 2008)

cMOOCs have no formal assessment, just informal feedback from knowledgeable others; the tutor is cast as a co-learner who shapes learning goals through collaboration with the participants. All users create content in the form of blogposts, tweets or discussion posts, and learners are encouraged to discover their own learning goals through navigating these distributed websites and engaging with other interested learners.

The web itself is constructed as a ‘smooth space’, building on Deleuze and Guattari’s concepts of smooth and striated (Deleuze & Guattari, 1988). That is to say, in an amorphous smooth space, one can rise up and move to any other space, whereas in a structured striated spaces options are more limited and closed (Bayne, 2004). The cMOOC is a nonlinear structure, so is a great learning opportunity for nomadic experiences.

However, this may be to overstate the ‘freedom’ of the learning space, which is necessarily restricted by the mechanics and ‘system’ of the learning platform (as mediating artefact): cMOOC platforms work by aggregating learner-created posts onto the main course webpage through RSS and the course leader writing a daily newsletter. Twitter is also a common tool used for community formation through use of hashtags, and this piggybacking on 3rd party technologies which were not primarily designed as learning tools relates back to Laurillard's claim that education does not design its own tools, but absorbs them from other contexts (Laurillard, 2012). The actual practice of participating in a cMOOC involves either being very active on Twitter (which is a realtime technology so requires an 'always on' commitment) or sifting through RSS feeds on the custom course platform. This makes them just as susceptible to problems of overload as the xMOOC format, as it is not possible to easily discover relevant things through the participation overload. While cMOOC advocates would suggest that students make sense of the information by connecting with other students and cyclically navigating networks, this idea overlooks research findings that many MOOC students are often overwhelmed by the peer-to-peer approach and the fragmented nature of the course materials (Andersen & Ponti, 2014; Knox, 2014; Mackness et al., 2010).

> It becomes clear that if people are learning on these vast, disparate information networks, they need the ability to understand the intricacies of the networks in order to negotiate their structures. (Kop, 2011)

Knox describes designing the Edinburgh University EDCMOOC (Educational and Digital Cultures), hosted on the Coursera platform, which attempted to use elements of cMOOC practice like writing blog posts and sharing OERs. He cites a learner who writes to the forum:

> "I am searching for some simplicity. Week One: Topic One: One Video, One Reading, One quiz, One discussion thread on Week One Topics."
> Here, the distributed arrangements of the course are considered to create unnecessary complexity, and reveal the strong desire for a centralised, logical, and linear pathway through the course material. (Knox, 2014, p. 171)

These critiques are not to say that the cMOOC style is impossible, rather it emphasises a key complexity with learning at scale: the diversity of learner preferences in a massive heterogenous cohort. The distributed nature of the cMOOC makes it resistant to the kind of learning analytics common to xMOOCs, which as described above, produce clean if slightly limiting datasets. The next section reviews an attempt to use Social Network Analysis (SNA) on a cMOOC and highlights key findings about learning and communities when taking this methodological approach.

2.6.1 Social Network Analysis on a cMOOC

cMOOCs represent a networked approach to self-directed learning, so Social Network Analysis (SNA) can be used to study, track and compare the dynamics of communities and the influence of individual contribution. Within a cMOOC context, as with learning analytics for xMOOCs, it is limited because it is not possible to collect ALL the data which may be pertinent to understanding learning processes. This is especially salient in the cMOOC context because the platform itself is distributed across the web, and participants may contribute to some platforms but not others. In this sense it is a macro level approach to understanding what relationships are present and absent within a social network, however, it isn't describing the process of learning itself or how individuals are meeting their personal goals.

Bozkurt et al. (2016) conduct a Social Network Analysis (SNA) on the Twitter interactions on the cMOOC: Rhizomatic Learning 2015 (#rhizo15), and makes several conclusions about learning in this context. The first is that communities in a connectivist network have chaotic relationships with other communities such that learners appear in different and overlapping communities; another finding is that hashtags have an important role as a social glue to keep learners together in a community and do not necessarily refer solely to the communities themselves, but also to ideas, discussions and other concepts. It is also important to note that interaction increases as numbers of learners decreases, so this supports the claim by Kop (2011) above that "learners need the ability to understand the intricacies of networks in order to negotiate their structures" and the claim that cMOOCs are just as susceptible to problems of overload. However, for those who do make this 'cut' and can work through these issues, there is a wealth of 'cognitive' presence in the subsequent conversations and the learning is conceived as participation or activity based, so fits in with the modern constructivist turn in pedagogical thought. The bar to making the 'cut' is problematic for making claims about democratising education; whilst this learning experience is technically 'free' and 'open', it is not actually 'open' to those outside the community of nomadic learning affiliates, and the finding regarding cognitive presence going up as people leave supports my claims of educational elitism, which actually further supports Sfard's support for pragmatism and avoiding a "trendy mixture of constructivist, social-interactionist, and situationist approaches".

Any social network analysis of cMOOCs necessarily will not include a complete data set, because the premise is that learners will create their own personal learning environments and own the data they choose to aggregate into the community. The analysis above is limited by its sole focus on Twitter; indeed one of the hashtags included in the SNA is #rhizoradio which suggests that some learners may be creating digital radio stations or podcasts on totally separate platforms, a practice first established in the DS106 (digital storytelling) cMOOC community[4].

It is also important to note that these analytics are also post-hoc. They are developed after the event for the purposes of the course design team, or educational researchers, which supports the argument that they cannot help learners at the point of their learning. In this sense, the withholding of this data from students is treating the learning environment itself in instrumental terms (as the neutral means of achieving objectives: 'learning communities' in the case of cMOOCs, 'engagement analytics' in the case of xMOOCs).

[4] http://ds106.us/ds106-radio

2.6.2 Summary of learning in cMOOCs

It is important to note that cMOOCs do not intend to measure fixed learning outcomes; it is the persistence of the community that evidences their value to the participants. They are an extension of open educational practices, so welcome nomadic learners who understand how to participate, and their participation is evidenced through creative endeavours like writing blog posts, so fulfil the higher order categories of Bloom's taxonomy.

cMOOCs are a collection of communities, creative endeavours and overlapping conversations, so are their own re-conceptualisation of the notion of a 'course' or a 'curriculum'. In traditional terms, a 'course' has a beginning and an end; cMOOCs do not. As the tagline for DS106 states: "Start any time, it never ends. Design it your way."[5]

The 'health' of the 'course' can be measured through proxies such as SNA but can never totally gather the complete data to track development of personal learning goals. They have a high barrier to entry, and many learners report that they are overwhelmed by the fragmentary nature of the course materials and the practice of peer support. In the language of usability designers such as Norman (1988) or

---

[5] http://ds106.us/

Neilsen (2012), cMOOC platforms are not very usable in terms of Learnability, Efficiency, Memorability, Errors, Satisfaction. They are not designed that way.

In summary, cMOOCs are representative of the knowledge of an open community, rather than being a ‘course’ as such and it relies on learner’s capability with digital platforms, and the features they afford to mediate the community participation. For example, blogs are a useful method of representing reflective accounts on a topic, and certainly interactive writing has been demonstrated to aid the learning process (Lapadat, 2006) but the central cMOOC platform uses aggregation technology (RSS) and this requires certain amount of manual effort to discover interesting things to read and reply. Therefore, they are susceptible to the same problems of overload faced by xMOOC platforms in that they are limited by their mediating artefacts.

2.7 Coordination and collaboration through discussion environments

The MOOC context for learning is self-defined as being unique from other distance learning models because of the potential for scale. This scale does not simply refer to the number of learners who are able to watch the videos (as this is no different from the TV based distance learning of the past), but also the ability to form communication networks and discover people and concepts who are able to help construct knowledge and cement understanding. I propose that new tools need to be introduced into the learning environment in order to scale discussion activity which both support self-directed learning and diversity of learner preferences, but also that are usable by as many people as possible. xMOOC tools are relatively easy to use but they mainly support a transmission pedagogy, and cMOOC tools are difficult to use, but have potential to support a participative pedagogical approach. The following section examines the problems with existing discussion toolsets in MOOCs and relates this to a line of literature which critiques discussion tools in the pre-MOOC era.

Rosé and Ferschke suggest that learning in today's MOOC environments is mostly a solitary experience and set out a challenge for work that may take a generation to accomplish:

> Currently available platforms for learning at scale are frequently impoverished on [the social] dimension … Though there are almost always discussion forums included in these environments, they are often just an appendage, and not effective in meeting the needs of learners. (Rosé & Ferschke, 2016, p. 661)

They cite an experimental design of a course delivered as an xMOOC and in parallel as a cMOOC; the xMOOC scaffolds learning content, then evolves into a more community-based design which encourages participation on the parallel platform.

Rosé and Ferschke challenge researchers and educators to design platforms and pedagogies where communities of practice can emerge with their own means of knowledge exchange; this includes methods for peer-assistance and navigating massive amounts of community generated data. This is set as a generational challenge for MOOCs, AI and pedagogical theory, and underlines that this type of MOOC research is still in its infancy.

Current research focusing specifically on MOOC discussion forums reveals that we have a long way to go before we can support dynamic and emergent community formation. There can be low levels of participation (Breslow et al., 2013) or low levels of responsive participation (Tubman et al., 2016), and disorganized, noisy posts are normal (Brinton et al., 2014; Mcguire, 2013). Posts are often not directly course-related (Wise & Cui, 2018) and increased scale creates different types of overload problems (Barricelli et al., 2015) which make it infeasible to navigate and discover interesting and relevant posts (Kop, 2011; Qiu et al., 2012). Whilst the above review focuses mainly on xMOOC platforms, the cMOOC platform cannot handle content discovery at scale either, as it is mediated by RSS feeds, Twitter conversations, and mass mailing lists. The distributed and fragmented nature of the conversation makes it difficult to follow, although the use of hashtags as 'social glue' is something which I explore in more detail in this project.

These findings with MOOC discussion tools are not altogether new or surprising as online forums have long been a site of contention. Hewitt simulated an asynchronous discussion forum years before MOOCs were established, focusing on the ‘unread posts’ affordance and the behaviour it subsequently encouraged. He concluded that the platform which prioritises new things also starves out the old (Hewitt, 2005). In design, a pull in one direction creates a push in another. Scardamalia has also recognised the limitations of online, threaded discussion as a pedagogically suitable tool for knowledge enhancement, and suggests that “discussions over the Internet show low levels of participation and a lack of continuity and moreover typically require a good deal of teacher direction” (Scardamalia, 2002, p. 72). She advocates the online collaboration tool, ‘knowledge forum’, which builds knowledge through mapping *relationships* between comments in a non-linear way but the cases she studies use this tool in small groupwork settings where there are few differences in learner behaviours and common motivations for participation. The point here is that certain affordances are necessary but not sufficient, and they need to be embedded within pedagogical frameworks which define or scaffold types of learner behaviour within the given context.

The use of pedagogical scaffolding to model social behaviour is not currently part of MOOC designs or research; other studies have taken a broader approach to explaining social behaviour in MOOCs, either a mechanism for learners to ‘feel’ less isolated or to ask for information (Kizilcec et al., 2013). This marks a distinction from discussion-based learning activities seen in more established online distance learning pedagogies, for example where ‘communities of inquiry’ are orchestrated through specific forms of engagement and 'cognitive presence' is scaffolded through guided 'teaching presence' (Garrison, 1997) or 'knowledge building' exercises as described by Scardamalia (2002).

Another area where MOOC pedagogy differentiates from previous online distance learning pedagogies is in the learning needs of MOOC learners who are a heterogeneous community and have diverse intellectual motivations for study; therefore, enabling and encouraging discovery of interesting and relevant content is of primary concern. It is important that MOOC platforms are designed so that learners have the agency and tools to personalise their participation to what is most meaningful to them.

2.8 Summary of the literature informing this thesis

Learning as a concept broadly can be traced back to 2 origin stories: the acquisition and the participative. Each approach has conceptual and practical issues and Sfard (1998) cautions on selecting just one when designing a pedagogical approach.

Earlier versions of e-learning lean more heavily towards the participative approach, and tools and learning designs are grounded in Vygotskian ideas and sociocultural theories, which place emphasis on structured discursive activities and meeting course learning outcomes 'through' participation in those activities.

Open Education discourses rely more strongly on the language of the acquisition metaphor for learning and are premised on a philosophy of ‘knowledge as commodity in the public sphere’. Structured learning designs relating activity to learning outcomes are eschewed in favour of a self-directed approach to discovery and engagement with open educational resources. The peer production strand of open education also emerges from an ‘acquisition’ metaphor for knowledge in its building of community-oriented resources such as open source software or coherent knowledge bases.

MOOCs can be divided into the xMOOC approach and the cMOOC approach and are often a hybrid of both approaches (Open + Online Course) but they ‘switch’ the focus of their metaphorical origin stories. That is to say, the xMOOC approach, which concerns structured courses uses an acquisition metaphor in terms of its efficient transmission of authentic academic knowledge and evidencing of learning outcomes through learning analytics. This contrasts with earlier uses of technology such as e-learning which focused on building knowledge through participation. The cMOOC is an extension of open education, which previously had been concerned with generation and licensing of knowledge objects or resources and is conceived as the production of communities or networks that challenge established institutional organisations.

Both these approaches contain problems, which centre on the concept of ‘massive’ in terms of participation overload and information overload. The xMOOC model handles the massive cohorts by simplifying the pedagogical approach, and therefore limiting the subsequent analysis to engagement statistics ("We have terabytes of data about what students clicked and very little understanding of what changed in their heads." (Reich, 2015)), whereas the cMOOC model ups the bar for effective participation ("…they need the ability to understand the intricacies of the networks in order to negotiate their structures." (Kop, 2011)).

In both cases the analysis assumes that the technology is a neutral mediator for the learning outcomes and restricts access to the analytics to designers and researchers. This project challenges this and attempts to develop social learning analytics which are available to learners as they are learning, and builds on the idea that learners are best placed to negotiate the emergent ecosystem which grows from the new affordances and roles brought about by the learning analytics (Buckingham Shum & Ferguson, 2012).

That is not to say that all analytics are immediately 'usable' by learners, as the case of the cMOOC proves that introducing complex ideas and tools raises the bar for participation rather than lowering the bar for access to higher education. The visualisation produced in this project attempts to find a balance between simplicity for the everyday user and their learning goals, and sophistication in terms of development of communities and active participation.

Chapter 3: Theoretical frameworks

3.1 Introduction

The exploration into recent and relevant literature highlights that there is a gap in research examining the relationship between affordances of MOOC platform tools and its relationship with designing for learning using a sociocultural framework. Several researchers (Bali, 2014; Bayne & Ross, 2014; Conole, 2015; Stacey, 2014) have suggested that xMOOC platforms are deficient in supporting collaborative activity, and encourage an individualistic learning design, which is highlighted as problematic because both narrowing the learning design possibilities and simultaneously expanding the reach through globalisation may result in less innovation overall in terms of teaching practice (Knox, 2014) or embed behaviourism as the normative approach to teaching and learning. There are clear resonances here with the analysis that 'the virtual university is . . . the university made concrete' (Cornford, 2000) in that the introduction of virtual platforms reify certain organisational practices towards greater standardisation.

The under-researched area into the relationship between the material affordances of the platforms and the possible learning designs is fundamental because current toolsets in xMOOCs tend to encourage consumption pedagogies. Current analysis 'black boxes' technological aspects with its focus on the analysis of clickstream data without fully acknowledging that it is the affordances (in terms of cybernetic qualities) of the toolsets which have created the forms of the data to be analysed.

> Technological variety and multiplicity is generally 'black-boxed' in these accounts: part-defined at best, sealed-off from interrogation via vague and homogenising terms like 'ICT' or 'online facility', technology is then positioned firmly as being 'in service' to the demands of the prior social activities of learning and teaching. This bracketing-off of technology from social activity is expressive of a more fundamental division of society from technology which is widespread within the field of digital education.
> (Bayne, 2015)

These ideas can be dated back to McLuhan's (1964) claims that 'the medium is the message' in terms of how 'media' develops the mindset for engaging in learning and social life, or Postman's (1969) claims about technology being intertwined with institutions whose organisation reflects a world view which is encouraged by the technology in a self-fulfilling cycle. Connections can also be made to Biesta (2005) who critiques the use of the term 'learning' at the expense of 'education'; Biesta claims that education is the discovery of what is required in order to solve authentic problems, as opposed to having those needs already pre-defined, pre-packaged and already met by an institution and its surrounding technologies.

I employ the 'Cultures of Participation' framework (Fischer, 2011, 1998) as a practical operationalisation of these ideas. This sets out specific levels of participation in a community or culture, ranging from passive consumer through to designer or community leader. Fischer's framework implicitly shares much of Postman (1992; 1969) and McLuhan's (1964) philosophical premises and warns against developing media which encourages users to be 'couch potatoes' as this will extend beyond learning into all aspects of how we interact in our social life more generally, such as democracy, echoing ideas from Illich (1971) about the decentralisation of authority. Importantly for Fischer, it is the affordances built into the media itself which encourage the sort of 'levelling up' in collaborative behaviours which are associated with higher levels of 'Cultures of Participation'. At its highest levels, user activities are directed at the community itself and contribute towards the overall mission of the group rather than an individual:

> The fundamental challenge for computational media is to contribute to the invention and design of cultures in which humans can express themselves and engage in personally meaningful activities…
> Unfortunately, a large number of new media are designed from the perspective of seeing and treating humans primarily as consumers.
> In *personally meaningful activities,* the possibility for humans to be and to act as *designers* (in cases in which they desire to do so) should be accessible not only to a small group of "high-tech scribes," but rather to all interested individuals and groups. (Fischer, 1998)

This analysis is important if MOOCs are to be used as vehicles of community learning, which would be a continuation of the trends highlighted by Conole (2013), Beetham and Sharpe (2013), Buckingham Shum and Ferguson (2012) towards more participatory and social methods of learning:

> There has been a paradigm shift in the field due to new thinking around learning theories from behaviourism, through to cognitivism and finally constructivism (Mayes and De Freitas 2004). These theories led to the development of particular uses of technology designed to support the underpinning principles of the theories. (Conole, 2013, p. 3)

It isn't new to make detailed claims about the design of online learning environments and the behaviours it encourages affects learning. Indeed, in 1999, Britain and Liber produced a report for JISC (Joint Information Services Committee) in the UK titled "A Framework for Pedagogical Evaluation of Virtual Learning Environments" taking a 'cybernetic' analysis to this problem. This report was updated in 2004 when institutional virtual learning environments had matured and were more widespread in higher education. Their 'cybernetic' analysis to the design of virtual learning environments examines not only the behaviours which are encouraged by the affordances, as Fischer's cultures of participation, but also several ways in which the features need to be extended to support the higher behaviours. They employ the 'viable system model' from Ashby (1956) which states that the variety of a controller must match the variety of the system to be controlled. In this sense, the 'controller' refers to the online platform (and its affordances or action potentials) and the 'system to be controlled' is the designed-for experiences of the MOOC cohort. Given that MOOCs have a large, heterogeneous cohort of learners with diverse learning needs and skillsets, the cybernetic answer to this problem is that the variety of educational provision must be increased to match the environment along the following communication channels: *resource negotiation, adaptation, self-organisation, monitoring and individualisation* (Britain & Liber, 2004). xMOOC platforms, such as the FutureLearn platform studied in this project, do not increase variety or complexity across any of those channels, instead delivering a single scalable product with few affordances for personalising experience.

Bayne and Ross (2014) review the MOOC phenomenon and note that pedagogical factors are under-represented in the discussions. They describe MOOC pedagogy as "something that emerges in complex negotiation between the platform, the teaching approaches of the academic team developing the course, and the pattern of learning interactions as the course is played out", which emphasises the relationship between artefacts, learning designs and behaviours which are subsequently encouraged. This aligns with Goodyear's ACAD learning design framework (Goodyear & Carvalho, 2016) which talks of the interaction between environment, epistemology and social activity, and this itself aligns with Vygotskian activity theory in terms of the relationships between artefacts (learning environment/ tools), rules (suggested learning activities) and division of labour (social activity) and is represented in Figure 2.

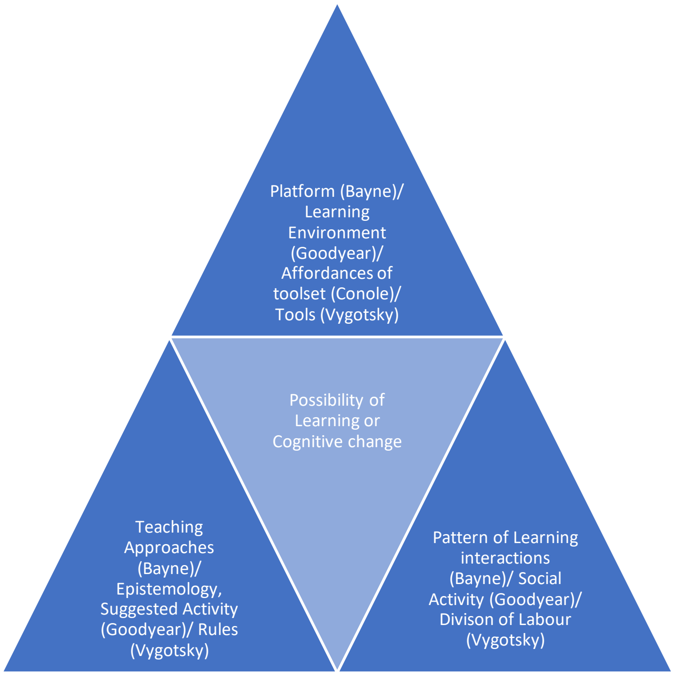


Figure 2 Interdependency between tools, teaching approaches or activity design and actual social activity or division of labour when designing for learning, as described by Bayne and Ross (2014) and Goodyear (2016) and aligned with Vygotskian activity theory. All 3 factors are intertwined in terms of understanding designing for learning and cannot be separated or 'black boxed'

In terms of analysing patterns of behaviour in MOOCs, I have critiqued methods of analysis which compiles individual engagement into forms of 'learning analytics' as these are insightful only for teachers, researchers or designers in aggregate and only focus on the 'pattern of learning interactions' in Figure 2; additionally, their form is at least partly defined by the technology itself. These patterns of behaviour are only disaggregated back to the level of the individual to evaluate performance, or 'progress', which can be problematic when equated to 'learning' as it can relate only to lower levels of criticality in Bloom's taxonomy using current toolsets. An alternative method of measuring learner behaviour is through engagement with the overall community as evidence of collaborative knowledge construction, alongside subsequent reflection on this process, rather than individualised performance. This is commonly found in sociocultural learning designs, grounded in Vygotsky's idea that higher level knowledge is necessarily socially created (Vygotsky, 1978).

I focus on learning designs which have emerged from sociocultural theory such as Laurillard's 'Conversational Framework' (Laurillard, 1993) which is the theoretical basis of the FutureLearn platform design (Ferguson & Sharples, 2014), and analyse the effect of introducing a new platform affordance in order to increase the potential for networked learning in the MOOC context, taking into account the 3 parts of Goodyear's ACAD framework (artefact or environment, learning activity, social activity) (Goodyear & Carvalho, 2016) and focusing analysis on one in each stage of design-based research. In researching the types of new activities and behaviours which are encouraged by this specific platform affordance, I ask how new mediating artefacts could provide a springboard for learning activities which embrace socially innovative pedagogy. It is important to note that the mediating artefact which I have developed actually allows learners to make use of their own 'analytics' which is a departure from the usual ways in which analytics have been deployed, especially at the MOOC scale, and that its use of visualisation technology resolves problems associated with information and participation overload in discussion forums (e.g. Brinton et al., 2013). Democratising analytics in this way also follows on from Buckingham Shum and Ferguson's recommendations that analytics "should be seen as tools to be placed in the hands of the very subjects being analysed – the learners" (Buckingham Shum & Ferguson, 2012).

I have identified that scaling up collaborative activities presents a major problem when the same learning designs used in small groups are applied. This is widely cited as being difficult to orchestrate, why discussion forums are used as an appendage rather than a core activity, and ultimately how they are not fit for purpose in their current form. A new way of looking at the design of social features is required, and I suggest 'Stigmergy', as a design paradigm to account for systems which thrive with scaled participation, can help understand how to manage scale on the learning platforms. Stigmergy is defined as communication through signs left in the environment (Dron, 2006; Elliott, 2016; Van Dyke Parunak, 2006) and is used to describe the processes by which large constructions can be built through micro-contributions from many actors. It employs the analogy of ants building complex hive structures through sensing the trails (signs) of pheromones left by those who came before them to coordinate activity.

Along similar lines, Haythornthwaite (2009) has developed theory to explain models of participation, utilising concepts of 'light and heavyweight peer production', but there are key differences in the forms of participation she cites and the practice of learning at scale. Lightweight peer production involves making a micro-contribution to a structured dataset with a weak-tie association with the overall community, and heavyweight peer production is attributed, has strong tie affiliation to the community and demands qualitative judgements about how it contributes to the social practice. A key difference between this and the stigmergic framework is that lightweight peer production is largely autonomous, and not reliant on *previous* *contributions to guide future action*. Contributions from learners in a MOOC do not fit neatly into either of these categories but are nonetheless peer produced 'signs' left in the environment for themselves and others.

This and similar theoretical frameworks have previously been used to describe the development of peer production systems such as 'Wikipedia', but have yet to be considered in terms of framing thinking about social learning at scale. A key difference between previous peer production systems and learning at scale lies in the end-product being developed: for example, a wiki system can handle scaled participation (indeed it thrives in this context), but it also has a very defined end point (i.e., a single, coherent page of information), and employs very strict rules of engagement in order to produce that product, therefore it is a highly structured cooperative process. On the other hand, social learning at scale requires that tools are developed which can process and coordinate the scale of participation and information, but it needs to be directed at the level of individual agency in terms of learners' self-directed motivations.

In this project, I have introduced a new mediating artefact for use on the FutureLearn MOOC platform, which uses an interactive 'word cloud' visualisation to represent the massive participation common to MOOCs. By using the raw computational power, the artefact offloads the requirement for learners to read through every comment in order to discover something which is relevant to their personal learning goals and in this sense it can simply act as a tool for filtering and discovery that is not available through the existing discussion tools. Arguably, the feature could have been replaced with a simple search tool, but by designing the artefact as an interactive visualisation of content, it is also layered and changes dynamically based on user choice, so upon selecting a single word, it also reveals which words were also frequently used in comments which used the same initial selection, encouraging serendipitous connections.

There is an implicit scaffolding affordance inherent in the design, dependent on learners being encouraged to 'play' with the tool in such a way that they perceive related concepts in the comment corpus that may not be immediately available through linear reading. For example, in a specific FutureLearn course on Dyslexia, if the word 'vocabulary' is clicked, then the next cloud displays only words in comments which used the word 'vocabulary' (sized according to frequency), and from there it is possible to see that the word 'grammar' has become more prominent, demonstrating that learners are talking about these concepts concurrently in their comments; other words are also more prominent after these initial selections, such as 'reading', 'practice', 'spelling' and so reveals further conceptual linkages or 'implicit scaffolds' (Lin & Puntambekar, 2019). It is also possible to link directly to the comments themselves, so it encourages further discussion and written interaction, which is related to higher levels of learning and critical thinking in a sociocultural sense (Lapadat, 2006; Vygotsky, 1978). These basic affordances of the tool extend the capabilities of the platform along the dimensions of resource negotiation, adaptation, self-organisation and individualisation as highlighted above in terms of cybernetics as described above. Whilst these affordances of the mediating artefact enable learners to discover relevant content for themselves and conduct their own 'taskwork' through reflecting on the emergent scaffolding and facilitating cooperative social behaviours by filtering to relevant conversations, this on its own is not levelling up to the highest levels in the cultures of participation framework, which actively supports and grows the community. In order to make this next step, the mediating artefact must be used in conjunction with the design of collaborative learning activities, which suggest models of participation which will enrich the visualisation, such as the use of hashtags which have been proven in cMOOC designs to act as 'social glue' for communities of affinity (Bozkurt et al., 2016). In defining this 'division of labour' (social tagging/ folksonomies) the learning design enhances collaborative behaviours, which are identified as lacking in MOOC pedagogies.

This demonstrates that a new mediating artefact, in conjunction with innovative learning designs which support and encourage self-directed enquiry, and divisions of labour which can work to form communities of affinity, can re-frame the dominant pedagogy used in MOOCs from narrow didactic towards a participatory, cooperative educational experience.

3.2 Cultures of Participation

A culture of participation, according to Gerhard Fischer, is one where all people are provided with the means to participate and contribute actively to personally meaningful problems. Technical innovation is necessary for a culture of participation because cultures are substantially defined by their media and their tools for thinking, working, learning, and collaborating. Technical innovation is not sufficient however, as cultures emerge out of the change in behavior and organization of the users which is brought on by the innovation.

> Our current educational institutions often treat learners as consumers, fostering in students a mind-set of consumerism rather than of ownership of problems, which they carry with them for the rest of their lives.
> (Fischer, 2011, p. 42)

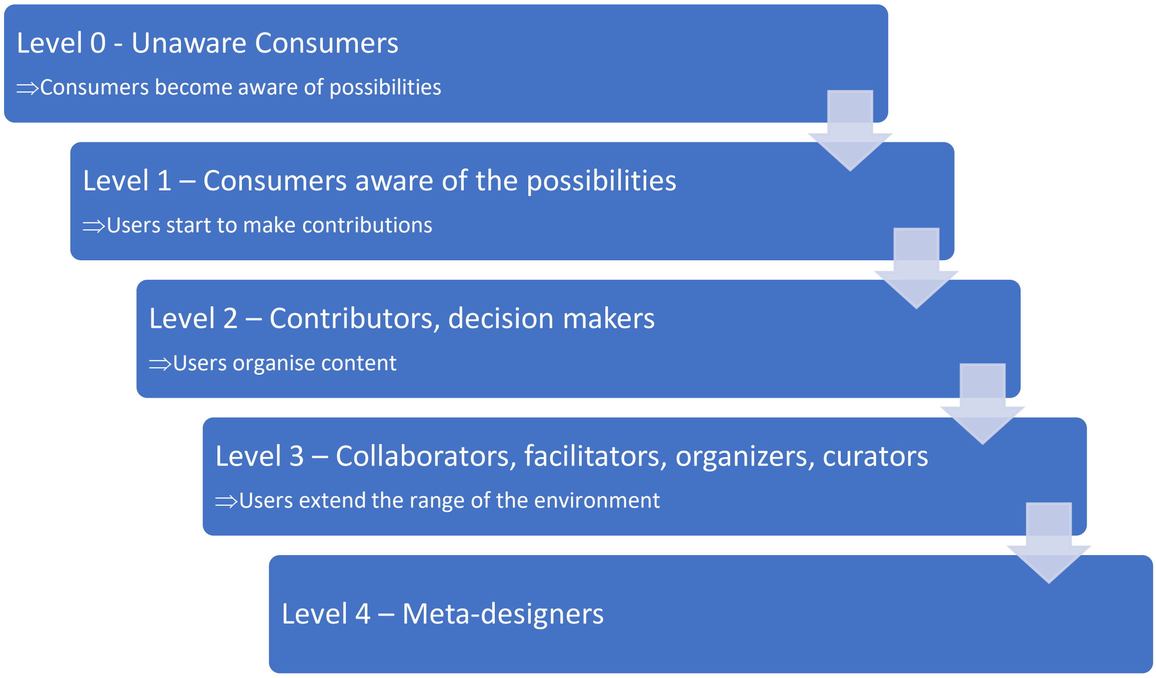


Figure 3 Cultures of Participation are represented into a 5-level process by Fischer, 2011 p. 42

Figure 3 represents Fischer's Cultures of Participation framework into a 5 level process. In terms of the designed for interactions in xMOOC platforms and tools, users do not have the capacity to extend beyond 'Level 1' in terms of utilising basic discussion tools which come with all the problems which have been identified in the literature. It is not possible to organise or personalise content, preventing movement to Level 2 and it is not possible to create self-directed learning pathways which would extend the range of the environment (Level 3). This encourages users to remain at 'level 0' or 'level 1' in terms of engaging with the published content and passively interacting with the media, as 'couch potatoes', as Fischer terms interactions on this level.

> In the past, the design of most media emphasized a clear distinction between producers and consumers (Benkler, 2006). Television is the

> medium that most obviously exhibits this orientation and has contributed to the degeneration of humans into “couch potatoes” (Fischer, 1998), for whom remote controls are the most important instruments of their cognitive activities. (Fischer, 2011, p. 42)

FutureLearn encourages users to contribute on any page of content, so has a ‘Level 2’ ceiling in terms of users making contributions which are relevant to that page of content. However, those contributions cannot be effectively organised, curated or extended, as per the highest levels of participation, and the typical interaction is to read only a few comments which are chronologically just ahead:

> Each teaching element (or “step”) is associated with a free-flowing discussion, which is intended to emulate a “watercooler conversation” about the immediate content. Each learner can see contributions from other learners and add a brief comment or reply. (Ferguson & Sharples, 2014, p. 101)

Note the use of ‘free-flowing discussion’; to extend this metaphor further, it could represent the fact that users which share your interests are ‘downstream’ or ‘upstream’ and there is no way to locate them through the toolset.

The levels within Cultures of Participation refer to behaviours within social computing at the level of the interactions afforded by and designed into the artefact, rather than speaking specifically about pedagogical methodologies, such as how learning activities are proposed. The framework is important to consider because it describes how media can be designed for collaboration and participation; it does not describe how collaboration can be specifically designed for learning. A pedagogical theory is required in order to bridge the design framework and learning design, and Laurillard's conversational framework (1993) has previously been used in order to examine the relationship between the designed environment and its potential for learning. Britain and Liber (2004) have employed concepts from cybernetics in order to understand the specific dimensions of this relationship. The following section explores their work with institutional Virtual Learning Environments and how this relates to MOOC platform design.

3.3 A Cybernetic account of Virtual Learning Environment design

In 1999, Britain and Liber produced a report for JISC (Joint Information Services Committee) in the UK titled "A Framework for Pedagogical Evaluation of Virtual Learning Environments". This report was updated in 2004 when institutional virtual learning environments had matured and were more widespread across the sector. In both reports they argued that there is a strong link between the way that a Virtual Learning Environment (VLE) is designed and the way that it is typically used, and presented 2 complementary models for understanding how learning can be understood: the Conversational Framework (Laurillard, 1993), as the pedagogical model, and the Viable System Model (VSM) drawn from management cybernetics, for the design and diagnosis of effective organisational structures. The VSM is used to determine whether the system will encourage or inhibit the pedagogical model overall. The VSM holds that it is variety that overwhelms systems organisations and needs to be managed effectively. Variety is a measure of complexity used in cybernetics and comes from the Law of Requisite Variety (Ashby, 1956) which states that *the variety of a controller must match the variety of the system to be controlled.* If this is applied to the increased variety and diversity of learning needs in a large, heterogeneous MOOC cohort, the cybernetic answer to this problem is that the variety of educational provision must be increased to match the environment along the following communication channels: *resource negotiation, adaptation, self-organisation, monitoring and individualisation* (Britain & Liber, 2004).

MOOC platforms, including FutureLearn, do not increase *variety or complexity* across any of those channels; they deliver a product for efficient scaling, such as a video or a multiple-choice quiz and offer few affordances for personalising experience. Therefore, it is difficult to see how the ‘conversational framework’ pedagogical model is encouraged at the level of the overall system. Britain and Liber give a concrete example about the design of a course:

> if [a course] is constructed on the basis that all learning activities have to be created and sequenced in advance of a course beginning, then there is no way that a teacher can create and add a learning activity to the course

> on the basis of a preceding conversation at the level of concepts with the students (a basic idea of the conversational model). Alternatively, if all the students in one group are treated as belonging a single ‘class’ object to which learning activities are assigned in the VLE, then there is no possibility for creating individual learning activities. (Britain & Liber, 2004, pp. 11–12)

Arguably, by designing discussions into every page of the course as described above, the FutureLearn platform affords generation of new content ‘at the level of concepts’ in the form of user comments which does align somewhat with the conversational model, albeit not in a direct feedback loop with the facilitator. In MOOC environments the teacher cannot be on hand to read, let alone facilitate all these conversations or provide effective feedback to everybody. There is also no search facility on FutureLearn (because there are so many areas where learners can comment), which as a design decision keeps users on their current page, rather than navigating them to the most personally useful comment on another page (this risks breaking individual progression and therefore course ‘completion’).

It is also a constraint of the primacy of the completion factor in the FutureLearn MOOC platform that educators cannot add new resources to the sequence of learning activities until the next occasion that a course is published (with new learners!). However, if the comments feature is taken as the designed opportunity for adding novel relevant resources in light of a conversation, there still remains a problem of co-ordination of these resources or managing 'information overload' or 'participation overload' in order to develop effective variety in resource negotiation or individualisation. Britain and Liber also acknowledge this constraint of the 'communication channel' as explained below:

> The design of the course – the sequencing and structuring of learning activities and resources (including assessments) – is how a teacher can provide a co-ordinatory framework. Like classrooms, many VLEs provide limited opportunity for flexibility here. Just as a 1-hour lesson in a lecture theatre encourages coordination by sitting still and being quiet so that the class can all learn together, many VLEs encourage a method that can be caricatured as "read this material, check the forum and do the test". It takes some effort on the part of the teacher to overcome these, but they do – and are supported to some extent by the design of the system. If you can move the chairs, you have more choice in a classroom; if you can adapt the workflow of a VLE, you can provide more flexibility in your learning opportunities. (Britain & Liber, 2004, p. 16)

3.3.1 "Read this material, check the forum and do the test": the tyranny of 'progression' in MOOCs

It is interesting how this critique of online platforms in 1999 and 2004 remains salient, more so as student numbers have increased, and platforms have both globalised and standardised, alongside decreasing levels of direct facilitation. MOOCs are fundamentally designed to value progression and completion – seemingly ‘good’ ideals – but potentially at the detriment of ‘learning’, especially as conceptualised in an open and self-directed sense. There is still hope though, in that “*if you can adapt the workflow of a VLE, you can provide more flexibility in your learning opportunities*” (and remember learning *opportunities* are naturally increased due to the nature of participation ‘overload’). In the next section I investigate the idea of engineering possibilities for serendipity or chance encounters in the learning platform design, demonstrating increased variety across the ‘self-organisation’ communication channel.

3.3.2 Engineering serendipity for the MOOC context – it’s all about the ‘discovery’!

> “Luck is what happens when preparation meets opportunity” Seneca the Younger (54-15BC)

A MOOC is a space where large numbers of similarly interested people congregate to learn and discuss related ideas. An analogy for this type of ‘crowd learning’ may be to imagine attending an event in a large stadium: you are physically constrained to make conversation with those who are in your vicinity, even though there are thousands of others around you who are similarly interested in the event. It may be that the person you need to engage with to meet your own learning goals is located somewhere else or came to the event yesterday. In this type of space, learning through conversations is constrained to narrow serendipity, or chance encounters.

When discussing learning in this space, and critiquing platforms in terms of how their affordances encourage or constrain the conversational model, it is important to briefly discuss how social learning may be defined and therefore measured. For connectivists learning happens cyclically; learners connect to new ideas through emergent and self-managed networks, formulate new beliefs and then re-connect the new ideas back into the network (Kop & Hill, 2008). Sociocultural theorists would suggest that learning is connecting with a more experienced other and working through a zone of proximal development or constructing situated knowledge with peers through the synthesis of ideas within different perspectives and experiences (Vygotsky, 1978; Wertsch & Bivens, 1992). Whichever theoretical lens this type of social learning is viewed through, the MOOC event is an ideal and unique place for it to occur; interested others of all scales of expertise are in and around the same virtual space with the freedom to explore. Some learners are already qualified and so their motivations for engagement may be simply to experience diverse viewpoints to further their own practice, meet similarly interested people, or share their expertise with novices.

Described in this way, it seems absurd that course platforms place so much value on individual progress and completion, rather than building tools which enable reflecting, connecting, investigating, discovery and discussion. The flip side to this argument is that learners can feel lost when there is no platform infrastructure or narrative within a course, so these experiences should be structured through platform affordances and pedagogical approaches (teaching activities). This project explores the space in between these viewpoints and the CDT artefact attempts to encourage serendipity through opening new possibilities for connecting with others and their ideas ("*adapt the workflow… provide more flexibility*"), thereby increasing the variety in the learning system as is the requirement of the viable system model as described above.

Serendipity is defined as a 'happy accident'; it occurs when people who are open to new things and willing to think tangentially can mentally make new connections and be creative with them. A MOOC is a hotbed of potential for 'happy accidents' as a global cohort of people with different specialisms and interests are able to connect. Jack Dorsey, CEO of Twitter, quoted in Forbes magazine, speaks about how he encourages serendipity and discovery in the physical environment of the workplace:

> We encourage people to stay out in the open because we believe in serendipity--and people walking by each other teaching new things…. I do a lot of my work at stand-up tables, which anyone can come up to… and I get to hear all these conversations around the company. I spend 90% of my time with people who don't report to me, which also allows for serendipity, since I'm walking around the office all the time. You don't have to schedule serendipity. It just happens. (Savitz, 2012)

Engineering serendipity in computing systems is about creating a platform in which these connections *might* occur; these connections are more likely to happen when the system moves from a model of *information search* to one of *information discovery*. Laura Larsell, information ontologist at Trapit (a content discovery, personalization and curation platform), writes for MashableUK blog about the problems with search, and the power of simply browsing:

> Search assumes a direct path between the seeker and the sought. Ironically, “search” works best when you have a pretty good sense of what you are looking for. But most people, most of the time, do not have concrete ideas of what they really want. Netflix knows this. Amazon does too. The two offer users the ability to browse collections by subject or genre. (Larsell, 2011)

FutureLearn has no designed affordance for browsing by subject or genre, preferring sequential navigation and temporal discussions; other MOOC platforms afford the ability to search forums for keywords, however both these approaches reduce the element of *discovery*, along with the associated serendipity which follows on from it.

Alex Krotoski explores the concept of serendipity in an episode of her ‘Digital Human’ podcast series. She states that there are 3 pre-conditions for serendipity to occur – firstly the chance encounter, then the insight, and finally the value to the connection

(Krotoski, 2012). Discovery is an important concept when designing a system where a chance encounter may occur; if a learner is ‘searching’ for something, they are only likely to discover what they are looking for, ranked according to impersonal algorithms such as Google page rank, which will return results based on what other people are searching and clicking. There is an efficiency to this, but it isn’t likely to present a huge number of happy accidents and may in fact drive down peripheral opinions and exclude people who do not know exact keywords. When learning through discovery and reflection, it is important to be open to all opinions and be presented with results that can be challenging and unexpected. It is in these unexpected results that insight can be formed, and value can be developed around the connection.

The CDT encourages browsing and seeing the words in the visualisation change when more terms are selected, acting as a kind of emergent, implicit scaffolding for possible connections. Mel Woods, head of the research project ‘Chance Encounters in the Space of Ideas’ (SerenA project), describes serendipity as “almost like a space, a serendipity space, it’s an ephemeral, moving shifting being” (Krotoski, 2012). The visualisation will never look the same twice because of the way in which it renders on the screen; even with the same corpus, words are in a different position. In this way, the CDT encourages a playful interaction and will change in ways unexpected to the user, providing a space in which serendipity can occur. However, it is important to prepare the learners for this type of nomadic, playful, surprising experience carefully, so as not to create confusion and this needs to be done through carefully designed pedagogy. The next section examines pedagogical models for collaborative learning and focuses on an article which attempts to provide a framework for learning, through which technology can be examined.

