Integration of Evaluation Processes into e-learning Environments.

Developing the learning of Practical Science with the Bradford Robotic Telescope

James Piers MACHELL

Submitted for the Degree of

Doctor of Philosophy

Faculty of Engineering and Informatics

University of Bradford

2017

Abstract

James Piers MACHELL

Integration of Evaluation Processes into e-learning Environments: Developing the learning of Practical Science with the Bradford Robotic Telescope

Keywords: e-learning, pedagogy, evaluation, conversational theory, confidence testing, robotic telescope, science education.

This thesis presents a novel framework for the collection and evaluation of data around e-learning. It shows how e-learning can play a positive role in empowering teachers in reflective practice through accessible statistical methods, as part of an evidence-based approach. Within this new framework data generated by pupils' actions in three levels of pedagogical activity: declarative content based, functional tools based and social functional are aligned with three levels of evaluation: satisfaction, learning and behavioural changes.

The framework is evaluated using the e-learning system for the Bradford Robotic Telescope. Analysis and assessment of the data by using the tools and concepts of statistics are performed. The first group of 78 clusters, mean size of 25.6 pupils, are examined to determine the effect of contextual factors. Pupils responded consistently to the same learning design across a range of contexts. National indicators for social/economic, academic achievement and group size are examined for bias.

Suitable measurements for interpretation by simple Gaussian distributions are identified: satisfaction through the frequency of use with a probability of $P>0.05$, learning through formative assessment ($P>0.3$) and behavioural changes through engagement with higher order activities ($P>0.2$).

The second group of 168 clusters, mean cluster size 25.9, demonstrates a meaningful effect size for a change in approach within the e-learning system in the areas of satisfaction, learning and behavioural changes with a probability of $P<0.01$.