3.4 Pedagogical models for collaborative learning

Collaborative learning models based in a sociocultural tradition follow on from the Vygotskian idea that higher level knowledge is necessarily socially created (Vygotsky, 1978). In terms of orchestrating the social creation of knowledge, various learning designs and toolkits as mentioned in the literature chapter have been developed as an attempt to construct and synthesise new knowledge through discussion and groupwork activities.

An example of how sociocultural learning theory can be applied to teaching practice can be found in Laurillard's Conversational Framework (Laurillard, 1993). The conversational framework embraces the notion that learning occurs through dialogue in a 'teacher constructed learning environment' between a 'teacher conception' and a 'learner conception' and 'at the level of concepts' which reflects the Vygotskian idea that optimal learning occurs in a learning environment which constructs a zone of proximal development. Activities are constructed in the conversational model to be between learners, peers, teachers, and concepts, as with the Figure 4:

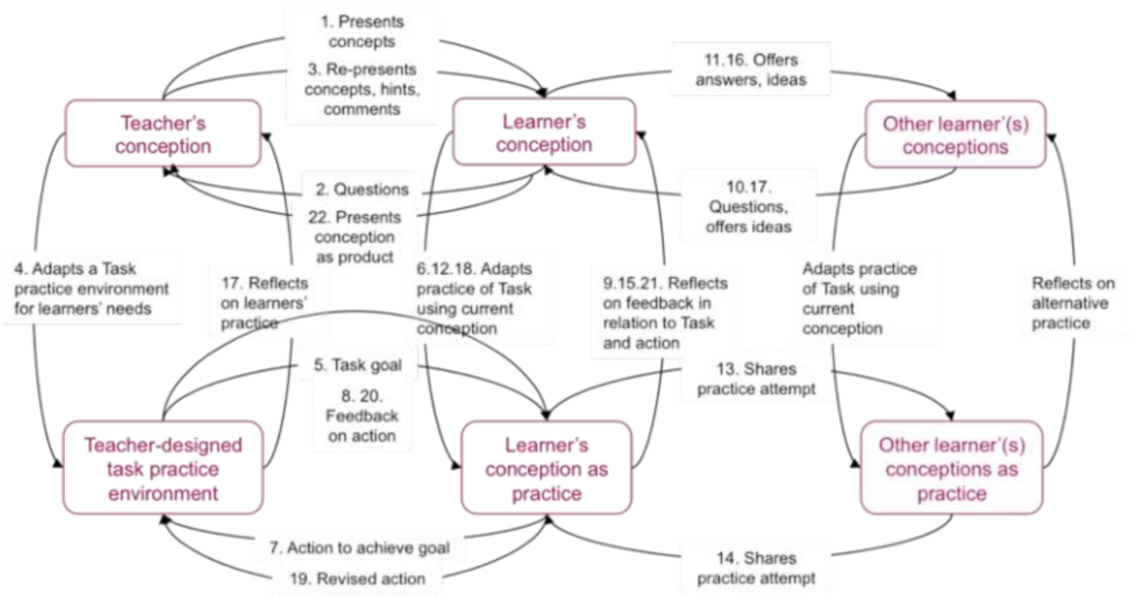


Figure 4 The Conversational Framework: instructionism, social learning, constructionism, and collaborative learning combine to provide a simplified representation of what it takes to learn. Numbers show a possible ordering of the successive activities of learner, teacher, and peers (from Laurillard, 1993)

In smaller groups and with direct facilitation, this model can map out different types and ordering of interactions and activities as a flow of conversations. It starts with a teacher presentation and learners' initial conceptions derived from the presentation, and follows through to learner practice, teacher feedback, learner reflection and peer discussion or groupwork until such a time when the learner has a firmer conception of the subject. However, when learning at scale, this highly collaborative flow of activity and discussion cannot be as strictly orchestrated because the large, heterogeneous cohort of self-directed learners will make their own way through the subject and want to discuss topics from a diversity of angles; it will not be possible to directly facilitate all the conversations or monitor all the creative products of 'practice' from a teaching perspective. This puts the focus on the right-hand side of the diagram, in terms of peer cooperation and the formation of communities of affinity within which more focused conversations can occur.

'Collaboration' is often used as a term to describe all kinds of mutual social activity, and in terms of smaller groups of learners the term can be used loosely like this. However, collaboration is dependent on 'coordination' and 'cooperation' as foundational constructs. In smaller groups, this coordination or cooperation can be established through direct facilitation and simple organisation, for example setting a specific reading and dividing the cohort into groups of 4 to discuss the themes is so basic in a seminar group setting that it does not warrant elaboration. In the MOOC context, learners cannot be expected to dive straight into collaboration activities, which refers to the organisation of social activity when there is already a shared understanding of the concepts (Laurillard, 2009; Wertsch & Bivens, 1992).

Laurillard (2009) presents the conversational framework above as a way in which to challenge digital technologies to deliver a genuinely enhanced learning experience against what we already know about what it takes to learn, as reflected in the processes in the diagram. In response to this challenge, the CDT enhances learning across the discursive and reflective components on the right-hand side of the diagram (the peer-peer connections) and primarily across the communication channels of resource negotiation, individualisation and self-organisation as described in the need for cybernetic variety. It can do this through utilising a design paradigm called ‘Stigmergy’ which is able to explain how it simplifies and coordinates the large numbers of contributions into a simple visualisation.

3.5 Stigmergic design as a conceptual framework for orchestrating scale and sustainability

The concept of ‘stigmergy’ (from stigma: sign, and ergon: work; literally ‘signs that do work’) underpins the design process of the project, as this conceptual framework centres the interactions between an agent and a changing environment. Stigmergic systems not only function at scale, but actually perform better with more agents and in an expanding environment which creates emergent ‘swarm intelligence’ (Van Dyke Parunak, 2006).

The concept of ‘stigmergy’ states that the agent changes the environment, which in turn changes the potential interactions of future agents (Elliott, 2016). In a system designed on stigmergic principles, the system itself does not make value decisions, but it provides structure through which value can be discovered. Specifically, in terms of sociocultural learning at scale, a stigmergic design supports the potential for dialogue where increasing participation also increases the possible connections a learner could make. However, Dron (2006) cautions that one should employ stigmergic systems with care:

> It is relatively straightforward to build systems that employ stigmergy, but there is a much greater challenge to build systems that are likely to lead to the kind of structures that we would recognise as educationally useful…. Stigmergic systems are part of wider social environments and will develop to help learners only if their users are there to learn. It is thus important to consider not only the stigmergic systems themselves but the contexts in which they are used. (Dron, 2006, p. 9)

Therefore, it is important to examine the context, and as with any learning design, it may help some learners more than others. MOOCs are unique in that they are large-scale public courses; some learners do enrol to consume content, in line with the 'acquisition' model for learning, while others contribute and indeed learn through 'participation'. It is important for platforms to support both Anna Sfard's metaphorical accounts of learning preference and not to banish teaching by telling by only supporting a "trendy mix of constructivist, social-interactionist, and situationist approaches" (Sfard, 1998, p. 10). My work with the FutureLearn MOOC platform has consistently revealed that around 50% of participants in any given course will not contribute any comments, therefore will not add to the system 'stigmergically'. This does not rule out vicarious learning and re-emphasizes the importance of understanding context when designing new tools and approaches for learning: some learners will not value others' comments where others will see value, but choose not to contribute, and others still will actively contribute and moreover learn *through* that interactivity. ALL these behaviours are legitimate and course designs should not focus in on one type of learner behavior or design for a low ceiling of interaction. It is through design for participation that the CDT artefact and the proposed learning activities raises this ceiling and enables a form of sociocultural learning.

Dron (2006) gives an example of a system designed on stigmergic principles, called ‘Co-Find’. Co-Find is a *collectively* generated database of resources (note: it is not *collaboratively* generated as this would require a *shared understanding* rather than simply a *common purpose*); communication in Co-Find is indirect and mediated through signs which are left in the form of ‘folksonomies’ (Sturtz, 2004), or tags that are created democratically by all users and attached to resources in the environment. A similar system is used in the social bookmarking platform ‘de.li.cious’. The collaboratively generated tags include categorisation of resources into topics, and into other ‘qualities’ such as ‘good for beginners’ etc. The tagging system means that resources can be selectable at the time and in the context they are needed so learners can choose what is appropriate to their immediate needs. This generates an environment focused on personalized pedagogical value rather than reinforcing pre-defined structural preferences in the system.

Stigmergic learning environments and mass collaboration platforms are still in their infancy and recent examples have included the ‘citizen science’ form of crowdsourcing, such as ‘iSpot nature’, ‘Wikipedia’ editing, and the ‘peer teacher’ approach which crowdsources questions, explanations, ratings and other metadata.

Common to all these approaches is the systematic coordination of large-scale participation and this is why they are described as 'stigmergic' designs, however differences lie in the 'structure' of the problem being designed for: iSpot nature, 'peer teacher', and some citizen science initiatives have a well-defined problem space where new contributions add to an overarching (database) structure.

3.5.1 'Light' and 'Heavyweight' peer production systems

Haythornthwaite (2009) has developed a conceptual framework to explain models of participation, utilising concepts of 'light and heavyweight peer production', but there are key differences in the forms of participation she cites and the MOOC context. Lightweight peer production involves making a micro-contribution to a structured dataset, like the well-structured stigmergic systems cited above. In these cases, contributions are rewarded with points, which acts as a motivator for the community, and gamification type systems can be developed. The example of heavyweight peer production cited is the production of peer reviewed academic work, where the community is much smaller, each contribution is attributed, contributions require more effort to produce and are negotiated with peers through publications so there is a reciprocal interdependence. User generated data in MOOCs is less structured and requires interpretation, attribution and negotiation, as with heavyweight peer production, but the contributions are small in term of meaning value within a single comment 'at the level of concepts' in terms of its placement on a page of a FutureLearn course but are structured by the platform into a relational dataset, as with lightweight peer production. In a sense, it is like heavyweight contributions within a lightweight structure, and the problem becomes how to extract those ‘heavyweight’ qualities and make them useful. One option is that individual contributions could be scored like in the reward models of lightweight production, but a new problem emerges in the tension between the development of scoring schemas and the intended sociocultural pedagogy, which undermines the interpretive aspects at the expense of ‘metadata’. For example, if league tables were developed based upon number of posts, number of ‘liked’ posts, or number of posts which stem a conversation, each would presuppose a fixed idea of value and reify certain social practice, which in itself would change the nature of the participation, restrict diversity, suppress niche themes and therefore restrict open serendipitous encounters. If these are included into a design, the affordances and limitations of these systems needs to be acknowledged, as their effect would be to place value on some aspect or other of the sociomaterial practice which would create pre-defined biases that alter or reflect the overall system or institution, in line with earlier comments from Postman (1992; 1969) and McLuhan (1964). Technology is not neutral and a design decision pushing one direction will also pull in another!

Stackoverflow.com is an example of an online discussion forum for answering technical questions and has become a massive repository of information on a range of subjects; it has developed a reputation score based on ‘upvoting’ by the community. However, stackoverflow.com values a very structured kind of knowledge such that questions and answers are concisely articulated (there is always an ‘answer’!), and posts are aggressively edited to avoid duplication. Wikipedia is another example of a peer production system which employs strict rules for participation in order to build a coherent knowledge base. The rule-based systems on both these platforms actually suppress the free flowing and interpretive nature of conversation in favour of structured, well-defined knowledge, and contributors accept or rapidly learn these baseline standards. It is certainly an option for MOOC designers to develop features which mimic these systems, but the danger, as suggested above, is that learners would cease to contribute if they see their comments being edited, deleted, and downvoted. This approach would also seem to go against the interpretive and knowledge building nature of sociocultural learning (common to other distance learning pedagogy) which values the pathway from sharing individual experiences through to application of theory and eventually finding synthesis and new knowledge (Gunawardena et al., 1997). Indeed, an upvoting, rule-based method ironically reinforces a canonical epistemology and positivist pedagogy, albeit not from a single teacher to a lecture theatre, but from groups of experts. There is value in this approach, perhaps in certain types of courses, in terms of organising contributions, but it is not the approach taken in this project.

3.5.2 A visualisation to coordinate peer production in MOOCs

The CDT combines the corpus of commentary (retried in the form of a structured database table) with the freeform agency of the learner to make decisions as to what is meaningful to their immediate learning goals at a given time. The realtime visualisation of this type of analytic demonstrates a novel way to both support self-directed learning and embrace massive latent participation. As stated in the literature chapter, Buckingham-Shum and Ferguson (2012) have developed a framework of 5 categories for 'social learning analytics' which is predicated on ideas for informal learning and self-directed study. They conclude that these analytics need to be in the hands of the learners so that as they use them, they can become aware of themselves as intrinsically motivated agents. My work makes a contribution to knowledge through taking this approach, and it differs radically from other approaches to learning analytics which use complete datasets of transcript analysis (Kovanović et al., 2017), automated text analysis (Reich et al., 2014), or social network analysis (Wise & Cui, 2018) in order to give after-the-fact analysis of what has already happened in a given course. This is of limited value to students, although is useful for course designers and researchers to iterate their product.

The CDT is interactive; learners click on terms in the visualisation, filtering the corpus and altering the visualisation. Typically, by selecting 3 words, the corpus can be filtered from 10,000 comments down to fewer than a dozen comments which use all the selected keywords. This workflow also presents the learner with 3 different visualisations along the way, implicitly scaffolding related concepts. Learners interact with this social learning analytics tool to discover other learners and tap into novel funds of knowledge which are interesting and meaningful to them yet hidden to them by the major functional affordances of the platform. This links to the conclusions drawn by Fournier et al. (2011) from their analysis of PLENK2010 cMOOC:

> the use of tags in the Moodle environment would have been helpful in linking various contents across weeks, allowing participants to search for relevant content and to *see how they were connected to various content and people with similar interests* (Fournier et al., 2011 my emphasis).

The CDT development employs a similar strategy for aggregating content and people around affinity; the project follows the development of this tool both technically and pedagogically using a design-based research methodology to iterate the design and implementation of the plugin based on participant feedback and theoretical reflection.

3.6 Summary of the theoretical framework

Technology is never neutral; it always has an impact on the social practices and is often reflective of the institutions where it is designed (McLuhan, 1964; Postman, 1992), and this includes the learning and teaching environment and the authoritative structures (Biesta, 2005; Illich, 1971). This is often understated in the literature on ‘technology enhanced learning’ and technology is cast in a ‘instrumental’ or ‘essentialist’ role which ‘black boxes’ the impact of the technology itself (Bayne, 2015; Oliver, 2011, 2013) in order to focus on predefined ‘outcomes’. This also applies to learning at scale, where the literature is often focussed on analysis of after-the-fact behavioural or engagement analytics. In this sense, the efficiency of the MOOC (i.e. progressing students through to completion) is being prioritised over the negotiation of knowledge which demonstrates that the authority of the content providing institution and the dissemination of canonical knowledge is more important than resolving learners' authentic problems, a fundamental aspect of an active learning or sociocultural pedagogy (see for example Beetham & Sharpe, 2013; Conole, 2013).

By focusing on systems design or technology, Cultures of Participation (Fischer, 2011, 1998) proposes a framework for interactions which can guide how computer systems can be design in order to enable users to be active contributors, rather than 'couch potatoes' in the passive mode. The extension of this kind of analysis onto *learning systems* is not actually new, and analysis from a cybernetic perspective has been conducted on virtual learning environments which examines affordances across various dimensions pertinent to learning such as *resource negotiation, self-organisation, monitoring, adaptation and individualisation* (Britain & Liber, 2004). 'Traditional' VLEs are demonstrated to be lacking across these dimensions, especially when the content is pre-loaded upfront without the means to change or adapt according to the specific learning situation, a basic construct of the pedagogical model (the 'Conversational Framework'). It follows onto the FutureLearn platform which also does not perform flexibly across these dimensions: content is published upfront and learners are encouraged to progress through the steps. This is reflective of a passive culture of participation, on the lower levels of the framework.

Laurillard (2009) proposes the conversational framework as a pedagogical challenge to technological innovations, and suggests that each new technology should be evaluated against the lines of interaction on the conversational framework. The framework, however, does not account for large scale participation where the major functions for cognitive change happens through peer discussions, rather it leans more towards a 'teacher constructed learning environment'. In this sense the technology for the MOOC context needs to be evaluated not only on how it may support the scaffolds of the framework, but also how it is able to coordinate or simplify the large scale of participation (the dual problems of 'information' and 'participation' overload) to support these conversations. The design paradigm of 'stigmergy' (Elliott, 2016) is suggested as this means of support. However, the stigmergic framework needs to be employed in conjunction with the suggested teaching activities that can encourage cooperation between learners, embracing the diversity of interests and experience found in large scale learning and engineering increasing 'moments of serendipity'.

Stigmergy is a theory for design which states that individual contributions change the overall system in a way which alters potential interactions in the future. By referring to this framework, it is possible to 'manage' the scale of participation and offer alternative pathways through content than those which are laid out by the initial course designers and the underlying platform, increasing variety across the dimensions identified in the cybernetic account of learning environments and also enable social practice at 'Level 4' of the cultures of participation, where users of a system are able to ‘organise content’ and also ‘extend the range of the system’. Peer production is analysed through the lens of Haythornthwaite’s (2009) constructs of light and heavyweight peer production, and the CDT interactive visualisation combines individual agency with the structured form of commentary analytics. The resultant design enables learners to visualise the concepts being discussed in a non-linear fashion but does not prescribe how those concepts should be analysed or interpreted, therefore increasing variety in the system, simplifying the participation overload and remaining in line with interpretative sociocultural pedagogy.

The Methodology chapter continues to describe how the artefact was designed and implemented through a process of design-based research and through practical and technical considerations.

Chapter 4: Methodology

4.1 Introduction

This thesis follows the implementation of the Comment Discovery Tool (CDT) intervention through 3 iterations. As stated in the theoretical framing of this project, the design of new media and the actions and activities which it affords can be defined according to the Cultures of Participation framework. The aim is to develop new media and tools which encourage and support higher levels of participatory behaviours in the MOOC context, enhancing current practice where the media encourages a consumerist, passive culture. The following chapters of this document presents empirical evidence as some pieces which make up the complex phenomenon of what it is like to learn in a MOOC, using design as a research instrument to improve practice.

In this chapter, I explore the feature set of the FutureLearn platform in detail, describing and demonstrating its challenges for pedagogical sophistication through a cybernetic analysis, as have been eluded in the theory chapter. I describe in detail how I came about with the design and implementation of the CDT and refer to the theoretical precepts explained in the theory chapter. I also show alternative viable designs and explain why the ‘word cloud’ implementation was the chosen mediating artefact for the affordance of discovery.

There are multiple methodological challenges when designing pedagogical interventions and developing new media, compounded by the large, heterogeneous populations under study. MOOCs learners are only weakly connected to one other, and to a certain extent, the learning designs and task proposals too as their actions are entirely self-directed and internally motivated. They have diverse motivations for study, levels of expertise, expectations or preferences and we should expect contradictions in the data. The challenge for the analysis is to be pragmatic, for example when considering tool design (i.e. to what extent learners are perceiving designed-for and unexpected affordances), the relationship between the affordances and learning preferences, as defined by acquisition or participation behaviours and how the tools encourage practices directed at the community. As with any design, a push in one direction will result in a pull in the other; for example, iterating the design to make the interface more intuitive may result in restricting pedagogically useful modes of interaction and complicating the design for sophistication may switch learners off. It is down to the researcher to collect data which demonstrates how affordances are being perceived and balance this with the theoretical aims of the project (developing sociocultural learning designs which embrace scale) and make appropriate iterations. There is an ongoing cycle of data collection, researcher reflection and technical iteration and therefore this type of project suits a design-based research methodology.

4.2 Design-based research (DBR)

Design-based research (DBR) is by nature a practical methodology enacted in real world settings. Design science is different from natural sciences as it seeks to *engineer* solutions to make the world a better place, rather than explain the phenomena of the natural world:

> the natural sciences are concerned with how things are … Design on the other hand is concerned with how things ought to be (Simon, 1969, pp. 132–133).

Design practices can be simple or complex and often create blueprints rather than an end-product; design experiments may involve prototyping several ideas with the goal of expanding or reframing the problem and are susceptible to failure. Not all designs work but by starting out from a theoretically sound foundation, and using processes of feedback, reflection and iteration, the worst scenarios can be mitigated.

In educational research, DBR is a relatively new technique that attempts to bridge the gap between theory and practice. In recent years, the journals Educational Researcher (vol 32), Educational Psychologist (vol 39) and Journal of the Learning Sciences (vol 13) have all devoted entire issues to the problems of DBR. This methodology owes much to the pioneering work of Ann Brown, who states that DBR is engineering “interventions that not only work by recognizable standards, but are also based on theoretical descriptions that delineate why they work, and thus render them reliable and repeatable" (Brown, 1992, p. 143). It is therefore important that such interventions are based in theory, and that they occur in naturalistic settings, ensuring that the observed differences are meaningful. DBR is a methodological response to the understanding of learning as situated in a rich system of social relationships and technologies, which has become axiomatic for many educational researchers. Learning is personal, it is relative to prior cognitive abilities, experiences and knowledge and such a complex conceptualisation is resistant to experimental testing. Indeed, experimental designs that look at the learner, the environment or the learning activity in isolation will inevitably lead to an incomplete understanding of the full context. When new methods or tools for education are developed, it makes sense to examine these tools in context and to reflexively use participant evaluation and researcher investigation to make improvements to the design.

Cobb et al. (2001) describe the integration of design into educational research in terms of developing a theory, deriving principles which are translated into concrete practice and then assessment to evaluate the theory. However, this is an idealised version of discourse rather than an exact prescription and it may not be realised in practice. Others (Brown, 1992; Kelly & Lesh, 2000; Richey & Nelson, 1996; van den Akker et al., 1999) suggest that design itself can play a role in the development of a theory, and not just its evaluation. This project takes the latter perspective on DBR, as the learners using the intervention bring with them a multiplicity of learning expectations (with associated theoretical underpinnings) and 'perceived affordances' (of technology) which cannot be adequately accounted for by appraising the intervention against a single theoretical viewpoint which has been designed for in advance. Instead, the intervention aims to develop heuristics for practice, which are theoretically accounted for, and emerge through the process of implementation.

4.2.1 Teaching as a design science

Diana Laurillard uses the term 'design science' to describe her framework for learning design which is predicated on practical methods such as DBR:

> A design science uses and contributes to theoretical science, but it builds design principles rather than theories, and the heuristics of practice rather than explanations, although like both the sciences and the arts, it uses what has gone before as a platform or inspiration for what it creates. Teaching is more like a design science because it uses what is known about teaching to attain the goal of student learning, and uses the implementation of its designs to keep improving them. (Laurillard, 2012, p. 1)

Goodyear states that there are 3 main categories of teaching that can be designed, which follow on from the basic precepts of Vygotskian activity theory: (1) good learning tasks for scaffolding an epistemology of the subject, (2) supportive physical and digital environments as mediating artefacts and (3) the division of labour as social organisation (Goodyear, 2015). This project divides the implementation of the designed solution into 3 phases of DBR, each of which represents a dimension of activity theory as suggested above.

4.3 The planned phases of Design-based research

This project develops the intervention design across 3 phases (DBR1, DBR2, DBR3), with each phase concentrating on an aspect of Vygotskian activity theory, in line with Goodyear (2015). That is to say: DBR1 is the prototype phase which introduces a new mediating artefact into the digital environment; DBR2 investigates the learning exercise in terms of its relationship to previous activities conducted in the week, so concentrates on the pedagogical and learning design aspects (the task proposals); DBR3 brings these pieces together to investigate how learner behaviour can be driven towards a cooperative paradigm in terms of making their writing on all the pages of the course more discoverable by others *through* the mediating artefact, thus increasing potential for social interaction and collaboration.

All these phases take an activity centric view of learning, in that what the students *do* is consequential for their learning. In DBR1 they are encouraged to use the artefact and give feedback on what impact it had on their learning *prima facie*. Their initial interactions with the tool expose their expectations and the most commonly perceived affordances, although this is at the expense of proposing a specific learning task, directing the activity. In DBR2, the suggested learning activity is proposed, which is related to the overall learning design and the perceived affordances. In DBR3, the proposed activity is not limited to the exercise containing the mediating artefact but works in concert with the technological affordances to encourage formation of knowledge communities which can be exposed through the tool.

> What people do is often best understood as an interplay between structure and agency (Giddens, 1984; Archer, 2000, 2007). People's action can be thought of as a self-directed journeying through a pre-existing landscape. Sometimes the landscape is flat and affords walking in any direction one chooses. Sometimes, the tilt of the land keeps surreptitiously edging one's feet downhill. Sometimes one can only walk where safe paths have been trodden by others … On this view, one of the ways that teaching can take place is through shaping the landscape across which students walk. It involves the setting in place of epistemic, material and social structures that guide, but do not determine, what students do. There is a beguiling recursiveness to this conception. The aim is not just to shape landscapes (learning environments, if you prefer) that help students become more capable agents—it is not just about increasing their personal agency. While this is important, it is also important to help students read the landscape and learn to (re)shape landscapes for their own future activity, and for the activities of others, including for future learning. (Goodyear, 2015, p. 34)

4.3.1 Prototyping the technology (DBR1)

Phase 1 is the prototyping stage, where the new technology is introduced into the course as a novel ‘mediating artefact’. The impact of the introduction of the new technology is measured quantitatively in terms of changes to the conversational structures, and qualitative responses are also collected and abductively analysed to evaluate broadly how the tool is impacting learning as determined by increased opportunities for interaction. The affordances that are designed for are also evaluated as to whether they are indeed being perceived by the end users, and whether there are unexpected ‘perceived affordances’ for learners, which can be theoretically explained and integrated into further iterations.

The aim of phase 1 is to gather evidence around how the design affects learning value in the course design and as a test for whether the technology will work in practice. The quantitative heuristics I have developed for this phase compare conversations in courses with the intervention with courses that pre-date the intervention using measures developed to represent overt sociality in conversations such as the proxies of length (in terms of number of posts in a conversation), unique participants in a given conversation, and a taxonomy I have developed which represents turn taking, based on the ‘Initiation, Response, Feedback’ (IRF) sequence, commonly cited in studies on classroom dialogue (see for example Littleton & Mercer, 2013; Mercer, 2003, 2007, 2010). The taxonomy is based upon the affordances of the discussion features within FutureLearn and builds on work by Chua et al. (2017) where individual comments are classified by their position in the

IRF sequence. My taxonomy uses the cumulation of IRF to classify the whole conversational sequence as ‘Lone’, ‘Question and Answer’ or ‘Extended Social’. In this way the taxonomy can provide a sense of the levels of overt sociality in the course, without reading and interpreting all conversations. When these quantitative measures are used in concert, they can give an idea of whether the intervention is performing as per the design: simply put, are learners ‘discovering’ conversations which previously were invisible to them within the context of participation and information overload, which are identified as major problems in the social learning environment of MOOCs.

Learner feedback is also taken in the form of a short 15 question survey which ascertains learner preferences (for example, whether they value social learning), to what degree they found the intervention useful (in terms of connecting ideas, developing thinking or general discovery) and whether it is of positive value to the course overall. These results are analysed using a Spearman’s rank to correlate learner preferences or styles of learning with the appreciation of the new affordances. There is also a free text area where learners can write any other impressions or suggestions for the development of the CDT artefact which is analysed in terms of the impact on learning and interpreted through a theoretical lens.

4.3.2 The epistemic dimension/ learning activity proposal (DBR2)

The findings from DBR1 are analysed theoretically in terms of the affordances which are most-oft perceived, aiding the development of a pedagogical form and moving away from the CDT being understood in the language of technology only. I analyse the learner feedback in terms of the relationship between the cognitive and social processes that the mediating artefact encourages and I combine this with a theoretically informed analysis of the types of learning activities which are excluded by the basic platform affordances to further develop the pedagogical dimension of the new learning activity in terms of the increased opportunities for serendipitous interactions.

DBR encourages reflexive iteration, therefore the researcher's developing understanding of the learning context should combine with empirical analysis to inform the iterative improvement of the intervention. DBR2 examines the pedagogical patterns of the course design using Diana Laurillard's framework for teaching activities: *acquisition, inquiry, collaboration, discussion, practice and production* (Laurillard, 2012), building the discovery affordances of the CDT into a scaffolded inquiry based task. In this way, the CDT activity can be evaluated holistically across the whole course and observed in combination with the other types of learning activities on the course (i.e.understood as a coherent piece of the overall learning design). In this respect DBR2 sets out how the intervention relates to the emerging practice of learning design in MOOCs, and how it supports important cognitive processes such as investigation, discovery, reflection and collaboration.

I measure its effectiveness through semi-structured interviews with participants to understand how the CDT has affected their experiences of learning with their peers. DBR2 sets out ‘rules’ or ‘suggestions’ for engagement which support activities such as reflection, rumination and developing a network. This is scaffolded onto the course platform in a manner which supports serendipity and reflection; this phase integrates these ideas within an overall learning design to make the learning experience more holistic. In the following sections of this chapter, I detail how the FutureLearn platform without the CDT artefact encourages a pedagogy which limits models of participation towards the individual and therefore an acquisition metaphor dominant view of learning, in line with Fischer’s comments about passivity and ‘couch potatoes’ (Fischer, 1998).

4.3.3 Cooperative learning and participatory pedagogy (DBR3)

DBR3 builds on the findings of the previous phases: the usability and perceived affordances in DBR1; the learning strategies for reading and writing which emerge from the pedagogical focus of DBR2 and the suggestions for improvements collected throughout both phases. In this final phase of research, I make several changes to the technology to improve discoverability and integrate the CDT artefact into all the social activities on the course by instructing learners to write in a style which works in concert with the new affordances. In this respect, this phase relates to the community and division of labour aspects of Vygotskian activity theory as learners are asked to change their behaviour to support the functions of the artefact, improving discoverability, creating communities of 'ambient affliliation' (Zappavigna, 2011) which work for vicarious and participatory learners, and overall encouraging a new social approach to learning at scale.

I achieve this by making the suggestion that learners include tags in their comments, demarked with a '#' symbol, so the terms can be separated by the CDT and act as a pedagogical method for community formation, where participants can exhibit cooperative behaviours throughout the whole course (coalescing folksonomically around certain tags), rather than acting as individual consumers of content. This new social behaviour is in line with the 'cooperative' dimension of stigmergic design paradigm and also develops 'designers' in terms of cultures of participation within large heterogeneous cohorts.

As stated in the theory chapter, community formation and active participation are important to weave into learning and more so in mass scale learning because the social practices developed in the educational domain will follow through into other areas of social life and educational expectation. MOOC learning is a first experience with higher education institutions for many people and it is important to demonstrate that it is not merely consumption of content. Therefore, the learning design must set an expectation that community effort is valued such that learners are able to self-identify as an active agent within the crowd, and that their contributions are a valid and important element to the overall knowledge community.

4.4 An analysis of the FutureLearn platform

This section takes a detailed analysis of the FutureLearn platform. It is important to theorise the platform and its existing affordances from a systems perspective, as the new intervention is not stand-alone so adds to the overall platform; that is to say the intervention is designed to work alongside the existing workflows on the platform, and aims to add value by opening new pedagogical approaches which embrace the possibilities for engineering serendipity, which are *unique qualities* of large scale participatory environments, and in the self-directed nature of learning in a MOOC.

The analysis below briefly introduces the platform and continues to examine it from the point of view of both its in-built affordances and its usability design, which together make up the possible types of interactions and which of these are inherently encouraged. The affordances of the new intervention encourage opportunities for interaction across various dimensions of learning and participatory behaviours, but this cannot be considered separate from the rest of the course or platform.

4.4.1 Introducing the FutureLearn platform

FutureLearn is a private company wholly owned by the Open University (OU) who work with their partners such as many UK and international universities, the British Council, the British Library, the British Museum, and the National Film and Television School to publish courses on their own proprietary platform. FutureLearn has approximately 8 million learners, or “FutureLearners”. On the FutureLearn website’s ‘About us’ page it states:

> our courses are delivered one step at a time, and are accessible on mobile, tablet and desktop, so you can fit learning around your life. We believe learning should be an enjoyable, social experience, so our courses offer the opportunity to discuss what you’re learning with others as you go, helping you make fresh discoveries and form new ideas.[6]

FutureLearn’s design aligns mostly with the xMOOC model, in that it uses rich media and quizzing technologies to deliver content and monitor progression (Conole, 2016; Yuan et al., 2014). These are measured by completing a sequence of individual webpages called ‘steps’; however, it differs from other xMOOC platforms in that discussion areas are included on every step (Figure 5). This is a very specific design decision and is based around the ‘Conversational Framework’ (Ferguson & Sharples, 2014; Laurillard, 1993). As described in the theoretical framing of this project, the conversational framework embraces the notion that learning occurs through dialogue in a ‘teacher constructed learning environment’ between a ‘teacher conception’ and a ‘learner conception’. In this sense the peer-peer and tutors-peers’

[6] https://www.futurelearn.com/about-futurelearn (accessed 2nd September 2019)

conversations will always be about the concepts being explained on that page/ step, making the resultant discussions (in theory) more focussed:

> Each teaching element (or "step") is associated with a free-flowing discussion, which is intended to emulate a "watercooler conversation" about the immediate content. Each learner can see contributions from other learners and add a brief comment or reply. (Ferguson & Sharples, 2014, p. 102)

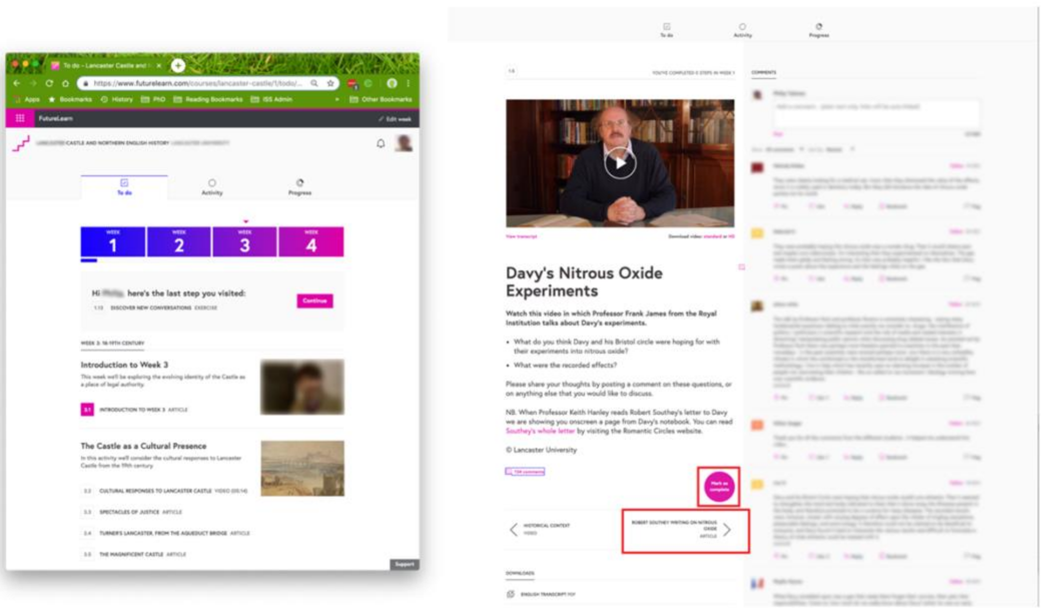


Figure 5 A FutureLearn 'week' complete with 'steps' (left), and a FutureLearn 'page' with inline comments (right)

4.4.2 The specific design of the FutureLearn platform

This section will discuss the design of the FutureLearn platform from the point of view of Usability and Human Computer Interaction (HCI), and link with a similar analysis on earlier e-learning management systems from the point of view of cybernetics. These analyses reveal an underlying tension in the design of the FutureLearn MOOC platform between its implementation of the conversational framework as its pedagogical model, and the affordances which have been designed for social learning which, as explained above, are on every page.

4.4.2.1 HCI and Usability

HCI is a process for designing a system which matches the needs and requirements of its users. This involves understanding the users and the technologies to match the interactional possibilities and interface design with the end users such that the system is appropriately deployed for its intended purpose. This can include external factors as diverse as health and safety when integrating a system into an organisational setting, such as a workplace. The language used by HCI researchers is very functional; it places great emphasis on understanding the wider context of a computing system and identifies interaction and interface as critical concepts. However, the design process is described such that users interact with machines to *further a goal or complete a task*. As stated by Head (1999):

> Good interface design is a reliable and effective intermediary, sending us the right cues so that **tasks get done** – regardless of how trivial, incidental, or artful the design might seem to be. (Head, 1999, p. 6, my emphasis)

If the ‘task’ at hand is ‘learning’, then it either requires a personalised set of interactions for each learner (increasing the complexity across the cybernetic dimensions of personalisation), or a narrower definition of ‘learning’, for example based around a homogeneous theoretical framework, such as behaviourism, or a systematic analysis of ‘learning analytics’. This is indeed the approach widely taken in instructional design methodologies such as ADDIE (Analysis, Design, Development, Implementation, and Evaluation) (Oh & Reeves, 2010) and the majority of the discourse on learning analytics in MOOCs (Fincham et al., 2019; Joksimović et al., 2018).

Usability is a *quality* of design, which assesses how easy a system is to use; in this sense it comes completely from the experience of the users and their perceptions of how straightforward it is to complete a task. Neilsen states that “usability is a necessary condition for survival. If a website is difficult to use, people **leave**” (Neilsen, 2012 his emphasis). He has developed a framework of 5 factors (Learnability, Efficiency, Memorability, Errors, Satisfaction) to determine how to test the usability of a user interface and suggests that these features are tested through observation of laypersons interacting with the system for pre-defined tasks. The most important concept in usability design, however, remains the *main function of the system* (and the ease with which this can be perceived and performed). Therefore, HCI and usability design need to be used in conjunction with each other; in the first instance focussed on HCI to ensure that a system is fit for the context of its use and appropriate for the chosen approach, then usability as an evaluation mechanism to ensure that it is effective for the users.

> HCI will assist designers, analysts and users to identify the system needs from text style, fonts, layout, graphics and colour, while usability will confirm if the system is efficient, effective, safe, utility, easy to learn, easy to remember, easy to use and to evaluate, practical visible and provide job satisfaction to the users. (Issa & Isaias, 2015, p. 19)

However, learning is necessarily situated in a rich system of social relationships and technologies. It is also relative to cognitive ability, personal experience and prior knowledge so therefore makes learning a complex conceptualisation that is resistant to the experimental testing seen in HCI and Usability. It is too complex to be reducible to a *goal* (in HCI terms) or *system function* (in usability terms) and needs to be personalised to any given user to meet the requirements of the viable systems model, described in the Cybernetic analysis. A more complex definition is required to describe the processes of an online learning system. Etienne Wenger states in 'Communities of Practice' that first and foremost, learning, nor activity or experience can be designed. They must be designed *for* (Wenger, 1998). This places emphasis on creating an environment which is *conducive* to activities and experiences which are related to learning. Cybernetics is a method for scientifically analysing systems in which complexity is outstanding and too important to be ignored, so an analysis from this perspective is useful when examining computer systems for learning.

Cybernetics has been used in the past to describe how this complexity plays out in the design of e-learning systems. If the FutureLearn platform is analysed through this lens, it can be seen to be less aligned with the conversational framework. In defence of their chosen designs it should be stated that any design is by definition a compromise between competing factors and the focus of a good design for Neilsen and Norman is the ease with which affordances are perceived (Neilsen, 2012). The FutureLearn website scores fairly highly in this respect as there are few affordances to learn which are presented within a clean and simple design, but this is at the cost of pedagogical sophistication in the social learning domain. My platform intervention creates a new affordance for the FutureLearn platform which could be analysed from an HCI perspective of how the affordance suits the context, and also from a usability perspective in terms of how well the new affordances are perceived by the users. This is necessary in order to make the new artefact appear as an environment where the activities and experiences conducive to learning are transparently perceived, and this is a necessary and important part of iterative improvement through the DBR process.

4.4.3 Analysis of FutureLearn platform from HCI/ Usability perspective

FutureLearn sequences weekly content into pages called ‘steps’. There are approximately 12 steps per week for a typical course, and learners are encouraged to click ‘Mark as Complete’ (3, Figure 6) on each step when they feel they are ready to move on. A sample ‘video step’ is displayed below (Figure 6).

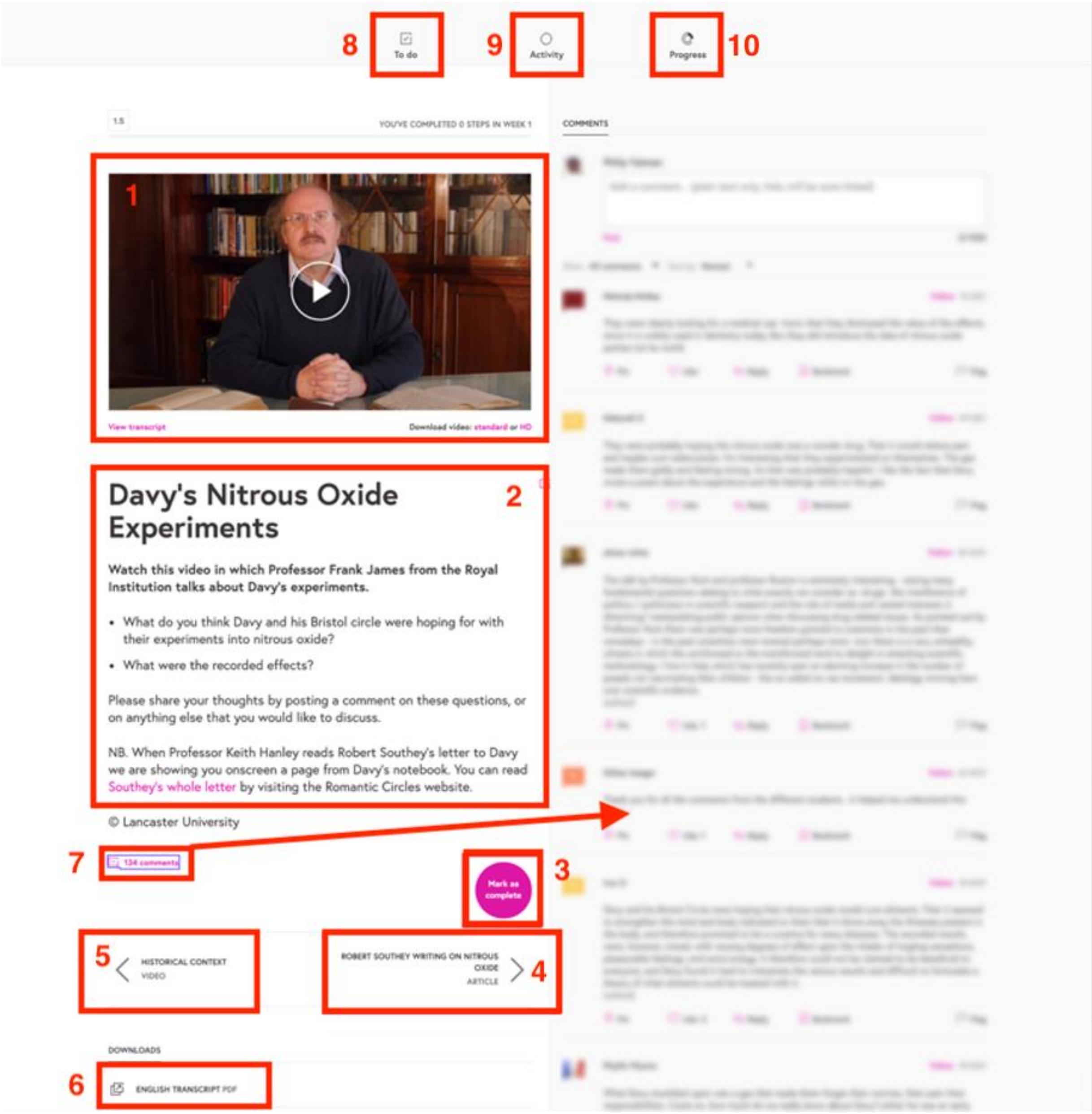


Figure 6 A sample FutureLearn 'video' step. The top 10 interactional opportunities are highlighted according to the breakdown above

The page is cleanly designed, the font is large and there are few areas for interaction. From an HCI perspective, the interface design is clearly sending the following cues as to what the 'goal of the task is' in terms of the prominence of each feature on the page:

1. Watch the video: *which is no longer than 6 minutes* (complete with subtitling and interactive transcript for jumping to specific parts)
2. Read the text: in a large, attractive, distinctive font (Europa), titling and subtitling made bold for emphasis
3. Mark this step complete
4. Go forward to the next step
5. Go back to the previous step
6. Download an associated file (if provided by the educator)
7. Open the comments thread for conversational possibilities (described below)

Other buttons on the page are not sending 'cues' as to what the task is, but are present nonetheless to aid other navigational options:

8. Go back to the weekly display
9. View all ongoing activity (arranged in a reverse chronological order)
10. View 'progress' page (showing what % of steps are 'completed')
11. Go to another FutureLearn course (the My Courses page)
12. Go to the FutureLearn homepage

The 12 opportunities for interaction designed into the interface are sequenced above in order of how much they stand out and annotated in Figure 6. By default, the comments thread is not shown, so an extra interaction (click) is required to engage with these affordances (cue 7). The largest buttons on the page are 'back one step', ‘forward one step’, and ‘mark complete’ and I interpret these as the main cues that FutureLearn wants users to make. This is not unusual for xMOOC platforms because the progression incentive is tied to certification sales which is a key part of the business model. There is also the button for viewing your own progress on a page with a large attractive graphic showing what percentage of the course you have ‘completed’ (cue 10), and this page includes information on what you need to do to qualify for a certificate and of course a link to purchase an ‘upgrade’ for ongoing access. In this sense, ‘business’ interactions are constraining pedagogical considerations.

It is likely that the 12 affordances above are designed with ‘usability’ in mind. Neilsen’s 5 categories for usability: *learnability, efficiency, memorability, errors, satisfaction* are well suited to the fewer interaction opportunities. The cues are easily learned, efficient, memorable, it is hard to make errors when there are few options for interaction, and the same options are repeated on every page. Remember also that Neilsen’s rules are to prevent users **leaving** a site and therefore not **buying** the product:

> On the Web, usability is a necessary condition for survival. If a website is difficult to use, people **leave**. If the homepage fails to clearly state what a company offers and what users can do on the site, people **leave**. If users get lost on a website, they **leave**. If a website's information is hard to read or doesn't answer users' key questions, they **leave**… The first law

> of ecommerce is that if users cannot *find* the product, they cannot *buy* it either. (Neilsen, 2012 his emphasis).

It is also notable that video content on FutureLearn is limited to 6 minutes in length (shorter videos are encouraged by their quality assurance process). This is where the average internet video user will *start* to lose interest; large scale studies in web marketing cite 5 minutes as a 'sweet spot' for optimum engagement metrics in web videos (Fisherman, 2016). It is noteworthy that these metrics are generalised for all internet video content (including 'lolcat' videos, music videos and general how-to videos). There are no such statistics for the 'learning' content of the 'lecture' type present in MOOCs or for viewers who are more 'engaged'. Therefore, the ceiling of 6 minutes seems to be a conservative approach towards maintaining engagement; it would be conservative even if discussing generic content, but also that it is not proven that higher level academic content will follow the same trajectory. The effect of this usability focus by FutureLearn is two-fold: firstly, the pedagogical variety is constrained by the limited number of functions which are not personalisable therefore treating users as homogeneous, and secondly, video content which goes in depth on a subject is required to be split across several steps, so the resulting discussion is either hyper-focussed to a particular segment of the content, or (more likely) fragmented across several steps. The latter is more likely because learners will comment when they feel they have something to say which could be at any point during the sequence. Arguably designers could place a 'discussion' page at the end of the pedagogical sequence (essentially another identical page, labelled "discussion") with questions that may encourage the meaningful conversations to happen there; in fact this would be just *another possible* area for conversation. Also, the sequential layout means it would not be visible to learners until they reach there by clicking the 'next step' button.

The comments are not threaded like on Coursera, edX or standard VLE platforms; rather they have a simple structure much like conversations on Twitter or Instagram and by default are ordered in batches of 100 in reverse chronological order. This makes them less 'noisy' which answers the 'information and participation overload' problem found in threaded forums (scale makes them 'knotty', dense and difficult to navigate), but ultimately FutureLearn *masks* the problems by restricting discovery. There are controls on the discussion section of the page to order by 'most liked', 'oldest' and 'newest' and to filter by 'bookmarked', 'following' and 'your comments'. The comments section on each page can grow to several thousand comments per step, paginated in lists of 100. The design is "intended to emulate a "watercooler conversation" about the immediate content" (Ferguson & Sharples, 2014) i.e. with those learners who are on the step *at the same time as you*, or who you have 'followed' previously, but these affordances push in one direction and pull in the opposite. It may be useful to be able to see who chronologically present with you, but this starves out the 'older' posts and indeed 'future' ones because the platform design (i.e. next/ previous buttons, and progress graphic) does not encourage you to return to a step when you have 'marked complete'. Also, whilst it may be useful to see posts from people you have 'followed' from previous steps, overuse of this filter constrains growing your network as the course progresses.

The problems with this design are that learners are encouraged to equate 'progression' with 'learning', where 'progression' is defined as an individual journey through a pre-ordered sequence of resources; there is no space for reflection, which is a key component in true cognitive change, in constructivist principles. There is also no allowance for increased variety in the system required by the huge cohort of learners with diverse learning needs; it would make sense to encourage learners to stop, reflect, discover and write, rather than consume media, make a short comment and move on. The affordances which are designed into the system which encourage the conversational model for learning (the comments stream) is ordered to value who is currently alongside you, rather than with whom in the crowd you share an interest or affinity. It requires a conscious break to the workflow to encourage engagement with those who are in a different 'time silo'. In this sense, chronological ordering is valued above affinity, driving the 'completion' and 'progression' business agenda. The bottom line is that users enter a 'bubble' of conversations and users which is hard to break out of using standard tools.

4.5 Development of the Comment Discovery Tool (CDT)

Sustained interaction is traditionally the marker of a successful sociocultural approach and this requires reflective activity which is not encouraged by the existing workflows designed into the platform. As I have stated previously, the CDT creates new affordances for learners such that they are able to navigate through the context of information and participation overload to discover new content and learners who may be invisible to them if they rely on the existing affordances and stay in their 'time silos'. A MOOC's defining feature is that it is learning at a scale not seen before in education, but its online platforms limit themselves to individualised consumption, rather than encouraging learners to define learning goals and engage in conversations and with ideas that are important satisfying them.

I make 3 assumptions about the nature of learning in a MOOC context in designing the CDT: firstly, that there exists information overload issue; secondly that there exists a participation overload issue and thirdly that the learner cohort is heterogeneous and self-directed, therefore self-selecting in terms of how to engage with new learners and content. The first and second problems are solved by using visualisation technologies for the user comments to enable learners to see the concepts that are relevant to their own learning goals and reflect on how the visualisation changes in response to their selections. I answer the third problem by allowing users to read and then join conversations based on their selections.

The tool has been designed to work in any FutureLearn course because it uses a layer of authentication called Learning Tools Interoperability 1.2 (LTI1.2) to integrate into the platform. LTI acts as a 'authentication launchpad' within the context of an 'exercise' step on the FutureLearn platform. When the learner clicks 'launch' from the web page, FutureLearn conducts authentication with the CDT and posts variables which the CDT consumes in order to visualise comments from the correct course. In this way, the CDT can be integrated into a range of FutureLearn courses and can afford the same features across the whole platform. Therefore, data is collected from a diverse range of learners and contexts, which strengthens conclusions about the types of learning it can encourage; this is useful for making claims about how the artefact relates to certain 'teaching types' within a learning design framework and that applies to all courses.

4.5.1 Technologies used by the CDT

During phase 1 of DBR, the CDT underwent a redesign in order that it had a more solid codebase from which to develop. It was initially developed using php and in the very first prototyping stage it was only possible to choose 2 words, although the main features (visualising 200 words, filtering after a selection, displaying 100 comments which use the selections) and use of d3.js and LTI1.2 libraries are consistent.

The CDT underwent a redevelopment using Python/ Django web framework which enabled further words to be selected and used the open-source Python 'natural language toolkit' (NLTK)[7] for language processing which placed the development on

[7] http://www.nltk.org/

a more solid codebase for future development as NLTK is a well-supported platform for language processing. The switch from php to Python was not entirely fluid however, and these technological challenges can be seen in the DBR1 data alongside the negative sentiments (i.e. learners are more likely to negatively appraise something if it doesn't work fluidly for them). These types of problems should be expected and accounted for when using a pragmatic and real-life methodology such as DBR. I describe below the main types of technologies which have been used in developing the CDT app and what potentials they could afford for future research.

4.5.1.1 Web Technologies

The CDT is integrated into the whole FutureLearn platform using Learning Tools Interoperability 1.2 (LTI1.2)[8] technology. As stated above, this means that the CDT can work across all courses and return results from a course specific corpus. This makes the artefact a piece of the learning design toolkit which can be deployed in concert with other artefacts and activities, rather than a niche tool for use in one course or another.

The tension between allowing all words to appear in the visualisation and restricting them to make it useful for a particular course is also a common theme across all the data, and as can be seen in the 'findings' chapters, learners frequently ask to remove

---

[8] https://www.imsglobal.org/specs/ltiv1p2

‘inessential’ words. However, the intention of this project is not to improve one course, but to introduce new toolsets at the level of the system affordances which can enable pedagogical techniques applicable to the entire system. In other words, I could have made the decision to remove words not directly relevant to the content of a particular course, but this could have affected other course topics, narrowing the range of affordances which can be perceived. For example, some learners used the cloud to click on verbs, such as ‘enjoy’, ‘disagree’, ‘love’ and ‘want’, because this narrowed the data to reveal other learners’ expectations or focus in on the areas of debate. Clearly this was not a designed affordance of the tool, and it took the ingenuity of users to devise these possibilities, which is in line with earlier comments from Buckingham-Shum and Ferguson (2012) around social learning analytics. The outcome of this tension was to design features which aggregate user generated tags, so creating a folksonomy which can be both course specific, and also introducing the possibility of designing the entire MOOC from the point of view of community participation, which is in line with comments from Stephen Downes, who suggests treating “a MOOC for what it is: a network” (Downes, 2013).

4.5.1.2 Language Processing Technologies

NLTK is a leading platform for building Python programs to work with human language data. It provides interfaces to over 50 corpora and lexical resources, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength natural language process libraries. The move between the php version of CDT and the

Python version implemented the tokenization library, to merge words with their capitalised versions, and to implement the stopwords library in order to exclude certain common words (such as “the”, “or”, “and”, “but”) from a standardised list, which has been agreed-upon by the NLTK community.

This CDT could afford a great deal more in terms of integrating specific corpus analysis techniques. For example, ‘noun phrases’ rather than single words could be processed and injected into the visualisation technology or words could be first analysed against log-likelihood in the British National Corpus (BNC) to establish ‘keyness’[9]. In practice, implementing this would require significantly more computing processing resource and add risk that the CDT runs either slowly or not at all with larger corpora, as comparisons with large corpora are notoriously processor hungry tasks. This would affect the user evaluations and make it more difficult to appraise the artefact against its theoretical precepts, as was seen in the technical glitch described above during the first phase and the transition to Python.

NLTK affords a significant amount of potential to the CDT, but each of these options needs to be first analysed carefully against the learning value for learners. They could

[9] http://www.thegrammarlab.com/?nor-portfolio=understanding-keyness. Keyness compares the frequency of a term against the British National Corpus, indicating its importance in the corpus in relation to typical usage across all language.

appear in future phases of DBR but this project examines how DBR can develop an intervention that affords new forms of behaviours in relation to a development of MOOC pedagogy; it is not a showcase for the potential of language parsing technology!

A limiting factor in integrating advanced NLTK libraries is that the more complex ones require a significant amount of computer processing prior to visualisation, so there is a balance between learner experience of the tool in terms of usability and amount of viable processing resource available. I have designed the CDT to provide a rapid, user-friendly experience to encourage take-up and feedback; for these reasons that NLTK is only used for word tokenization and stopword exclusion in this project.

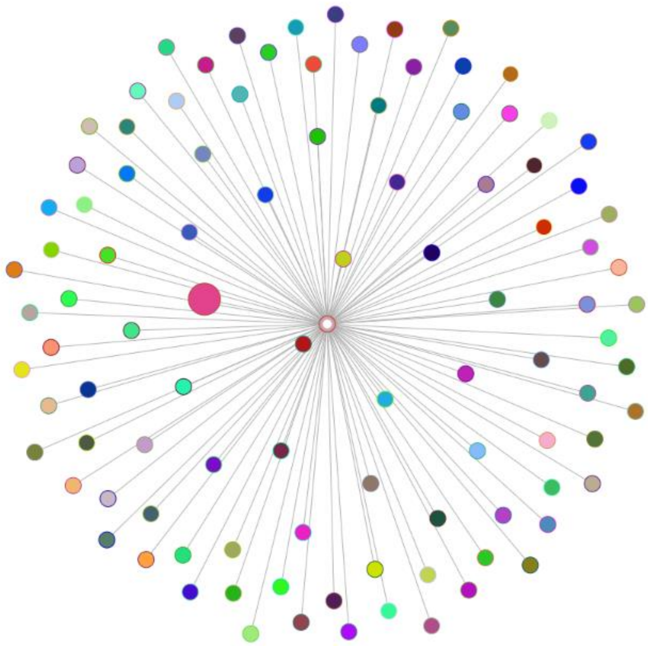
4.5.1.3 Visualisation Technologies

D3 is a javascript library for visualising data with html, SVG and CSS. I integrate the open source 'word cloud' project[10] into this project, which processes text data into a word cloud. The word cloud visualisation is chosen for the design as it is commonly used to visualise data and so should be familiar to participants (indeed several learners reported having seen word clouds before, but had not understood their value until presented with this interactive version). It is also intuitive to use and

[10] https://github.com/jasondavies/d3-cloud

importantly, it affords new methods of interacting with other learners, their comments and therefore their ideas.

D3 itself is a platform for visualising many different types of data. Another experiment with the MOOC comment corpus (which I developed offline) visualised user comments into a ‘force directed’ graph, where the nodes on the graph represent learners, and their proximity to each other are determined by the similarity of comments. Figure 7 below demonstrates this type of visualisation:



Comment Link	Comment Text
Link	Hello! I'm Arianna, I come from Italy but I'm currently living and working in Spain. I'm an Italian language teacher to adults and I'm taking this course because sometimes we have adult students with dyslexia and I'd like to learn how to help them learning foreign languages. Moreover, I'm interested in dyslexia and I would like to improve my knowledge and understanding of it.

Figure 7 A force-directed graph visualising comment similarity. The central node is the learner, other nodes are their peers, where their proximity represents similarity of comments. The larger node is the selected 'peer' and highlighted text below displays matched words

As with NLTK, D3.js visualisation technology is very powerful for manipulating data and demonstrating the power of raw computing for processing large amounts of text, but it comes with the same limitations. Processing all the comments in the corpus to determine a similarity factor between learners is very intensive, so it is not practical for a dynamic online system which could potentially have to process more than 20,000 comments for a single course, which could be up to 4 million words. This example of visualisation technology is also not as intuitive to use as the word cloud, and I made the judgement that the affordances of this artefact (matching users with each other based on shared words in their comments) wouldn't be immediately clear to the average user. That is not to say that from a theoretical perspective it doesn't make sense to match users, as it is also a method of handling participation overload and creating sub-communities of interest. Ultimately, in a DBR project such as this, the designer must make judgements and defend their choices using theory.

4.5.1.4 Overview of the development framework used in the development of the CDT

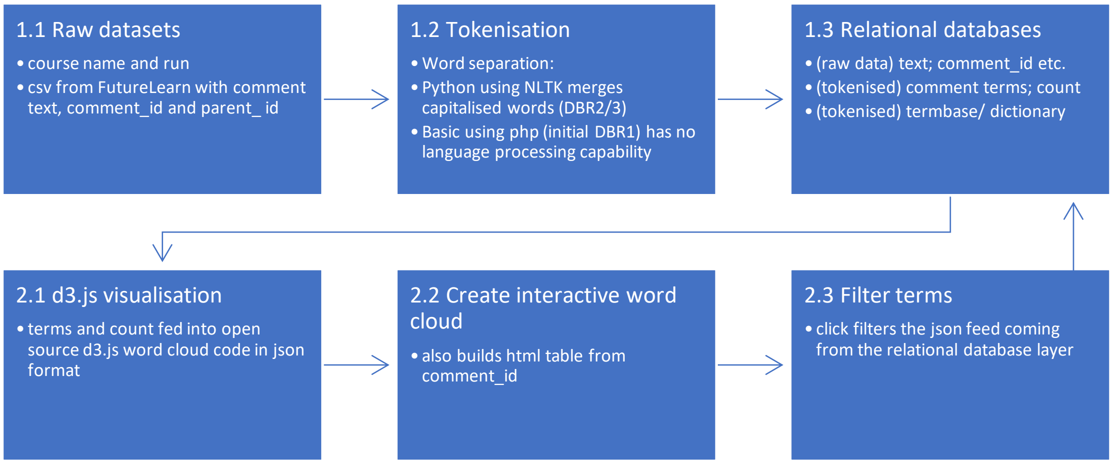


Figure 8 A schematic diagram of how the technologies and frameworks work together to create the CDT. The top row (1.1 – 1.3) signifies data structure and server processes and the second row (2.1 – 2.3) signifies the visualisation technologies used on the web page. LTI posts course name and course run parameters to the server on launch which creates the initial word cloud

The raw comments dataset from FutureLearn is uploaded to the CDT 'admin page' which creates a data structure that can be fed into the visualisation technologies (Figure 8). At first, I used php to simply break the 'text' column of the comment spreadsheet into composite words, but by utilising Python later in the development of the CDT I was able to use NLTK to merge capitalised words with their non-capitalised versions. Further development could use NLTK to find 'noun phrases' which could be used in a table and fed into the word cloud d3.js page for a different type of cloud. This demonstrates the extensible nature of Python as a technology for this type of language processing. LTI is not represented on Figure 8 because it simply posts initial parameters to the server to filter the comments to the ones written on the host course. It is not required after this and the cycle of filtering words for new clouds comes from users' word selections. Hashtag terms are treated as the same as other words, but are filtered out by the webpage before the terms are requested from the databases.

4.5.2 Ethical considerations arising from the development of the CDT

The major ethical considerations in the development of the CDT concern the user data and the informed consent. FutureLearn's terms of research ethics[11] states that participants are informed that activities may be monitored for research purposes on sign up to the site, so opt-in consent from each participant is not required. Regarding users' comments on the site, these are considered to be subject to a Creative Commons Licence (Attribution-Non Commercial-NoDerivs; BY-NC-ND). It is also submitted to the site on the understanding that there is an "irrevocable, worldwide, perpetual, royalty-free and non-exclusive licence to use, distribute, reproduce, modify, adapt, publicly perform and publicly display or otherwise such Learner

[11] https://www.futurelearn.com/info/terms/research-ethics-for-futurelearn

Content"[12] . The CDT is not able to attribute commentary directly within the application itself because the attribution is not made available in the FutureLearn datasets (a condition of the research ethics) so whilst it reproduces content which has a CC BY-NC-ND licence without the attribution, this is mitigated by the fact that the commentary is only made available within the host course itself and so is merely a different representation of data that otherwise would have been available to other learners if they had read through all the comments posted on that course. Furthermore the decision to licence users' content in this way was intended to give them agency over their own commentary and if they choose to remove their commentary from the site itself, it is also removed from the CDT on the next dataset upload. Other ethical issues arising from this research, such as the feedback of user experience and participation in surveys does include specific participant information and consent forms prior to participation.

4.6 Summary of the methodological approaches to the project

I have identified that there is a gap in the literature which relates to the platform design of MOOCs and have connected this to the commonly cited problem that MOOCs, especially xMOOCs, serve a didactic pedagogy, and are solely framed in behaviourist and instructivist theory (Bali, 2014; Downes, 2012). This means that from a systems perspective, MOOCs offer a limited menu for pedagogical

[12] https://www.futurelearn.com/info/terms (point 6.3)

sophistication and create a space for new research into tools and methods which are able to create more opportunities for interaction. In order to follow on from recent trends in learning design and sociocultural pedagogy, learning is framed as a process of discovery, through scaffolded interactions with peers and tutors (Beetham & Sharpe, 2013; Conole, 2013; Haythornthwaite & Andrews, 2011; Laurillard, 1993) and systems are needed to offer opportunities for these types of interactions. It also needs to align with developments in 'learning design' which seek to encourage a social pedagogy through scaffolded interactions. Bayne and Ross (2014) encourage us to think about MOOC pedagogy less in terms of the cMOOC/ xMOOC binaries but on a micro level, in terms of individual learning design. This includes the mediating artefacts which support the interactions designed for.

I have analysed the FutureLearn platform in terms of its scope for social interactions and concluded that temporality is valued, in terms of surfacing *recent* comments and encouraging users to stay in time-based cohorts. The system's cues are to keep moving forward to the next page. This values personal course 'progress' over taking time to explore, inquire and reflect on others' thoughts. Ultimately, this values *your own* comments because you are encouraged to make a short comment on every page, rather than engage in depth with learners who are interested in the same thing. This does not support a pedagogy which values community development or seeks to engineer serendipitous encounters.

I have designed a platform intervention which attempts to take the valuable parts of learning at scale, such as the diversity and increased potential for serendipitous encounters, combined this with the self-directed nature of learning and individualised measures of success, as supported by Liyanagunawardena et al. (2017). This 'mediating artefact' is introduced into the platform in 3 iterations of DBR, with each iteration moving the concept forward, allowing time to take participant feedback and reflect on the previous phase. Design-based research has enabled this narrative to emerge through its focus on theoretical premises, naturalistic settings, participant feedback and researcher reflexivity in between the phases.

Chapter 5: Findings from DBR1

5.1 Introduction

The methodology chapter describes how the research was conducted with 3 iterations of DBR. In this first phase of DBR, the intervention is developed as 'prototype' and measured against practical factors such as technical viability as well as through a theoretical lens in terms of identifying the gaps between designed for and perceived affordances. The overall corpus of peer produced content is quantitatively examined using an ANOVA comparison of means, demonstrating statistical significance on the dimensions of conversation length and unique contributors to a conversation. An extra heuristic framework is also developed which measures turn taking in conversations, building on work by Chua et al. (2017) by extending their taxonomy onto whole conversations, and combining with measures based on the 'initiation, feedback, response' framework which has previously been used to analyse classroom interactions (see for example Littleton & Mercer, 2013). These measurements can be viewed in combination with the 'unique participants' attribute to develop a heuristic taxonomy for describing all conversations on the platform which adds richness to the simple comparisons of length or unique members as per the ANOVA. These data are used comparatively with courses which did not use the intervention and as with the pilot study are able to provide a macro level analysis determining the intervention's impact and can answer the first research question of the design-based research:

DBR1-RQ1: What are the changes to the conversational structures in the platform when the visualisation intervention is deployed and are these changes statistically significant when compared to courses without the intervention?

This analysis can measure the change in overall patterns of overt behaviour but it cannot answer other questions regarding learners' experiences of the activity or how it should be integrated into the existing learning design of the courses. This is a reformulation of the overarching question of how learners are interacting with both the standard discussion features and with this tool as an extension of those features as part of their self-directed study. The major narrative of this thesis explores the complexities around this question and seeks to understand how the CDT can be used to enrich the social experience when learning at scale.

The foundation to this narrative is set using a survey instrument which can produce quantitative and qualitative analyses and the learner responses on the specific page of the course where the CDT is located. On the quantitative dimension, a Spearman's rank correlation analysis is used to identify correlations between learner preferences with regard to social learning, perceived value of the tool, and ease of which the designed affordances are perceived (i.e. basic usability of the tool). Therefore, the second research question of the DBR1 phase is:

DBR1-RQ2: What is the correlation between learner preferences and perceiving the designed affordances of the new comment discovery visualisation tool?

On the qualitative dimension, the survey collects free text learner responses which are analysed in terms of how the learning experience has been specifically affected by the tool, how it was used in terms of designed affordances, unexpected perceived affordances, and suggestions for improvement in the subsequent iterations of DBR. The comments dataset on the CDT page is considered as an extra open text response in this analysis. These responses inform the next stage of DBR which integrates the CDT artefact into the course structure in such a way that the designed affordances (and suggested uses of the tool) are more apparent and the suggested learning activity more concrete. These survey responses alone cannot reveal rich descriptions of learners' experiences and specifically their use of the discussion features as this would be sought by an interview method, but it is able to collect a wide array of ideas and suggestions which can be considered for implementation in the next phase of DBR.

DBR1-RQ3: How do learners' free text responses relate to the conceptual frameworks detailed in the theory chapter, and how should the learning activity be scaffolded into the course in order to more the affordances more perceptible?

5.2 Quantitative analysis

Tubman et al. (2016) use conversation length as the sole indicator of sociocultural knowledge construction, which is useful for analysing basic conversational structures but further dimensions should be considered to represent conversational learning more holistically. Henri (1992) highlights 5 dimensions which describe the holistic nature of sociocultural learning: *Participative; Interactive; Social; Cognitive; Meta-cognitive*. The conversation length proxy indicates basic interactivity and comment counts indicate basic participation, but these dimensions can be represented with more precision by factoring turn taking within a conversation. Furthermore, the social dimension can begin to be represented by counting unique participants in a conversation, indicating the diversity of opinions and experiences (therefore wider funds of knowledge) in the conversational unit.

The FutureLearn comments stream lends itself to operationalising these dimensions quantitatively as the conversations are not 'threaded' in terms of supporting 'replies to replies', which is the likely cause of the 'thread density' problems reported in the Coursera platform by Brinton (2014). This means that each comment within a conversation can be associated with the initial post or original poster and turn taking can be measured in terms of the same relationship.

5.2.1 A new taxonomy for turn taking and unique participants attributes of a conversational unit

Chua et al. (2017) propose a taxonomy of comments based on the affordances of the FutureLearn platform in which each comment on the platform can only be one of 5 categories (*initial, lone, first reply, further reply, initiator reply*). I extend these categories onto the whole conversation in order to place focus on the turn taking necessary for 'going further' in a conversation, which is supported by other frameworks used in 'classroom talk' such as the IRF, or Initiator, Response, Feedback framework (Littleton & Mercer, 2013; Mercer, 1995, 2007, 2010). In terms of the FutureLearn platform, this taxonomy for conversational units can be represented by Table 1, below:

	Initial Post (IP)	**First Reply (FR)**	**Further Reply (FurR)**	**Initiator First Reply (IR)**	**Initiator Further Reply (IFurR)**	**Heuristic Type**
1	✓	✓	✓	✓	✓	Extended Social
2	✓	✓	✓	✓		Extended Social
3	✓	✓	✓			Extended Social
4	✓	✓				Q&A
5	✓	✓		✓		Limited Social
6	✓	✓		✓	✓	Extended Social
7a	✓			✓		Lone
7b	✓			✓	✓	Lone
7c	✓					Lone

Table 1 The complete breakdown of possible conversational structures afforded by the FutureLearn platform with 'heuristic groupings' which provide an indication as to what level of knowledge construction is implicit in the conversation

Each conversation on the platform necessarily falls into one of the 9 types details above because of the nature of the possible conversational structures afforded by the system. This taxonomy is useful because the ability to quantitatively categorise every conversation into one of the types above demonstrates certain qualities about the conversation which adds more rich detail than attributes such as length. For example, in type 3 the initial post is 'hijacked' by others as the initial contributor never returns to share their summary or further opinion based on others' contributions, and in type 6 the initiator is the only contributor 'going further' so perhaps is fielding several questions from different learners. Using these heuristic groupings, it is possible to forecast what types of knowledge construction may be occurring in these units, as well as providing a framework for highlighting conversations worthy of detailed study using a more detailed content analysis instrument.

When the 'turn taking' dimension is combined with the conversation length attribute, it is possible to derive the levels of sociality. For example, if we only know that a conversation has 10 replies, we do not know whether all those replies are from the initiator (type 6), from other learners (type 3) or indeed with no input from peers (type 7b). In this respect, the taxonomy fills in some detail about the main contributors, and the quality of initiation, feedback, and response.

When the 'turn taking' dimension is combined with the unique contributors attribute of a conversation this can represent the social dimension as described in Henri's framework, and also indicate the range of diverse opinions which is important for a sociocultural understanding of conversational learning.

5.2.2 From full taxonomy to heuristic modelling

The turn taking dimension is further reduced to 4 categories, so that it can be used more effectively in combination with the unique participants taxonomy as described below. These 2 factors in combination can demonstrate levels of diversity in conversations, as well as 'going further', so can quickly measure more factors pertinent to sociocultural learning.

The more generic categories are labelled "*Lone*" (no turn taking or conversations 'with oneself'), "*Q&A*" (a single comment and a single reply with no 'going further' by any either the initial poster, or any repliers), "*Limited Social*" (where the initiator comes back but cannot trigger more 'going further') and "*Extended Social*" (to describe all conversations where either the initiator, the repliers, or both go further). Extended Social is where conversations indicative of social learning is likely to be discovered, but as is seen in the results (below) these conversations only make up a small fraction of the total, which further highlights the findings of Tubman et al. (2016), namely that participation is high, but levels of sustained interactions are much lower.

Participation in conversations is divided into 3 distinct groups in the heuristic framework and there is scope to expand upwards. I have given each category a 'nickname' to distinguish them from each other. These are:

1. Conversations with 1 member ("*Lone*"), a single posting, or a reply to self;
2. Conversations with 2 members ("*Watercooler*"), borrowed from the 'watercooler conversation' quote from Sharples and Ferguson and represents 2 people standing over a watercooler. These conversations may be a single post and reply but may also display 2 people 'going further' into an extended sustained interaction;
3. Conversations with 3-9 members ("*Cocktail Party*") represents a small grouping of learners who are likely to be able to bring a diverse range of experiences to the conversation and may be able to construct a shared understanding to develop. This is in line with traditional ideas of knowledge construction and models how groupwork activities are usually coordinated.

As with the social dimension, these categories can represent all conversations on the course, with no exceptions, so are useful for macro level quantitative analysis. It also allows further categories to be established, for example "*Conference*" for 10+ members, and perhaps "*Festival*" for multiple dozen members. The 3 categories above emerge from actual data in DBR phase 1, in that 1 member, 2 members and 3-9 members conversations are clearly defined whilst 10+ members do occur but are outliers (>0.05% of all conversations).

5.2.3 A working example of heuristic modelling

As a demonstration of the utility of the heuristic model, I have created a table with counts of all the conversations in the Dyslexia and Foreign Language Teaching course which can be seen in Table 2.

	Lone (1)	Watercooler (2)	Cocktail Party (3-9)	Conference (10+)
Lone	30792			
QA		2986	650	
Limited Social		553	170	
Extended Social		175	313	1

Table 2 Counts of conversations by type in the Dyslexia and Foreign Language Teaching MOOC

These descriptive statistics demonstrate that most social activity are ‘Lone’ or single comments (30792); they could also be longer threads but only containing self-replies. Most conversations with 3-9 members are actually QA (650). This means that no individual learner in that conversational unit is ‘going further’ in the conversation, so it is likely that these types of conversation constitute an original post, followed by several learners adding a single reply, and nobody returning to the conversation to engage or attempting to synthesise the content. Only 489 (313 + 175 + 1) conversations count as ‘extended social’ which demonstrate a degree of ‘going further’ in a sustained interaction by either the original poster or the repliers, which is approximately 1% of the total comments. However, the majority (313) of these are conducted in groups of 3-9 people (“Cocktail Party”) which demonstrates that there is evidence of interactive writing with a diversity of opinions present in the data.

Therefore, the heuristic model can describe the form of the interactive writing to a certain level of detail across a scale which is not practicable manually.

The limitation of the quantitative heuristic model is that it assumes that conversational units are linked through their 'reply' structure, so does not account for adjacent 'Lone' posts which are related in content, but not through the structure form of the discussion tool. It also does not adequately describe the potential levels of diversity from the point of view of a learner who uses the comments thread to read a range of opinions and then writes a single/ Lone post which is informed by the reading. However, it would not be possible to quantitatively model this type of behaviour, neither would that type of behaviour relate to previous modelling of knowledge construction, predicated on sustained interactions. These behaviours can be uncovered through a qualitative method and in DBR1 these data are presented as survey responses and free text remarks on the CDT 'step'.

The following study uses this quantitative framework to compare levels of conversations between courses which have the CDT deployed, and those which do not, also conducting an ANOVA analysis on conversation length and unique members. This study is a first attempt to demonstrate the impact of the CDT intervention on the level of the whole course, and to provide an evidential basis for future developments.

5.3 Studies to measure the impact of the CDT artefact

5.3.1 Introduction to study 1

The CDT has been designed in this phase to occupy a single step in week 1 of the FutureLearn course. On its most basic level it is designed to allow learners to filter comments by the words used in them (indicating thematic content) but on a more sophisticated and reflective level, it also displays connections between ideas and could be perceived as a 'concept map' of related terms. I have made the assumption that the filtering features are self-evident, although the description of the tool does not explicitly scaffold any suggested ways of using it which presents a risk that some learners will not able to begin with the exercise without explicit instruction, but in order to harvest more use cases without biasing the ways of using the visualisation, I have considered this an acceptable risk for this stage of the project.

This approach means that I am encouraging learners to make their own meanings for the technology, and this is deliberate to harvest a wide range of opinions and descriptions of what learners do with the tool. I consider this important at this stage of the project because although there are some key affordances which are 'designed', such as the ability to filter and join different conversations, there may also be affordances which are 'perceived', and the rich descriptions of how learners approach the tool will be useful in order to separate these different uses.

Learners are simply prompted to click on words and see what happens, then report feedback through the page comments. These comments are analysed later in this chapter in terms of how learners respond to the CDT, how they are using it, what this might mean in terms of learning and what suggestions are made for how it could be improved in the next phase. A total of 590 comments are gathered and analysed against these criteria over the course of DBR phase 1.

It is especially important to note in this phase of DBR, the tool is introduced in the first week only, and it is introduced to learners as a new technology without prescribing how it should be used, or what value to expect. It is simply described to learners thus:

> *The tool searches ALL comments from this course and visualises individual words in the form of a word cloud. The most frequently used words are larger.*
>
> *By clicking on a word, the tool will list the most recent 100 comments which use that word and you will be able to select a second word to further filter those results to comments which use both your selected words.*

The choice to position the tool on a single 'step' in week one is driven from practical and research related reasons and common to any design project, this choice creates its own limitations. On a pragmatic level, any new technology runs the risk that it simply will not function as expected, or that for some unforeseen technical reason it may break. These risks are compounded when developing a tool that is formed from peer produced data, as it requires many contributions to function properly, so testing cannot be done locally on a small scale. Therefore, the technology itself is a 'production beta', with appropriate steps that mitigate the worst possible outcomes. An example of the CDT step itself and its position within the weekly structure is displayed in Figure 9:

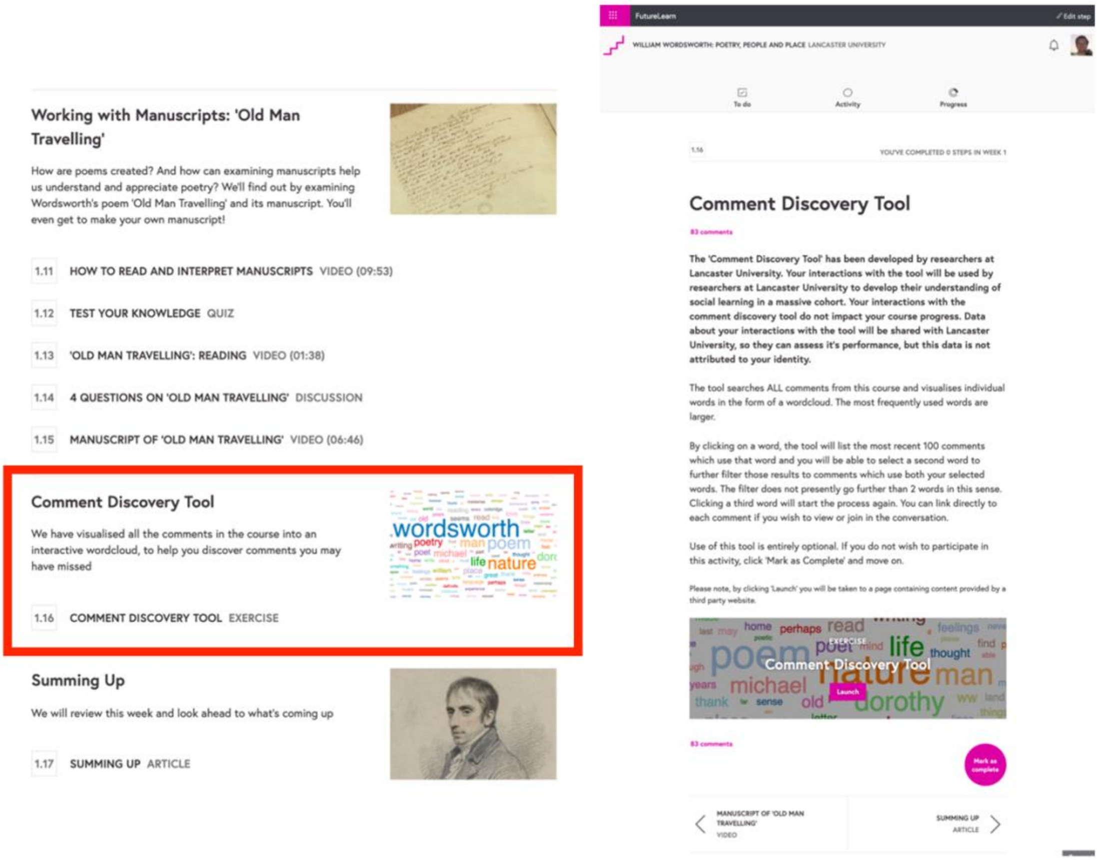


Figure 9 Position of the CDT activity on the weekly view of the course page (L), and the step itself (R)

At the level of research, the major aim of the DBR1/ prototype stage is to collect a wide variety of opinions on how it is used in practice, to further develop the CDT and explicitly scaffold the most perceived and beneficial for learning affordances into a learning activity. The choice of positioning in the first week (and linked back from other weeks) is to create maximum impact and visibility; it is well reported in the literature that attrition in MOOCs is high so if it were positioned in the final week then fewer learners would encounter the step and have the opportunity to feedback on the impact for their learning.

However, this creates a limitation because the corpus of peer produced content is smaller in the first week, compared to the final week, so learners who do not follow the link-backs in subsequent weeks will only encounter the tool once. I have already suggested that the 'stepped' design of the platform, and its visual cues to keep moving forward and 'complete' go against the idea of returning to an already completed step or re-engaging with an 'already completed' activity. For these reasons, in this phase of DBR, the technology can only be appraised from a point of view of 'mediating artefact' or tool, in terms of the perception of the designed affordances, rather than as an activity which is finely tuned in terms of overall learning design of the course.

5.3.2 Comparative studies on different instances of a single course: quantitative analysis and heuristic modelling

The first comparative study (Tubman et al., 2018) looked at 2 versions of the same course (William Wordsworth: Poetry, People and Place) where only one version had the CDT exercise step included.

This study investigated the impact of the CDT intervention which was embedded into the platform on the 2nd run. The total units of analysis are: 10,515 conversational threads (6005 non-intervention/ 4510 with intervention) which are analysed for length and unique participants with an ANOVA comparison of means test, and a social dimension is inferred from the turn taking in each conversation in line with the taxonomy based on the platform affordances, and modelled according to the heuristic described above. The research questions for this study are:

1. *Does the CDT have a statistically significant impact on the length and unique members of conversations?*
2. *How are these conversations modelled using the heuristic measure?*

5.3.2.1 Results quantitative study 1

Does the CDT have a statistically significant impact on the length and unique members of conversations across 2 instances of the same course?

10,515 conversations were analysed. An ANOVA analysis showed that the unique learners variable was significant, $F(1, 10513)=116.47$, $p=0.00$, and also that the conversation length variable was significant, $F(1, 10513)=87.82$, $p=0.00$. Cohen's d scores were also calculated for a measurement of impact, and generated a score of 0.21 for unique learners, 0.18 for conversation length. This suggests the CDT has had a modest impact in its first iteration. The descriptive statistics of these calculations can be found in Table 3:

	Course run	*N*	*M*	*SD*
Unique Learners	1 (no CDT)	6005	1.35	0.803
	2 (CDT)	4510	1.55	1.054
Conversation Length	1 (no CDT)	6005	1.51	1.36
	2 (CDT)	4510	1.80	1.88

Table 3 Descriptive statistics in DBR1, quantitative study 1

How are these conversations modelled using the heuristic measure?

These breakdowns demonstrate that the Run 2 (with the CDT) has a larger proportion of the heuristic groupings associated with higher levels of social constructivist learning: extended social conversation, conversations with more members, and fewer Lone conversations (seen in Table 4). This turn taking dimension is not taken into account by the length variable on the ANOVA test, which does not look at whether people 'go further' or whether Lone posts are discovered and replied to only once (i.e. the increase in Q&A conversation types).

		Run 1	Run 2	Difference
Social dimension	Lone	76.65%	69.00%	- 7.65%
	Q&A	15.89%	19.65%	+3.76%
	Limited Social	3.78%	4.83%	+1.05%
	Extended Social	3.68%	6.30%	+2.62%
Unique participants	Lone	76.65%	69.00%	-7.65%
	Watercooler	16.18%	17.20%	+1.02%
	Cocktail Party	7.16%	13.54%	+6.38%
	Conference	-	0.02%	+0.02%

Table 4 Percentage of conversations in each heuristic grouping; DBR1, quantitative study 1

5.3.2.2 Discussion of the initial quantitative results

These results are encouraging for demonstrating the success of the CDT intervention, given the limitation that the tool was embedded in a single step in the first week. There is a statistically significant difference between the 2 datasets across the dimensions of conversation length and unique members of a conversational unit, which may indicate that some learners are firstly perceiving the affordance of linking directly to new conversations through the CDT, then discovering and contributing to threads which occur outside of their 'time-based' cohort of learners, therefore using the tool as a strategy for choosing conversations to engage with and coping with information overload. Recall my analysis of the discussion tools in the methodology chapter presenting the idea there is a deficit of cybernetic and pedagogical variety at the level of resource negotiation, self-organisation and individualisation because learners are presented with the most recently posted comments *prima facie*, with the opportunity to use in-built platform affordances such as the 'most liked' and 'following' filter. This creates silos of information, primarily based on time, and the results of this study suggest that this new affordance could be making a statistically significant difference on levels of overt sociality.

The heuristic modelling of the conversations is supportive of this conclusion as it demonstrates that there is a decrease in 'Lone' conversations, and an increase in 'QA' and 'Extended Social' ones, supporting the idea that the tool is increasing the opportunities for interaction by leading learners to conversations outside their time-based silo of information and making an impact in terms of subsequent interactive writing. Further evidence to support this claim is demonstrated by the unique participants measure because the largest increase comes in terms of 'cocktail party' conversations, providing evidence for the claim that use of the CDT encourages learners to join in with existing conversation, thus increasing the diversity of opinions and experience. Both these results demonstrate a difference in measurements related to ideas of knowledge construction such as the participative, interactive and social dimensions as referred to by Henri (1992).