Acknowledgements

I would like to thank

My Supervisors

My Family

Everyone on the Telescope Team

The Schools involved

For my son, may he reach for the stars

Table of Contents

Abstract i

Acknowledgements ii

List of Figures x

List of Tables xiii

Acronyms xiv

Chapter 1. Introduction 1

1.1 Introduction 1

1.2 Aims and Objectives of this Thesis 7

1.3 Definitions of the Learner and the Teacher their Actions and Roles 8

1.4 The Bradford Robotic Telescope 10

1.4.1 Telescope Outreach Programme 10

1.4.2 The Telescope Setup 11

1.4.3 Website and Associated Technologies 13

1.5 How this Thesis is Structured 14

1.6 Chapter Summary 18

Chapter 2. E-learning 19

2.1 The aim of this Chapter 19

2.2 Definition of e-learning 19

2.2.1 Theory and Practice 21

2.2.2 Design 21

2.2.3 Management 22

2.2.4 Utilisation..22

2.2.5 Evaluation 23

2.2.6 Networked learning Technology 23

2.3 Technology in e-learning...27

2.3.1 Asynchronous Technologies and Synchronous Technologies 28

2.3.2 Virtual Learning Environments, Managed Learning Environments, Learning Management Systems 29

2.3.3 Narrative and interactive media...29

2.3.4 Support and Communication Tools 30

2.3.5 Adaptive Teaching Systems and Intelligent Tutoring Systems......30

2.3.6 Learner Authoring Tools and Collaborative Technologies.............31

2.3.7 Massively Open Online Courses 32

2.3.8 Mobile Learning...32

2.3.9 Learning Analytics...32

2.4 Adoption of e-learning..33

2.4.1 E-learning as a Solution to Business Training..............................33

2.4.2 E-learning Adoption in Higher Education......................................34

2.4.3 E-learning Adoption in Schools 35

2.4.4 Common Themes...36

2.5 Chapter Summary and Conclusions 37

Chapter 3. Pedagogy of e-learning Environments 39

3.1 The aim of this Chapter 39

3.2 The Impact of Pedagogy on the Design of e-learning Environments ... 39

3.3 Pedagogy in e-learning 41

3.4 Content Centred Learning Environments 43

3.4.1 Student Activity in the Environment 44

3.4.2 Support for Students 45

3.5 User-Generated Content Centred Learning Environments 46

3.5.1 Learning Through Tools 47

3.5.2 Learning Through Communities and Social Interactions 48

3.5.3 Reflective Content Systems 50

3.5.4 Support for Students 50

3.6 Motivation of Students 51

3.7 Adoption of an E-learning Framework 52

3.8 Implementation of Ideal Learning Environment 58

3.9 Chapter Summary and Conclusions 61

Chapter 4. Alignment of Pedagogy and Assessment 62

4.1 Chapter Aims and Objectives 62

4.2 The Aim of Assessment 63

4.3 Alignment of Assessment 66

4.4 The Use of Multiple Choice Questions 68

4.5 Measurement of Confidence 69

4.6 Assessment to Promote Deep Learning Outcomes 71
4.7 Chapter Summary and Conclusions 72
Chapter 5. Evaluation 75
5.1 Aims and Objectives 75
5.2 The Need for Evaluation Processes in e-learning 75
5.3 Who Should be Doing the Research? 78
5.4 How can Teachers Engage in the Research Process? 79
5.5 Evaluation Methodology...81
5.5.1 Selection of a Research Approach 83
5.5.2 Developing an Experimental Method..85
5.6 Alignment of Learning Analytics with Pedagogy and Learning Design 89
5.7 Measuring Satisfaction...90
5.8 Measuring Learning 91
5.9 Behavioural Change a Move Towards Deep Learning 92
5.10 Evaluation Conclusions 93
Chapter 6. Implementation of the e-Learning Environment...............................95
6.1 Outline of this Chapter 95
6.2 Implementation of The Telescope e-learning System 95
6.2.1 Learning Objects and Content Centred Virtual Learning Environments...97
6.2.2 Lesson Plan Route System 103
6.2.3 Multiple Choice Question System...103

6.2.4 Storage of User Response to Multiple Choice Question 105
6.2.5 Telescope Systems 105
6.2.6 Social and Reflective System: Spacebook 105
6.2.7 Objective System 107
6.2.8 User Interface 107
6.2.9 Implementation of Student Model 108
6.2.10 Teacher Model 110
6.3 Summary of Chapter 111
Chapter 7. Method of Data Analysis and Evaluation 114
7.1 Evaluation of the delivery methods 114
7.2 The Context of the Study 115
7.2.1 National Curriculum Requirements for Teaching 115
7.2.2 Profile of UK Primary School Teacher 116
7.2.3 The Setting of Learning and Teaching in most UK Schools 116
7.2.4 Bradford and the Surrounding Area 117
7.3 Experimental Design 120
7.3.1 Delivery of Introduction to the Bradford Robotic Telescope 121
7.3.2 Development of Learning Materials for Comparison 121
7.3.3 Pilot Study 125
7.4 Clustered Method of Analysis 128
7.5 Measuring satisfaction 130