However, this analysis cannot be comprehensive because it is only based on 2 versions of the same course, which may share several similarities such as teaching style, content or demographic of the learners. In order to demonstrate efficacy as a general platform tool these results will need to be repeated across several courses which do not share these commonalities. Therefore, the following section presents results from a study where the CDT was implemented in 8 further courses in the same manner over the course of an academic year (September 2017 – July 2018) in order to test for repeatability, which is a key part of conducting design-based research:

> [engineering] interventions that not only work by recognizable standards, but are also based on theoretical descriptions that delineate why they work, and thus render them reliable and repeatable (Brown, 1992, p. 143).

5.3.3 Extending the comparative studies: quantitative study 2

I deployed the intervention in 9 courses and analysed all the 31621 conversations which occurred in these 9 intervention-enabled courses using the same method described above (ANOVA and heuristic modelling based) (Tubman, Benachour, et al., 2019; Tubman, Oztok, et al., 2019). These were compared with 26 courses which had previously been published prior to the development of the CDT consisting of 225618 conversations and asked the same research questions as above:

1. *Does the CDT have a statistically significant impact on the length and unique members of conversations?*
2. *How are these conversations modelled using the heuristic measure?*

5.3.3.1 Results

Does the CDT have a statistically significant impact on the length and unique members of conversations across a range of different courses and instances?

257239 conversations were analysed. An ANOVA analysis showed that the unique learners variable was significant, $F(1, 257239)=496.265$, $p=0.001$, and also that the conversation length variable was significant, $F(1, 257239)=601.703$, $p=0.001$. Cohen's d scores were also calculated for a measurement of impact, and generated a score of 0.15 for unique learners, 0.12 for conversation length. This suggests the CDT has had a small but noticeable impact across the courses in DBR phase 1. The descriptive statistics can be seen in Table 5.

	Courses (n=35)	*N*	*M*	*SD*
Unique Learners	no CDT	225618	1.33	0.80
	CDT	31621	1.46	0.91
Conversation Length	no CDT	225618	1.48	1.43
	CDT	31621	1.67	1.70

Table 5 Descriptive statistics from DBR1, quantitative study 2

How are these conversations modelled using the heuristic measure?

The breakdown of conversations by type according to the heuristic modelling shown in Table 6 demonstrate that courses with the CDT have a larger proportion of the heuristic groupings associated with higher levels of social constructivist learning: extended social conversation, conversations with more members, and fewer Lone conversations.

		No CDT	CDT	Difference
Social dimension	Lone	78.15%	71.06%	- 7.09%
	Q&A	14.93%	19.53%	+ 4.6%
	Limited Social	3.29%	3.88%	+ 0.59%
	Extended Social	3.63%	5.53%	+ 1.9%
Unique participants	Lone	78.15%	71.06%	- 7.09%
	Watercooler	15.08%	19.16%	+ 4.08%
	Cocktail Party	6.75%	9.76%	+ 3.01%
	Conference	0.02%	0.02%	----

Table 6 Percentage of conversations in each heuristic grouping; DBR1, quantitative study 2

5.3.3.2 Discussion of quantitative study 2 results

The results from quantitative study 2 further demonstrates a statistically significant relationship along the dimensions of conversation length and unique participants. This confirms that there is an impact in terms of interaction as measured by additional contributions to conversations when the CDT intervention is deployed. The Cohen's d scores also demonstrate that there is an impact and although this is smaller when more courses are included (and of course more contingent and unfactored factors), it is still statistically significant. This suggests that the intervention is worthy of further investigation, especially when limitations of the design in this phase are accounted for, such as the minimal scaffolding of the learning activity, and the single-step implementation, limiting the re-use by some learners. Indeed, the survey data below reveals that a large number of respondents used the tool once in this phase (141/308 responses); further analysis of the survey instrument is explored below.

In terms of the heuristic modelling of conversations, both analyses saw increases and decreases in the same areas, which supports the hypothesis that the intervention is contributing towards a change in social behaviours. The levels of difference are also similar: for example, there is approximately a 7% drop in 'Lone' conversations in both analyses. This may imply that 'Lone' comments are being discovered through the tool, and learners are replying to some of them which puts them into a different heuristic category. In terms of the 'turn taking dimension' this 7% drop is taken up by Q&A (no learner 'goes further' in the conversation by returning for a second reply), and Extended Social (one or more learners 'goes further' by making further replies, indicating a sustained interaction), and the proportion of the conversations moving into these categories is also similar (Q&A: 3.76% vs 4.6% and Ext Soc: 2.62% vs 1.9%).

However, in the unique participants dimension, the levels of difference are more different, in that the wider analysis saw a smaller increase in 'Cocktail Party' (i.e. 3-9 participants) conversations than in the 'Watercooler' (i.e. 2 participants) conversations. This could be explained by examining other commenting behaviours in the William Wordsworth course; for example, perhaps this demographic of learners are more keen to join conversations that are already underway than the overall population, and further comparative analysis would determine where the William Wordsworth course lies in relation to the others. However, it is not necessary to make these calculations at this stage to answer the question of impact, as the heuristic modelling only seeks to understand conversational forms, so has a limited explanatory function. The most important part of this analysis is that the similarity of the drop in 'Lone' conversations which implies that the CDT is increasing the discovery of Lone comments and encouraging learners to reply, thus shifting them into another heuristic category.

It is important to note that all the results above demonstrate the impact in terms of contributions to conversations and differences in participation through use of the CDT. It **does not and cannot** represent where the CDT is being used as a strategy for vicarious learning; learners who prefer to read others' perspectives but do not actually leave comments or reply themselves are excluded from these data. These data can only be established through a closer interview method, which is conducted in the next phase of DBR when the CDT can be embedded into all weeks of the course, and I have a better understanding of how to formulate and explicitly scaffold suggested uses based on learner feedback. In this phase of DBR, my own biases are sacrificed in terms of the types of activity that I believe the designed affordances can support for learning to collect the widest pool of feedback in terms of the value that learners have perceived. This wide pool of feedback is qualitatively analysed later in this chapter.

5.3.4 The DBR1 survey instrument

I provided a short survey of 15 questions at the end of the course in order to analyse the correlation between learning preferences and behaviours which are related to learning. For example, I ask how much learners value social learning, whether the CDT helped them make conceptual linkages, and also whether it encouraged them to comment more. 308 users across 9 courses responded to this survey.

The research question that this aspect of the research seeks to understanding is:

What are the relationships between learning preferences, experiences, perceived affordances and evaluation of the CDT as a means to develop thinking?

The full set of questions are listed below:

1. (Confirmation of consent to participate and link to the participant information sheet)
2. How many FutureLearn courses have you enrolled on? (1, 2-5, 6-9, 10+)
3. How important is social interaction in the discussion areas for developing your thinking (not at all, slightly, moderately, very, extremely [important])
4. Please give further details… (free text)
5. How many times did you use the word cloud application? (1, 2-5. 6-9. 10+)
6. Has the word cloud visualisation helped you develop your thinking about the course content? (No, Maybe, Yes)
7. If it had an impact on your learning, please describe what kind of impact… (free text)
8. Has the word cloud visualisation helped you discover new conversations you might have otherwise missed? (No, Maybe, Yes)
9. Has the word cloud visualisation helped you discover new people you might have otherwise not found? (No, Maybe, Yes)
10. Has the word cloud visualisation encouraged you to comment more on the course? (Definitely not, Probably not, Not sure, Probably yes, Definitely yes)
11. How useful did you find the word cloud visualisation on this course (Definitely not useful, Not useful, Slightly useful, Extremely useful)
12. What features (existing and potential) of the FutureLearn platform might help you learn better? (free text)

13. Are there any changes you could suggest specifically for the word cloud visualisation? (free text)
14. Any other comments (free text)

5.3.4.1 Results from the DBR1 survey instrument

308 learners fully completed the survey instrument, and I conducted a Spearman's Rank Correlation analysis to establish relationships between the questions about experience and learning preference (Qs 2-5 above), and the questions about the types of perceived affordances and perceived value of the CDT exercise. Spearman's correlation coefficient is a measure of the strength of monotonic relationships between paired data where 1 would be a perfect positive correlation, 0.6 would be considered strong and 0.4 would be considered moderate (Weir, n.d.). The full descriptive statistics for this analysis can be seen in Table 7:

		Q2	Q3	Q5	Q6	Q8	Q9	Q10	Q11
Q2	Correlation Coefficient	1.00	-0.03	0.10	-0.116	-0.03	-0.06	-0.07	-0.04
	Sig (2-tailed)	-	0.55	0.09	0.04	0.65	0.28	0.25	0.54
Q3	Correlation Coefficient	-0.03	1.00	0.179	0.343	0.343	0.263	0.359	0.329
	Sig (2-tailed)	0.55	-	0.00	0.00	0.00	0.00	0.00	0.00
Q5	Correlation Coefficient	0.10	0.179	1.00	0.389	0.431	0.358	0.354	0.479
	Sig (2-tailed)	0.09	0.00	-	0.00	0.00	0.00	0.00	0.00
Q6	Correlation Coefficient	-0.116	0.343	0.389	1.00	0.703	0.611	0.717	0.766
	Sig (2-tailed)	0.04	0.00	0.00	-	0.00	0.00	0.00	0.00
Q8	Correlation Coefficient	-0.03	0.343	0.431	0.703	1.00	0.720	0.683	0.742
	Sig (2-tailed)	0.65	0.00	0.00	0.00	-	0.00	0.00	0.00
Q9	Correlation Coefficient	-0.06	0.263	0.358	0.611	0.720	1.00	0.683	0.677
	Sig (2-tailed)	0.28	0.00	0.00	0.00	0.00	-	0.00	0.00
Q10	Correlation Coefficient	-0.07	0.359	0.354	0.717	0.683	0.683	1.00	0.761
	Sig (2-tailed)	0.25	0.00	0.00	0.00	0.00	0.00	-	0.00
Q11	Correlation Coefficient	-0.04	0.329	0.479	0.766	0.742	0.677	0.761	1.00
	Sig (2-tailed)	0.54	0.00	0.00	0.00	0.00	0.00	0.00	-

Table 7 Descriptive statistics of the correlations and significance (n=308); DBR1, survey instrument

5.3.4.2 Spearman's Rho correlation coefficient

Relationship between previous experience and the CDT

There is no statistically significant relationship or strong positive correlation between Q2 (How many FutureLearn courses have you enrolled on?) and any other question on the survey. This suggests that experience with the FutureLearn platform is not a determining factor for ability to use and evaluate the CDT, therefore the CDT is as easy or difficult to use regardless of previous experience; in other words, there is a level playing field for all learners.

Relationship between importance of social interaction (learning preference) and the CDT

There is a very small (0.179) but significant correlation between Q3 (How important is social interaction in the discussion areas for developing your thinking?) and Q5 (How many times did you use the word cloud application?) which indicates that some learners who value the social learning experience will return multiple times in the course, but this is not the most common practice, and many will use the tool only when it is presented to them as part of the sequence. In this phase of DBR, this equates to a single use at the end of the first week. This is not the same as not perceiving the CDT as supportive of social interactions or developing thinking, because there is a much stronger correlation between Q3 and Q6 (=0.343), Q8 (=0.343), Q10 (=0.359) and Q11 (= 0.329), which is a clearer indication of *how* the learners who value social interaction in the discussion area are using the CDT step. Therefore, there is a moderate correlation between learners who value social interaction and learners who report that the CDT helped them develop their thinking (Q6), discover new conversations (Q8), and comment more (Q10). This is good evidence that the CDT is seen as a useful extension to the discussion toolset and means there is a correlation between valuing social learning and finding the CDT useful, but it is limited in that many learners will only experience the tool once, and make their evaluation on that solitary experience, even if they value social interaction generally. The correlation between Q3: learners who value social interaction and Q9: discovering new people is lower, which indicates that discovering new people is not seen as important and that the value of social interaction is not perceived in terms of connecting on a personal level, rather to access a diverse fund of knowledge to further personal understanding.

Relationship between amount of use of CDT and perceived value in terms of learning

The correlations between Q5 (How many times did you use the CDT?) and Q6 (=0.389), Q8 (=0.431), Q10 (=0.354), Q11 (=0.479) supports that claim that that further use of the CDT will enable learners to develop their thinking (Q6), discover new conversations (Q8), comment more (Q10) and find the tool useful overall (Q11). It is not surprising to find the strongest correlation here between number of times using the CDT and finding it useful, as learners who find it as useful are also likely to return. However, greater use of the tool is also correlated with the activities for which the affordances are designed, such as discovering new conversations, providing further opportunities for commenting and developing thinking. This supports the claim that the CDT could be planned into the learning design where reflection and discovery are important activities. It is interesting that Q10: the further opportunities for commenting (=0.354) correlates less than Q8: discovering new conversations (=0.431), which indicates that there are more learners using the tool as a reading strategy rather than those who view it as an opportunity to write.

Indeed, the correlation between using the tool more and Q9: discovering new people (=0.358) is like that of Q10: seeing it as an opportunity to write more comments (=0.354) which confirms that this affordance is perceived by some but less than those who use the tool solely as a reading strategy, and this may also go some way to explaining why only a moderate impact can be demonstrated in the quantitative analysis of overt contributions above.

Relationship between perceiving the designed affordances and developing thinking

There are very 'strong' correlations in Table 7 between learners who respond positively to perceiving the designed affordances (Qs 8-10) and Q6: experiencing conceptual changes, strongly related to 'learning'; for example there is a very strong correlation between Q6: those who develop their thinking through using the CDT and Q8: discovering new conversations (=0.703), Q10: commenting more (=0.717), again less so for Q9: discovering new people (=0.611). This indicates that discovering different conversations and engaging more with others is strongly correlated to conceptual change, which is well supported by sociocultural learning theories (Lapadat, 2006; Littleton & Mercer, 2013; Vygotsky, 1978) and this is evidence that the CDT is creating the new opportunities for sociocultural learning and useful serendipitous encounters. However, the conceptual change is more related with exposure to a range of diverse opinions rather than engaging personally with the people who post those opinions. Learners do not join MOOCs specifically to meet new people, however they see a great deal of value in others' contributions. In this sense the MOOC is not a space of social networking, but it requires a mass of individuals to generate enough content such that each individual can find value in the 'swarm'. In this sense, the theoretical framing and design of the tool on stigmergic principles which take advantage of the weak tie connections and peer production is coherent with users' experiences.

5.3.4.3 Discussion of the Spearman's Rho correlation coefficient

These results demonstrate that the CDT is having a positive impact on a number of learners and that integrating it further into the platform (for example at the end of EVERY week) may increase the overall number of times that learners use it; indeed, making a relevant discovery is more likely as the comments corpus grows, so there is a greater chance of a serendipitous encounter in the final week than the first.

A key finding is that there is a level playing field in terms of being able to perceive the affordances of the tool, so learners who have done more FutureLearn courses are not more likely to see value than those who have only done one course. Perceived value is slightly more associated with discovering new conversations than with questions related to interaction and writing which suggests that there is a population of learners who are using the tool solely as a strategy for further reading, but for those who perceive the most value in using the tool and report that the tool has helped them develop their understanding, there is a strong correlation with writing more, which is well supported by sociocultural theories. The CDT does not strongly associate with discovery of new people, which may be explained by learners not wanting to engage deeply with others on a personal level (preferring just the traces of their presence to guide them), or that the CDT itself is just not a useful tool for this practice. These questions are explored further using an interview method in DBR2.

It is a known limitation that the tool is only made available in the course once, and the strongest conclusion from these results is that the CDT needs to be embedded more into the course, with further explicit scaffolds outlining the affordances: starting with use as a tool for further reading, but also to discover new conversations in order to engage more with others and write more, which has the strongest correlation with learning.

The final data analysis in DBR1 is the free text comments which were written on the step itself combined with the free text comments in the survey. These are analysed abductively which is to say within a developing conceptual framework. The next section seeks to understand more about how learners respond to the CDT step and what types of learning activity they associate with it.

5.3.5 Qualitative analysis of learner feedback

The CDT step, as with all the steps on the FutureLearn platform, allows learners to leave a comment, or engage in discussion. I encouraged learners to write some short feedback about this step and gathered 590 short comments over the 9 courses, which I coded using NVivo 12 against the following criteria which emerged alongside the analysis. All comments quoted below are anonymized and appear in either the survey 'free text' responses, or the comments on a page related to the CDT activity within the 9 courses studied during DBR1. Quotes have been selected where they concisely describe a common theme and I have chosen several comments to illustrate a dimension to represent the range and scale of the feedback.

1. **Sentiment analysis**: mixed, negative, neutral, positive
2. **Overall evaluation of the step**: difficult, interesting, hard to understand, time consuming, problems with technology, useful or not useful
3. **Perceived learning value of the step**: connecting [ideas and people], discovery [general and time-based], learning through extension
4. **Perceived affordances**: concept map, filtering mechanism, types of words clicked [noun phases vs grammar phrases]
5. **How the step is scaffolded into the course**: explanation and frequency of occurrence in the course structure
6. **Other comments and feature suggestions**

5.3.5.1 Sentiment analysis

589/590 items were coded with an exclusive sentiment code, representing 99.83% coverage of the dataset and forming the basis for understanding the other codes. 245 responses were positive, 203 responses were neutral, 105 responses were negative and 36 were coded as mixed. Many negative comments cite technical problems (83/105), but this doesn’t account for all negative responses, merely demonstrating that the CDT is not working technically for some learners, for example this response stating technical problems:

> *I am not allowed to use the Comment Discovery Tool, there is a message saying: 'your connection is not secure" - any idea why this is?*

The ‘technical problems’ type comment is useful feedback, but does not answer the research question of how useful the CDT activity is for learning. Other negative comments challenge the idea of the CDT which suggests that some learners do not make the step from perceiving the basic affordances to using them effectively for learning. I cite 4 different responses below also coded for negative sentiment, with a discussion as to what I may learn from them with relation to the major research questions:

> *Tried it out, just for fun. Not very useful though, as the 100 comments listed consisted of only 1... Room for improvement, I suppose :-)*

> *I didn't really understand how the tool could be useful. I would like to have been able to enter a word to see how often it was used and what words were used with it, rather than just see some words that were repeated.*

Not sure about the usefulness of this. It's like an index of most-used words. Many of those words are pretty mundane, directing us to random comments. More useful if words more relevant to poetry or Wordsworth in particular are picked out, surely.

Wordclouds are quite interesting but not sure if the instructions were entirely clear, the second word I clicked on, Coleridge, flagged up only 2 vague comments. I think this activity could be more clearly thought out to be more meaningful to learners who have a variety of different learning styles. I already think the [suggested learning hours] associated with the course are a little unrealistic, I've spent more than 4 hours on this section this week.

These comments suggest features that they would like to see (*being able to enter a word or view concordance; making the word choices more 'relevant')* or challenging the way in which the idea is scaffolded onto the course (*I didn't understand how the tool could be useful, could be more clearly thought out*). This demonstrates that there needs to be more effort devoted to describing how to use the CDT for optimal learning in a context of scale; it is also a known risk of displaying the CDT without close instruction, intended to harvest a broad range of descriptions of resultant activity, indicating perception of affordance.

A notable finding is the way in which other comments cite time as a factor for negatively appraising the step (*I already think the [suggested learning hours] associated with the course are a little unrealistic*). Time acts as an 'invisible hand' and can create wildly differing interpretations of the step, depending on how the learner is self-regulating and self-directing their learning (i.e., dependent on their goals and amount of time at their disposal). On one hand, time can restrict the effort expended into learning a new tool or technique which reduces efficiency of 'progress' or 'course completion', as is evident from the comments above. On the other hand, some learners perceive the tool as either a time saver if they wish to concentrate on a particular theme, or as a method of viewing new and interesting comments outside their immediate timeframe leading to successful and even joyful learning experiences. I cite 4 comments coded positively for sentiment below, with a discussion on how these may be interpreted:

> *Really great way of personalizing the exploration! I could spend hours reading different threads, but I must continue playing catch up, is this feature available at the end of every week?*

> *Something different but definitely a time waster! Interesting picking up on some comments that obviously pre-or post-dated my own comments and weren't read before.*

This is a great tool. I rarely look at comments older than those shown on the first page of comments, and I suspect this is common among learners. However, this tool allows you to follow ideas that take you to comments posted on earlier pages, and also you can go back to it and check on what has been said after you have completed a week's activities.

I enjoyed this exercise; it would be interesting to see how it can help with other courses. For me it triggered other thoughts on what I have read and wanting to read these passages again.

These comments relate to designed affordances which extend the capacity for learners to engage in 'resource negotiation', 'individualisation' and 'self-organisation', necessary dimensions for the affordances of the platform to meet the pedagogic model following Britain and Liber's framework; these comments demonstrate that some learners are indeed perceiving this limitation of the default platform and finding the CDT step useful to overcome it. There is an important point in the 3rd comment above which is related to this ("*I rarely look at comments older than those shown on the first page of comments, and I suspect this is common among learners*") which supports the claim that the design of the FutureLearn platform limits the resource negotiation, individualisation and self-organisation by displaying comments chronologically, rather than having an affordance to discover comments based on affinity. It also supports my claim in the methodology chapter that the possibilities for interaction on any given page is limited to 12 possible 'clicks' ("*… and*

I suspect this is common among learners"); this claim can only be made if it is self-evident to the learner how the platform is encouraging interaction.

It also supports a key finding from the correlations analysis of the survey data: that learners do not use the tool to connect with other learners (as people), rather to expose themselves to a diverse fund of knowledge, as the final comment explicitly states and this supports the idea of swarm intelligence, and the choice of a stigmergic design paradigm for the visualisation. The following comment invokes the metaphor of a river, which is not uncommon when discussing FutureLearn courses ('comments stream', 'free-flowing discussion'…) and also raises an interesting point about the value of the tool being related to the learners' own investigative questions which trigger selecting a certain word (my emphasis):

> *The powerful potential of this tool, which is l see now is a form search engine, as an aid to navigating through the* ***rapidly running river of comments*** *that is the lifeblood of an active MOOC, is emerging as time passes. [Writing on Thursday of Week 3]. Like any database (i.e. the comments posted in discussion threads in sections of each week)* ***it's value will derive from the clarity of the line of investigative questions/enquiries*** *which provide the rationale for selecting particular words from which to generate fresh clouds which may either offer answers or provide leads for further enquiry. I have a sense that use of a*

tool like this could even be a springboard to(wards) discovery of fresh ways of getting to grips with the literary canon that is WW's legacy.

This comment suggests that the step may require closer scaffolding for learners, for them to reflect on *'the clarity of the line of investigative questions'* and selecting words which may answer these questions. The tool 'affords' this by design, as is in the final comment, but it is not immediately 'perceived' by learners who are concerned by the ticking clock or their progression/ completion status.

5.3.5.2 Overall evaluation of the CDT activity

386/590 comments were coded as evaluating the CDT activity, representing 65.42% of the data. These were broken down into 7 dimensions as seen in Table 8 below:

Evaluation dimension	No. of comments
Technical problems	135
Useful	112
Interesting	78
NOT useful	20
Did not understand	16
Time consuming	15
Difficult	10

Table 8 Codes of evaluative commentary about the CDT activity; DBR1, qualitative analysis

135 of these are related to technical problems accessing the step. This is different from the 83 "negative/ technical problems" cited in the sentiment category above, because technical problems across the evaluative dimension are also 'neutral' in

sentiment (=43), mixed (=2), or even positive (=7). The positive ones tended to report that they had overcome the technical problems and were now finding the tool useful.

The next most popular codes are ‘useful’ (=112) or ‘interesting’ (=78). The short form of the comments in the dataset limits the ability to gather rich feedback in this analysis. Many learners simply wrote: “*Love this. Very interesting*” which does not give much detail on the specific way in which it may be helping their learning value although is noted as an evaluation of the CDT step. This limitation on the detail is pervasive across many comments coded for evaluation, but it is useful to note both the proportions and the emergence of these specific codes. The codes were not fixed before analysis, as per an abductive method, but were emergent through analysis, so the fact that these codes emerged and not others is pertinent. These codes are also not exclusive, for example it is possible for a single comment to be BOTH time consuming AND interesting.

These results are not conclusive, although it is important to note that ‘difficult’ is the least cited category, supporting the results of the Spearman’s Rho correlation of the survey instrument that the CDT is not challenging to operate (it does not require experience with FutureLearn), at least at a basic level. The next section delves more deeply into the types of learning value that is reported from learners.

5.3.5.3 Perceived learning value of the activity and perceived affordances of the technology

Learners' perceptions of learning value and the affordances are related in that affordances should give rise to learning value. Therefore many comments cited both in their feedback, but at the same time new affordances were revealed which were not originally intended.

The CDT has 2 main designed affordances: 1. filtering or discovery (of comments and people) and 2. connecting ideas or concepts (through 'reading' the visualisation as a concept map). 82 comments are coded as recognizing the filtering affordance, and 81 comments are coded for improved learning through discovery of new comments, demonstrating the closeness of the affordance (as a cybernetic factor of the system) and the learning value (as a pedagogical factor). The second designed affordance replicates this closeness of cybernetics and pedagogical value, in that 32 learners perceive the affordance of 'concept map' and overall, 34 learners cite connecting ideas as specifically improving their learning experience. These results also support the finding from the correlation analysis, that users perceive the tool to be a 'concept map' less frequently than as a tool for general discovery. The following sections discuss the 'concept map' affordance and the filtering affordance in relation to theory.

The 'Concept Map'

The concept map affordance is a feature of the changing visualisation as terms are selected. This acts as a 'conceptual scaffolding' tool to support learning through displaying related terms which may challenge or surprise the learner, triggering new ideas, as demonstrated in the following comments. As stated above, only 32/590 comments were coded as having explicitly perceived the 'concept map' affordance, representing only 5.42% of the data. Feedback is not prompted or structured so it is possible that more learners did perceive this affordance, but didn't explicitly reference it in their comments.

> *I enjoyed this exercise; it would be interesting to see how it can help with other courses. For me it triggered other thoughts on what I have read and wanting to read these passages again.*

> *Like poetry, read it, and read it again, focussing on the different meanings of certain words and expressions. It is interesting how it can direct me towards understanding a certain aspect of the writer writing about, in this case, Davy.*
>
> *This seems an intuitive way of finding interesting links among all those thousands of comments.*

This feature is unique to visualisation technology in that the aggregation of the corpus brings a new dimension to the platform which is impossible to experience through linear progression. In DBR1, this affordance is 'implicit' (not explicitly scaffolded through pedagogic instruction), meaning that the epistemic implications are not made obvious to learners through direct instruction on the page.

> Implicit epistemic scaffolding refers to the support (including tool, activities, resources, etc.) that has epistemic implications but is not made obvious to learners. Explicit epistemic scaffolding refers to the support that intentionally makes epistemic ideas explicit to learners to promote their epistemic understanding. Many of the CSCL environments have been designed with epistemic underpinnings, which could serve as implicit epistemic scaffolds. We propose that embedding explicit epistemic scaffolds in CSCL environments that are designed with implicit epistemic scaffolds is a promising way to maximize the power of support for epistemic growth. (Lin & Puntambekar, 2019, p. 597)

In DBR2, this affordance needs to be more explicitly referenced on the course page, as suggested by Lin and Puntambekar (2019), because it has not been recognized by enough learners, although the feedback above demonstrates a high level of cognitive change can be achieved through using it in this way. This is also related to the dispositions of the learners because learners who can frame their investigative questions, be open to new ideas through serendipity (rather than seeking confirmation) and *only then* operate the visualisation will gain the most from the CDT exercise, but this requires quite a high level of academic sophistication, and crucially the time to reflect and return to the step.

> *When I first read about this activity I had the same idea as [Learner A], I thought Why are we doing this? What is the purpose behind this task? , so I thought: well, it something I have to do and that is all, but when [I] started reading the different ideas and opinions of people and how different our interpretations of the same word could be I got really interested. Now I see this task more like [Learner B] does. This task opens your mind and lead you to see other points of view. As [Learner B] says: sometimes you agree and sometimes you don't but it is interesting and enriching either way. I have enjoyed and I have learning while doing this task and I guess that is the purpose of all this process.*

This comment concisely summarises many of the points made above about utility, time and sentiment as well as how the platform encourages certain modes of activity: it starts with the problem of lack of instructions for using the step ("*Why are we doing this? What is the purpose behind this task?")* but continuing to give on a more reflective and sophisticated account. The learner's journey starts out with 'this is something I have to do', which reinforces the idea that the platform is strictly guiding you through a sequence rather than following your interests, steps that you must complete it to progress. However, under guidance from another learner, the commenter recognizes that the CDT '*opens your mind and leads you to see other points of view'*, which is something the platform is less capable of doing on its own because of the forward-focused, 'stepped' pathways to completion, and the lack of affordances for filtering comments based on affinity or supports for reflective practice, which is essential to learning. The fact that another learner can give support to the commenter is important when we consider that this support indirectly extends the range of the environment for others, which is a high-level behaviour in the cultures of participation framework. It is a challenge of this project to create affordances which allow users to reach these higher levels, and guide each other through the process, because when learning at scale, peers are as important, if not more so, than tutors.

Filtering by affinity

The FutureLearn platform orders comments chronologically in a 'stream', with the most recent 100 comments listed per page. Learners can scroll to the end of the first page, and if there are more comments, click on the second page etc. but it has been noted in the comments above that learners may only read a few comments at the top of the page without incentive to read everything on every page.

The ability to filter based on chosen words is well received by many learners, and 81 comments are coded as citing this affordance as beneficial to their learning experience. However, these instances can be divided between learners who generally find filtering and discovery useful, and those who specifically cite time as a factor affecting their experience (the 'time' code emerged through the abductive coding). The following comments describe how this capability for further discovery has generally improved the learning experience:

> *I love this little tool; it has highlighted comments I hadn't seen before and so gave me lots more to think about. So, I have learnt more and read some fascinating ideas that I would never have known were there otherwise.*

> *What an excellent way to reach comments that explore themes and issues you are drawn to. I often have to give up reading comments due to time constraints, but the links provided by this tool were very interesting. I found the comments more appealing. Good activity.*

Learners who cite time as a factor can be divided into those who enjoy reading comments on pages they have already 'completed', and those who started the course 'late' and felt it was useful as a time saver to access comments they may not have usually seen. Of course, late starters will see a more developed visualisation as the course develops so does the corpus of available comments. I have cited some comments from the late starters below:

I found this tool very good, I started late in this course and it gave me access to comments that I would not have previously seen.

I like this tool. Really useful to catch up for late arrivals on the course. Appreciated.

Others cited time in terms of being able to look back or revise and perceived the CDT step as a ‘time saver’ because ‘there is a limit to how many comments you can go back and read’. In fact, this ‘limit’ is often cited in the feedback and this highlights something fundamental to the way in which the FutureLearn platform guides learners through the steps, which is not coherent with a ‘reflective’ or ‘inquiry based’ approach that is central to literature on self-directed learning:

Interesting picking up on some comments that obviously pre-or post-dated my own comments and weren't read before.

I like the idea of it. I never have time to read all the comments and I rarely go back to pages I've finished to see what people have added since I was there, so a search tool like this towards the end of every course is a neat idea.

I rarely look at comments older than those shown on the first page of comments, and I suspect this is common among learners. However, this tool allows you to follow ideas that take you to comments posted on earlier pages, and also you can go back to it and check on what has been said after you have completed a week's activities.

So useful for looking back on previous comments, a great time saver for those of us who enjoy these courses around a busy lifestyle!! There is a limit to how many comments you can sometimes go back and read, but an ability to read comments specific to what you are thinking and wish to add informed comment to is a valuable enhancement.

5.3.5.4 Theorising the CDT in relation to self-directed learning

The need for adult learning to be self-directed in learning is well cited, for example Malcolm Knowles writes about the problems of employing the same pedagogical methods with children and adults in the 1920's (when adult education had started to become more systematically organized):

> [Adults were] resistant … to the strategies that pedagogy prescribed, including fact-laden lectures, assigned readings, drill, quizzes, rote memorizing, and examinations. Adults appeared to want something more than this, and drop-out rates were high. (Knowles, 1980, p. 40)

Knowles also writes that the most important lessons for adults are the skills of self-directed inquiry and continues to speak about when adults turn to tuition for their learning needs, which is where we can draw parallels with the FutureLearn platform. On one hand FutureLearn is allowing access to learning resources, making it easier for adults to turn to tuition, but on the other it is lacking in its affordances which support reflective and investigative activities (key to adult self-directed learning). I have emphasized the quote below to relate this to the discussion of FutureLearn:

> [guided instruction tends to] interfere with their learning by substituting their **own pedagogical sequence of steps** rather than flowing with the learners' natural sequence (Knowles, 1980, p. 42)

When these concepts are applied to the FutureLearn platform, it could be argued that exactly the opposite type of pedagogical experience is being designed-for, in line with the didactic drill and test methods cited above, and this is being cited as a gap in the platform by the learners who positively appraise the CDT activity as allowing for individualized inquiry and recognize that a lot is missed through following a pre-determined sequence. Knowles concept of 'Andragogy' (adult learning) evolves into the concept of 'Heutagogy' or the study of 'self-determined learning', defined below by Hase and Kenyon, and emphasized for relevance to the pedagogy of FutureLearn:

> the need to be flexible in the learning where the teacher provides resources but the learner designs the actual course he or she might take by negotiating the learning. Thus, learners might read around critical issues or questions and **determine what is of interest and relevance to them.** (Hase & Kenyon, 2001)

It is clear from the learner feedback above that the CDT step is enabling learners to negotiate their own learning in line with 'heutagogic' principles and is enhancing the pedagogical dimensions outlined above of 'resource negotiation', 'self-organisation' and 'individualisation' in a way that is not present in the default platform's cybernetic qualities. Furthermore, I have argued that the platform is in fact restricting these dimensions through its chronological ordering of comments and its design which encourages learners to 'mark as complete', which is also a disincentive to look back (literally: reflect). This is also demonstrated in the learner feedback above ("There is a limit to how many comments you can sometimes go back and read"; "I rarely look at comments older than those shown on the first page of comments, and I suspect this is common among learners"; "I never have time to read all the comments and I rarely go back to pages I've finished" etc.).

5.3.5.5. Scaffolding of the CDT activity

I scaffolded the CDT activity in a 'light touch' manner to prevent my own biases for how it could used being made explicit and this decision created some benefits and challenges. I wanted learners to play with the visualisation and formulate their own questions and pathways of inquiry. The system is engineered to support serendipitous encounters with related content as the visualisation changes to reveal related terms which may be unexpected. The final step of the learning process is to discover something specific in the new set of focused comments and follow the link to 'join the conversation', thus adding new funds of knowledge to the system's 'swarm intelligence' which can be picked up and encountered by another learner in a virtuous cycle.

These ideas are challenged by several factors such as learners' need for clearer instructions in terms of operating the tool (perceiving it as a concept map for instance) and in terms of the reflective work that needs to be done prior to choosing a term. Overshadowing all these complexities in the learning activity is time: thinking time to gain a helpful insight, time to read more comments, time to learn a new tool, discovery time to find something useful, and separately the exposure to comments which occur in the same timeframe as them (or exclusion from comments outside this timeframe such as comments which are posted after they have left a step).

A frequently coded comment which relates to how the step is integrated with the rest of the platform is that it needs to appear at the end of every week:

I could spend hours reading different threads but I must continue playing catch up, is this feature available at the end of every week?

I'll try to remember to use it throughout the course.

Scaffolding on the course page, including frequency of CDT inclusion is related to a wider issue with the learning design of FutureLearn courses in terms of the toolsets to support the types of teaching types suggested by Laurillard (2012). Laurillard cites 6 activities as broadly descriptive of all teaching and learning scenarios: *acquisition, investigation, discussion, practice, collaboration, production*. The FutureLearn platform's toolsets relate mainly to 'acquisition' in terms of videos and article steps, and discussion in terms of the commenting features, although my findings of this chapter disputes how well this actually grows into synthesis and what would be recognized as cooperative knowledge construction (Tubman, Benachour, et al., 2019; Tubman et al., 2016; Tubman, Oztok, et al., 2019). The 'investigation' learning type is not well supported through a specific toolset and until learners encounter the CDT step learners may not have been asked to do an activity relating to investigation or inquiry, rooting them firmly to the bottom/ consumption levels of the Cultures of Participation framework (Fischer, 2011). For effective use of the CDT, learners must first reflect on what they have already learned and think about how they could extend this interest through a cycle of selecting of terms and further reflection on the new terms which are surfaced by the visualisation. This will necessarily involve both time and preparation, neither of which the FutureLearn platform generally affords for: time is seen as a diminishing resource unless a premium is paid for ongoing access, and the platform moves learners forward through the cues in the design of the pages and the 'completion' or 'progress' incentives. The next phase of DBR will need to acknowledge these factors, and specifically suggest that learners stop and think before they interact with the step, to reduce quick evaluations and improper use of the tool.

5.3.5.6 Feature suggestions

The most common suggestions for extra features were more instructions for use and for there to be fewer, more focused words. Sometimes learners would ask for words to be removed from the cloud, to make the corpus more focused on the course at hand. This presents a challenge for the designer because the CDT is intended to act as a system level tool to increase the range of interactions. Learners do not see this as they only view the visualisation on their own course, so it is natural to ask for a slimmer, more thematic corpus. This is not an easy problem to solve and it invites suggestions from two separate viewpoints: that of computing, in terms of topic modelling from a computational perspective, for example using Latent Dirichlet Allocation (LDA)[13] or using social techniques to support collective topic modelling, such as the use of hashtags, or 'searchable text' (Zappavigna, 2012).

---

[13] https://en.wikipedia.org/wiki/Latent_Dirichlet_allocation: LDA is an example of a topic model and belongs to the machine learning toolbox and in wider sense to the artificial intelligence toolbox.

The basic problem for LDA is that the algorithm requires a set number of topics to 'look for'. LDA would be a useful processing method if we already know that we are just looking for the top 10 topics mentioned in the corpus. This requires a complete corpus, and a preset value for how many topics to count for, neither of which is possible for realtime visualisation or the word cloud format. This technical solution places too much emphasis on the work of the computer, rather than developing a community-based pedagogy which I have highlighted as being deficient in xMOOCs and is essential to encouraging behaviours which are at the higher levels in the Cultures of Participation framework. In other words, having the computer determine what topics are important would diminish the possibility of serendipitous encounters and reinforce a passive consumerist pedagogy.

This project takes a social perspective in terms of investigating the ability to build 'tags' in the form of hashtags into the visualisation; this is something which can be scaffolded throughout the whole course, and this will both speak to the challenges highlighted above about the CDT step appearing as unusual and unclear to a proportion of learners and also allow learners to define their own set of topics. In this way I can begin the course by inviting learners to theme their comments using hashtags and if I suggest a few in the first introductory step of the course, this will set the expectation and allow the emergence of a community-based approach to commenting, which could radically tip the balance for levels of communitarian practice. This will afford the ability to use the cloud to focus the discovery of topics at the end of the week and encourage behaviours which are related to higher levels in the cultures of participation framework, such as the ability to ‘extend the range of the environment’, perhaps by tagging other learners’ comments, which will improve everyone’s ability to organize content by theme. In this way learners will not view themselves as ‘passive consumers’ in the learning process but as an active part of a large community who cares about the big questions that the course is tackling, whether that be the makeup of soils, environmental or social issues, or the interpretations of poetry.

5.4 Summary of findings from DBR1

There is a lot in this chapter because it seeks to add empirical data to the suggestions in the methodology chapter about the ways in which the FutureLearn platform reinforces a certain type of interaction and this evidence needs to be layered carefully. It also frames the questions to be asked in the DBR2 about the general experience of learning at scale, and the ways in which the new set of affordances benefit and challenge this. DBR2 focuses on how the mediating artefact can support reflective and investigative activities and I have harvested a wide range of evaluations and uses to determine how learners perceive this tool’s affordances in relation to their own experiences of learning at scale so it can be embedded further into the general practice of learning in this context.

I have provided quantitative and qualitative evidence to support the claims that the CDT activity has a benefit for learning, using a novel framework for scripting the structure of all conversational units in a course, and followed this up with a survey and abductive analysis of free-text responses. Courses with the intervention displayed a greater number of conversational types which can be related to social learning, and fewer 'Lone' instances, which indicates that comments are being found and replied to, possibly due to the extension of the discussion affordances granted by the CDT. These quantitative results are statistically significant, but the impact is moderate. It is important to note that these quantitative results only measure overt contributions, as opposed to a 'vicarious' use of the tool where no new comments are added; this activity would not show up in any of these quantitative datasets.

The survey tool linked from the CDT-enabled courses gives more detail into how the tool is used and a Spearman's Rho correlation to demonstrates several findings: firstly, that learners who are more experienced with the FutureLearn platform are not more likely to find benefit so there is a level playing field for all. Learners who value social learning are more likely to develop thinking and find the tool useful so it is perceived as an extension to the commenting features, and greater use of the tool is correlated with greater cognitive benefits (discovering new comments, developing thinking) so some learners see enough value to keep returning. Discovering new people is less correlated with cognitive benefits which suggests that either the tool is not able to bring individuals together, or that learners are more interested in discovering comments than the people who left them, which is coherent with ideas of lightweight peer production (Haythornthwaite, 2009). Some learners are encouraged to comment more as a result of using the tool, and this is most strongly related to those who develop their thinking as a result of using the tool, which is a finding theoretically supported by sociocultural learning theories (e.g. Lapadat, 2006; Littleton & Mercer, 2013; Vygotsky, 1978). It will be important to emphasise through explicit scaffolding that learning is most strongly related to this type of interactive writing.

The free-text responses and feedback demonstrate a broad range of sentiment, evaluation and use cases, and show that the tool is largely viewed positively, and as a means to discover new comments and increase exposure to a diverse set of viewpoints. It is not clear what proportion of these users are actually adding to the corpus vs those who are purely using the tool as a reading aid, but the use as a reading aid is not represented at all in the quantitative analyses and this type of use adds further evidence that the platform by itself is not able to make best use of the total funds of knowledge, and that the tool is perhaps having more than the moderate impact on learning than is suggested by the Cohen's d scores in the quantitative analysis of the overt conversational patterns. The use of the tool as a 'concept map', to see related concepts throughout the comments is the unique affordance that only a visualisation can offer, but it is less well perceived by learners, and this is perhaps related to how the CDT activity is explicitly scaffolded into the course structure, which is feedback that can be carried forward into the next phase of DBR.

Other factors which cannot be measured purely by quantitative means are the range of effects that time has on the experience of doing a FutureLearn course. This is in terms of the arrangement of comments which places all learners into time-based cohorts and which the tool does a decent job of breaking up, exposing a variety of opinions to learners and giving more opportunities to make new conceptual linkages. It is also in terms of the feeling of countdown that many learners seem to allude to when they negatively appraise the tool for not getting them to where they want to be quick enough or describe it as a 'time waster'; other learners have the same experience and actually positively appraise the tool for the same thing, praising its ability to allow playful interactions and surfacing serendipitous encounters which were made invisible by the default affordances of the platform. This seems to be related to the degrees to which learners are motivated by open self-directed inquiry or focused on progression through the course. Learners who are focused on progression and completion are less likely to spend time following their interests and do not see learning this tool as important towards the goal of completion (indeed it is not designed to aid completion, rather the 'slow learning' of reflection and discovery!). This emphasis on completion which is common in MOOCs does not concur with previous literature on adult education (Knowles, 1980), where following interests and not taking a pre-determined pathway is suggested as the best way to achieve positive outcomes from a point of view of deep learning. However, it is the invisible affordance of the ticking clock on the platform which motivates learners to keep moving, and this may be detrimental to the actual learning experience, as it does not encourage the reflective or investigative teaching types (Laurillard, 2012) and prioritizes only elements which can be counted towards progression, such as acquisition through viewing videos or reading articles, or practice through computer aided assessments.

The CDT activity design can take on board some of the feedback around time to learn, by attempting to model expected behaviours with the tool, such as framing questions, reflecting on the changing clouds, and joining new conversations. I will attempt to scaffold some of this using peer produced or crowdsourced modelling in order to create a bank containing a broad range of possible uses, which are relevant to the course in hand, rather than generic examples and walk-through demonstrations. The crowdsourcing element relates to the Cultures of Participation framework and the development of cooperative production, as it can be seen from this first phase of DBR that many learners are still occupying lower levels in terms of identifying mostly as passive consumers. The most common suggestion for a feature is to remove 'irrelevant' words, and I also intend to speak to this in future iterations of the CDT step by creating new visualisations of folksonomic terms, preceded with a hashtag, which can both serve to scaffold the idea of the communitarian peer production into all steps of the course, prepare learners for the visualisation when they get there, and therefore encourage cooperative behaviours which occupy higher levels in the Cultures of Participation framework.

In the next phase of DBR I embed the CDT exercise into all weeks of the course, and prepare a video walk through of how it can be used technically. This should familiarize learners with the workings of the tool to reduce the initial learning curve and give more time in between each use to reflect on it and how it can fit into their learning patterns and strategies. I will also ‘merge’ the exercise into a more general step around reflection in order to encourage learners to reflect on the week’s content before they launch the visualisation, in order to prepare learners’ investigative questions.

I use a semi-structured interview method to collect richer descriptions about the experience of learning at scale, discussion tools, and how the CDT works as an extension of the discussion features. This phase has identified that there are limitations to any quantitative analysis which examines contributions, as many learners do not add their own contributions, therefore are invisible in the data, yet still appreciate the tool for its ability to surface a range of opinions.

Chapter 6: Findings from DBR2

6.1 Introduction

In this chapter, I expand on the Activity Centred Analysis and Design (ACAD) theoretical framework (Goodyear & Carvalho, 2016) which is used to analyse learning networks. This proposes that affordances are best understood as relational to the proposed learning activities and explicitly includes the physical environment as a factor for teaching design, including the digital learning environment or tools and wider sociomaterial factors. In the case of this project, an 'invisible' factor impinging on learners' ability to perform activities is time and can be described in terms of:

1. Learner self-regulation;
2. A structural factor of the platform, silo'ing learners and available knowledge;
3. A finite resource because 'non-premium' or paid access to the courses is time-limited.

These factors form invisible boundaries within which the learning activities can be proposed. Therefore, the 'physical' situation can be understood as limited in terms of 'things-that-learners-are-able-to-do' with the platform tools (i.e. affordances), and the 'time-they-have-to-do' these things (i.e. the invisible sociomaterial environment). These factors combine in point 2 as learners are physically situated in 'virtual silos of content', pre-determined by structural and temporal factors. Any design must work within these boundaries, which will inform the learning strategies that learners adopt when navigating the platform and its funds of knowledge (the static teacher-published, and dynamic peer-produced content).

Therefore this chapter pays attention to the proposed tasks which are mediated through the CDT, and how the instructions or scaffolding of these activities can encourage a deeper, reflective learning within a context of self-directed study (Knowles, 1980). Activities are proposed which encourage reflection and inquiry, with the problem space is neatly summed up by the learner in the previous chapter who said: "*[the CDT's] value will derive from the clarity of the line of investigative questions/enquiries*". I draw on the concepts in 'Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology' developed by Laurillard (2012) which are expanded in the ABC learning design framework (C. Young & Perović, 2016) to describe the pedagogical patterns of proposed activities in terms of their 'teaching types' to take a wide view on the learning design of the course and how the reflective and investigative activities which are supported by the CDT helps develop a participation based pedagogical approach.

The previous chapter (DBR1 findings) highlighted that comments are perceived by learners as opportunities for sharing experiences rather than explicit opportunities for knowledge construction through dialogue. This is contrary to ideas of knowledge construction in the learning sciences literature, for example the framework by Gunawardena et al. (1997) would place this type of discussion on the lower levels (sharing experiences rather than synthesising content into new ideas). These are also important findings in terms of understanding individual experience of FutureLearn MOOCs within an acquisition-based metaphor for learning, where others' comments are seen as emergent resources for extended consumption, rather than opportunities for social interaction and collaboration. This is strongly supported by the quantitative analyses of Tubman et al. (2016) and the findings from the previous chapter (Tubman, Benachour, et al., 2019; Tubman, Oztok, et al., 2019). The CDT is designed as a visualisation to coordinate peer produced content such that interaction is made easier for learners, and there is an increased possibility of collaboration. Collaboration in the sociocultural sense requires commonalities that are not easy to discover in a context of information and participation overload. I explore how the pedagogical scaffolding of the activities in FutureLearn encourages an individualistic approach, which further supports the results of the quantitative analysis above.

In this chapter, I examine the tasks which are mediated with the CDT using the ABC learning design framework, specifically framing them in terms of 'investigative' or 'reflective' learning activities which most reflect the feedback in DBR1 from learners who found this activity useful. The types of activities which are mediated by the tool affords learners the ability to reduce information and participation overload by filtering out irrelevant comments and allows them to *focus in and therefore reflect and extend or increase* their opportunities for making relevant connections. In terms of the Cultures of Participation framework (Fischer, 2011), the new affordances encourage opportunities for self-organisation of content and provide more agency for learners to decide how to personalise the course to their individual learning goals (Levels 2 and 3 in the framework). Resource negotiation, individualisation and self-organisation are required cybernetic dimensions to create a viable system for a social pedagogic approach (Britain & Liber, 2004). These dimensions are explored further in this chapter using an interview methodology which gathers rich insights into how learners work in the MOOC context, focusing on individual learning strategies and motivations for study. The CDT is evaluated by interview participants against these descriptions of their own practice. Semi-structured interviews have been chosen as the most suitable because it enables me to guide participants somewhat to the topic of interest, without influencing the outcome.

The interviews reported in this chapter act as a qualitative reflection on how learning at scale is experienced at the level of the learner and the focus on the learner adds a great deal of complexity as they each bring with them diverse experiences and motivations for study, and a range of expectations and learning strategies. In a MOOC, the learner is extremely self-directed and independent, so uses the course and its tools as they find them with little teacher involvement (in terms of modelling zones of proximal development, for example). This creates a complex environment in which to plan and deploy known learning design frameworks, but it is important that these tools of learning design are considered because they provide the framework for thinking beyond the passive, acquisition based pedagogy. This emphasises the importance of the ACAD framework, which can make some sense of the complexity by identifying the known boundaries of the sociomaterial context and also validates the utility of the ABC framework which adds descriptive clarity and transferable conceptual terms from which I can deduce the levels of active learning suggested by each activity. As a reflexive practitioner, this work does not have a fixed endpoint, and decisions such as the ordering and design of content, learning activities, and selection of tools is under constant reflection, evaluation and change.

I start by detailing the changes which were made to the implementation of the CDT exercise in the course, relating these concepts back to principles of learning design and using Diana Laurillard's 'teaching types' as a common language. I also use the same quantitative techniques developed in the previous chapter to demonstrate how the intervention's outcomes can be reproduced across new courses, which is key to a design-based research methodology (Brown, 1992).

The main research questions in this chapter seek to develop a framework around learners' interactions with both the CDT and with the whole platform, and I use a semi-structured interview method with 10 participants who give rich descriptions of their general challenges and opportunities of learning at scale and also their specific interactions/ workflows with both FutureLearn as a whole and also specifically with the CDT artefact, focusing on how it affected their experiences. I conclude by making suggestions for changes to the CDT exercise which makes the suggested cognitive activities more explicit and suggest crowdsourcing practices where learners share their own workflows and reflections. This builds on FutureLearn's broad pedagogical approach in that learners' comments act as 'breadcrumbs' for other learners and by asking learners to leave descriptions of their own interactions with the tool, it emphasises a community based pedagogical approach which is associated with the social aspects of knowledge construction (Henri, 1992).

6.2 Introducing discovery and reflection into the FutureLearn platform

In the first set of courses (analysed in DBR1), the CDT exercise could only be located on a single step. This approach collected a sizable and broad range of learners' perceived affordances, and evaluations. DBR1 established a quantifiable and statistically significant result in terms of demonstrating that where the CDT was present, overt sociality and conversation length are increased. Qualitative analysis added further depth to the data and described what effect the CDT has on learning and what other elements of the sociomaterial environment are important factors for learning at scale. Overall, this demonstrated that for some learners, when used well, the CDT acted as a supportive artefact for a reflective, investigative and discursive learning activity. Therefore, the CDT should be embedded as this type of activity within a formal pedagogical pattern, supporting a pedagogical model based in self-directed learning which follows from Knowles' (1980) theories of andragogy that learners should be directed by their own inquiry, rather than a predefined set of content and "fact laden lectures". In short: where the aim of DBR1 was to *discover* the specifics of these tangible benefits, the aim of DBR2 is to *uncover* these benefits to more learners through developing a better understanding of how to use "explicit scaffolding" (Lin & Puntambekar, 2019) to describe the proposed learning activity.

6.2.1 Learning Design

Learning Design is a relatively new research field and is neither a theory nor a methodology but focuses on the specifics of designing learning materials and artefacts which can meet certain learning objectives through development of toolkits and frameworks.

Learning Design can be defined as:

> a methodology for enabling teachers/ designers to make more informed decisions in how they go about designing learning activities and interventions, which is pedagogically informed and makes effective use of appropriate resources and technologies. This includes the design of resources and individual learning activities right up to curriculum-level design. (Conole, 2013, p. 8)

In this sense, learning design is useful to this project because it provides a framework where courses can be analysed in terms of how the mediate artefacts support the activities. I employ 2 learning design frameworks and toolkits as an analytic basis: Laurillard's 'teaching types' (Laurillard, 2012) as expanded into the ABC framework (C. Young & Perović, 2016) and the ACAD framework (Goodyear & Carvalho, 2016) which allows for emphasis to be placed on other sociomaterial aspects of the learning environment (such as the emergence of ‘time’ as an important concept) as part of the situated nature of learning.

Laurillard suggests 6 activities which make up most teaching situations. These involve learning through *acquisition, inquiry, discussion, practice, collaboration, and production*. As described previously, the CDT artefact is designed to support learning through inquiry, focusing on self-directed learning and discovery.

6.2.2 Learning Design in FutureLearn

The pertinent questions for creating useful pedagogical patterns of activity are at which point the task should be introduced, and it should be scaffolded such that it is supports learning (i.e. where to position it and how to develop instructions or explicit scaffolds). FutureLearn tools, such as videos, written articles and quizzes mainly support learning through acquisition or practice, and the activities of the week generally involve several acquisition activities and then some practice in the form of a multiple-choice quiz. Discussion activities are problematic on the FutureLearn platform in that questions may be asked in the text of a ‘step’ to encourage users to leave a single comment, but this type of informal knowledge sharing is distinct from the types of knowledge construction activities found in learning sciences literature or from sociocultural methods of participation based learning where the desired endpoint is creation of new knowledge through synthesis of experiences (e.g. Gunawardena et al., 1997; Henri, 1992; Scardamalia, 2002). As I have demonstrated through quantitative analysis in DBR1, the comments on FutureLearn are mostly 'Lone' and do not generally develop new knowledge through synthesis, but they are extensive and diverse, which gives them a unique quality that is exploited by the CDT artefact.

A recurrent theme through my analysis is that the FutureLearn platform itself creates incentives for 'progression' in the form of 'marking [a step] complete', but not 'reflection' in terms of suggested note taking, investigation and further reading, and synthesising what has been learned through discussion. Therefore, I chose to place the reflection/ inquiry exercise on the penultimate 'step' of each week, just before the tutor's written summary 'closes' the week, and propose the task as a means for learners to think about what they have learned, revisit themes to discover new ideas, and relate this to their personal motivations for study, developing through the learning process. The CDT does not occupy a whole 'activity' called 'Comment Discovery Tool' as is the case in DBR1 but is scaffolded as 'reflect and extend' which suggests the pedagogical intentions of the activity rather than the novelty of technology. The exception is in the first week where the step is called 'Discover New Conversations' where the instructions on the page relate to pedagogic intentions (inquiry, reflection) but the title familiarises learners with the new technology.

6.2.3 Humphry Davy: Laughing Gas, Literature and the Lamp

The following section gives a concrete example: the course "Humphry Davy: Laughing Gas, Literature and the Lamp" (2019)[14] and I explain both how the new 'inquiry' activity fits within the overall learning design of the week and also how it is described on the step. I use the language of Diana Laurillard's teaching types to describe the pedagogical intentions of each activity in the week. The example cited is very typical of a FutureLearn week and contains all the commonly used FutureLearn tools. Figure 10 shows the weekly structure of the course on the FutureLearn platform and also the scaffolding of the 'Reflect and Extend' task and Table 9 shows a breakdown of the weekly structure into Laurillard's teaching types.

[14] https://www.futurelearn.com/courses/humphry-davy/3

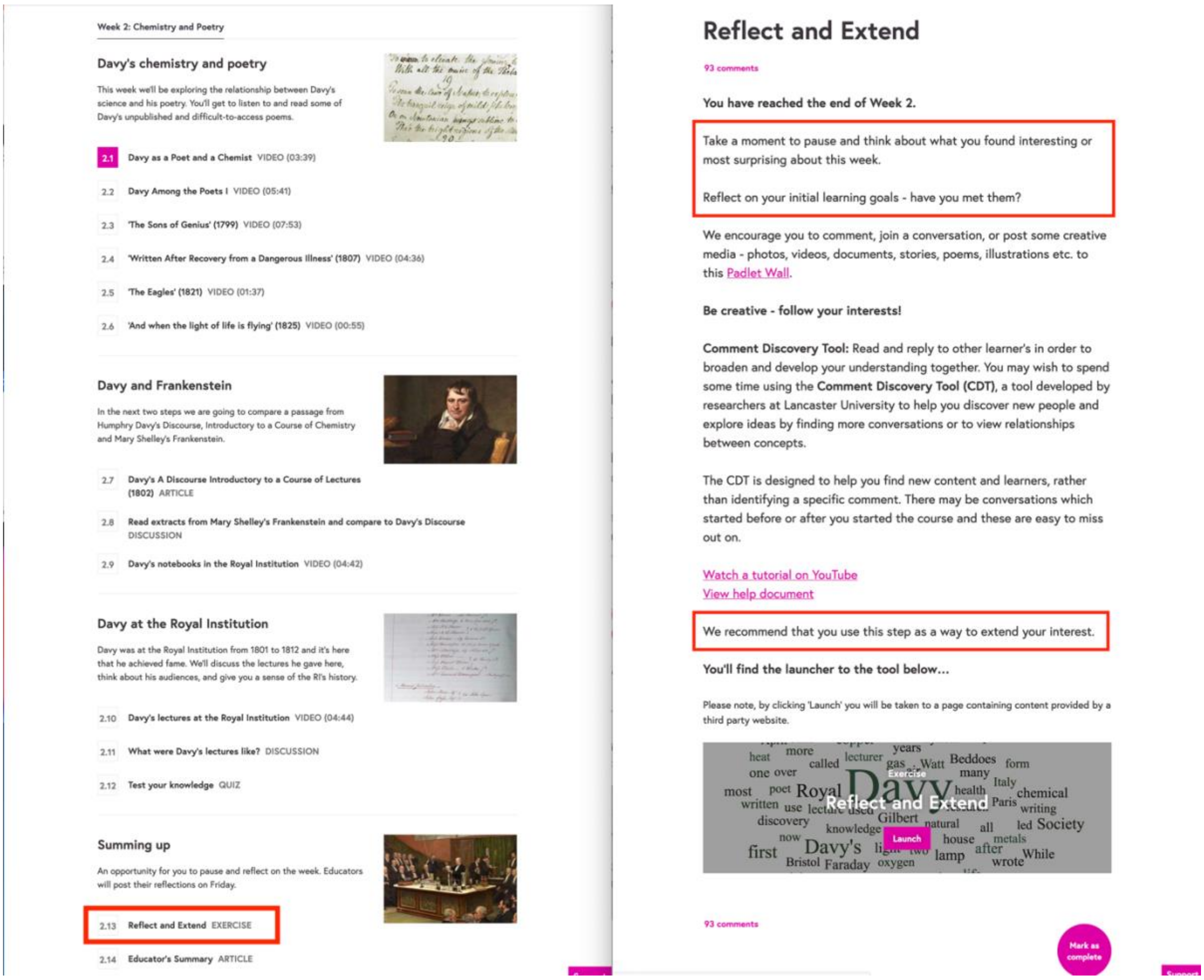


Figure 10 The weekly structure of the course (L) and the task proposal for the 'Reflect and Extend' step supported by the CDT artefact

Step	Description	'Teaching type'
	Learning Activity 1 – Davy's chemistry and poetry	
2.1-2.6	Short video lectures introducing Humphry Davy as a poet and scientist and readings of his poetry	Acquisition
	Learning Activity 2 – Davy and Frankenstein	
2.7-2.8	Reproduced passages with questions inviting commentary	Acquisition, Discussion
2.9	Short lecture about Davy's notebooks	Acquisition
	Learning Activity 3 – Davy at the Royal Institution	
2.10	Short lecture about Davy's own lectures	Acquisition
2.11	Questions inviting commentary about Davy's lectures and learners' own experiences of the lecture format	Acquisition, Discussion
2.12	Multiple Choice Quiz	Practice
	Learning Activity 4 – Summing up	
2.13 [CDT step]	(i) Learners invited to reflect on the week's lessons and how this relates to their own personal learning goals (ii) Learners invited to share their own files/ creative works on a 'Padlet' website (iii) Learners invited to follow their interests by using the CDT as a tool for personal inquiry, to discover new conversations they may have missed or which the platform excluded them	(i) [self-directed] Reflection (ii) [optional] Production (iii) [investigative] Inquiry
2.14	Educator Summary	Acquisition

Table 9 Pedagogical patterns from the example course broken down into teaching types

In DBR1 I demonstrated empirically that commenting mainly takes the form of single ("*Lone*") comments. It is also evident from short form feedback that few participants will read further than the first dozen or so on any given page and there is no encouragement to return to a previous step so will not see what people 'behind them' are writing. These are structural flaws in the platform affecting self-organisation and personalisation, which are important to levelling up the platform to afford interactions which are on the higher levels in the Cultures of Participation framework. The CDT artefact acts as a scaffold for self-directed inquiry and the content of the page is framed around reflecting on individual learning goals. The instruction on the page reads "Reflect on your initial learning goals – have you met them?" which encourages thinking about the "*clarity of the line of investigative questions/enquiries*". I also include a video walkthrough of the CDT which describes how to use the tool to refine and filter through comments[15]. The video walkthrough demonstrates the Humphry Davy course and the words I chose were "women", "science" and "poetry", to emphasise niche interests within the course content (that Davy's lectures were attended by women which would have been unusual in the early 1800's), and I demonstrate practical tips such as how to remove words and select new words. I also chose "Frankenstein" as a smaller word, to filter down to the comments which talk of Davy's influence on this novel; these themes are not explored in detail in the course, so will only be of interest to a subset of learners. In the earlier courses of DBR2 the video could only be a link on the page, whereas in the later courses of DBR2 the FutureLearn platform was developed further so I was able to embed the video inline using markdown coding; this 'affordance' generated exponentially more views, further demonstrating how structural factors of the platform can affect learning.

[15] https://www.youtube.com/watch?v=JKvue5ou1E0 – 7300 views, November 2020

The reflect/ extend activity encourages learners to take an active approach to their own learning goals and the CDT provide some affordances to support the development of the platform, which can be described by the Cultures of Participation framework. In the example week above, as with all FutureLearn pages, learners can comment at every single step, and are even invited to give some commentary in the 'discussion' activities detailed above in Table 9 (2.7, 2.8, 2.11). However, they are never actually required to personalise their reading and writing to their own needs or level of knowledge by revisiting themes or exploring every comment around their specific interests, or pursue an investigative path of self-directed learning, all of which relates to Level 2 of the Cultures of Participation framework ("Users organise content") (Fischer, 2011). The CDT tool within this exercise allows users to engage with a diversity of opinions as it filters the entire corpus into a focused list of comments related to a specific term, surfacing all variations of opinion around selected terms. It has been demonstrated in both DBR1 and DBR2 that most learners will typically use the comments area on a learning step to read between 10-20 comments before moving on, which severely limits their ability to focus in on a range of opinions. This is summed up by an interview participant (P009) in this phase:

> *The whole internationalisation of [MOOCs] is really useful… It is really interesting to hear different viewpoints from around the world and it is often surprising to hear things that you've never thought of before, or in that way… that's what makes the wordcloud tool potentially really awesome. (P009)*

The explicit scaffolding in this exercise/ learning activity proposes tasks which direct learners to take control of their own learning and intends to help them develop skills of investigation and research that are related to self directed learning and theoretical notions on andragogy (Knowles, 1980), which is opposed to the quiz type activity which acts as a test of memory or the passive consumption of video content.

6.3 Quantitative Results from DBR2: counting all the conversations

I conducted the same tests on courses in DBR2 as I did on those in DBR1, which as explained above differs due to the increased exposure to the tool, and the pedagogical scaffolding and attention to learning design principles. It is important for a DBR methodology that the interventions are based in theory so their impact can be reproduced (Brown, 1992), and the subsequent analysis provides empirical evidence for this. Descriptive statistics are seen in Table 10.

	Condition	N	M	SD
Unique Learners	No CDT	225618	1.33	0.80
	DBR1	31621	1.46	0.91
	DBR2	31587	1.42	0.92
Conversation Length	No CDT	225618	1.48	1.43
	DBR1	31621	1.67	1.70
	DBR2	31587	1.59	1.66

Table 10 Descriptive statistics; DBR2, quantitative analysis

6.3.1 ANOVA analysis

Conversations were analysed using the same variables as in DBR1 (unique learners and conversation length) using an ANOVA comparison of means. However, the extended dataset allows for the analysis to combine both phases for comparison against a no intervention condition, and just the new DBR2 data against the no intervention condition (Table 11).

		Unique Learner	**Conversation Length**
		No CDT	No CDT
DBR1	ANOVA	$F(1, 257239)=496.265$, $p=0.001$	$F(1, 257239)=601.703$, $p=0.001$
	Cohen's d	0.15*	0.12*
DBR2	ANOVA	$F(1, 257205)=285.470$, $p=0.001$	$F(1, 257205)=146.273$, $p=0.001$
	Cohen's d	0.10*	0.10*
Combined	ANOVA	$F(1, 288826)=743.680$, $p=0.001$	$F(1, 288826)=508.509$, $p=0.001$
	Cohen's d	0.13*	0.10*

Table 11 ANOVA and Cohen's d results across all conditions; DBR2, quantitative analysis

*A Cohen's d score of less than 0.2 denotes a small effect size

In all cases both variables were significant, and the Cohen's d score which measures impact demonstrates slightly less impact in DBR2 than in DBR1, although there is still a small but noticeable impact. The following analysis examines all the conversations according to the heuristic categories of conversation types.

6.3.2 Heuristic Analysis of the conversation types

All the conversations were also categorised against the heuristic categories (Table 12) which relate to unique learners (Lone, Watercooler, Cocktail Party, Conference) and also to turn taking (Lone, Q&A, Limited Social, Extended Social), and which as described in the previous chapter demonstrates the development of the "initiation, response, feedback" conversational patterns (Littleton & Mercer, 2013) and creates categories relating to diversity within a conversational unit. The difference from the condition of no CDT is displayed in parentheses next to the result in Table 12:

		No CDT	DBR1	DBR2	Combined
Unique Learners	Lone	78.15%	71.06% (-7.09)	74.35% (-3.80)	72.71% (-5.44)
	Watercooler	15.08%	19.16% (+4.08)	16.40% (+1.32)	17.78% (+2.77)
	Cocktail Party	6.75%	9.76% (+3.01)	9.20% (+2.45)	9.48% (+2.73)
	Conference	0.02%	0.02% (-)	0.05% (+0.03)	0.03% (+0.01)
Social Dimension	Lone	78.15%	71.06% (-7.09)	74.35% (-3.80)	72.71% (-5.44)
	Q&A	14.39%	19.53% (+5.14)	18.33% (+3.94)	18.93% (+4.54)
	Limited Social	3.29%	3.88% (+0.59)	3.20% (-0.09)	3.54% (+0.25)
	Extended Social	3.63%	5.53% (+1.9)	4.11% (+0.48)	4.82% (+1.19)

Table 12 Breakdown of conversations into heuristic groupings; DBR2, quantitative analysis

6.3.3 Discussion of quantitative analysis

These results demonstrate that whenever the CDT is present there is a modest 'bounce' in sociality but increasing the number of times it is made available in the course does not dramatically increase the levels of overt sociality. This may suggest that the links on the CDT which allows learners the ability to interact with each other are used by a similar proportion of learners (those predisposed to 'join in') in DBR1 and DBR2, which is not to say that the CDT is not fulfilling its promise for discovery, moreover that the discovery elements cannot be solely understood through quantification of all conversations. Furthermore, the CDT is not *specifically* designed to increase overall replies, rather for grouping comments by affinity and giving the option of individualising the learning experience. Therefore, any reproducible change is arguably good enough to demonstrate that the intervention is having a similar impact. A qualitative method is required to discover how the cybernetic qualities of the CDT increases personalisation for vicarious learning strategies, which extends these quantitative findings.

6.4 Analysis of survey responses

6.4.1. CDT: tool or support for learning design?

In DBR2, each course linked to the same survey instrument as courses in DBR1, and 727 participants completed it in full, which is more than double the responses in DBR1. There is a significant difference from DBR1 in terms of the number of participants who used the CDT "2-5 times" which relates to 'once per week' (i.e., when the CDT step is explicitly surfaced in DBR2 within a 'reflection' activity as described above). In DBR1, only 30% of respondents used the tool in this 'once per week' range, but in DBR2 this increases to 50%, which indicates that learners generally follow the linear pathways encouraged by the platform structure, validating the cybernetic analysis and the importance of using pedagogical patterns for creating active learning. There is also a decrease in students who are using the tool more than once a week (6-9 times, or 10+ times) from 18% to around 5%, indicating that framing the exercise around reflection and self-directed discovery rather than technology results in learners focusing on the activity, rather than seeking out the tool as a novelty piece of technology, and this is a significant finding in terms of learning design in the context of the whole course. This is also demonstrated in the free text comments on the 'reflect and extend' pages in that less than 20% of comments are related to technology *per-se* with the remainder providing a commentary on the process of reflection and self-directed learning, such as what they found surprising or interesting in the week's content. This commentary acts as cooperative production as it is subsequently read by learners who follow behind.

This enables the research to focus more closely on the CDT artefact as a support for the epistemological and cognitive dimensions of learning and demonstrates how the activity can be contextualised according to the pedagogical approaches of an individual course. For example, educators could deploy the artefact at the beginning of a week's study, as a tool for looking back to brainstorm ideas for the upcoming content, or alternatively it could be deployed at the end of the week in order for students to revisit content that they have already interacted with in order to write a synopsis of their learning, as is the case in this study. FutureLearn scaffolds a linear experience and learners do what is asked of them on each step; in the case of the DBR2 courses and this now includes a period of reflection and investigation at the end of each week which is a type of activity is not encouraged by the default platform which values progression or step 'completion' above everything else.

6.4.2 Spearman's Rho correlation coefficient

Spearman's correlation coefficient is a measure of the strength of monotonic relationships between paired data where 1 would be a perfect positive correlation, 0.6 would be considered strong and 0.4 would be considered moderate (Weir, n.d.). In DBR1, the strongest relationships were between learners who perceived the designed affordances in terms of discovery and experiencing conceptual changes (=0.7) and these results are replicated in DBR2 with a 236% increase in survey completions. This reinforces the finding of DBR1 that discovering new conversations and engaging more with others is strongly correlated with sociocultural learning, and that the CDT facilitates these interactions by creating increased capacity for individualisation and personalisation of the learning environment, which relates to Level 2 ("users organise content") in the Cultures of Participation framework (Fischer, 2011).

There is increased exposure to the CDT in DBR2 and the 'number of times the CDT is used' question is correlated to factors relating to learning (developed thinking, conversation discovery, commenting more) by an average factor of 0.45; this is an increase from DBR1 (where the average correlation is 0.39). The result is more significant than the 0.06 difference may indicate due to the increased population size of DBR2 and increase in weekly use which could have reduced the correlative effect. That is to say, increased exposure does not necessarily have to correlate to increased learning value if the learning step is perceived as incomprehensible or not useful; however these results demonstrate that the 'Reflect and Investigate' exercise makes sense to learners and that they are able to use the CDT effectively to produce something valuable to their learning needs.

The learning preference where social interaction is valued for developing thinking also demonstrates similar results to DBR1, in that there is an average of 0.35 in correlation between valuing social interaction in the discussion areas, and ability to use the designed affordances of the CDT for learning (developing thinking, discovery of conversations, commenting more and overall usefulness). This reiterates the finding from DBR1 that the CDT is perceived a useful extension to the social affordances, but it does not build on this finding significantly. However, the difference in population sizes between the 2 studies combined with both populations consisting of different individuals increases the validity of the claim that the CDT is a useful extension for a broad range of learners who find similar value in peer produced content which increases with further use.

In summary, whilst these results validate the claims from DBR1 that the CDT is perceived as a useful extension of the social features, the interactions themselves need to be broken down further into behaviours which are theoretically related to cognitive change, for example the relationship between the CDT and vicarious learning or increased participation. It is the intention of the semi-structured interviews to begin this analysis.

6.5 Moving beyond counting conversations

A challenge for overall quantitative studies on MOOCs is that it does not adequately describe the richness of individual experience, rather reduces it to an average. The quantitative analysis above counts each comment as the same, creating an ‘average’ or ‘mean’ which is useful when drawing high level comparisons between CDT and no CDT conditions but it does not paint a nuanced picture of learners’ diverse interests and motivations. In the MOOC context, learners are necessarily self-directed, meaning that they are intrinsically motivated towards meeting unique and self-defined outcomes and will develop individual learning strategies towards that end. 50% learners never make a comment but many of them will read others’ contributions; most commenters only make new comments, as evidenced by the high proportion of ‘Lone’ comments in the quantitative analysis; fewer again will actively interact and reply, although this is most strongly associated with cognitive change, as evidenced from the survey results. In terms of course level interactions, some learners will progress to completion whilst others will merely sample and this has been demonstrated by quantitative analyses of the Coursera platform (Kizilcec et al., 2013) and the FutureLearn platform (Ferguson & Clow, 2015).

The challenge for the quantitative analyses of conversations that I have developed is that it makes an average of this contradictory activity, which is of limited descriptive value in terms of the *experience* of learning at scale, although it is able to give a broad indication of the impact of the new toolsets I have developed. Comparison of means analyses such as ANOVA or heuristic groupings have a limit in terms of what they can explain about individual interactions with the CDT as an extension of the discussion features and how this affects their strategies for learning.

> Activity is key, but activity cannot be designed. Rather, design for learning has to work indirectly by proposing tasks – suggestions of good things to do – which may stimulate and otherwise influence the real-world activity that eventuates, but which cannot prescribe or actually generate that activity. In other words, there is no mechanism linking proposed tasks directly to learning outcomes; the relationship is mediated by activity. This necessarily indirect relationship also applies to other components of design, such as the design of learning tools and other kinds of learning resources – which physically situate learning activity – and proposals to people about how they might best work together – which are part of what makes learning activity socially situated. (Goodyear & Carvalho, 2016, p. 221)

The beginning of this chapter describes the activity in which the CDT is embedded, proposing reflective and investigative tasks. The quantitative evidence which followed was able to describe to some degree the outcomes of this learning design but the missing piece in this analysis is the activity of the learners, as described in the quote above. The remainder of the analysis in this chapter investigates the activity of the learners, and this will be able to determine how well the task itself is defined and interpreted, and how successfully this translates into desirable emergent activity. This is a complex task of evaluation, and may not result in coherent answers, as active learners can contribute to the learning design themselves by explicitly suggesting ways in which the tools should be used or by giving personal reflections on how the task is interpreted, thus altering the physical, social and epistemological situation, and therefore the emergent activity and outcomes (Figure 11).

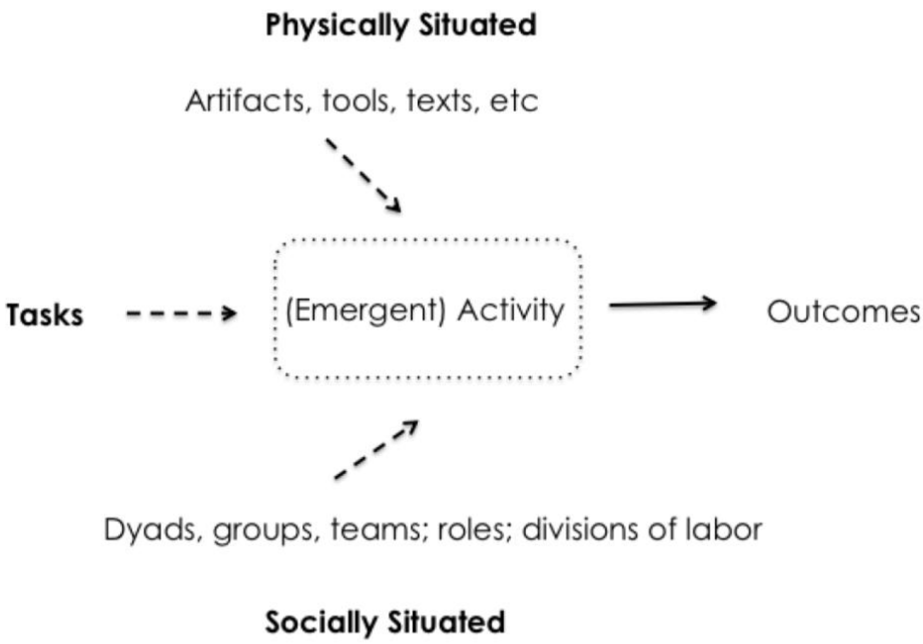


Figure 11 The ACAD Framework,(from Goodyear & Carvalho, 2016, p.221)

Tasks in the epistemic space need to be mediated through affordances in the physical and social spaces, so I focus on discovering where the task can be improved to surface the affordances of the tool (in the physical space) to more learners and to affect the social situation such as the divisions of labour or group formation. Analysis from the previous chapter suggests that the 'social context' is not entirely cooperative as most learners view their learning as an individual pursuit, with others' comments as additional resources rather than opportunities to interact and build knowledge. The aim of this project is to move MOOC pedagogy towards community-based activity (cooperation and collaboration), through reconfiguring the design of the physical environment and epistemological situation (the task propositions) towards a participation based pedagogical approach.

6.6 Experiences of "Learning" in a "FutureLearn crowd"

I conducted 10 semi-structured interviews to gather rich descriptions of learners' interactions, in order to precisely connect these experiences with concepts of learning, and to relate these to the task proposals (epistemic situated-ness), sociomaterial context and cybernetic qualities of the CDT (physical context).

6.6.1 Demographics of interview participants

10 Participants (coded as P001-P010) were asked some basic demographic questions to establish age, gender, highest formal qualification, and number of FutureLearn courses taken and the results are displayed below:

- Age: 18-34 (2); 35-44 (3); 45-54 (1); 55-64 (1); 65+ (2); Prefer not to say (1)
- Gender: Female (7); Male (3)
- Highest formal qualification: None (1); High School (1); Undergraduate degree (5); Postgraduate Research (2); Prefer not to say (1)
- Number of FutureLearn courses enrolled: 1 (1); 2-5 (3); 6-9 (2); 10+ (5)

These demographic data demonstrate that levels of comparative experience with the FutureLearn platform are relatively high and most participants have a formal education which is of university standard or higher. This is important to note because the context of the interview is to discover their interactions with the platform and the CDT exercise, so the ability to make comparisons with other FutureLearn courses is important. Also, levels of previous education indicate experience with critical thinking skills required to formulate “[a clear] line of investigative questions/enquiries”, which is required for the deepest engagement with the affordances of the CDT. All participants had encountered the CDT exercise in a course from the DBR2 phase, with the CDT scaffolded at the end of each week of the course.

6.6.2 The interview structure

I chose semi-structured interviews as the most suitable method because it enables me to guide participants somewhat to the topic of interest, without influencing the outcome. The first section of the interview concerns broad experiences with MOOCs or online learning and participants are asked to describe their perceptions of the opportunities and challenges of learning in this context to explain their preferences. Then the questioning specifically focuses on interactions with discussion boards to describe the opportunities and challenges specifically in the social aspects with online learning and learning at scale online. This sets the scene for the interview and creates the space to further explore what might be understood by the question in the survey about the ‘importance of social interaction in the discussion areas for developing your thinking’. I have established that this question requires further definition because the results of the correlation analysis suggest that learners may not be specifically relating to discovery or inquiry behaviours when they answer this question, rather generally appraising the value of reading some peer commentary.

The second section of the interview explores specific uses of FutureLearn and probes participants' workflows when using the platform. In this section, I explore how the features and design of the platform, which I theoretically appraised in the methodology section are enacted in practice. For example, interview participants detail their specific interactions with the discussion features revealing the number of comments that participants read on each page is generally quite low, and supporting the short form comments qualitatively analysed in DBR1. This is important in terms of understanding the capacity for individualization and self-directed inquiry afforded by the discussion area. In relation to perceptions of the discussion area, this line of questioning reveals specific reading and writing behaviours on FutureLearn in relation to how tasks are proposed, which helps develop a richer understanding of how the CDT relates to these tasks (the epistemic situated-ness) through its affordances and the wider sociomaterial context (the physical situated-ness).

The third section of the interview specifically examines the CDT exercise: whether participants watched the introductory video which describes the tool, what words they clicked on specifically and what this revealed to them, and what type of learning activity they most strongly relate to the exercise and their use of the novel tool. This explores how the CDT was used in practice and whether the proposed activity is clear enough to stimulate the appropriate activities that may result in learning. Finally, participants are asked to suggest further improvements to the exercise.

6.6.3 Continuous platform engagement

Learning in MOOCs inhabits the space between professional development and 'leisure' learning (Walji et al., 2016), and FutureLearn continuously engages learners in a manner similar to an online retail experience, through marketing mailshots and recommendations. Learners may either enrol on a course to fulfil a professional need or out of general interest. P006 describes her motivations for study:

> *The first one I did was on dentistry because I'm a medical secretary for a maxillofacial surgeon … I've been a medical secretary for 20 years, but I'd never done maxillofacial - a lot of that is dentistry. So, I decided to do [a MOOC] on dentistry. Then I've just picked a lot of literature and some art and just culture and general interest things really. (P006)*

P006's first interaction with FutureLearn was to prepare for a new chapter in her career, but course suggestions moved the participant from this professional goal towards literature, arts and culture courses which she enrolled on for personal development or general interest; it is open ended whether she will return to a subject to support her professional life. The point here is that FutureLearn is able to foster a 'discovery space' at the widest possible context of the platform (choosing a course) which is not developed further *within courses* because their structure dictates a linear pathway through to 'completion' of predefined content, limiting space for the "natural sequence of inquiry" in the words of Knowles (1980).

FutureLearn acknowledges that learners demonstrate their desire to learn through discovery when they choose a course because they surface options, recommendations, and related collections in the portfolio through marketing efforts, but these options are constrained *within* the courses. There is a dissonance here as this level of discovery at the context of the learning is also indicative of how learners may wish to engage with the content of the courses to further their own self-directed development needs. Indeed, P001 reflected on his use of the CDT that "I liked the idea of serendipity and lining things up, because you never know what two things are going to bump together and cause the spark of interest for you", suggesting that his preference for learning is most aligned with theories for self-directed inquiry, such as Knowles' comments that adults want more than fact-laden lectures and that learning should flow with the natural sequence of inquiry.

All the interview participants had familiarity with online learning, either through participating in previous MOOCs or Open University style formal study, so recognised the benefits of using online platforms to satisfy their learning goals. For example, P005 says "*I feel like I missed a lot of content in my education. MOOCs are kind of filling it in"*, P008 says "*I just sign up when I see them being advertised or when I get an email [from FutureLearn]*" and P006 says that she has "*done loads, I do one all the time*". P001 goes into the most detail about his motivations for study and the fluidity between formal and informal learning and says, "*I'm almost addicted to FutureLearn courses because they keep coming up with some really interesting slices of courses to take and it keeps my mind going*". He continues to give a more detailed explanation of the point P005 is making above (about filling ‘gaps’ in education): “*I’m an engineer and I do project management… We were STEM, not ‘STEAM’… So, any of that kind of work [in arts] I’ve always had to do on my own because it wasn’t part of the core structure for the engineering side of the equation*”. He also says “*I try to teach my daughter that [learning] doesn’t stop with a degree and with the way the world’s working with jobs moving, jobs disappearing, new jobs forming that require people to be more flexible, you need to have a broad interdisciplinary approach to education*” (P001). In other words, he is saying that the skills of learning online will develop capacity to discover solutions to problems using the internet, which will be more important in the future and is also at the heart of Connectivist principles for learning (Siemens, 2005). I also interpret these words in the sense that being exposed to different world views increases your ability to work with others across the globe who see things differently (“new jobs forming *that require people to be more flexible*”) and that exposure to these world views is not necessarily something that can be taught directly, but that it can be *experienced* through learning in a heterogeneous cohort.

Discovery and recommendation systems are leading P001 from one course to another in order to ‘fill the gaps in his education’ through self-directed inquiry but (ironically) on choosing a course, the structure is rigid and the questions are pre-determined. This project aims to navigate the space in-between the open/ ‘smooth’ platforms for self-directed enquiry such as cMOOC platforms, which can be overwhelming for novice users (Mackness et al., 2010), and the rigid/ ‘striated’ sequences of content of the xMOOC platforms such as FutureLearn which emphasise the publisher/ consumer hierarchy (for example Fischer's 'couch potatoes' (Fischer, 1998)). Serendipitous encounters with others to grow those "interdisciplinary approaches" is fundamentally to do with peer interactions, and the CDT artefact promotes peer produced content by making it more available and customisable than is determined by the rigid platform structure.

6.6.4 Strategies for learning

The 'completion criteria' inhibits learning designs by limiting delivery of a fully participative pedagogical approach. It is an intransient feature of fixed online courses that tutors cannot add new content part way through the course in reaction to the dynamics of the current cohort and their emergent interests, as would be required to follow a conversational pedagogic model. Therefore, contributions in the discussion areas are the sole method of extending the course contents. Every participant described the value of learning from others, such as discovering new and surprising perspectives or interpretations, but all acknowledged that it was difficult if not impossible to read the full range of comments so many were lost. Four distinct approaches of engaging with FutureLearn emerged from the interviews, as can be seen in Figure 12 and I have interpreted these as responses to the overarching sociomaterial factors (i.e., the rigidity of the platform in terms of discovery, and diminishing time). The first 2 approaches are passive to the platform placing them in 'time-based silos' and accept content will be 'lost'; 3 and 4 take an active approach to bending either the platform or available time towards their own learning goals or preferences to fully exploit the peer produced content.

Figure 12 The 4 major approaches to engaging with others on the FutureLearn platform

I have described approach 1 as ‘forming a pack’ and ‘learning in-step’ and this prioritises making social connections with the same others who happen to be on the same learning step at the same time. This reduces the massive and diverse cohort to a small and manageable group, who learners get to know by name. These learners go through the weeks in order and at the pre-determined time. This approach prioritises completion, using social connections as a motivator. These learners are usually the first ‘silo’ to arrive at any learning step, so most peer production happens in their wake. As they prioritise completion, they do not ‘look back’ or reflect on previous steps according to individual inquiry which limits diversity or exposure to viewpoints outside the group. In terms of the physical environment (platform affordances and time), this approach encourages surface level reading within their grouping which is a reductive answer to the problems of overload. The CDT exercise encourages these learners to step out of their ‘time-based silos’ and extend their reading, and therefore their ‘pack’ to one based on affinity as well as chronology. Without the CDT activity, this is the most reductive and passive strategy for learning at scale as it actively ignores many comments to prioritise completion.