7.6 Measuring learning 131

7.6.1 Confidence Measurement 132

7.7 Measuring Behavioural Changes 134

7.8 Analysis Against Known Characteristics of Subject Clusters 135

7.9 Analysis of Correlations Between Evaluation Tools 136

7.10 Comparison of Different Approaches 136

Chapter 8. Results 137

8.1 Modelling Against Predicted Distributions 137

8.2 Satisfaction 138

8.2.1 Satisfaction Method 1: Number of Visits 139

8.2.2 Satisfaction Method 2: Period of Use 140

8.2.3 Satisfaction Method 3: Frequency of Use 142

8.2.4 Analysis of Known Factors for satisfaction measurements 143

8.3 Analysis of Learning Measurement for Clusters 146

8.3.1 Analysis of Known Factors for Learning Measurements 149

8.3.2 Confidence-Based Testing 152

8.4 For Behaviour 155

8.4.1 Analysis of known factors for behavioural Measurements 156

8.5 Summary of Validation of Measures for Evaluation 158

8.6 Comparison of Approaches 159

8.6.1 Satisfaction Test Comparison 159

8.6.2 Learning Test Comparisons 161

8.6.3 Behavioural Test Comparison 162

8.7 Evidence of Relationship: Satisfaction -> Learning -> Behaviour 163

Chapter 9. Discussion and Conclusion 167

9.1 Discussion of Tools for Measuring Key Factors in Educational Interventions 168

9.1.1 The Measure of Satisfaction 168

9.1.2 Measure for Learning 170

9.1.3 Analysis of Confidence-Based Questions 172

9.1.4 Measures of Behavioural Change 173

9.1.5 Analysis of Known Factors 174

9.2 Comparison with Alternative Delivery Method 175

9.3 Comparison to Other Studies 178

9.4 Conclusions from Multi-dimensional Evaluation Framework Tools 188

9.5 Further work 191

9.5.1 Measures of Satisfaction. 192

9.5.2 Measures of Learning 193

9.5.3 Behavioural Measures 196

9.6 Conclusions 197

References 200

List of Figures

Figure 1.1 Structure of Thesis..16

Figure 2.1 Relationship in terms of e-learning definition, adapted from Reiser & Ely (1997) analysis of the definition of educational technology........................26

Figure 2.2 Grouping of e-learning research areas, Conole and Oliver (2007) ..27

Figure 3.1 Common learning design categories with example adapted from Lockyer et al (2012)..53

Figure 3.2 Laurillard's conversational framework (2008) 55

Figure 3.3 Framework for e-learning environment...58

Figure 3.4 Model of the First level e-learning Environment: Content Centred Environment. Learningscope 59

Figure 3.5 Second level of e-learning environment: tools level.........................59

Figure 3.6 Social level of the learning environment: Spacebook 60

Figure 5.1 Data use and applications from Greller & Drachsler (2012).............77

Figure 5.2 Model of transformative reflection presented by Biggs and Tang (2007) p49 79

Figure 6.1: Example of learning object data structure.....................................102

Figure 6.2 Website interface 107

Figure 7.1 Number of visits per week per user following initial contact via different learning routes (Machell and Baruch, 2015) 125

Figure 7.2 Spread of groups answer questions in the pilot study (Machell and Baruch, 2015) 126

Figure 7.3 Confidence question 621 (Machell and Baruch 2015) 127

Figure 7.4 Confidence question 623 (Machell and Baruch, 2015) 127

Figure 8.1 Plot of the mean number of visits per user grouped by clusters 139
Figure 8.2 Histogram of the frequency of the mean number of logins per user grouped by cluster 139
Figure 8.3 Mean number of days use for clusters 140
Figure 8.4 Histogram of the mean number of days use per user in a group ...141
Figure 8.5 Mean frequency of visits per person for clusters............................142
Figure 8.6 Effect of cluster size on the frequency of visits 143
Figure 8.7 Effect of percentage of students achieving level 4 and above in maths and English on the frequency of visits 144
Figure 8.8 Effect of percentage of students achieving level 4 and above in Science on the frequency of visits..145
Figure 8.9: Effects of percentage disadvantaged pupils on the frequency of visits...145
Figure 8.10 Raw data for clusters mean route facility 146
Figure 8.11 Results for cluster mean route facility. 147
Figure 8.12 Histogram frequency of groups achieving a given route facility ...147
Figure 8.13 Cronbach's alpha for learning measurement 148
Figure 8.14 Route facility against the number in a cluster 149
Figure 8.15 Effect of Math and English level 4 and above on the route facility 150
Figure 8.16 Effect of percentage achieving science level 4 and above on the test scores 151
Figure 8.17 Effect of percentage disadvantaged students in school on route facility...151
Figure 8.18 Effect of item facility on confidence grouped per question 152