The most ‘active’ strategy (3) is a ‘read everything’ approach where learners will recursively read each previous step before moving forward, to ensure that nothing has been missed out. This is an arduous task and those who adopt this strategy spend many hours skim reading comments to complete the course ‘in time’. It is a reaction to the physical environment in that it actively requires extensive effort to maintain, and development of a learning method (skim reading). This approach is ultimately unsustainable even for the most dedicated learner when faced with a massive level of peer contributions. The CDT artefact helps this group because it creates an efficiency to the skim reading method; in a sense the CDT was ‘reading’ all the comments and visualising them which aids the skim-reading strategy.

There are 2 further derivative strategies described below which fall in between these approaches: Serendipitous Encounters (2) and Sampling (4). Serendipitous

Encounters involves dropping in and reading a few comments at the top of the page at any given time. It differs from the first strategy as there is no preference for maintaining active social connections with people but is similar in surface reading and valuing progression. The Sampling strategy involves identifying individual learning steps as important and recursively reading those pages which is closer to the latter approach in terms of extensive reading, differing in that progression is not an aim of the sampling strategy at all, rather engaging with peer produced content on specific steps. This can be interpreted as a response to the time variable, in that samplers do not have the extensive time to devote to the course, so take a pragmatic approach to their learning.

Learning approaches 1 and 2 share a surface reading technique and could be described as 'horizontal pathways' in that they value moving through the course to the final step at the expense of looking back and extensively reading; peer produced content is valuable, but not more so than completion. These passive strategies are fluid in that a learner who starts in the first wave (strategy 1) may not find time to learn at the same pace and find themselves adopting a learning strategy more akin to 2. The important point is that both these strategies passively respond to the cues and workflows encouraged by the platform.

Learning approaches 3 and 4 value peer produced content actively resist the 'flow' of the platform cues. These could be described as 'vertical pathways' through the content in that they value the full length of the comments stream, going down the page. They are also fluid in that a learner who attempts to read everything will eventually be overwhelmed by the overload of information and will need to decide whether to bend time or completion. The important point is that the peer produced content is perceived as of equal importance as the published activities and is prioritised over completion.

The strategies are described in more detail in the following sections with extended quotes from the interviews to support the categorisations.

6.6.4.1 Learning Strategy 1 – forming a pack

This strategy is delineated from other surface reading strategies by the active formulation of a specific 'time-based silo' where groups of learners do the course 'in-step'. The problems of overload are reduced by scaling back the interactions to a smaller group of known individuals. P002 describes how she uses this strategy and the value it adds to her learning by comparing 2 courses which she thought were vastly different in terms of the quality of the published content (such as videos etc.) but were both made worthwhile through making social connections with a 'pack' of learners who were grouped through time. This punctuality in terms of engaging with the course in-sync limits flexibility valued by others, however this strategy proves successful in this case as it provides motivation:

[Those courses] were both fantastic because of all the work that the tutors were putting in the comments, answering questions as we went along… That makes it more real and satisfying. I would hate it to have done it on my own [and] it was the other people commenting that [made it worthwhile]. You recognise the names of the people who are working at the same pace as you and you reply to each other's comments as well. (P002)

[That (other) course] was the only one that's been disappointing. The content wasn't good [and there was no tutor involvement]. The only thing that kept me going was ... a fairly constant group of people and they all seem to have much the same kind of attitude to the content, which was scanty, to put it mildly and had a very definite political agenda. That group of people (and we go to know each other's names) from America, New Zealand, Australia … they talked about what they knew about the subject related to where they lived [and it was these other learners] who made it worthwhile going onto the end. (P002)

P002 "tried to work week [by week] … and really keep in contact with the group of people who are moving through at the same pace as me". She quickly realised with learning at scale that even though she had more time in retirement to do the courses, "*there's too much [to read]. I don't know what you're meant to do; I feel I would have become bogged down if I had attempted that [to read all the comments]*" indicating that this is also a fallback position because there is a paucity of available options. She also recognises that "*reading [other people's comments], taking on board what they said, thinking about it in the next step and reading links*" is very important (in terms of learning) which creates a dissonance between the strategy she uses for progression and the value of peer produced content which she also perceives, but feels she must reject.

In the case of the course with the CDT activity, this dissonance is resolved because the CDT is used as a focusing tool, because "*you typically go along with the same group of learners, and then when you did the word cloud, it allowed me to focus my mind on one aspect and read what people had said about it… I picked up things about certain aspects that I might have missed because I was going through week by week with a similar group*" (P002). This strategy for social learning is also articulated by P006 who says: "*at the beginning when everyone starts together, I think that you build up a better relationship [as opposed to] down the line and there are loads of comments and you can't read them all*" and also: "*[I would use the cloud] if there were something that I hadn't picked up enough about during the course or I wanted to know more*" (P006). For these learners, the CDT is a useful scaffold for an activity of investigation and further extensive reading at the end of the week which wraps up the weekly content and gives an opportunity to focus on affinity rather than simply 'keeping up'.

This highlights how designs push and pull at the same time. The default chronological ordering of comments scaffolds a ‘supportive social network’ based around time (i.e., a ‘silo’ of learners and information), but the cost of this is a reduction in the exposure to diverse opinions. These participants present a richer description of the theme from DBR1 which cited time as an invisible sociomaterial factor in terms of the learning environment. It takes on a new dimension in P002’s second quote above (about a non-CDT enabled course) because the small grouping of people “from America, New Zealand, Australia” had “much the same attitude to the content” indicating that it was their shared attitude which reinforced their social connections, rather than exposure to new and different ideas. The creation of time-based silos also can have the effect of *reducing* exposure to new and challenging ideas. However, in both cases, this limitation was alleviated by the filtering affordance provided by the CDT tool supporting an inquiry and reflection exercise.

This strategy is a social method for learning and provides an external motivator in terms of maintaining connections and staying with the pack. However, this strategy is subject to the forces of scale because as more learners enrol on a course, the size of any cohort will increase, and it will become harder to maintain meaningful personal connections. This may increase the sense of isolation in the crowd but would have no effect on learners’ capacity in terms of time to read comments, so brings with it the same problems of overload. Indeed, most interview participants could not maintain the punctuality to regulate their learning in-step with a group of others and used a related strategy which I term Serendipitous Encounters: learning from others' comments but not attempting to connect on a social level.

6.6.4.2 Learning Strategy 2 – Serendipitous Encounters

This strategy does not involve the formation of a group but reads the comments at the time of participation, which is different depending on how much time they have for study and when they can devote this time as well as the general issues of participation overload. P005 quoted Heraclitus to emphasise this point: "No man ever steps in the same river twice, for it's not the same river and he's not the same man", by which she means that the flow of comments has a random and serendipitous quality which is exciting, but not structured or coordinated in a way which would enable structured learning. She said: "That's what makes the wordcloud tool potentially really awesome" (P005) which indicates that through using the CDT tool she was able to perceive a new method of organising peer production around a stigmergic paradigm, or to use another Roman philosopher's quote: "Luck is what happens when preparation meets opportunity" (Seneca the Younger (54-15BC)).

Six participants describe this method of engagement, so it is the most common approach to learning at scale. In this strategy, other participants do not support each other socially or act as a motivating factor for each other but value peer produced content as enriching the course contents. These learners read about the same number of comments as those describing approach 1, so this is also a surface reading/ horizontal pathway through the content, and an accepted part of study that many comments are lost in the overload. Most interview participants reported reading between 3-100 comments (100 being 'the first page'), although the median is at the lower end of this spectrum, depending on interest and time (e.g., P002: "Typically I'd go down 10 or a dozen"). It is a structural limitation of the platform and requires an active approach to bending the platform or time constraints to overcome, so most learners will simply accept this limitation. As the most common form of participation, this approach emphasises flexibility afforded through online study, cited as the best thing about learning online, but it is important to note that this form of participation remains an individualistic pursuit based around intrinsic motivations:

> *[live discussion] is a challenge just because of the volume of people and you get overwhelmed fast. I almost prefer doing MOOCs in archive state, so I can kind of avoid that … But I would say especially in MOOCs that there is an international perspective that's really fascinating - but it is more time consuming. So, but once you engage with it, you realise that somebody over in that corner of the world will say something about their own environment which comes from their own history in terms of reading and [understanding the] content and it's given them a perspective that's completely different than yours. That's always good. It's eye opening. (P005)*

The whole internationalisation of it is really useful. Like, it is really interesting to hear different viewpoints from around the world and it is often surprising to hear things that you've never thought of before, or in that way… [however] the reading is much more extensive that you imagine. Like it says you only need 3 or 4 hours in order to complete the course, but that is [just a baseline] – I mean, in order to really make best use of the limited time [in MOOCs], you have to spend a lot more time. Also thinking about what you want to get out of it, because it isn't necessarily what is in the videos. (P009)

This second comment (P009) emphasises the aspect of time when studying MOOCs, and that the dedication required to complete increases as more comments are added. She continues to say that these emergent resources are often the most important parts of learning, because you don't necessarily want only what is in the videos.

it is possible to do this course really superficially, and probably still get a lot from it just by watching the videos, but as a teaching professional, I have more deeper questions and these are not always sufficiently explained in the videos – they are aimed at a general audience … I would need to read through the comments to see if others agree and if they have links or more to say about that thing (P009)

P009 wants to engage at a higher level than the course materials, so reading peer produced content is 'next level' for her. The platform is not really designed for this approach as she has specific questions that go beyond the video introductions. She wants to engage with others 'like her' to see how they have made sense of the problems. A 'horizontal' strategy of superficially reading and prioritising progression through published course materials isn't enough for her, and she needs something to help her engage more actively with the peer content, whether that is a learning strategy or a novel artefact. Time is an overarching and diminishing aspect of the sociomaterial environment, running parallel to an increase in peer produced resources. As she says in the second quote, peer produced resources provide further links and ideas which help her situate her learning and strongly reflect her desire for a more sociocultural approach to learning. In this case, P009 perceived the CDT tool as a time saver in that she was able to collate all the comments about a particular topic together which increases the capacity for 'individualisation'. She appreciates the instruction to stop, think and reflect, although found some of the words in the cloud too generic. In summary, P009 adopts a Serendipitous Encounters strategy on non-CDT enabled courses, as she perceives this as the only way to 'do' a FutureLearn course, but she perceives this as unsatisfactory and wants to take a more active approach to individualising her learning. Learning strategies 3 and 4 actively bend the sociomaterial environment to their learning preference; P009 is not quite there yet but appreciates the CDT exercise as a method of individualisation without needing to 'hack' the platform's workflows.

In order to increase the range of commentary and resources (universally described as the most useful part of online learning), it is necessary that comments are thoughtfully written and discoverable, so leaving a comment has a dual purpose: firstly to aid the author in concisely articulating their thoughts, and secondly as a cooperative strategy because they want others to benefit. They are not written to maintain friendships as in the first learning strategy, but simply as an attempt to add knowledge or experience to the pool. P001 always looks at the notifications alerts when he opens FutureLearn, "to see if anybody's commented [replied to me] or liked my post". This demonstrates that he seeks the interaction (of random individuals) to learn and improve his understanding on the level of knowledge construction. When he comments, he tries to "move up to the next level", rather than say something like 'that's interesting'" (P001). It is not the people who he is connecting with that is important, but their reflections on his own commentary and he always replies to "close the loop or say something back" which is a form of interaction closely related to levels of knowledge construction that is defined in learning sciences literature (e.g. Gunawardena et al., 1997). The platform's affordances (notifications etc.) for extending conversations as described by P001 above are useful *only if* other people can 'find' his comment, and it is this discoverability which is deficient on FutureLearn, for reasons of 'time-based silo'ing' of content, and also to manage 'participation overload' by de-prioritising older comments.

In P001's description of his use of the FutureLearn platform, he takes a methodological approach and leaves a breadcrumb on the page (or 'Lone' comment) before moving onto the next page. This breadcrumb may get 'liked' or replied to by somebody behind him, sparking a series of interactions mediated through the notifications feature which may result in higher levels of knowledge construction. This is exactly the type of interaction that P001 intended when he left the comment, and is also the type of interactive writing which is strongly associated with learning in sociocultural literature (Lapadat, 2006). It is also clear that the diversity of opinions in the discussion areas is perceived as the most valuable part of learning at scale, so discoverability is extremely important to prevent thoughtful and useful comments being overlooked. There are no existing social bonds to spark interaction, merely the contents of the commentary and it is a limitation of the platform that his comment (or breadcrumb) can only be picked up by a learner who is a few dozen 'comments' behind him in the trail (or linear pathway). The CDT goes some way to increasing discoverability of this comment/ knowledge and P001's description supports the need for discoverability tools.

The emergent interactions of learning strategy are dependent on 2 factors – firstly encouraging learners to leave a thoughtful comment, and secondly making that comment available to as many other learners as possible. I describe this strategy as 'serendipitous encounters' because it is dependent on having interactions which spark a conversation and potential knowledge construction, but as learners are not getting to know each other personally, it is more down to luck whether any comment seen is replied to. FutureLearn has a feature in terms of the notifications alert to keep an individual interaction going, but it does not have the feature set to enhance discoverability from 'luck' to 'coordination', which is the first level of a stigmergically designed system (Dron, 2006; Elliott, 2016).

6.6.4.3 Learning Strategy 3: Attempt to read everything

This approach is the extreme of an active strategy to handle the problems of participation overload and diminishing time. This approach takes an active choice to prioritise peer produced content (as described by P009 above) and also to bend time constraints. An extreme version of this strategy is described by P004:

> *[I read] not 100%, but I would say 95% of them. And what I did was like 'diagonal' reading - which is what we call it in Spanish – like go through the post really fast, and when you find something interesting that you want to focus on, that is when I will slow down and take some notes (P004).*

She also reported that she would study from her workplace and "stayed very late each day. I took 3-4 hours and came back [home] at 1AM, in order to read or listen to everything" indicating that in order to 'beat the clock' she was willing to stay at work until very late each day. There are some obvious challenges to this approach, like P009 above who lamented the value of the increasing peer produced resources was suffocated by the sands of time slipping through the hourglass towards the end of free access. This approach can only ever work with MOOCs of a certain size, and there will be a critical mass when reading everything becomes unsustainable. P004 adopted the method of manually going through each page up to the one she was on, looking for new and interesting comments before moving on, and employing the strategy of 'diagonal reading' or 'speed reading' and note taking.

There are 2 important points to draw out of this description of this learning strategy: 1. the visualisation is created by the CDT 'reading' (processing) the entire corpus, regardless of how large the course is, so creates an efficiency for this approach; 2. That the strategy of 'diagonal reading' is the same activity required for analysing the visualisation to find a term. Indeed, P004 did not perceive the cloud as a filtering or focusing mechanism like many other participants (as she had already read most of the comments anyway!), rather was interested in the way the visualisation is altered on each selection.

> *the word [you have chosen] changes the cloud and it's like everything is moving and changing and narrowing or broadening the ideas depending on what terms you choose – so it takes you to another cloud with a related concept, and the comment [list] changes (P004)*

This demonstrates that P004 had moved beyond merely filtering comments towards perceiving the affordance of the 'concept map', but this came after the realisation that she had read most of the comments anyway through employing her own radical strategy for coping with overload. P004's had already read 95% of the comments so did not need to use the tool for discovery but took time to think about how the words related to each other much like when P001 described how the different words "lined up.. and you never know what 2 things are going to bump together and cause that spark of interest for you".

P004 explored her capacity to speed or scan read ("diagonal reading") which could be related to her ability to use the CDT as a tool for making conceptual connections through analysing the visualisation. This reflection could inform part of the scaffolding of the exercise in terms of modelling how learners should use the visualisation and speaks to a common piece of feedback from the interviews and the shortform comments in DBR1 that learners were not sure what to make of the visualisation when they first saw it.

P004 was the only participant to speak about note taking/ making external notes, which isn't specifically related to this strategy of reading everything but concerns the general idea of reflection. P006 also said that she "went out of her way to comment for the benefit of myself" and agreed that her writing of new comments was reflective and as a way of organising her thoughts, as did P005 who said, "it helps you articulate and refine your thoughts carefully". This is useful to consider when considering an overall learning design for MOOCs and something that could be integrated throughout the whole course. The CDT activity is titled 'reflect and extend', but whilst extension of knowledge through discovery is well cited through all the interviews, the reflection element is not as well supported but is something which could enhance the whole course, perhaps by the development of note taking tools which do not leave breadcrumbs but organise thoughts in a way which can encourage reflection.

6.6.4.4 Learning Strategy 4: Sampling

This strategy is also defined in MOOC literature (Ferguson & Clow, 2015; Kizilcec et al., 2013). P007 describes this approach by saying "I jump around the content to see what is of interest to me, rather than do the course in a linear fashion. Sometimes I read the comments, but don't watch the video, because of the devices I have to hand" which highlights that peer production is perceived of equal importance as the course materials and reinforces comments from other learners about the value of socially produced content. P007 says, "someone will share something random which is just like a little nugget of knowledge that you're never going to read in a book – you know, something about where they come from or something. It's not 'bookish' knowledge, you know, but I find it interesting". In this sense, having the comments below each page is a useful feature of FutureLearn for P007 as she compares using a Coursera forum (which is threaded and separate from the main content).

> *[in Coursera forums] …there are lots of threads listed which is quite random. It is hard to scan down them and see what may be of interest, especially as the course progresses. To be honest it is a little overwhelming*

although there is a search feature to find things which relate to individual topics. In a way that is what I like about the FL forums – they are on the bottom of the pages, so you do not need to find 'topic related' threads. It has its own problems though because a lot of people just write 'thanks' or 'thank you', which dilutes it a little. (P007)

This comment reinforces that the skill of 'speed reading' or 'scanning' are important for efficient navigation of content. Where learning strategies 1 and 2 reduces discoverability to the content which is in and around them at the time of participation (LS1 organising this time to work with the same people), learning strategies 3 and 4 really mine the resources to answer their predefined questions using the technique of scanning in both cases or selection and sampling in learning strategy 4, and this focussed approach to studying enables a deeper recognition of the 'advanced' affordances of the CDT. For example, P007 says "when I got the hang of the 'categorisation' it actually worked quite well for me and actually saved me a lot of time in terms of how I could read and join in with conversations based on my interests" which relates to using the CDT in a 'tiered' fashion, so firstly selecting general words, in order to filter the corpus which surfaces niche words in the second cloud. Learners may not find the term they hope for in the first visualisation and this problem is cited by many learners in DBR1 as the reason they found the exercise challenging or not useful. The first cloud often has very general words and some learners found them irrelevant, but after selecting a general word relating to the chosen topic, many other words will appear as the corpus is shrunk to only comments with the first term, therefore new and more insightful words are surfaced.

6.6.5 Reading and Writing on FutureLearn

6.6.5.1 Writing: sharing is caring

These 4 strategies detail different approaches for interacting with the same platform. LS1 and LS2 take a 'horizontal' approach to learning focused on completion at the expense of depth, where LS3 and LS4 takes a 'vertical' approach focused on mining peer production for new resources to read, at the expense of time (LS3) or other content and completion (LS4). These approaches are differentiated mostly through the activity of reading. There is less differentiation across all learning strategies concerning writing new content, which revolves around 'answering the question set by the page'.

> *In my last course, I skipped [making comments] on a few pages where there were no questions from the course to answer. I prefer to do new comments; I do not reply [to] comments unless people ask [me a] question. (P003)*

> *If the step asks for a comment, I will write a new one, and if I find someone who is talking about the same thing as me, at the level I want, then I will engage in longer conversations with them, and the notification*

feature is good for seeing this. Also, the endorphin boost when you get one saying someone has 'liked' one of my comments, although this is more for the ego than useful from a learning point of view. (P009)

This highlights that the actual writing of comments is directly in response to whatever 'question' has been asked on the page. P001 tries to 'level up' the conversation to something which may spark a conversation as described in the Serendipitous Encounters (LS2) section above, and P005 describes how she reacts to pages where there is no 'question' to answer.

..if the page didn't ask for comments, I would notice the comments section and I would be like: "Oh, everybody has to leave a comment all the time." Like that's so 'modern narcissism' and I wouldn't want to do it. I would look at them though, and maybe I would look the top five just to see what we're talking about. But I wouldn't leave a comment [just because I could]. Although I would if [the page] did prompt us to leave a comment. (P005)

The comments demonstrate that writing is prompted from a question on the page, asking them to share some experiences or perspective, and it is the diversity of this shared opinion that is considered valuable universally. It is necessary to consider learning design in the context of the page of content and consider questions which require reflection and would also expose a diversity of opinion and experiences.

P005 (above) says that she would comment if prompted by the text on the page, highlighting that the ‘stepped’ learning design is adhered to relatively strictly by most participants. An alternative pedagogical strategy might be to ask learners to reply to something which they disagreed with which would increase overall interactive writing although may present problems in the context of social interaction:

> *I'm very careful with who I would engage with on the comments. And I think depending on what it is; for example: antisemitism, because I did that quite recently. I think you have to be quite careful with what you say and also quite careful to react to what other people have said - it’s very important. There was one that I did on mental health in South Africa…. I think that that was one where people for obvious reasons could get very upset in the comments. I'm very careful and I deliberately don't become part of a long discussion if there are a lot of different opinions in that learners should be asked to leave a comment which will expose this diversity. It is best achieved when the question is framed in a reflective manner. (P006)*

P006 is suggesting that the ‘question’ should be framed in a reflective manner, in order to get a deeper account of personal experience but suggests a preference for ‘Lone’ comments rather than engaging in a longer discussion, which is at odds with P001 who left comments which ‘go further’ and which he hopes will elicit a reply. However, both these accounts share the idea that high-quality reflective commentary is key, although they differ on methods of input and resultant interaction. P006 sees ‘answering the question’ as an individualistic task and she identifies strongly with learning strategy 1 (forming a pack), so will mainly interact with learners with whom she has already formed a social connection; P001 does not form personal social connections and identifies most strongly with making random encounters to further his understanding, finding a certain eclectic and serendipitous quality in the randomness. Both acknowledge that high quality writing increases diversity and therefore enriches course content, but their methodological differences will alter the quantitative results in terms of counting and categorising conversations according to overt sociality and affect how the in-built platform tools (such as notifications) enable discovery, follow-up and ‘knowledge construction’ in the traditional learning sciences sense.

This is an important finding because it demonstrates that learners readily acknowledge that novel contributions lead to a greater diversity in peer produced course content, and suggests that the ‘questions’ on the page should be written in a way to encourage the types of interactive writing which are demonstrated both theoretically (Lapadat, 2006), and empirically through the results of both surveys (DBR1 and DBR2) to be important for learning. The sociomaterial qualities of the platform do not inherently encourage or support writing because ‘Lone’ comments are susceptible to being lost in the overload and engaging with existing commentary increases discoverability (through notifications) but comes at the risk of offending someone, as described by P006 above. The CDT helps with discoverability, but on its own does not create a culture of cooperative writing which is open for responses. Extensions to the CDT may improve cooperative writing through the folksonomic qualities of hashtags, but this culture of writing needs to be developed through the course in the way that questions are posed on the page. Some options could be to encourage learners to end their own comments with a question to invite further interactive writing, and to emphasise throughout that the course should be a ‘safe space’ for all to learn and talk without prejudice or *ad hominem* confrontation.

6.6.5.2 Reading: Horizontally or Vertically

The interviews demonstrate that reading is more common than writing, although learning analytics derived from clickstream data will not be able to quantify this, in line with the findings of Veletsianos et al. (2015) who report that learners make off platform notes, combine notes from other MOOCs and spend a lot of time reading, all interactions important for learning but not represented in clickstream data.

Time is an all-encompassing sociomaterial dimension and affects all learning strategies in terms of a diminishing resource, making learners choose between surface level interactions, which prioritise progress, or deeper level strategies which view all peer production as valid content and present the choice between devoting unlimited time to the course or the sampling of specific content. I term these differing reading strategies as ‘horizontal’, meaning surface level and moving towards a completion, and vertical which represents mining ‘down’ the comments list. This metaphor is also consistent with P004, who called her reading strategy as 'diagonal', representing mining the comments stream downwards (vertical) for maximum exposure to content, but skim reading to select appropriate content (horizontal). Taking this metaphor further, the CDT affords a 'diagonal' reading strategy in skimming the visualisation to select words, but that the vertical aspect of this (analysing everything) part is done by the computer, which is a time saver.

This suggests that from a pedagogical perspective, the activity can be scaffolded into the course in terms of defining these strategies more specifically and leading the learner towards adopt a more active learning strategy. However, learners must reflect on what they want to know more about before they use the tool for most impact, so the task of reflection should be scaffolded throughout the whole week. This itself must sit on the individual pages asking appropriately articulated questions in order that learners will write high quality reflective comments which invite further interaction, sparking new ideas and a reflective mindset. When these 3 things work together then learners will be able to make best use of the artefact by using higher order cognitive skills (reflection, interaction, discussion) and then be able to 'mine' this content in a time-efficient manner.

6.6.6 Suggestions for improvement to the working of the CDT

The third section of the interview explores participants' use of the FutureLearn platform and probes specifically into the ways in which they had used the CDT exercise. I asked if they had any suggestions for improvement, either of the exercise (the discovery and reflection activity) or the technology (the CDT tool). P007 asked: “is it possible to @ someone who is not in the thread, but you think they may be interested” which is a general point about the FutureLearn platform and not something that could be affected with this project, but highlights that she found the discovery elements lacking and also that she took interest in the people who were commenting, not just the comments themselves. P002 also asked “[is it] possible to actually give the names too?” which re-emphasised that she used the platform to ‘find her pack’ (learning strategy 1) and wanted the CDT tool to aid in this respect too.

6.6.6.1 Too many irrelevant words

A majority (6/10) of participants commented that there were too many irrelevant words in the cloud, for example P007 said “it was disappointing to see so many boring words in the cloud” and asked if it were “possible to show related concepts [close together in the visualisation]” (P007) which highlights that she found the process of scanning the page for a relevant term arduous. P009 gives a clear use case of this problem and sketches a way out of this problem (she used the tool in the Dyslexia for Foreign Language teaching course):

> *Many of the words were not relevant – they were very generic, and it was hard to find what I was looking for. In my case it would have been useful*

to see words like ‘phonological’, or ‘tokens’, or ‘pronunciation’ as these all relate to learning through sounds, but I didn’t see this – the only word I saw was ‘sounds’ which is the generic form of this concept. (P009)

In this description, P009 could not find the specific technical terms she was looking for but a more generic version of the term she was looking for (“sounds”) did appear, most likely because many of the people who spoke of the other technical terms *also* used the word ‘sounds’ in their comment. This relates to the ‘tiered’ use of the CDT which a minority of participants perceived. If P009 chose ‘sounds’ as her first term, it is very likely that the other technical terms would appear in the secondary cloud. P007 (above) says that when she “got the hang of the ‘categorisation’ it actually worked quite well for me and actually saved me a lot of time in terms of how I could read and join in with conversations based on my interests”. This feature is included in the video explanation of how to use the tool, but not entirely obvious from merely looking at the page.

P001 suggested another method out of this problem might be through use of tagging:

I was wondering if there was a way to integrate tagging, like coming up with a concept tag. Is there a way that you could have categories by class and just have that be the theme? (P001)

In this way, the ‘uninteresting’ words could be filtered out as they are not ‘tagged’. P001 leaves it open as to whether the computer could aggregate comments and apply a theme or whether users should be asked to self-tag their comments. I intend to develop the tool to display folksonomic terms prefixed with a hashtag (#) so I will suggest to learners that they add a tag or two to each comment they write and then visualise the tags, which will hopefully relate to topics or themes.

Indeed, both the ‘tiered’ approach and the ‘tagging’ approach are valid methods for understanding how to make the CDT more intuitive and are non-exclusive, meaning they both could be employed. However, both methods rely on a clear introduction to the task and the tool.

6.6.6.2 A clear introduction to the task

P005 says she “had no intuitive idea what to do, even though I'd seen word clouds before, I don't remember interacting with them a lot… mostly they're just like a graphic so at first maybe it didn't even occur to me to click on it.” P004 went even further than this and described how her colleague who was studying the same course in parallel “thought she might break something” as it was such a novel technology. P004 suggested to tell people that it isn’t possible to make a mistake, and to just play around with it.

This highlights that there is a disconnect between the technology and the task: the explicit scaffolding on the activity page is not clear enough to propose how using the CDT might aid with reflection, discovery and inquiry. In later courses in DBR2, I changed the wording of the reflection task slightly, to ask learners to say which word they clicked and what this revealed:

> Activity: Which words did you click on? What story does this tell? Did you choices reveal surprising connections or help you focus on something specific?

This simple change enables cooperative commentary, as it asks learners to scaffold the task for those who are 'behind them', speaking directly to the concern that of the intuitive-ness of the task.

6.6.6.3 Further technical suggestions for the CDT

Another technical change emerged the interview with P007 wort carrying forward into the next iteration of the CDT is to to 'locate' the conversations or comments in terms of the pages or conversations which they came from, and provide more detail on the thread, such as how many learners were involved. P007 said: "[I'm not sure] …whether the tool takes you to the first comment on the thread or jumps straight in where the keyword is mentioned, because often ideas develop over the course of the thread."

It also became apparent whilst doing the interviews that learners really valued the comments left by mentors and staff members, so I intend to implement a further feature to highlight those comments made by the teaching team and give a clear indication of the learning step.

6.7 Summary of findings from DBR2

This chapter takes a deep view on pedagogical factors relating to learning design in terms of the 'proposed task' and examines the whole week of study and not just a single page with a novel piece of technology. The task is proposed as one for reflection and extension of understanding and complements the other learning types of the week, which are mainly acquisition or practice tasks (C. Young & Perović, 2016). All pages invite learners to write some commentary, but this is quantifiably demonstrated in DBR1 to consist of 'Lone' comments which do not show development into new knowledge through dialectic (Tubman, Benachour, et al., 2019; Tubman et al., 2016; Tubman, Oztok, et al., 2019), the usual marker for discursive learning in the learning sciences literature (De Wever et al., 2007; Gunawardena et al., 1997; Scardamalia, 2002).

DBR2 quantitative analysis replicates the findings of DBR1, but this only adds empirical data to how overt sociality is affected by the exercise and the CDT when the unit of analysis is a large grouping; it does not provide rich descriptions of how individual learners interact with the platform, or with the CDT activity. A series of semi structured interviews provide this rich qualitative data. The interviews revealed that there were 4 main learning strategies for interacting with the platform overall, with each strategy resolving one problem of learning at scale but uncovering others. For example, the first strategy is to learn in a small pack of those who start at the same time. The social interaction provides motivation for completing the course at the expense of flexibility and exposure to a diversity of viewpoints. A variation is to concentrate on the knowledge, increasing flexibility and self-regulation of study, but again losing the ability to interact with a diversity of viewpoints, leaving the exposure to chance or serendipity. There is a certain eclectic quality to random exposure to commentary, but it restricts the capacity for learners to take an active approach to their goals and individualise their learning pathways. The 'active' learning strategies (LS3 and LS4) bend the sociomaterial factors to their will by 'beating the clock' or disregarding the progression criteria and merely sampling content. The active strategies place an equal value on peer produced content as the teacher provided content.

All interviees report a similar approach towards writing content, in that they 'answer the question on the page'. This could be exploited in terms of the questions which are asked on each page, and in terms of developing cooperative behaviours on the intervention exercise itself. In later courses during DBR2 I made a change to the wording for the activity and this appears to attract a different kind of writing, so this finding can be carried forward into further iterations of the exercise (and technology which mediates it).

Reading strategies can be divided into 'horizontal' and 'vertical' where the former describes a surface level reading of a few comments before moving to the next step, and the latter describes reading the whole stream to the bottom and returning to view new content in the case of LS3 (read everything). A parallel can be drawn between the vertical reading strategy and the CDT technology, in that the CDT can 'read everything' without exception, and represent it in a visualisation, so the task of 'skim reading' is altered to look for words in a visualisation, rather than across the commentary. This parallel should be explicitly scaffolded into the activity, to emphasise the methodologies of using the CDT (also using the 'tiered approach' to term discovery) as presented as a 'time saver', which would also help the horizontal readers who accept a certain number of lost comments.

Chapter 7: Findings from DBR3

7.1 Introduction

DBR3 builds on the finding from DBR2 that learners should write in a cooperative style, to encourage formation of interest groups, and explicit scaffolding of CDT use, and I draw on the concept of 'Stigmergy' as a framework in order to describe the development and exploitation of this 'swarm intelligence' (Elliott, 2007; Van Dyke Parunak, 2006). Throughout DBR1 and DBR2, learners often reported that words in the visualisation are 'irrelevant', which I interpret to mean that they do not relate to clear concepts relevant to the course. The pretext for this type of comment could be that 'irrelevant' linking terms edge out more 'niche' conceptual terms which do not occur frequently enough to be included. This presents a challenge for the designer, which I attempt to overcome by steering all writing behaviours towards a 'cooperative' endeavour. The method I use to do this is by encouraging learners to use hashtags in their commentary, which act as 'searchable text' (Zappavigna, 2012) and aids in the development of communities of 'ambient affiliation' (Zappavigna, 2011). In this respect DBR3 focuses on the learners' behaviours to examine how this can be harnessed to affect the overarching research question of how to make MOOC pedagogy more participatory, especially given the challenges presented by scale.

I made changes to the CDT's codebase (henceforth: CDT 2.0) so hashtagged terms are visualised into a specific area of the CDT tool. The folksonomic and bottom up nature of hashtagging means that learners can act collectively in a way which adds value to the community itself by 'extending the range of the environment', thereby acting on the highest levels of the Communities of Participation framework (Barricelli et al., 2015; Fischer, 2009). This draws heavily on the concept of Stigmergic collaboration and design (Dron, 2006; Elliott, 2016) which is described in more detail below, and has the practical side-effect of reducing 'irrelevant' terms because learners are most likely to tag conceptual terms rather than the 'linking' words in sentences; it is these 'linking' words (e.g. 'how', 'way', 'well', 'got', 'may', 'often' etc.) which stuff the original clouds and squeeze out more niche conceptual terms (e.g. #dyspraxia or #dyscalculia in the Dyslexia course). The CDT 2.0 has a tabbed structure, so the 'original' clouds are not removed or changed, but new 'weekly' and '#hashtags' tabs are added, shown in Figure 13:

Figure 13 The CDT application with the 'tabbed' structure, including hashtagged terms

7.2 Cooperation and Stigmergic design

7.2.1 Collaborative 'interthinking'

In order for an activity to meet the criteria of collaboration in a sociocultural sense, all members of a group must have a common understanding of the problem or object of inquiry from which to build or construct new knowledge (Vygotsky, 1978). Littleton and Mercer (2013) describe several collaborative scenarios in their book "Interthinking: Putting talk to work", from jazz musicians 'jamming' in rehearsal spaces to school children working in small physical and online group environments and develop the concept of "interthinking" and collaborative activity. They offer insights as to why collaborative activity isn't always productive or successful, by distinguishing between concepts of 'cumulative', 'disruptive' and 'collaborative' talk. Successful collaborative talk occurs when all members of a group share the same foundational understanding and listen to each other as much as they contribute. In large groups, this is much harder to orchestrate, although a limited form of collaboration is achieved in the case of 'lightweight' peer production (Haythornthwaite, 2009). This collaborative activity is achieved by narrowing the shared foundational concepts and format of participation, for example, Haythornthwaite cites 'NASA clickworkers', who decide what an object in a picture might be from a small range of possible options and a decision is made by an AI algorithm choosing the most likely of many responses.

In a MOOC, learners' contributions take a freeform format and cannot be restricted to the same extent as Haythornthwaite's examples of 'lightweight peer production'.

However, I describe below how this idea of 'narrowing' options may also make sense to enhance the effectiveness of hashtags, and avoid 'splitting' terms through spelling or choice of term. The previous phases of this project have demonstrated quantitatively that peer commentary mostly takes the form of 'Lone' comments and that FutureLearn's attempts to manage 'participation overload' through pagination restricts the ability to find and reply to others in a sense conducive to self-directed study. Learners are effectively placed into silos based around the time they encounter the step, and I have demonstrated in the previous chapter that this is the most defining feature of both modes of passive participation. This strongly points towards contributions being 'cumulative' in Littleton and Mercer's terms for 2 reasons: firstly, most interview participants reported that they read no more than 20 comments before making their own post; secondly, I have quantitatively demonstrated throughout the previous phases that the overwhelming majority of posts are 'Lone' or 'original posts'. This suggests that most posts are not synthesising points made in a thread of conversation, rather they are expressing the author's original opinion which is at best guided by the most recent 20 or so posts. In this sense, Littleton and Mercer's terms provide useful language to contextualise how changes to the discovery and writing affordances of the platform could encourage 'collaborative' talk, measured through the ability to 'listen to a wide range of opinions' (i.e. discovering more than the latest 20 posts), and contributing either to the conversations directly, or creating a new post based around that initial inquiry activity.

7.2.2 Stigmergic design: distinguishing coordination, cooperation, collaboration

In larger contexts of co-production, changing the cybernetic dimensions of the system and encouraging different forms of writing through the pedagogic scaffolding are the most appropriate interventions to steer user interactions towards a model of collaborative 'interthinking'. Littleton and Mercer's terms are useful for describing less productive collaboration and to understand how the modes of participation driven by the platform encourage 'cumulative talk', but we also need a conceptual framework to describe how collaborative activity works functionally to alter the learner behaviours towards this. Stigmergy provides a useful theoretical framework to break down the collaborative activity into its components and acts as a useful design paradigm to draw on to understand how massive peer production can be leveraged towards productive collective behaviours. Mark Elliot defines stigmergy as "a form of mediated communication where signs placed in the environment by agents serve as stimuli to other agents to further transform the environment" and in a learning context "stigmergy helps explain how collaboration scales from small group settings to large online communities, thereby shattering the 'glass ceiling' of face-to-face collaboration" (Elliott, 2016). These principles ensure that platforms can be designed to grow sustainably through peer co-production.

In his earlier work defining a theoretical basis for stigmergic collaboration, Elliot distinguishes between the related concepts of 'coordination', 'cooperation' and 'collaboration', placing them in a hierarchical arrangement where coordination is a necessary condition for cooperation, which in turn is a necessary condition for collaboration (Elliott, 2007). In smaller groups, it isn't always necessary to explicitly make this distinction as the social dynamics and roles adopted by group members will ensure that coordination of material and cooperation between agents makes for a successful collaborative activity.

However, in the context of this project, I have demonstrated through previous phases of DBR that FutureLearn does not encourage longer form conversations, and in DBR2 I demonstrated that the inability of the platform to coordinate peer produced content has consequences for engineering serendipitous encounters and self-directed learning. I demonstrate that the CDT activity can increase the access to new content which is selectable by learners' affinity and interview participants described the positive impacts this had on learning at scale, both for wider reading and increased interactive writing. Framed in the language of stigmergy, this can be explained as managing the problems of participation and information overload by focusing on the individual components of collaborative activity and designing an artefact which offers a coordination affordance, functionally achieved through the computational analysis of words in the corpus. There are challenges in this method because language is separated by brute force into the component words and visualised according to frequency; this itself creates a certain amount of dissonance and false positives which require interpretation to leverage learning. This becomes a barrier to participation in the activity and some learners are not able to perceive the positive affordances of the artefact, rather seeing it as a confusing cluster of 'random words'.

DBR3 refines the coordination affordances. This could be achieved in several ways, for example through further brute force computational linguistic analysis, or corpus linguistics, but this would not alter the pedagogy towards collaboration, nor is it likely to require any less interpretation of which computationally determined terms hold the most importance. I have chosen to draw further on the stigmergic framework of social activity and have developed the coordination affordance of the CDT exercise to leverage cooperative peer participation across all the steps of the course, thus becoming a pedagogical as well as technical intervention. The intention is also to level up learners' participation in terms of the Cultures of Participation framework by affording them the agency to define and search on important themes. This is primarily achieved by encouraging learners to use 'searchable text' (Zappavigna, 2012) in the form of hashtags in all their comments, which attempts to foster a culture of cooperation and make learners think about the others on the course, and to act in a way that creates sub-communities of 'ambient affiliation', which is the starting point for future collaborative activity. In other words, hashtags are a starting point for the participatory pedagogy on a MOOC scale rather than an end in themselves.

Building on DBR2, this phase reconceptualises writing within the CDT activity as cooperative. This is achieved by 'asking a direct question' as the findings of DBR2 suggest that when a direct question is asked on a learning step, learners will respond with a direct answer. The questions on the CDT activity asks learners to both describe the process of using the CDT for that particular course ("which words did you choose and why?") and expand on the results of using the CDT for learning ("what did you find surprising or new through this process?"). By taking this cooperative approach, the CDT can continue to provide the basis of an investigative exercise, but learners model their use of the tool for those who do not perceive the affordances, drawing on a key finding of DBR1 that "[the CDT artefact's] value will derive from the clarity of the line of investigative questions/enquiries." Taken together, these 2 new features of the course design are intended to level up MOOC platforms and pedagogies according to the Cultures of Participation framework to give users the agency to create an informal and bottom-up classification system for course themes and the means to demonstrate to others how to create critical and investigative questions.

7.2.3 What is cooperation?

A traditional definition of cooperation in a small group activity system is for agents to play bespoke roles to accomplish a defined task, and this is a form of collaboration due to the shared understanding of the defined task. A much cited example is the activity of a beater taking part in a primaeval hunt, described by Leont'ev (1981). The individual activity of the beater alone does not satisfy his need for food, but as part of a cooperative endeavour, his actions are a necessary condition for the overall activity. Therefore, the actions of the beater are not sufficient for the activity of hunting, but in combination with different actions from different actors, the hunt can be realised. This definition of cooperative behaviour (i.e., bespoke roles for different people within an activity system) is a top-down description of the concept of cooperation. In larger activity systems such as managing traffic on the road, cooperative behaviour (stopping, giving way, overtaking safely) refers to a general adherence to the highway code, rather than individuals taking on bespoke roles to accomplish a shared task. It is not through performing different roles that the activity is fulfilled, but only when all agents always perform all the roles appropriately. This form of cooperation is bottom-up and is achieved through each agent aligning their individual activity for the benefit of the wider group.

In small groups, it would be common for members of the group to cooperate by dividing the work of the activity into roles and responsibilities, in a top-down, managed way, for example dividing a discussion topic into themes and allocating each member a theme to research. In the case of lightweight peer production described above, large scale cooperation is achieved through limiting the possible formats of contributions so it is also an example of top-down management. The CDT can provide some top-down management of the peer produced content, but it does not encourage users to align their behaviours towards enhancing the overall activity system. The intention of DBR3 is to establish some new ‘rules’ or ‘suggestions for participation’ that can shift the learners and the learning platform to create emergent groupings of content and people, defined around affinity rather than temporal positioning, which is beneficial to ideas relating to self-directed learning. This gives learners the agency to take an active stake in the development of a learning community, extending the range of the environment and laying the foundation for a virtuous cycle of cooperation, in conjunction with the affordances of the mediating artefact

I am encouraging the use of hashtags throughout all the comments in this phase of DBR and modelling personal use of the CDT within a reflective commentary on the CDT activity page. It is not necessary for ALL learners to display these behaviours for the overall system to take on a cooperative dynamic, but for just enough learners to do these, to varying degrees, to cross a threshold for large scale, bottom-up cooperative activity.

7.2.4 A short history of the #hashtag

The hashtag was first used on Twitter on August 23rd, 2007 by Chris Messina and followed up by a blog post a few days later in which he proposed the hashtag could improve "*contextualisation*, *content filtering* and *explanatory serendipity* within Twitter" (Messina, 2007). Indeed the "#" or 'pound' symbol had already been used in previous chat software such as IRC (internet relay chat) to create new 'channels' or 'groups', but Messina advocated for the symbol to represent "ad hoc assemblages of people with similar interests" rather than delineated shared-interest groups (such as the IRC channels), or self-managed groups (such as a contacts list); these latter 2 types of groups imply that the group is 'managed' in some 'top-down' way. He argues that a hashtag is completely decentralised, user controllable and compatible with any text-based comments system without extensive development, meaning any user can create any hashtag at any time, thus taking a folksonomic approach to grouping conversations or content. He also makes the point that using hashtags on Twitter enables other users to eavesdrop informally on a topic without having to join the conversation thread for notifications (also the case for FutureLearn conversations). It is important in online learning to have the option of minimally participating in a conversation or community before acting, and to move from 'legitimate peripheral participation' towards active contribution (Lave & Wenger, 1991), and hashtags are intended to highlight important themes in the community's emergent knowledge base.

This idea is followed up by Zappavigna (2011) who suggests that hashtags on Twitter function as a community-building linguistic activity, and frames this in a broader shift of social activity online towards 'ambient affiliation':

> This is a shift from searching purely for content, to searching what other people are saying online and forming communities of shared value. In popular terms, it is becoming increasingly useful to search the 'hive mind'. (Zappavigna, 2011, p. 789)

More general tagging has a much richer history across the web, being used on sites like de.li.cious and Flickr much further back than 2007. Indeed, in a learning context, Dron (2006) describes an e-learning environment called 'CoFIND'. This system is stigmergic in that all content and metadata is generated by users and the system is self-sustaining. The metadata is folksonomic which is to say created from the bottom up rather than from a list, and other learners can use the metadata as filters later. The major difference between this system and the 'ad hoc assemblages' described by Messina or 'ambient affiliation' described by Zappavigna is that the metadata relating to the resources in CoFIND are to some degree defined, in terms of ratings, classification in topic groups, or 'qualities' such as 'good for beginners', 'reliable', 'detailed' and feed into an algorithm that encourages or discourages the use of any resource, like page ranking on a search engine. Dron acknowledges that CoFIND is best suited to relatively small groups of learners with coherent needs, who can create self-determined pathways through an ever-growing body of resources, so is not directly applicable to the MOOC context (large groups of learners with diverse learning needs). It does, however, demonstrate that learning systems can be built with sustainable design principles and co-produced by learners without resorting to top-down methods of adherence (as in the case of StackOverflow threads being deleted for duplication, for example, or the format of participation limited in the cases of Haythornthwaite's lightweight peer production).

DBR3 combines these ideas by suggesting useful hashtags to learners, to kick-start the process and to allow for new tags to be added by the community itself, as in the case of Twitter. The next section shows how the idea of hashtags are introduced in the first learning step on the FutureLearn platform to scaffold the cooperative behaviours, and how the reflection/ investigation task itself is further described to encourage sustainable learning practices by asking learners to model their use of the tool for their peers ("what did you click on?", "did it reveal anything surprising or new?")

7.3 Enhanced CDT, Enhanced Learning Design?

The CDT in DBR3 (henceforth: CDT2.0) represents a change to the codebase to support additional visualisations based on what week a comment was made, and to support hashtag terms, for example in the 'Dyslexia and Foreign Language Teaching' course in Figure 14. The existing affordances of filtering and displaying a new filtered cloud remained and learners could still pick several words to filter as appropriate.

Figure 14 CDT2.0 with changed codebase to separate comments per week and by hashtag terms

The idea of using hashtags is introduced in the very first step of the course (Step 1.1) and suggested some words that learners could use, specific to the Dyslexia course:

Tagging your comments

Learning together is a strength of learning online with a large number of peers, but it will be impossible to read all the comments which interest you, so we have developed an interactive search tool to help you discover new conversations that you may have missed.

In order to make this tool sort through comments effectively, we ask that you #hashtag important words in your comments. For example if your comment is about #dyslexia in #primaryschool settings you may wish to add these as tags. Feel free to make up your own tags. Be creative!!

Figure 15 Explicit scaffolding of the use of hashtags throughout the course, with examples cited

In the following step (1.2), learners are asked to introduce themselves in the commentary sections, I suggest they use their location as their first hashtag term, and it is possible to see several place names in the cloud in Figure 15 above. This serves as an easy-to-understand introduction to the idea of ‘ambient affiliation’. It is important to note that at this point of the course (step 1.1. and 1.2), most learners are neither familiar with CDT, CDT2.0, nor with the behaviour of making hashtags in their comments on FutureLearn. The CDT2.0 tool and investigation task are introduced in step 1.14 so there are several videos, texts and discussions before the benefits of using hashtags are revealed. I did add some suggested tags on some of the steps between 1.2 and 1.14 to serve as a reminder. My initial expectations were that learners would add a single hashtag term at the end of their comment, as a piece of descriptive metadata, but there is no limitation as to learners' choice of term/s, in line with the hashtags origins, as described above in 7.2.4.

The CDT activity was positioned at the end of the week, in line with the other phases of DBR, to encourage learners who had been working in their silos or time-capsules to break out, discover new conversations or people, and take time to reflect on their learning goals. I made a change to the description of the activity to ask a direct question on this page which builds on the finding from DBR2 that learners write in response to questions. I also tried to encourage supportive writing by suggesting that the course team would 'pin' their favourite comments to the top of the page, to introduce an extra incentive for participation (Figure 16).

Activity

Which words did you click on? What story does this tell? Did your choices reveal surprising connections or help you focus on something specific? Tell us in the comments section and we will pin our favourite stories to the top!

Figure 16 Explicit scaffolding on the CDT activity learning step, to ask a question and provide a small motivation for cooperative participation

7.4 Results

7.4.1 Research questions

DBR3 aims to examine the overall behaviours of the cohort, so uses a new survey instrument, quantitative analysis of the user generated corpus, and a qualitative analysis of the learners' comments, rather than an interview method. This is chosen because it is important to see how the change to the learning design and pedagogical affordances affect the entire cohort, rather than a self-selecting group. DBR3 investigates cooperative behaviours in producing peer produced content, so it is important to analyse all comments for this, rather than elicit rich descriptions from a smaller population. This is coherent with the overall methodological approach of the pilot studies and DBR1 and is considered sufficient for discovering insights to these questions:

RQ1: How are hashtags used across the corpus?

RQ2: To what extent do learners model their own use of the CDT tool to help others?

RQ3: Do learners adapt writing through using hashtags *because* they are used by others?

RQ4: How useful is the new CDT exercise overall for learning at scale?

The questions investigate cooperative writing behaviours, and how useful these are perceived for social and vicarious learners: RQ1 asks how social learners respond to doing something new; RQ2 asks whether cooperative behaviours can affect learning communities; RQ3 asks how social learners adapt their writing style explicitly towards cooperative behaviours and RQ4 examines how the new exercise is evaluated overall in terms of learning at scale.

The new CDT activity was deployed across 4 MOOCs (3 repeated courses, and 1 newly designed course) with a total of 11,546 learners. 493 learners completed the survey instrument which included a total of 31 questions, although some questions were conditional on previous answers and all were optional. A total of 1247 comments were left on the CDT exercise learning steps.

7.4.2 How are hashtags used across the corpus?

48747 comments were created over the 4 courses analysed, producing a total of 17862 words preceded with a "#" symbol. Many comments use multiple hashtags because these 17862 terms are used across 6562 comments, making an average of 2.72 hashtagged terms per comment containing a hashtag. Each hashtag is repeated between 2-3 times on average (hashtags/ unique hashtag terms). This is broken down in Tables 13 and 14.

Course	Comments	Comments with hashtag	Hashtag terms	Average term per comment	Unique hashtag terms
Dyslexia (6)	22228	3514	10289	2.93	3426
Penshurst (1)	7167	322	674	2.09	281
Quakers (5)	15928	2290	5845	2.55	2376
Soils (6)	3424	436	1054	2.42	511
ALL	48747	6562	17862	2.72	-

Table 13 Descriptive statistics: comments and hashtag terms

9780 learners participated in the 4 courses in total, 4736 made at least one comment, and 1784 used hashtags in their comments. I interpret the use of hashtags as demonstrating cooperative behaviours, because the repeated use of a single tag is a positive behaviour which enhances the functionality of CDT2.0 and also creates the ambient affiliation as described above (Zappavigna, 2011).

Course	Total 'active' learners[16]	Total 'social' learners[17]	Total #'ers	Avg term per #'er
Dyslexia (6)	5120	2602 (51%)	1122 (43%)	9.17
Penshurst (1)	722	320 (44%)	52 (16%)	12.96
Quakers (5)	2675	1192 (45%)	420 (35%)	13.93
Soils (6)	1263	622 (49%)	190 (31%)	5.54
ALL (avg)	9780	4736 (48%)	1784 (38%)	10.01

Table 14 Descriptive statistics: learners and hashtag terms

[16] An active learner is someone who marks at least one step complete

[17] A social learner is someone who makes at least 1 comment

The CDT2.0 can produce a visualisation up to a maximum of 200 terms. The smallest course in terms of enrolments, social learners, hashtag social learners and hashtags (Penshurst (1)) produced 281 unique hashtag terms, demonstrating that the application itself does not require a large proportion of the cohort to participate in the cooperative behaviour to produce a relevant, course-specific list of terms. In this course, 7% of the total number of active learners or 16% of social learners used hashtags in their comments, but the range was varied enough to meet a baseline requirement for the CDT2.0, producing 281 unique terms. However, clicking on hashtag in the cloud can only return comments of this small subset of learners (52/320), which is limiting for exposure to diverse opinions. The course with the greatest number of hashtag social learners (Dyslexia) can aggregate comments from a much wider pool of learners (1122 or 43% of the social learners), even though the hashtag social learners in this course do not have to be as hard working individually, i.e., they do not write as many hashtags on average (9.17 as opposed to 12.96). Therefore, techniques for encouraging wider use of hashtags should be developed to reveal its benefits.

Hashtags are a bottom-up cooperative behaviour. They require learners to both identify relevant themes and use the syntax appropriately. In the first course to trial the CDT2.0 (Dyslexia), some learners did not use the syntax properly, for example using spaces between terms. For example, "#Sri Lanka", which would only return "#Sri" in the cloud. Another common mistake was to do the opposite and write terms together with no space bar, e.g. #dyslexia#ADD#ADHD which returns the whole string and makes it unlikely for the terms to be repeated. A third error was for learners to use very similar terms for the same theme, for example in the Penshurst course using “#sirphilipsidney” and “#philipsidney” which refer to the same person, but do not work with the visualisation code and divide a single theme into 2, halving its frequency in the cloud. Some of these mistakes are possible to remedy using clearer scaffolding and examples on the FutureLearn steps, and indeed in later courses there was a noticeable reduction in learners making syntax errors. However, it is not possible with the current level of intervention on the FutureLearn platform to coordinate the hashtag terms such that similar themed terms can be grouped together, along the lines of the ‘lightweight peer production’ (Haythornthwaite, 2009). On other social media platforms that use hashtags, the text editor recommends terms when it recognises the syntax and displays the measures of popularity of the terms and derivatives. Figure 17 is a public domain posting from Instagram[18], annotated to display the feature for coordinating terms.

[18] https://www.instagram.com/p/BrRwwoOn4Do/

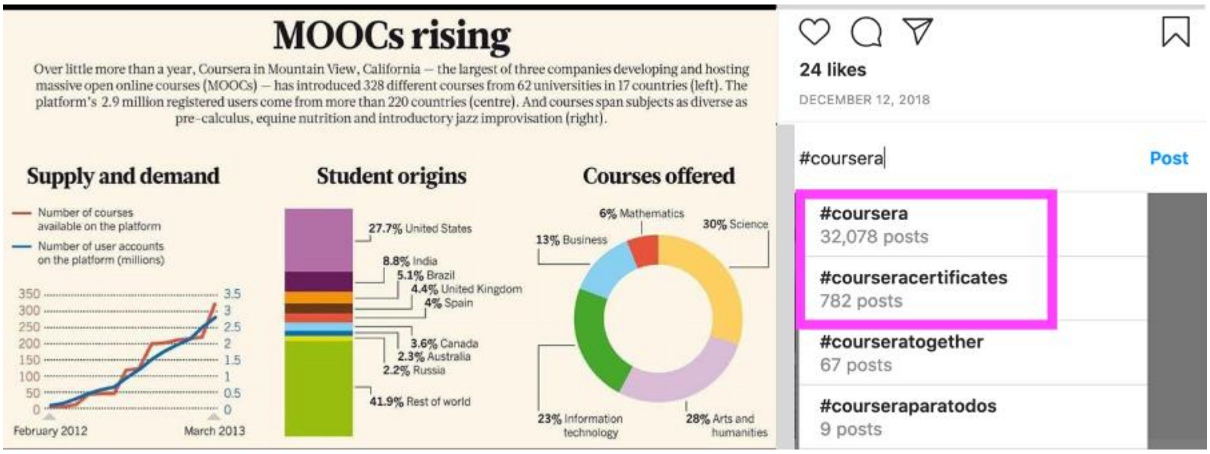


Figure 17 Instagram post, highlighted to display techniques used by the text editor for coordinating hashtag terms

This effect could be described in terms of stigmergy by saying that cooperative behaviour of the users can also be driven by the coordination affordances developed in the platform. This is also the case for the first iteration of the CDT on the raw words themselves, but the use of hashtags limits the terms that are coordinated to a set of course-specific themes which are deemed important by the learners, so it is easier to scan for the themes in discussion and find further information on the themes of interest. This thesis makes the strong recommendation to FutureLearn that this affordance is integrated into the text editor.

Learners are encouraged to use the hashtag syntax in the course on the first step, but they cannot know the reasons for this until later in the week when they are introduced to the CDT2.0 activity. This behaviour is also not common for other FutureLearn courses because the intervention is only used on courses designed by Lancaster University. As a result of this, many learners forget to use the syntax when they write their comments and then realise the importance of it later in the week when they start to use the CDT2.0 activity to discover comments they had missed.

7.4.3 How do learners exhibit cooperative behaviours by modelling their use of CDT for others?

I analysed 1247 comments from CDT2.0 activity. The explicit scaffolding of the pages asked the learners which words they clicked on, what this revealed and whether they found it surprising or not, as described above. I classified the responses to these questions into 3 levels of modelling behaviour; sometimes these descriptions would include hashtag terms:

1. Simply writing the words clicked on;
2. The chosen terms and a minor reflection on either the process or the findings;
3. A longer reflective piece on what the words revealed and what connections they were able to make.

In total 316/1247 comments could be described as modelling both how to use CDT2.0 and some level of reflection; this means on average around a quarter (or 1 in 4) comments would be of this type. Combined with the strong findings from both the survey results and the interviews that learners tend to read approximately 20 comments from each page as their participation in the social elements of the course, it follows from this that most learners would see around 4-5 comments which model a use of the CDT2.0 if using the site in the most common way. This demonstrates that by asking a simple yet direct question it is possible to encourage cooperative behaviours as the learners who answer, even in the simplest form, are modelling for the learners who follow behind them, and that it does not require most learners to answer to achieve this result.

43/316 modelled in a minimal way and literally just left the words they clicked on, for example on the Soils course, a learner wrote “my words were agriculture, sustainable, permaculture, regenerative and policy”. Whilst this level of modelling is minimal and does not involve much reflection on what impact reading conversations including these words may have had on learning, it highlights words to watch out for, and simply following in this learner’s steps and clicking on these words will reveal a set of comments and an ‘ambient affiliation’ with the community of learners who are interested in sustainable agriculture.

155/316 comments went further than this and added some short reflective commentary either about the tool or about what the tool helped them find, for example on the Penshurst course:

> *I clicked on Urania and that led me to lots of interesting comments on the character and the work in which she features which I wouldn't have found otherwise.*

This comment goes further than simply writing the words that are clicked on by positively appraising the tool and demonstrating the impact it had on their self-directed learning – that the tool led to interesting comments which wouldn't have been found otherwise, so encourages other learners to try the tool and to think carefully about which terms to select to further their own learning.

118/316 comments answered the questions most fully, by writing very reflectively about what the words meant to them, how it surprised them which words were or were not in the CDT2.0, or what conceptual connections were made from seeing the commentary in this format.

> *I started off by focusing first on life and then death and then [clicking] woman. When I connected death to women, I found a conversation about death being portrayed by a female figure in their writings. When I connected woman with life, I found more interesting threads about woman and free will in this time period and the ways in which women express love and also desire. Very interesting experiences.*

This comment highlights that through using CDT2.0 and through seeing the comment corpus in this format, connections are highlighted that are not possible to make through linear reading, and through 'answering the question' this modelling is apparent to other learners, creating a cooperative and sustainable pedagogy. I found it encouraging that such a large proportion of the comments which modelled use of the CDT2.0 had so much depth to their reflective content. Earlier findings demonstrate that learners tend to read around 20 comments, so an ‘average learner’ would likely see at least 1 comment of this type of depth. This begins to speak to the comment made by a learner in DBR1 that “[the CDT2.0 artefact’s] value will derive from the clarity of the line of investigative questions/enquiries”, and demonstrates that even a low level of this type of high-quality response is enough to create a cooperative framework for others to work in.

Further to the 316 comments which modelled both use of the CDT2.0 and types of reflection, there were also 134 comments which were simply reflective in nature, without directing other learners towards terms which yielded those reflections. For example, in the Soils course:

> *I found most of the things very interesting - for example cow dung beetles and their impact on environment. Since childhood have been visiting farms and living there but this relationship I could not see.*

This demonstrates that even without the CDT2.0 artefact to extend learners’ reading and interactive writing, it is useful to have a step in the learning design asking learners to reflect on what they have learned this week, and to articulate those reflections on the page.

If all these comments are aggregated as ‘modelling either reflection or use of the CDT2.0 to support reflection’, more than one third (450/1247 or 36%) of all the comments left on the pages model the intended learning type of ‘inquiry’ in the ABC learning design framework. If very short comments such as ‘Thank you’ or ‘interesting’ and comments which are totally not related to the course are removed from the corpus, this increases the percentage to nearly half of all commentary (46%). This learning type is not supported by any other tool or activity in the FutureLearn platform and learners are in fact teaching each other how to do this, using platform tools and pedagogical design which sustainably support the learning type. I consider this to be one of the most important findings of this project in terms of the pedagogy of MOOCs.

7.4.4 Do learners adapt writing through using hashtags *because* they are used by others?

This research question aims to build on RQ3 in terms of learners modelling commenting behaviours for others. The premise of RQ3 is that learners will read comments with hashtags, and either start to use hashtags themselves, or start to use *certain* hashtags to build the visualisation in a structured manner. I can discover insights into this question by examining the survey, which was completed by 493 learners.

The survey instrument differed from the ones used in DBR1 and DBR2 as I wanted to discover more about the use of hashtags. Certain questions were dependent on the answers to previous ones, for example, "did you use hashtags in your comments?" and 173/493 (35%) responded positively. Other questions about writing hashtags were not displayed unless this question was answered positively. The proportion of learners who answered positively to this question is almost the same as the average number of hashtaggers/ social learners (38%) in the overall statistics in RQ1, demonstrating that the survey is a representative sample with respect to the hashtagging behaviour.

64/173 (37%) of the hashtaggers continued to respond positively to the question "Did the CDT2.0 encourage you to use *certain* #hashtags in your comments?" and 40/64 (63%) of these learners responded positively to the question "how useful did you find the hashtag tab on the CDT2.0?" which differs from 93/493 (19%) learners overall who responded positively to this question. This demonstrates that when hashtags are used and behaviour is modified through selecting certain hashtags over others, the tab itself is perceived as more valuable for learning; in other words, when learners perceive the value of cooperative behaviours, it inclines them to cooperate further. The challenge for course designers is to further encourage these cooperative behaviours, and platform developers to create affordances which allow learners to conform their hashtags more easily to ones already in the system, as outlined above and is a core feature of the text editor in Instagram and Twitter.

Many learners commented on the CDT2.0 pages that they had forgotten to use hashtags, or that they will try to remember to do so going forward. For example, one learner wrote "A great way to not miss any interesting comments. I must confess I forgot more or less to put in the hashtags, but I will do it more from now on." It is a limitation of this project that the new affordances of the CDT2.0 are only available on courses which are designed by Lancaster University, whereas learners will identify more closely with FutureLearn and do courses from different providers. Although this was advantageous during the interviews of DBR2 as learners could compare their interactions and social behaviours between different course designs, in this phase it is a disadvantage because learners are not primed for this pedagogical technique in other FutureLearn courses. Hashtagging needs to be a latent commenting behaviour because suggesting tags too often would bias the bottom up and folksonomic nature of the tagging, which is the most valuable and cooperative aspect of creating 'ambient affiliation' within a community.

In terms of modelling behaviours for other learners, it was encouraging that one learner in week 1 of the 'Soils' course in suggested that they would go back and edit their previous posts to add in locations as hashtags, to create a better cloud of where learners are from:

> *Something that I learned clicking around in that tool was that very few of us (including me) have #hashtagged locations. As the students are so international it would be great if we could do more hashtagging of place*

names. I am going to go back to my older posts and edit them with more hashtags.

There is no way of knowing whether this encouraged other learners to do the same, but it is a good reflection of cooperation through 'ambient affiliation' and that some learners can perceive that collective activity in terms of choosing certain hashtags over others can have a positive impact on the entire cohort.

7.4.5 How useful is the new CDT2.0 exercise overall for learning at scale?

This question draws on both the 1247 comments on the CDT2.0 activity and the responses to the survey. The data is divided into a sentiment analysis of the comments, to determine how learners feel towards the CDT2.0, and the responses to questions regarding whether the CDT2.0 enables learners to discover new content, discover new people and encourages interactive writing or further reading. These responses will determine whether the CDT2.0 is having an impact on factors related to learning rather than simply being perceived positively.

543/1247 comments spoke directly of the CDT2.0 artefact, with the remaining 704 either being reflections simply about the content of the course without reference to the CDT2.0, or comments that had no reflective content at all. Only 32/543 or 6% of CDT2.0 comments were coded as negative. Some of these referenced technical problems of access, mirroring the findings from DBR1. Others implied that it was difficult to see relevant words such as a learner who wrote that they "could not always find a word; like someone said they clicked on 'silence' but I could not find it in the cloud." Aside from technical issues one learner wrote that he "had found valuable conversations but was disappointed that the words of interest [to me] … did not show up". This comment mirrors a finding from DBR2 that learners may have to learn to click on related words before the word they are thinking of shows up in the visualisation and realise that this tool is about discovery and serendipity rather than focused search, but this is useful feedback for the development of the instructions for using the artefact.

126/543 or 23% of comments were coded as neutral, and these mostly were comments that just wrote the words that they had clicked on, without giving any reflection on usefulness of the exercise. This specific behaviour is an example of cooperation in terms of modelling how to use the CDT2.0 for others, so these comments are useful because this type of comment helps other learners to understand the activity and is a factor for a sustainable learning design-based around cooperation.

385/543 or 71% of comments were positive comments about the CDT2.0 exercise, ranging from "cool tool", to "clever and useful way of honing in on the area(s) that interest you" right through to extremely reflective comments which answered the questions on the page exactly (i.e. they said what words were clicked on, what story this told, how it related to personal learning goals and also how seeing the comments broken down like this gave a new perspective on the course itself.)

These proportions differ vastly from the same analysis in DBR1 where far fewer learners wrote positive comments about CDT (245/590 positive; 203/590 neutral and 105/590 negative). Negative comments that also reference technology problems are a common finding, but far fewer learners (almost nobody) said that they didn't understand what to do, which is a significant difference. This demonstrates that the exercise has matured both technically and as a learning exercise. Learners are modelling use of the CDT2.0 for each other, and the scaffolding on the page text is asking a direct and specific question, which can be answered regardless of whether any value is extracted from the tool itself.

In terms of how useful the CDT2.0 exercise is *for learning at scale*, it is necessary to relate answers in the survey to dimensions which specifically relate to learning, such as development of thinking about course content, an increase in interactive writing (Lapadat, 2006) or increases in exposure to new conversations or new people. In the previous phases of DBR, I conducted a Spearman's Rho correlation analysis on the results of the survey, to correlate perceptions of the affordances with learning behaviours, but it is not possible to do this type of statistical analysis with this survey instrument due to the branching and conditionality of the questions. Also, each question is optional, so every question is answered by a different proportion of the total population.

363/493 or 74% of learners used CDT2.0 once or more, with the most common answer being 2-5 times (254/363 or 70%) which is once per learning week. This supports the finding from DBR2 that learners will interact with the course in a linear fashion and do the suggested activities as they appear along this pathway and this further emphasises the importance of learning design and using a wide variety of learning activities to match different learner preferences. 63% of learners responded that they read between 1-20 comments per page and 23% responded that they read more than 20 but not all. Only 7% read all the comments, with two thirds of these never returning to the step once they had read them all. These results also verify the findings of DBR2 that most learners passively flow with the pathways of the platform and will not return to a step to extend their reading, so inclusion of the CDT2.0 tool which affords this marks a different behaviour from the usual mode of participation. For 330/363 or 91% of survey respondents, this was the first time they had encountered the CDT2.0 activity and 128/363 or 35% tried the hashtag tab on the visualisation. Table 15 shows the breakdown of the responses related to learning and teaching, and the hashtag features of the CDT2.0.

	Yes/ Positive	No/ Negative	Maybe/ Unsure
The CDT helped me discover new conversations	125/296 (42%)	84/296 (28%)	87/296 (29%)
The hashtag tab was important for discovering new conversations	65/90 (72%)	25/90 (28%)	-
The CDT helped develop my thinking about course content	77/297 (26%)	89/297 (30%)	131/297 (44%)
The hashtag tab was important for developing my thinking	61/90 (68%)	29/90 (32%)	-
The CDT helped me discover new people	83/297 (28%)	135/297 (45%)	79/297 (27%)
The hashtag tab was important for discovering new people	49/71 (69%)	22/71 (31%)	-
The CDT encouraged me to read more	158/285 (55%)	127/285 (45%)	-
The CDT encouraged me to write more	117/294 (40%)	177/294 (60%)	-
I found the CDT useful overall	216/294 (73%)	78/294 (27%)	-
I found the hashtag tab useful overall	85/111 (77%)	26/111 (23%)	-

Table 15 Breakdown of the proportions of respondents who answered the questions relating to learning and evaluation of the hashtag features; DBR3

More than 70% of respondents reported that they found CDT2.0 useful overall, and 77% of the subset of respondents who used the hashtag tab (35% of total CDT2.0 users) found this useful overall. This pattern is repeated throughout the questions on the survey (e.g., 42% found the CDT2.0 helped them discover new conversations, but 72% of the subset of hashtag tab users found it important for discovering new conversations). This demonstrates that learners who spend time using all the features of the visualisation including the hashtag tab will derive more value specifically in terms of discovering new conversations, people or developing their thinking about the course.

Learners who followed up the question about the CDT2.0 helping to develop thinking referenced the ability of the tool to structure content, and discover conversations they had missed, for example:

> *It is structured to lead you in directions that are expansive. I clicked on phosphorus and found many comments on different types of soil contamination and their effects, some of which I had never heard of.*

This comment suggests that the learner understood which words would generate comments of interest, and that this led to further comments which they could not read when doing the course in a linear fashion. In this respect, they could personalise or individualise the course to their interests, which is a key cybernetic quality required for a sociocultural pedagogic model (Britain & Liber, 2004). It is also Level 2 in the Cultures of Participation framework: Users organise content (Fischer, 2011).

117/ 294 or 40% of learners responded that the CDT2.0 encouraged them to write more on the course. This is consistent with findings from the other phases of DBR, where many learners use the CDT artefact to extend their reading, rather than as an opportunity for further writing. However, interactive writing is strongly associated with learning in a sociocultural sense (Lapadat, 2006), and an increase in writing by 40% of learners is likely to have a very positive impact on the overall corpus, and on the cooperative and structural aspects of the commentary such as the use of hashtags for creation of the ambient affiliation.

Overall, the CDT2.0 has a positive effect on learning, especially for those learners who persevere with the tool and use the hashtag tab to extend their reading and writing. It is important to note that no learner reported that attempting to use the CDT2.0 activity decreased their learning, so it either had a neutral or a positive effect. Furthermore, it can be seen from the qualitative coding of commentary on the CDT2.0 activity that learners would often write reflective commentary but not reference the CDT2.0 as a supporting artefact for this reflective activity. Therefore, inclusion of a step which proposes 'reflection' as an activity is a useful addition to the learning design and supporting this activity with the CDT2.0 extends their ability to reflect and extend their understanding.

7.5 Summary of findings from DBR3

The major focus of this chapter is to bring together the findings from DBR1 and DBR2 to level up the platform to foster cooperative participative behaviours in learners. DBR1 found that learners were interested in the CDT but many needed more instruction for using it effectively; DBR2 found that integrating it further in the course to give users a few attempts made a significant difference and embedding it in a reflective exercise encouraged some learners to use it to scaffold their reflections. DBR2 also discovered that most learners interact with the courses in a linear fashion and although they find the diversity of commentary incredibly important, they did not know how to exploit it further. For some learners, the CDT was able to provide a means to support their inquiry and self-directed study in terms of mining the 'hive mind', but there remained a significant proportion who did not understand how to use it effectively and found the words too random or not useful.

In DBR3, I collected all this analysis and developed the activity in 3 major ways:

1. A tab for comments in each week
2. A tab for terms preceded with a # symbol
3. Instructions on the page to ask a specific question for learners to model their use of the CDT2.0 in the commentary.

The focus of the analysis in DBR3 is on the second and third points above and how these have changed the ways in which people interact with the CDT and the writing process in general. From a theoretical point of view, by changing the system to include hashtag terms, I have demonstrated that learners can move to higher levels of the Cultures of Participation framework, as they are able to organise content for themselves and extend the range of the environment (Fischer, 2011). This gives them the agency to develop sub-communities or communities of practice around a framework of 'ambient affiliation' which is not possible with the default affordances of the FutureLearn platform. However, it would not be possible to level up participation solely through these technical changes; there needs to be a mechanism for learners to scaffold behaviours for each other, through either cooperating on a choice of hashtag term, or modelling their use of the artefact for reflection and self-directed inquiry in their commentary, to develop a mode of participation which is self-sustaining.

The changes made to the platform and the pedagogy in DBR3 has positively impacted on both these factors. Over a third of social learners converted to 'hashtag social learners', meaning they used hashtags in at least one comment, and around 15% of all comments had a hashtag term. From a standing start, and with around 90% of learners having never seen any version of the activity before, this should be interpreted as positive, and something to build on further in terms of learning activity proposals and scaffolding on the course pages. However, there are limitations to how much further the CDT can go technically without adding extra affordances to the FutureLearn text editor (e.g. suggesting hashtag terms based on previous frequency), which is beyond the scope of this project. These additional features would present new affordances and create the basis for the types of lightweight peer production behaviour which is described by Haythornthwaite (2009), but purely technical developments such as an enhanced text editor would not be sufficient for fostering cooperative behaviours from learners; this requires an examination of the types of learner interactions that FutureLearn encourages on a broader, cultural level. DBR2 demonstrated that modes of participation are affected by platform affordances and limitations, such as the inability to 'deep dive' into areas of interest and affinity but also highlighted invisible factors such as completion and timeliness of progression which sit above the platform affordances and affect learners in the way that they think of FutureLearn, and their learning within FutureLearn. The learners who took an 'active' mode of participation, as described in DBR2, forfeited either the recommended time on task ("read everything") or the idea of 'completion' ("sampling") so that they could take the course on their own terms. The 'hashtag behaviour' needs to be embedded within the course (and indeed the whole platform) from the point of view of being a social good as 'individual work' (doing 'work' to extend the range of the environment and streamline its resources for other learners) but should also be promoted as a means to create sub-communities of interest, which could form the basis for further collaborative activity. For example, it is not a huge leap to see how the learner above who discovered different types of soil contamination through using CDT2.0 could create a hashtag to bring together other interested learners in a community of 'ambient affiliation' and start to collaborate on discovering new resources and case studies around this particular area of the course. Indeed, this type of behaviour could result in anything from a group assessment, a new database of knowledge, or even direct social action. I propose that these possible outcomes for collective work should be woven into the fabric of social behaviour in a MOOC and encouraged through the pedagogical method and means of assessment, rather than the current emphasis on progression, timely completion, and individual attainment in multiple choice questionnaires, which drives a behaviourist pedagogical method. It would take a leap of faith to integrate such activity into the learning and assessment design of a course, and a move away from determinist measurements of activity such as logging individual progression, but this project and CDT2.0 demonstrates that there is the beginnings of a method towards harnessing a MOOC for collective endeavour, using hashtags as a glue for communities of ambient affiliation and a means for learners to extend the range of the environment (Fischer, 2011).

Hashtags are an easy-to-understand method of fostering a participatory pedagogy and create ‘virtuous cycles’ of participation. The activity of writing hashtags also encouraged other hashtag social learners to select their hashtags carefully, in order that the visualisation would reflect the emergent themes accurately, and indeed many hashtags were repeated several times to amplify their importance as an emergent theme, for example in the Dyslexia course, there were 3426 unique hashtags out of a total of 10289 hashtag terms, meaning on average each term was repeated 3 times. The survey also demonstrated that learners who understood most how to use hashtags and the CDT in support of reflection and inquiry were able to extract the maximum value, and their reflections on the commentary page served as models for learners who were behind them.

In terms of modelling reflective behaviours or use of the CDT, around 25% of all comments on the CDT exercise page were coded as demonstrating some modelling behaviour, which means that the average user who reads between 1-20 comments would read 4 comments that model their use of the CDT when they arrive on this page. This is important because the intervention is made at the level of the platform, rather than the course, so for the instructions for use to be self-sustaining, peers are required to do some work to help each other learn how to use it for that particular course. This modelling does not just extend to the use of the CDT artefact, but into narratives on how the CDT artefact has supported their self-directed learning and therefore models how to ask critical questions to peers, which is a form of cooperative behaviour that is not prevalent in the other types of learning activity which the platform can offer. This is not altogether altruistic work from an individual learner, as interactive writing in its own right is a useful method for reflection and in order to articulate one's thoughts (Lapadat, 2006). To reiterate the insightful quote from the learner in DBR1: "[the CDT artefact's] value will derive from the clarity of the line of investigative questions/enquiries."

In summary, it is important to allow space for cooperation and extended reading and writing in a MOOC, to switch the 'problems' of overload into a 'positive'. It does not require that all the learners demonstrate these cooperative behaviours, but this phase of the research has demonstrated that there is a desire to give something back in all courses, regardless of their subject domain or populations of learners. This could be by using hashtags to create ambient affiliation, or to model reflective activity and inquiry for others to replicate. These developments are able to shift learning in a MOOC from one of individualism and passivity towards activity which is community focused, revealing the idea of the 'hive mind' to learners, and flipping the problems of information and participation overload (Barricelli et al., 2015) into an advantage. Further research should investigate how the learning activities, such as the assessment models could be used to reward collaborative activity, such as the emergence of groupings of ambient affiliation and the various types of new knowledge that could be created through further collaborative activity, such as the creation of new databases, social networks, or community action groups.

Chapter 8: Conclusion

8.1 Looking back

Interactive writing in FutureLearn courses is not proportional to the levels of participation (Tubman et al., 2016) and these empirical data seemingly position the pedagogy of learning at scale in contrast to existing literature on online distance learning where the online network is leveraged for flexible, cooperative activity (Beetham & Sharpe, 2013; Conole, 2013; Garrison, 1997). It does not necessarily follow, however, that learning at scale represents 'supersized behaviourism' (Baggaley, 2014), but it is evident that the pedagogy of current MOOC platforms – Coursera, EdX, FutureLearn – is reliant on interactive tools which prioritise completion, consumption and basic computer aided assessment (Bali, 2014; Downes, 2012), termed 'acquisition' and 'practice' in the language of Laurillard's teaching types (Laurillard, 2012). I leverage peer produced content by creating new tools which support learning activities based around inquiry, reflection and discovery and which act as the basis for collaborative activity within a large cohort. The development of the tools uses an iterative design-based research methodology (Barab & Squire, 2004; Brown, 1992) which stands in contrast to other studies of MOOCs which use learning analytics and statistical models to describe the learning taking place (Fincham et al., 2019; Kizilcec et al., 2013; Reich, 2015). These other narratives and analyses focus on post-hoc clickstream data and on completion as the primary factor for engagement, whereas a design-based methodology focuses on intervening directly to discover how an intervention affects learning at the level of the learner. The new visualisation, a type of social learning analytic, is made available at ‘learntime’ as a guide for navigating the scale of participation, which is often problematised as ‘information’ or ‘participation’ overload.

The critique that MOOCs do not naturally follow on from the active learning, sociocultural pedagogies commonly found in institutional distance learning, or the reduction of learning to clickstream data analysis seems to me to be a misunderstanding of how we should understand a MOOC on both fronts. MOOCs should not be wholly rejected because they do not naturally follow on from previous models of distance learning; these pedagogies are typically deployed with small groups of motivated students taking on specific roles to do defined tasks within an institutional framework. Conversely, the reduction of activity to clickstream data cannot inform the actual learning processes taking place; indeed, masses of clickstream data are most likely to present an amplification of algorithmic biases which are baked in: pedagogic and cybernetic.

MOOCs require a different conceptualisation: they have a large, heterogenous cohort, students have self-directed motivations for study and their interactions occur both inside and outside the platform (Veletsianos et al., 2015). These factors cannot be underestimated because a radically different space emerges from the diversity of peer produced content and measures of engagement such as completion are not keystone markers of success, because learners are self-motivating – therefore, MOOCs are a different kind of online course. A sociomaterial perspective would suggest that there is a different assemblage of human and non-human factors which make up the environment (Knox & Bayne, 2013) and that it is important to describe this in order to understand the important features of a MOOC. This thesis is an investigation of the material nature of the discussion features and the social dynamics which result from interacting in this space.

8.1.1 MOOC participation – a material perspective

Participation in a MOOC is typified by its chaotic nature, with the scale of participation overloading system architectures. Learners write about radically different things right next to each other, all at different levels of expertise, and whatever their reflections were at the time. The consequences of this type of 'information overload' phenomenon is clear in the modelling of Brinton et al. (2014), who demonstrates that MOOC forums get less and less activity as information piles up in them, suggesting that learners are less and less likely to rifle through the miscellany. On the other hand, this information pile-up brings in a certain amount of serendipity which is unlikely to emerge in smaller, structured models of online education, and this is demonstrated through the interviews in DBR2. Rosé & Ferschke (2016) state: "Currently available platforms for learning at scale are frequently impoverished on [the social] dimension … Though there are almost always discussion forums included in these environments, they are often just an appendage, and not effective in meeting the needs of learners" which supports Brinton's modelling.

Analytics grounded in engagement-as-completion are also thrown into disarray as not all students want or need to complete the course. This is a commonality shared with other forms of informal open education (Stracke, 2017) and this learner has been identified in the literature as a 'sampling' student (Ferguson et al., 2015; Kizilcec et al., 2017). In short, MOOCs represent something new in terms of types of student engagement and measures of that engagement (DeBoer et al., 2014) and this supports my claim that new platform tools and pedagogical interventions are required in order to mine valuable content from these deep seams of peer produced content.

MOOCs provide a context of diversity which has never been seen before in formalised education, but which has never been exploited through either the design of platform tools or learning design. Indeed, from a sociomaterial point of view, the platform itself is 'black boxed' across most analysis and considered in instrumental terms as the neutral means of achieving educational outcomes (Hamilton & Friesen, 2013) - completion outcomes for xMOOCs and social network outcomes for cMOOCs – which are taken as proxies for successful learning. FutureLearn organises peer produced commentary chronologically, so in a context of high participation, learners are encouraged by the platform cues to only view comments made at the same time as they are working on that learning step, and it is through 'luck' or 'serendipity' that connections are made. In other words, many original contributions were lost due to participation overload and a platform design encouraging a one-way direction of travel through the course, with no pedagogical cues to look back! Furthermore, the presentation of comments on the learning step encourages learners to read only a dozen or so comments, and this is demonstrated through the interviews in DBR2 and verified in the survey in DBR3. I therefore designed a platform tool (the "CDT") which enabled learners to 'make their own luck', drawing on the ancient proverb that "luck is what happens when preparation meets opportunity", the CDT affording a new opportunity through visualisating of the corpus. I want learners to be able to gain something 'in real-time' from the intervention, following the recommendations that analytics "should be seen as tools to be placed in the hands of the very subjects being analysed – the learners" (Buckingham Shum & Ferguson, 2012). The CDT is an open-source tool which has been made available from GitHub[19] and licenced with a GNU General Public License v3.0. This licence permits other users to modify, distribute or patent for private or commercial use and is only limited in terms of liability or warranty.

8.1.2 MOOC participation – a theoretical perspective

This project is framed theoretically by drawing on the Cultures of Participation framework (Fischer, 2011), which is based on the premise that interactions in computational media should contribute towards the design of cultures in which humans can express themselves and engage in personally meaningful activities. This contrasts with media which sees humans primarily as consumers; xMOOC platforms

[19] https://github.com/philtubman/comment-discovery-tool

put learners in the position of being consumers of content, and this project has demonstrated through both an analysis of the major affordances, and through rich descriptions of learning pathways, that consumption and progression is the primary designed factor on the FutureLearn platform. Learning cannot be reduced to the same type of activity as watching a 'Netflix box set', and learners should have the agency to propose problematic areas within a defined course space, find interested others, and discuss their own solutions to these problems. Current xMOOC platforms do not provide this agency, and cMOOC platforms can only provide it to the most digitally literate learners (Mackness et al., 2010).

The Cultures of Participation framework proposes 5 levels of agency, and FutureLearn is only able to reach the first level: Consumers aware of the possibilities. A learner using FutureLearn can see that useful information is posted by peers and is able to make contributions, but they cannot organise this content which is Level 2. Indeed, as evidenced in the DBR2 interviews, the learners who did organise content around their interests had to subvert the platform design or their own time constraints and use a method of manually skim reading, or focused sampling.

Innovative technological developments are not sufficient to foster and support cultures of participation because "cultures of participation are not dictated by technology; they are the result of changes in human behaviour and social organization, in which active contributors engage in the innovative design, adoption, and adaptation of technologies to their needs and in collaborative knowledge construction.” (Fischer, 2011, p. 42). Therefore, it is not enough to simply create a new technology, and the phases of DBR in this project examine and create new technical and pedagogical interventions which level up learners’ activities in the framework, focusing each phase of DBR on an aspect of the sociotechnical environment and following the tenets of Vygotskian Activity Theory by concentrating on mediating artefact (DBR1), rule of engagement or pedagogical factors (DBR2), and new learner behaviours (DBR3).

8.1.3 Phases of research

The CDT in DBR1 gives learners the ability to organise the peer produced content towards their own learning goals, and this is tested experimentally against a no CDT condition to determine whether it made a significant impact on the quantitative levels of overt sociality. I developed a novel framework for analysing conversational patterns on the platform in terms of how they grow from individual comments to longer interactions, taking account of the diversity (number of unique participants in the conversational unit) and length (number of posts in a conversational unit). This novel framework extends previous analysis of the FutureLearn platform by Chua et al. (2017) by applying their constructs of comment-placement onto whole conversational units, creating a novel schema for describing conversations. The advantage of this approach is that hundreds of thousands of conversational units can be analysed to increase the significance of the findings. In DBR1 and DBR2 I analysed 288,826 conversations across both conditions using this technique.