Figure 8.19 Effect of confidence on item facility 153
Figure 8.20 Effect of mean residual confidence on item facility 154
Figure 8.21 Mean Number of Spacebook Entries for Class Clusters 155
Figure 8.22 Frequency of clustered mean space book entries per pupils 155
Figure 8.23 Number of pupils in a cluster against the number of space book entries 156
Figure 8.24 Effect of percentage achieving Maths and English level 4 and above on behavioural changes 157
Figure 8.25 Effect of percentage achieving science level 4 and above on behavioural changes 157
Figure 8.26 Effect of Percentage of Disadvantage Students on Behavioural Measures 158
Figure 8.27 Comparison on satisfaction Method 3 Frequency of use Table 8.1 Comparison of Frequency of Visits in the First 30 Days of Use 160
Figure 8.28 Comparison of Learning Measure...161
Figure 8.29 Comparison of Behavioural Changes 162
Figure 8.30 Graph of the Correlation of Satisfaction Method 1 and Learning .163
Figure 8.31 Correlation of Satisfaction Method 3 and Learning Measures. 164
Figure 8.32 Graph of the correlation between learning and behavioural changes 165
Figure 8.33 Graph of the correlation between satisfaction and behavioural changes 165
Figure 9.1 Framework for evaluation and learning relationship. 191

List of Tables

Table 4.1 Confidence categories from Kampmeyer et al (2015) 70

Table 6.1 Database Table for Sections in Learning Objects Implementation..100

Table 6.2 Database Table for Layout in Learning Object Implementation 101

Table 6.3 Example database table for text content 101

Table 6.4 Database Table for Storage of Question Response.........................105

Table 6.5 Database Table for Student Core Information.................................109

Table 6.6 Database Table for Student Access..110

Table 7.1 Percentage of student achieving English and Maths level 4+ at key stage 2 Department for Education (2015) 118

Table 7.2 Science performance by district 2010 – 2013, Department for Education (2015)..119

Table 7.3 Percentage of schools seen as good or outstanding by district (Ofsted n.d.)..119

Table 7.4 Learning Design Metadata for Route 65 123

Table 7.5 Learning Design Metadata for Route 84 124

Table 8.1 Comparison of Frequency of Visits in the First 30 Days of Use 160

Table 8.2 Comparison of Learning Measure...161

Table 8.3 Results from Behavioural Comparison..162

Table 9.1 Results and Methods of Analysis of Comparable Studies...............184

Acronyms

BBC	British Broadcasting Corporation
BRT	Bradford Robotic Telescope
CBT	Confidence Based Testing
cMOOC	Constructionist massively open online courses
CSS	Cascading Style Sheets
DfES	Department for Education and Skills
HE	Higher Education
HTTP	HyperText Transfer Protocol
IAC	Instituto de Astrofisica de Canarias
IEEE	Institute of Electrical and Electronics Engineers
KS	Key Stage
MCQ	Multiple Choice Questions
MOOC	Massively Open Online Courses
n.d.	No Date
Ofsted	Office for Standards in Education, Children's Services and Skills
PHP	PHP: Hypertext Preprocessor
SQL	Structured Query Language
SOLO Taxonomy	Structure of the observed learning outcome
UK	United Kingdom of Great Britain and Northern Ireland

Chapter 1. Introduction

1.1 Introduction

This thesis examines the role of evaluation in e-learning and how through the integration of evaluation into of e-learning environments it is possible to enable teachers' engagement in a reflective cycle in order to improve the process of teaching and learning.

E-learning has four overlapping areas of research: technology, pedagogy, organisational and socio-cultural (Conole and Oliver 2007). Here it is suggested that these areas must be looked at together and the relationships between these areas considered if it is going to be possible to build effective frameworks for evaluation of learning and teaching involving e-learning environments.