In DBR2, I extend the availability of the tool and re-frame it within a specific learning activity which encourages learners to use the tool for inquiry and reflection, making the designed affordances easier to perceive. I conducted semi-structured interviews with 10 learners to further develop my understanding of how learners interact with the FutureLearn platform, and repeated the statistical analysis of DBR1 in order to test for replicability of results, which is key to the practice of design-based research (Barab & Squire, 2004; Brown, 1992).

In DBR3, I take stock of the findings of DBR1 and DBR2, and attempt to level up the pedagogy and the affordances of the tool, drawing further on ideas of 'stigmergy' as a design paradigm for scaling up participative behaviours (Dron, 2006; Elliott, 2016), and 'ambient affiliation' (Zappavigna, 2011) as a description of the weak-tie relationships fostered between learners in FutureLearn courses. The alterations to the intervention in DBR3 encourage learners to use hashtags to self-aggregate content according to the emergent concepts in the course, and the tool is developed so these tags can enhance discoverability and reduce the number of 'irrelevant terms'. This allows users to 'extend the range of their environment', levelling up in the Cultures of Participation framework, and encouraging cooperation in terms of choosing tags, aligning with Mark Elliot's concepts of Stigmergy (Elliott, 2007), where large-scale behaviour is divided into coordination, cooperation and collaboration. It sets the scene for new pedagogical models to emerge and for educators to create courses which can truly claim to have scaled collaboration, in the sense of people working together on problems where all participants have the same type of understanding (Vygotsky, 1978). This project does not claim to have proven large-scale collaboration in the higher educational context, but the tools and methods I have developed set the groundwork for exploiting mass collaboration to this end.

Therefore, the research questions I asked in each phase of the research built on the previous:

1. DBR1: How can a new platform tool affect levels of overt sociality, as defined by the novel conversational unit analysis, and do learners intuitively perceive the designed affordances of the intervention?
2. DBR2: Can the effects of the intervention be replicated in terms of the novel conversational unit analysis? How can the affordances and benefits of the platform tool be made more visible to learners through task proposals? How do learners qualitatively interact with the platform and the new tool?
3. DBR3: How can the pedagogical approach be enhanced through use of the intervention by affording learners the agency to extend their use of the platform?

8.2 Research as design; design as research

> Design is not a panacea. Designers often get things badly wrong … but crystallising good pedagogy into designed artefacts … is also a way of turning recurrent expenditure (of time, effort, and cash) into durable assets. (Goodyear, 2015, p. 28)

I start with the concern that the intervention would not work technically, that learners would not understand how to use the visualisation, or that it would disconnect learners from the course by showing how much they haven't read. To a certain extent, all these things can be found in the data, and their opposites, so the role of the researcher in the DBR methodology is to balance the evidence and develop replicable schemas so a clear narrative or trend can emerge. For this reason, my starting point is to develop a schema to define overt sociality and to test the results of the intervention between CDT-enabled courses and courses without intervention, a quantitative and repeatable measure.

The important findings from this quantitative analysis are that "Lone" posts decrease in all courses with a CDT intervention condition, by an average of 7% in DBR1 and 4% in DBR2; this means that more comments were replied to than in a no-CDT condition. This result is statistically significant and demonstrates that there is a likelihood that the CDT artefact has a noticeable and similar impact across all courses in terms of general overt sociality. The other measures from the conversational unit analysis split this difference across dimensions of unique learners and conversational development, with similar trends across all the categorisations and this indicates that the intervention is having a consistent and noticeable effect specifically in terms of discovery and subsequent interactive writing, particularly for the 'Cocktail Party/ Q&A' condition. This type of conversation refers to a multi-person/ single response conversation, or multiple learners finding an original post and making a single comment, indicating an increase in diversity of opinion within the conversational unit, but not strong enough to suggest any synergy between responses. To a smaller degree these conversations were growing into multi-person/ multi-response conversations, which does indicate both increases in diversity and synthesis in terms of knowledge construction. These results demonstrate a trend of increased social behaviour from learners, and their replicability suggests that the designed affordances of the CDT artefact are perceived by some learners in terms of the trend towards increased interactive writing. 288,826 conversations were analysed and categorised across both phases of DBR. The strength of this methodology is that it can be scaled indefinitely, and it does not just 'count posts' to measure social activity, rather uses the materiality of the platform in order to categorise conversational units into those which are likely to contain levels of sociocultural knowledge construction (Tubman, Benachour, et al., 2019; Tubman et al., 2018). The full breakdown of conversational unit categorisation, as extended from Chua et al. (2017) could be used to identify 7 different types of conversation, such as one where the original commenter does not return to the thread (the original ideas are 'hijacked' by other learners), or ones where the original commenter makes further replies but other learners only reply once (‘a multi-dimensional question and answer’) and this full breakdown could be used in further research to establish the types of conversation on the platform.

The quantitative framework described above is one part of the original contribution that this thesis makes, alongside the design-based research methodology for extending MOOC platforms (Tubman et al., 2018), and the development of new cooperative behaviours in FutureLearn courses towards sustainable peer-production and community formation (Tubman et al., 2020).

The quantitative results of DBR1 are further supported by learners who commented that the CDT was a welcome addition to the FutureLearn suite of tools because it increased their exposure to different voices, which demonstrates that the creation of this new visualisation enables learners to pursue their own learning ends and discover affinity-based pathways through the peer-produced commentary. This also suggests that the quantitative framework for measuring overt sociality does not tell the whole story of the design and that the CDT provides affordances not just for increasing digital traces, but for extended reading, and for increased perceptions of conceptual connections through the visualisation, where digital traces are not available for quantitative analysis. This is pursued in DBR2 in terms of interviewing participants and eliciting rich descriptions of their interactions on the platform and with the tool.

I interpret these results as being a positive indicator that the CDT can add value to the platform structure, but this value is not universal as some learners were confused and saw it as a distraction to their progression. DBR2 aims to understand how learners approach the course, and how the CDT can use learning design principles (Laurillard, 2012; C. Young & Perović, 2016). A learner in the DBR1 phase of data collection said:

> *it's [the CDT's] value will derive from the clarity of the line of investigative questions/enquiries which provide the rationale for selecting particular words from which to generate fresh clouds which may either offer answers or provide leads for further enquiry.*

This supports the premise of the design as scaffolding an affinity based search which cycles around learner needs, and is theoretically supported by ideas of self-directed learning such as Knowles (1980) who claims that learners should be directed by their own inquiry, rather than "fact laden lectures". The challenge for learning design is to introduce the inquiry step at such a stage when learners have had a basic introduction to the course, and to propose the activity in such a manner as to inspire learners to look back at the peer produced content through this self-constructed filter of inquiry and reflection.

The interviews of DBR2 produce rich descriptions of how learners approach the FutureLearn platform, and 4 main approaches emerged from these descriptions to describe how peer produced content is used for learning. The most common approach is to use the platform at the most convenient time and to read a few dozen comments at the end of each page; other learners are perceived in less interpersonal terms and learning from the commentary is a random and serendipitous activity. The survey results in DBR3 further supported this model of interaction as 63% of respondents read 1-20 comments per page and a further 23% read more than 20 comments but not all of them. This provides further evidence to support the cybernetic platform analysis of the methodology chapter, which concludes that progress is deemed as more important than reflection (looking back) or inquiry (discovery based on affinity), and my claim that these categories of the ABC learning design framework are missing from FutureLearn.

These findings support the claim that learners are able to level up in the 'Cultures of Participation' framework in terms of being able to "organise content to suit their individual learning needs" (the third level), which moves on from a linear consumption-based model, but it does not enable learners to work *together* more so than increasing the ability to discover and join conversations based on affinity. The work of Zappavigna (2011) provides a conceptual model for understanding the weak-tie relationship between MOOC participants, in terms of her descriptions of 'ambient affiliation' on Twitter. For Zappagivna, the hashtag functionality on Twitter creates a space where users can self-create and affiliate around themes without necessarily contributing to the thread. Ambient affiliation is related to the models of participation described within the DBR2 interviews, that learning from peer participation is a common, yet serendipitous experience. It is common because of the weight of overall participation, but not structured, and many learners bemoan the lack of a 'search' box as their solution to this problem. The CDT design takes a different approach to solving this problem, in that it encourages 'browsing' rather than 'searching' (Larsell, 2011) and that it presents words in the new form of conceptual connections in the hope of sparking a new and unexpected line of inquiry. In DBR3 I embed the practice of hashtagging throughout the whole course, to create self-sustaining categorisations at the level of 'concept' and encouraging cooperative activity where learners help each other find important content through the tool by identifying themes. In the Cultures of Participation framework, this is 'Level 3: extending the range of the environment', the penultimate level of agency proposed by this framework (Fischer, 2011).

Another aspect of cooperative behaviour proposed in DBR3 comes from the finding in DBR2 that learners answer any question that is posed on the page, so the wording of the task is altered to include *"Activity: What words did you click on, and what story does it tell?"* This simple alteration to the task proposal encourages around 25% of learners to write either a detailed modelling of how they used the CDT within this particular course, with a narrative of their inquiry and reflection on the resultant cognitive processes, a shorter albeit reflective modelling, or simply writing the words they clicked on as a guide for future learners. In encouraging 25% of learners to display this behaviour, the result is that most learners following the most common 'I read 1-20 comments' mode of participation are likely to see 4-5 comments which model how to use the CDT artefact. These comments focus on discussions of content relevant to the course and support the over-arching claim of this thesis that a mediating artefact plus pedagogical method can create new types of experience in MOOCs, where scale pivots to a position of advantage.

The CDT and CDT2.0 could be used across different large scale learning environments and MOOC platforms in terms of the way in which it enhances inquiry tasks through visualising large corpora, but any specific implementation needs to consider the material affordances of the platform, the database structure of the discussion board corpus, and the options for pedagogical scaffolding. For example, separating words and visualising them in an interactive word cloud is interoperable as it uses common web technology, but 'link-backs' to full conversations depend on how commentary is recorded and displayed. In FutureLearn, each comment has a number which can be appended to the base URL for a direct link-back. The integration of the mediating artefact into the pedagogical scaffolding of the course is also an interchangeable feature of the development methodology, as the inquiry and reflection tasks that it affords are common to all learning designs. I recommend that the use of hashtags should be generalised across all large scale learning platforms even if visualisation technology is not deployed because it enables learners to cooperate and create sub-communities which is a pedagogical step up in terms of the Cultures of Participation framework.

Further DBR research on MOOC platforms should investigate how to integrate longer form reflective writing into MOOC platforms through bookmarking and saving existing peer commentary. In this way the comment from DBR1 suggesting a clear line of inquiry takes you into the CDT activity, which presents new ideas and new lines of inquiry can be extended into longer form writing, collecting important comments and developing into the emergent practice of reflective blogging. The important point here is that learners need to start small with easy to understand processes (a proven disadvantage within the cMOOC model), but that this is just the launchpad for further interactive writing, most associated with higher levels of learning (Lapadat, 2006). All learners do not need to blog for the practices to emerge, but if each interaction is directed towards a cooperative effort, each small contribution acts as a signal towards the bigger picture. Current xMOOC platforms are not able to pedagogically level up and scale at the same time because they focus on individual progress, rather than group effort, as conceptualised through stigmergic principles of sustainable co-creation.

8.3 The importance of future research in this area

It is more important now to create new models of peer participation, than back in 2012 (“The Year of the MOOC”) because MOOCs have not only captured the mainstream, but have also positioned themselves into providers of formalised, accredited higher education[20] without radically altering the platform tools or pedagogical approaches. Indeed, in 2020 ("The Year of the Pandemic"), MOOC platforms saw a growth in enrolments of up to 640% during Spring (Impey, 2020), and this growth is likely to continue as 'online' is considered a safer and more flexible mode of education. Cooke (2020) further evidences this growth in 'online' by quoting industry figures from TechNation stating that venture capital investment in the EdTech sector is growing by 22% this year and is now worth £224bn. He continues:

> Demand for online learning has been rising exponentially across the world, even before the pandemic started. Learners of all ages are taking advantage of the opportunities afforded by a more accessible and flexible approach to education and access to the UK's finest academic institutions. (Cooke, 2020)

This echoes much of the initial hype around MOOCs from 2012, emphasising the flexibility of working online, and the access to elite institutions. A different vision may emphasise different models of learning, global cooperation and knowledge sharing through large scale educational practice. The technologies deployed in MOOC platforms are currently unquestioned as the neutral means of achieving these

---

[20] See https://www.futurelearn.com/programs for an example of how FutureLearn is moving into the mainstream market for accredited higher education.

outcomes; ‘learning’ is not analysed or critiqued in these accounts, moreover it is seen as the natural consequence of having access to material. This reflects both the ‘instrumentalist’ and the ‘essentialist’ critiques of technology given by those who view technology as a social force (Bayne, 2015; Hamilton & Friesen, 2013) and reinforces Knox’s claims that given “the international scope of MOOCs, continued research may need to explore the extent to which massive enrolments might reduce the diversity of instruction in particular disciplines” (Knox, 2014). This ‘continued research’ has never been more important as investment and enrolments outpace pedagogical innovation by such a significant margin; the repetition of claims from Pappano in 2012 and Cooke in 2020 suggest that MOOCs solely create efficiencies in terms of their economies of scale rather than creating a beautiful diversity and space for interdisciplinary, cooperative knowledge construction.

This thesis has demonstrated that FutureLearn, and specifically the cybernetic makeup of the platform, encourages an individualistic mode of participation, centred on a progression incentive. This is in opposition to a learning design where the course content is problematized and discussed in detail by a diverse range of voices. The progression incentive, coupled with behaviourist assessment in the form of quizzes, serves to reify existing knowledge through the massification of ‘broadcast education’. The major contribution I make in this thesis is to suggest that it doesn’t have to be this way; by using a flexible and practical methodology and treating learners as co-creators, it is possible to create new artefacts which afford a cooperative pedagogical approach rather than one centred on individualistic activity.

The CDT should not be the only new artefact designed for MOOC learners but should inspire learning designers and theorists to develop new media which transcends the base levels of consumption.

Marshall McLuhan famously said “the medium is the message”; this statement holds more importance in the 21st century socially digital age than when it was first printed, as platforms take increasing control not only of how we interact with each other, but also how we think about larger (societal level) concepts such as education. Online education should not have to settle for the digital chalk and talk efficiencies of scaled broadcast, rather it should strive for cooperative and collaborative co-creation of knowledge across a large-scale heterogeneous cohort of interested and active agents (‘we get what we give’ rather than ‘we get what we are given’).

8.4 Reflections on this research journey

The production of this thesis has taken me on a long journey which has only served to emphasise the increasing importance of treating students as co-creators, not subjects for data collection. Post-hoc analytics treat students as lab rats in the labyrinth and serve the learning designers into making preferred pathways more favourable, prioritising completion above all else to maximise profits for platform owners and shareholders. Learning at scale can mean 2 things: exposure to diverse voices, or an economy of scale for accreditation. I believe that it is through hearing diverse voices that we can challenge ourselves to think differently about concepts, following the sociocultural paradigm; in a massified context this needs to be organised to prevent being classed as 'overload'. Every single interview participant in DBR2 spoke of an experience when they came across a different viewpoint from reading how others had interpreted the content, which made them think again about their own understanding to further prove the strength of this type of learning. Kop and Hill describe learning as cycling through this process of discovery, modification and sharing:

> The learning process is cyclical, in that learners will connect to a network to share and find new information, will modify their beliefs on the basis of new learning, and will then connect to a network to share these realizations and find new information once more. (Kop & Hill, 2008)

I have learned that existing methods and frameworks can be used to develop pedagogical approaches which encourage this type of activity, in design-based research and the treating of learning as a design science. Following the thinking of 'The Wisdom of Crowds' (Surowiecki, 2005), scale does not mean that every single learner has to make the same action or level of participation, rather that participation should be channelled towards a cooperative effort, as opposed to individualistic, so everyone can benefit from the others' experiences and voices. The narrative arc is of MOOC design not complete, and whilst MOOCs have usurped much of open education and the lifelong learning sector, we still can refocus this land grab towards cooperative and collaborative activity through the continual development of mediating artefacts and associated pedagogical innovation. It is the role of the university partners of these platforms to bring together educational and computer science research to rebalance the pedagogical effort towards the sociocultural learning paradigm which has dominated online and distance education and defines ‘higher education’. This thesis intends to serve as a blueprint for how this can be achieved in the future by highlighting existing theoretical and methodological positions which support this ongoing research agenda.

Reference List

Andersen, R., & Ponti, M. (2014). Participatory pedagogy in an open educational course: Challenges and opportunities. *Distance Education*, *35*(2), 234–249. https://doi.org/10.1080/01587919.2014.917703

Ashby, W. R. (1956). *An Introduction to Cybernetics*. Chapman and Hall.

Attwell, G. (2007). Personal Learning Environments-the future of eLearning. *Elearning Papers*, *2*(1), 1–8.

Bady, A. (2013, May 15). *The MOOC moment and the end of reform*. The New Inquiry. https://thenewinquiry.com/blog/the-mooc-moment-and-the-end-of-reform/

Baggaley, J. (2014). MOOCS: Digesting the facts. *Distance Education*, *35*(2), 159–163. https://doi.org/10.1080/01587919.2014.919710

Bali, M. (2014). MOOC pedagogy: Gleaning good practice from existing MOOCs. *Journal of Online Learning and Teaching*, *10*(1), 44–56.

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a stake in the ground. *Journal of the Learning Sciences*, *13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Barricelli, B. R., Fischer, G., Mørch, A., Piccinno, A., & Valtolina, S. (2015). Cultures of Participation in the digital age: Coping with information, participation, and collaboration overload. In P. Diaz, V. Pipek, C. Ardito, C. Jensen, I. Aedo, & A. Boden (Eds.), *End-User Development* (Vol. 9083, pp. 271–275). Springer.

Bayne, S. (2004). Smoothness and striation in digital learning spaces. *E-Learning and Digital Media*, *1*(2), 302–316. https://doi.org/10.2304/elea.2004.1.2.6

Bayne, S. (2015). What's the matter with 'technology-enhanced learning'? *Learning, Media and Technology*, *40*(1), 5–20. https://doi.org/10.1080/17439884.2014.915851

Bayne, S., & Ross, J. (2014). *The pedagogy of the Massive Open Online Course (MOOC): The UK view*. Higher Education Academy. https://www.heacademy.ac.uk/resources/detail/elt/the_pedagogy_of_the_MOOC_UK_view

Beetham, H., & Sharpe, R. (2013). *Rethinking pedagogy for a digital age: Designing for 21st century learning*. Routledge.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Biesta, G. (2005). Against learning. Reclaiming a language for education in an age of learning. *Nordisk Pedagogik*, *25*(1), 54–66.

Biggs, J. (2003). *Aligning teaching for constructing learning*. Higher Education Academy. https://www.heacademy.ac.uk/sites/default/files/resources/id477_aligning_teaching_for_constructing_learning.pdf

Bissell, A. N. (2009). Permission granted: Open licensing for educational resources. *Open Learning: The Journal of Open, Distance and e-Learning*, *24*(1), 97–106. https://doi.org/10.1080/02680510802627886

Bloom, B. S. (1956). *Taxonomy of educational objectives. Vol. 1: Cognitive domain*. McKay.

Bozkurt, A., Honeychurch, S., Caines, A., Bali, M., Koutropoulos, A., & Cormier, D. (2016). Community tracking in a cMOOC and nomadic learner behavior identification on a connectivist Rhizomatic learning network. *Turkish Online Journal of Distance Education*, *17*(4). https://doi.org/10.17718/tojde.09231

Breslow, L. B., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Research & Practice in Assessment*, *8*, 13–25.

Brinton, C. G., Chiang, M., Jain, S., Lam, H., Liu, Z., & Wong, F. M. F. (2014). Learning about social learning in MOOCs: From statistical analysis to generative model. *IEEE Transactions on Learning Technologies*, *7*(4), 346–359. https://doi.org/10.1109/TLT.2014.2337900

Britain, S., & Liber, O. (2004). *A framework for pedagogical evaluation of Virtual Learning Environments* (No. 2; Educational Cybernetics: Reports, pp. 1–69). Bolton University. http://ubir.bolton.ac.uk/301/1/iec_reports-2.pdf

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Buckingham Shum, S., & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society*, *15*(3), 3–26.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Christenson, S. L., Reschly, A. L., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagement*. Springer.

Chua, S.-M., Tagg, C., Sharples, M., & Rienties, B. (2017). Discussion Analytics: Identifying conversations and social learners in FutureLearn MOOCs. *Workshop at the 7th International Learning Analytics and Knowledge Conference, Simon Fraser University, Vancouver, Canada*, 36–62.

Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences*, *10*(1–2), 113–163. https://doi.org/10.1207/S15327809JLS10-1-2_6

Conole, G. (2013). *Designing for learning in an open world*. Springer.

Conole, G. (2014a). A new classification schema for MOOCs. *INNOQUAL - International Journal for Innovation and Quality in Learning, 2*(3).

Conole, G. (2015). Designing effective MOOCs. *Educational Media International*, *52*(4), 239–252. https://doi.org/10.1080/09523987.2015.1125989

Conole, G. (2016). MOOCs as disruptive technologies: Strategies for enhancing the learner experience and quality of MOOCs. *Revista de Educación a Distancia*, *50*(2).

Conole, G. (2014b). The 7Cs of Learning Design—A new approach to rethinking design practice. *Proceedings of the 9th International Conference on Networked Learning*, 502–509.

Conole, G., & Dyke, M. (2004). What are the affordances of information and communication technologies? *Research in Learning Technology*, *12*(2). https://doi.org/10.3402/rlt.v12i2.11246

Conole, G., Dyke, M., Oliver, M., & Seale, J. (2004). Mapping pedagogy and tools for effective learning design. *Computers & Education*, *43*(1–2), 17–33. https://doi.org/10.1016/j.compedu.2003.12.018

Conole, G., & Fill, K. (2005). A learning design toolkit to create pedagogically effective learning activities. *Journal of Interactive Media in Education*, *2005*(1), 9. https://doi.org/10.5334/2005-8

Cooke, J. (2020, October 23). As the world embraces online learning, can the UK become a global hub for edtech? *The I*. https://inews.co.uk/opinion/comment/online-learning-uk-global-hub-ed-tech-735085

Cormier, D. (2008). *Making the community the curriculum* [EPub Book]. https://pressbooks.com/. https://davecormier.pressbooks.com/front-matter/introduction/

Cornford, J. (2000). The virtual university is… the university made concrete? *Information, Communication & Society*, *3*(4), 508–525. https://doi.org/10.1080/13691180010002314

Cuffe, P. (2015). Design patterns for promoting peer interaction in discussion forums in MOOCs. *ELearning Papers*, 64–79.

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2007). Applying multilevel modelling to content analysis data: Methodological issues in the study of role assignment in asynchronous discussion groups. *Learning and Instruction*, *17*(4), 436–447. https://doi.org/10.1016/j.learninstruc.2007.04.001

DeBoer, J., Ho, A. D., Stump, G. S., & Breslow, L. (2014). Changing "course": Reconceptualizing educational variables for Massive Open Online Courses. *Educational Researcher*, *43*(2), 74–84. https://doi.org/10.3102/0013189X14523038

Deleuze, G., & Guattari, F. (1988). *A thousand plateaus: capitalism and schizophrenia*. Bloomsbury Publishing.

Donald, C., Ramsay, E., & Joerg, I. (2017). Designing for learning in a MOOC: a pedagogical model in disguise. *Journal of Perspectives in Applied Academic Practice*, *5*(3), 90–101.

Downes, S. (2012, August 16). Connectivism and connective knowledge. *Connectivism and connective knowledge essays on meaning and learning networks*. http://www.downes.ca/files/books/Connective_Knowledge-19May2012.pdf

Downes, S. (2013). Evaluating a MOOC. *Half an Hour*. https://halfanhour.blogspot.com/2013/03/evaluating-mooc.html

Dron, J. (2006). The way of the termite: A theoretically grounded approach to the design of e-learning environments. *International Journal of Web Based Communities*, *2*(1), 3–16. https://doi.org/10.1504/IJWBC.2006.008611

Elliott, M. A. (2007). *Stigmergic collaboration: A theoretical framework for mass collaboration* [Doctoral Thesis, The University of Melbourne]. http://hdl.handle.net/11343/39359

Elliott, M. A. (2016). Stigmergic collaboration: A framework for understanding and designing mass collaboration. In U. Cress, J. Moskaliuk, & H. Jeong (Eds.), *Mass collaboration and education* (pp. 65–84). Springer.

Engestrom, Y. (2014). *Learning by expanding: An activity-theoretical approach to developmental research* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139814744

Ferguson, R., & Clow, D. (2015). Examining engagement: Analysing learner subpopulations in massive open online courses (MOOCs). *In Proceedings of the Fifth International Conference on Learning Analytics and Knowledge (LAK15)*, 51–58. https://doi.org/10.1145/2723576.2723606

Ferguson, R., Clow, D., Beale, R., Cooper, A. J., Morris, N., Bayne, S., & Woodgate, A. (2015). Moving through MOOCS: Pedagogy, learning design and patterns of engagement. In G. Conole, T. Klobučar, C. Rensing, J. Konert, & É. Lavoué (Eds.), *Design for Teaching and Learning in a Networked World* (Vol. 9307, pp. 70–84). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-24258-3_6

Ferguson, R., & Sharples, M. (2014). Innovative pedagogy at massive scale: Teaching and learning in MOOCs. In C. Rensing, S. de Freitas, T. Ley, & P. J. Muñoz-Merino (Eds.), *Open Learning and Teaching in Educational Communities: 9th European*

Conference on Technology Enhanced Learning, EC-TEL 2014, Graz, Austria, September 16-19, 2014, Proceedings (pp. 98–111). Springer International Publishing. https://doi.org/10.1007/978-3-319-11200-8_8

Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge - LAK19*, 501–510. https://doi.org/10.1145/3303772.3303775

Fischer, G. (2009). End-user development and meta-design: Foundations for cultures of participation. In V. Pipek, M. B. Rosson, B. de Ruyter, & B. Wulf (Eds.), *End-user development* (Vol. 5435, pp. 3–14). Springer.

Fischer, G. (2011). Understanding, fostering, and supporting cultures of participation. *Interactions*, *18*(3), 42–53.

Fischer, G. (2014). Beyond hype and underestimation: Identifying research challenges for the future of MOOCs. *Distance Education*, *35*(2), 149–158. https://doi.org/10.1080/01587919.2014.920752

Fischer, G. (1998). Beyond 'Couch Potatoes': From consumers to designers and active contributors. *Proceedings. 3rd Asia Pacific Computer Human Interaction (Cat. No.98EX110)*, 2–9. https://doi.org/10.1109/APCHI.1998.704130

Fisherman, E. (2016, July 5). How long should your next video be? *Wistia*. https://wistia.com/learn/marketing/optimal-video-length

Fournier, H., Kop, R., & Sitlia, H. (2011). The value of learning analytics to networked learning on a Personal Learning Environment. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 104–109. https://doi.org/10.1145/2090116.2090131

Garrison, D. R. (1991). Critical thinking and adult education: A conceptual model for developing critical thinking in adult learners. *International Journal of Lifelong Education*, *10*(4), 287–303.

Garrison, D. R. (1997). Computer conferencing: The post-industrial age of distance education. *Open Learning*, *12*(2), 3–11.

Gaver, W. W. (1991). Technology affordances. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 79–84.

Gibson, J. (1979a). *The ecological approach to visual perception*. Taylor and Francis.

Gibson, J. (1979b). The theory of affordances. In *The ecological approach to visual perception* (Vol. 8).

Gillani, N., Yasseri, T., Eynon, R., & Hjorth, I. (2014). Structural limitations of learning in a crowd: Communication vulnerability and information diffusion in MOOCs. *Scientific Reports*, *4*(6447). https://doi.org/10.1038/srep06447

Goodyear, P, & Carvalho, L. (2016). Activity centred analysis and design in the evolution of learning networks. In S. Cranmer, N. Dohn, M. de Laat, T. Ryberg, & J. Sime (Eds.), *Proceedings of the 10th International Conference on Networked Learning 2016* (pp. 218–225). Springer.

Goodyear, Peter. (2015). Teaching as design. *HERDSA Review of Higher Education*, *2*, 27–50.

Gunawardena, C., Lowe, C., & Anderson, T. (1997). Analysis of a global on-line debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*, *17*(4), 395–429.

Hagel, J., Seely Brown, J., & Davison, L. (2010). *The power of pull*. Basic Books.

Hamilton, E., & Friesen, N. (2013). Online education: A science and technology studies perspective / Éducation en ligne: Perspective des études en science et technologie. *Canadian Journal of Learning and Technology / La Revue Canadienne de l'apprentissage et de La Technologie*, *39*(2). http://cjlt.csj.ualberta.ca/index.php/cjlt/article/view/689

Hase, S., & Kenyon, C. (2001). Moving from andragogy to heutagogy: Implications for VET. *Proceedings of Research to Reality: Putting VET Research to Work: Australian Vocational Education and Training Research Association (AVETRA), Adelaide, SA*.

Haythornthwaite, C. (2009). Crowds and communities: Light and heavyweight models of peer production. *42nd Hawaii International Conference on System Sciences*, 1–10. https://doi.org/10.1109/HICSS.2009.137

Haythornthwaite, C., & Andrews, R. (2011). *E-learning theory and practice*. Sage.

Head, A. J. (1999). *Design wise: A guide for evaluating the interface design of information resources*. Information Today Inc.

Henri, F. (1992). Computer conferencing and content analysis. In *Collaborative learning through computer conferencing* (pp. 117–136). Springer.

Hewitt, J. (2005). Toward an understanding of how threads die in asynchronous computer conferences. *Journal of the Learning Sciences*, *14*(4), 567–589. https://doi.org/10.1207/s15327809jls1404_4

Hylén, J. (2006). Open educational resources: Opportunities and challenges. *Proceedings of Open Education*, 49–63.

Illich, I. (1971). *Deschooling society* (1st ed.). Harper & Row.

Impey, C. (2020, July 23). Massive online open courses see exponential growth during COVID-19 pandemic. *The Conversation*. https://theconversation.com/massive-online-open-courses-see-exponential-growth-during-covid-19-pandemic-141859

Issa, T., & Isaias, P. (2015). Usability and Human Computer Interaction (HCI). In *Sustainable Design: HCI, Usability and Environmental Concerns* (pp. 19–36). Springer London. https://doi.org/10.1007/978-1-4471-6753-2_2

Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A. C., & Brooks, C. (2018). How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*, *88*(1), 43–86. https://doi.org/10.3102/0034654317740335

Jones, C. (2015). *Networked Learning*. Springer International Publishing. https://doi.org/10.1007/978-3-319-01934-5

Kelly, A., & Lesh, R. (2000). Trends and shifts in research methods. In *Handbook of Research Design in Mathematics and Science Education* (pp. 35–44).

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 170–179.

Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in MOOCs. *Science*, *355*(6322), 251–252. https://doi.org/10.1126/science.aag2063

Knowles, M. S. (1980). *The modern practice of adult education: From pedagogy to andragogy* (Rev. and updated). Cambridge Adult Education.

Knox, J. (2014). Digital culture clash: "Massive" education in the e-learning and Digital Cultures MOOC. *Distance Education*, *35*(2), 164–177. https://doi.org/10.1080/01587919.2014.917704

Knox, J. (2015). *Posthumanism and the Massive Open Online Course: Contaminating the subject of global education* [Doctoral Thesis]. University of Edinburgh.

Knox, J., & Bayne, S. (2013). Multimodal profusion in the literacies of the Massive Open Online Course. *Research in Learning Technology*, *21*(21422). https://doi.org/10.3402/rlt.v21.21422

Kop, R. (2011). The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *The International Review of Research in Open and Distributed Learning*, *12*(3), 19–38.

Kop, R., & Hill, A. (2008). Connectivism: Learning theory of the future or vestige of the past? *The International Review of Research in Open and Distributed Learning*, *9*(3). http://www.irrodl.org/index.php/irrodl/article/view/523

Koper, R & Olivier, B. (2004). Representing the learning design of units of learning. *Journal of Educational Technology & Society*, *7*(3), 97–111. JSTOR.

Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., Čukić, I., de Vries, P., Hatala, M., Dawson, S., Siemens, G., & Gašević, D. (2017). Exploring communities of inquiry in Massive Open Online Courses. *Computers & Education*. https://doi.org/10.1016/j.compedu.2017.11.010

Krotoski, A. (2012, November 6). *Chance* (No. 7) [Mp3]. https://www.bbc.co.uk/sounds/play/b01jqfk8

Lapadat, J. C. (2006). Written interaction: A key component in online learning. *Journal of Computer-Mediated Communication*, *7*(4), 0–0. https://doi.org/10.1111/j.1083-6101.2002.tb00158.x

Larsell, L. (2011, June 9). Why browsing is so important to content discovery. *Mashable.Com*. https://mashable.com/2011/09/06/browsing-content-discovery/

Laurillard, D. (1993). *Rethinking university teaching: A conversational framework for the effective use of learning technologies*. Routledge.

Laurillard, D. (2009). The pedagogical challenges to collaborative technologies. *Internationa lJournal of Computer-Supported Collaborative Learning*, *4*(1), 5–20. https://doi.org/10.1007/s11412-008-9056-2

Laurillard, D. (2012). *Teaching as a design science: Building pedagogical patterns for leaning and technology*. Routledge.

Lave, J. (1988). *Cognition in Practice: Mind, mathematics and culture in everyday life*. Cambridge University Press. https://doi.org/10.1017/CBO9780511609268

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. https://books.google.co.uk/books?id=CAVIOrW3vYAC

Leont'ev, A. N. (1981). *Problems of the development of the mind*. Progress.

Lin, F., & Puntambekar, S. (2019). Designing epistemic scaffolding in CSCL. *A wide lens: Combining embodied, enactive, extended, and embedded learning in collaborative setting*, *2*, 597–601. https://cscl2019.com/

Littleton, K., & Mercer, N. (2013). *Interthinking: Putting talk to work*. Routledge.

Liyanagunawardena, T. R., Parslow, P., & Williams, S. A. (2017). Exploring 'success' in MOOCs: Participants' perspective. In *Massive Open Online Courses and Higher Education* (pp. 92–108). Routledge.

Mackness, J., Mak, S., & Williams, R. (2010). The ideals and reality of participating in a MOOC. *7th International Conference on Networked Learning*, 266–274.

McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and evolving a concept. *Proceedings of the Graphics Interface 2000 Conference, May 15-17, 2000, Montr'eal, Qu'ebec, Canada*, 179–186. http://graphicsinterface.org/wp-content/uploads/gi2000-24.pdf

McGuire, R. (2013, September 3). Building a sense of community in MOOCs. *Campus Technology*. https://campustechnology.com/articles/2013/09/03/building-a-sense-of-community-in-moocs.aspx

McLuhan, M. (1964). *Understanding media: The extensions of man*. Gingko Press.

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners*. Multilingual matters.

Mercer, N. (2003). The educational value of dialogic talk in whole class dialogue. *New Perspectives on Spoken English in the Classroom: Discussion Papers*, 73–76.

Mercer, N. (2007). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics and Professional Practice*, *1*(2), 137–168.

Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *British Journal of Educational Psychology*, *80*(1), 1–14.

Messina, C. (2007, August 25). Groups for Twitter; or A proposal for Twitter tag channels. *Factory Joe*. https://factoryjoe.com/2007/08/25/groups-for-twitter-or-a-proposal-for-twitter-tag-channels/

Neilsen, J. (2012, January 4). Usability 101: Introduction to usability. *Neilsen Norman Group*. https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Norman, D. (1988). *The psychology of everyday things*. Basic Books.

Norman, D. (1998). *The invisible computer: Why good products can fail, the personal computer is so complex, and information appliances are the solution*. MIT Press.

Oh, E., & Reeves, T. C. (2010). The implications of the differences between design research and instructional systems design for educational technology researchers and practitioners. *Educational Media International*, *47*(4), 263–275. https://doi.org/10.1080/09523987.2010.535326

Oliver, M. (2005). The problem with affordance. *E-Learning and Digital Media*, *2*(4). https://doi.org/10.2304/elea.2005.2.4.402

Oliver, M. (2011). Technological determinism in educational technology research: Some alternative ways of thinking about the relationship between learning and technology: Educational technology and determinism. *Journal of Computer Assisted Learning*, *27*(5), 373–384. https://doi.org/10.1111/j.1365-2729.2011.00406.x

Oliver, M. (2013). Learning technology: Theorising the tools we study. *British Journal of Educational Technology*, *44*(1), 31–43. https://doi.org/10.1111/j.1467-8535.2011.01283.x

Plato, & Grube, G. M. A. (1976). *Plato's Meno*. Hackett Pub. Co.

Postman, N. (1992). *Technopoly: The surrender of culture to technology*. Vintage Books.

Postman, N., & Weingartner, C. (1969). *Teaching as a subversive activity*. Dell Publishing Co.

Qiu, M., Hewitt, J., & Brett, C. (2012). Online class size, note reading, note writing and collaborative discourse. *International Journal of Computer-Supported Collaborative Learning*, *7*(3), 423–442. https://doi.org/10.1007/s11412-012-9151-2

Reich, J. (2015). Rebooting MOOC research. *Science*, *347*(6217), 34–35.

Richey, R., & Nelson, W. (1996). Developmental research. In D. H. Jonassen (Ed.), *Handbook of Research for Educational Communications and Technology* (pp. 1213–1245). Macmillan.

Rodriguez, C. O. (2012). MOOCs and the AI-Stanford like courses: Two successful and distinct Ccourse formats for Massive Open Online Courses. *European Journal of Open, Distance and E-Learning*, *2012*(1).

Rosé, C. P., & Ferschke, O. (2016). Technology support for discussion based learning: From computer supported collaborative learning to the future of Massive Open Online Courses. *International Journal of Artificial Intelligence in Education*, *26*(2), 660–678. https://doi.org/10.1007/s40593-016-0107-y

Savitz, E. (2012, May 11). Jack Dorsey: Leadership secrets of Twitter and Square. *Forbes Magazine*. https://www.forbes.com/sites/ericsavitz/2012/10/17/jack-dorsey-the-leadership-secrets-of-twitter-and-square/#668a6b285e2b

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal education in a knowledge society* (pp. 67–98).

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, *27*(2), 4–13. https://doi.org/10.2307/1176193

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, *2*(1), 3–10.

Siemens, G. (2012, July 25). MOOCs are really a platform. *Elearnspace: Learning, Networks, Knowledge, Technology, Community*. http://www.elearnspace.org/blog/2012/07/25/moocs-are-really-a-platform/

Simon, H. A. (1996). *The sciences of the artificial*. MIT Press.

Stacey, P. (2014). Pedagogy of MOOCs. *International Journal for Innovation and Quality and in Learning*, *3*, 111–115.

Stewart, B. (2013). Massiveness+ openness= new literacies of participation. *Journal of Online Learning and Teaching*, *9*(2), 228–238.

Stracke, C. M. (2017). Why we need high drop-out rates in MOOCs: New Evaluation and Personalization Strategies for the Quality of Open Education. *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, 13–15. https://doi.org/10.1109/ICALT.2017.109

Sturtz, D. N. (2004, December 1). *Communal categorization: The folksonomy*. http://davidsturtz.com. http://davidsturtz.com/drexel/622/communal-categorization-the-folksonomy.html

Surowiecki, J. (2005). *The wisdom of crowds* (1. ed). Anchor Books.

Tubman, P., Oztok, M., & Benachour, P. (2016). Being social or social learning: A sociocultural analysis of the FutureLearn MOOC platform. *The 16th IEEE International Conference on Advanced Learning Technologies*. The 16th IEEE International Conference on Advanced Learning Technologies, Austin, Texas.

Tubman, P., Benachour, P., & Oztok, M. (2018). *Design based research approaches towards enhancing social learning practices in MOOC platforms*. 13th International Conference of the Learning Sciences, London, UK.

Tubman, P., Benachour, P., & Oztok, M. (2019). *Interactive visualizations to enhance social learning practices in MOOC platforms*. *2*, 925–926.

https://www.scopus.com/inward/record.uri?eid=2-s2.0-85073321565&partnerID=40&md5=b707239777f0439470c3fe7c2ddb18b5

Tubman, P., Oztok, M., & Benachour, P. (2019). New platform affordances for encouraging social interaction in MOOCs: The 'Comment Discovery Tool' interactive visualisation plugin. *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, *2161-377X*, 34–36. https://doi.org/10.1109/ICALT.2019.00014

Tubman, P., Oztok, M., & Benachour, P. (2020). Can we afford it? The cybernetic determinants for pedagogical models in MOOCs. *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, 11–15. https://doi.org/10.1109/ICALT49669.2020.00010

van den Akker, J., Branch, R. M., Gustafson, K., Nieveen, N., & Plomp, T. (Eds.). (1999). *Design approaches and tools in education and training*. Springer Netherlands. https://doi.org/10.1007/978-94-011-4255-7

Van Dyke Parunak, H. (2006). A survey of environments and mechanisms for human-human stigmergy. In D. Weyns, H. Van Dyke Parunak, & F. Michel (Eds.), *Environments for Multi-Agent Systems II* (Vol. 3830, pp. 163–186). Springer Berlin Heidelberg. https://doi.org/10.1007/11678809_10

Veletsianos, G., Collier, A., & Schneider, E. (2015). Digging deeper into learners' experiences in MOOCs: Participation in social networks outside of MOOCs, notetaking and contexts surrounding content consumption: Digging deeper into learners' experiences in MOOCs. *British Journal of Educational Technology*, *46*(3), 570–587. https://doi.org/10.1111/bjet.12297

Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature Published in 2013–2015. *International Review of Research in Open and Distributed Learning*, *17*(2). http://www.irrodl.org/index.php/irrodl/article/view/2448/3655

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Walji, S., Deacon, A., Small, J., & Czerniewicz, L. (2016). Learning through engagement: MOOCs as an emergent form of provision. *Distance Education*, *37*(2), 208–223. https://doi.org/10.1080/01587919.2016.1184400

Weir, I. (n.d.). *Spearman's correlation*. Retrieved 15 December 2018, from http://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge university press.

Wertsch, J., & Bivens, J. (1992). The social origins of individual mental functioning: Alternatives and perspectives. *Quarterly Newsletter of the Laboratoryof Comparative Human Cognition*, *14*(2), 35–44.

Wilson, B. G. (1996). *Constructivist learning environments: Case studies in instructional design*. Educational Technology.

Wise, A. F., & Cui, Y. (2018). Learning communities in the crowd: Characteristics of content related interactions and social relationships in MOOC discussion forums. *Computers & Education*, *122*, 221–242. https://doi.org/10.1016/j.compedu.2018.03.021

Young, C., & Perović, N. (2016). Rapid and creative course design: As easy as ABC? *Procedia - Social and Behavioral Sciences*, *228*, 390–395. https://doi.org/10.1016/j.sbspro.2016.07.058

Young, J. (2019, November 7). How Udacity could return to its higher ed roots. *EdSurge*. https://www.edsurge.com/news/2019-11-07-how-udacity-could-return-to-its-higher-ed-roots

Yuan, L., Powell, S., & Olivier, B. (2014). Beyond MOOCs: Sustainable online learning in institutions. *Cetis. White Paper. Recuperado de Http://Publications. Cetis. Ac. Uk/2014/898*.

Zappavigna, M. (2011). Ambient affiliation: A linguistic perspective on Twitter. *New Media & Society*, *13*(5), 788–806. https://doi.org/10.1177/1461444810385097

Zappavigna, M. (2012). *Discourse of Twitter and social media*. Continuum.