The social and cultural changes that have taken place as part of the information revolution have changed the way in which information is accessed and shared. Ideas and technologies previously only available to research groups in universities are now available to the general public (Beetham and Sharpe 2013). Google through its Google Home (Google n.d.) and Amazon through their echo (Amazon n.d.) technologies are attempting to supply voice controlled lifestyle assistance that can answer our verbal questions. This is something that would have been the dreams of intelligent tutoring systems of the 1980s (Luckhardt Redfield and Steuck 1991) who were taking the first steps into web-based technologies. In parallel Facebook with two billion users (British Broadcasting Corporation 2017) has changed the social world for a generation. However, the same changes haven't taken place in e-learning, Laurillard (2007) points out that these changes have been held up by institutional systems and not moved at the same speed as technological or wider social changes. This has resulted in technology being used at the fringe of learning rather than as a transformative force. The information revolution needs to be thought of as more than just the use of this new technology in traditional teaching, in the way previous technologies might have been, instead, it represents a paradigm shift in education into the information age (Beetham and Sharpe 2013). These changes are happening at all levels of culture and will impact learning in both Higher Education all the way down to Primary Schools and so it is important to study the changes and possibilities at all levels.

E-learning has progressed from a technology-driven area, which was seen as synonymous with distance learning, through an approach that saw it as the a personalised computer-based experience (Tavangarian et al 2004) into a tool that is seen as a part of everyday teaching to be utilised in the way any other teaching strategy might be (Biggs and Tang 2007). The focus has moved towards what the learners do (Beetham 2013). New Technology must be approached with pedagogy in mind, reviews of learning technology have shown that in and of itself it does not improve learning (Alexander 2001), and often has a smaller effect on teaching than many other factors (Biggs and Tang 2007).

There is a need to be open with pedagogical and data collection methodology as this is a building block of good e-learning research (Oliver et al 2007). It will affect the design of the technology (Alexander 2001), the way in which it will be used and how existing technology is evaluated (Britain and Liber 2004), as our approach towards pedagogy changes, so will the technology to reflect this (Laurillard 2013). If data is to be analysed correctly the teaching strategies that are used to collect the data must be understood and aligned (Lockyer et al 2013). It will then be possible to unlock the potential in these new technologies and the data on learning that they can make available.

There are two contrasting pedagogical approaches to e-learning and these result in very different designs for e-learning environments. The first approach is a content centred approach which builds out of associationist principles (Dyke et al 2007) that ultimately ends up in the principles of learning objects (IEEE 2002). The second approach builds on social and constructionist principals and sees pupil activity at the centre of learning using the experience to build up conceptual understanding (Mayes and de Freitas 2004; Oliver et al 2007; Beetham 2013; Keengwe 2013). These are not the only viewpoints available as many including Laurillard's conversational framework (2002) or the more popular in traditional practice of Bloom's revised taxonomy (Krathwohl 2002), provide opportunities to combine ideas within learning environments (Dyke et al 2007), as is the case in many Higher Education institutions where a mixed approach is taken (Mayes and de Freitas 2004). Chapter 3 will look more closely at these different approaches and will build a framework for an e-learning environment.

Laurillard (2008) points out that a framework is needed in order to test new approaches through action research. Adoption of either of these pedagogical approaches will affect every following stage of research, as they not only place a subjective stance on what good learning is but they change the opportunities and methods for data collection (Oliver et al 2007).

Further to this, the use of the information and the stakeholders that will be using it must be considered. There are a number of different levels where learning analytics can be used to inform e-learning (Greller and Drachsler 2012). Feedback can inform an individual student and can inform teachers of the actions to take with that individual and feedback from the teaching of a group can inform the actions to take for that group going forward, this is in line with the principles of assessment for learning (Black and William 1998; Black et al 2007). The same may not be true for the next time the subject is taught, what worked for one student may not work for the next and the same is true for the next group. However, this does not mean that it is not possible to plan for the next group or that our approach cannot be informed. Beetham and Sharpe (2013) outline the idea of design for learning, and Biggs and Tang (2011) talk about curriculum alignment in planning our structures for activity.

The framework presented in this thesis aims to show how a data-driven reflective approach can allow us to see past the individual context but at the same time does not ignore its influence. A key research question of this thesis is whether this is possible within the contextual restrictions of multi-institution primary school teaching environments, an area much less closely examined in the teaching and learning in Higher Education or in Massively Open Online Courses (MOOC). The null hypothesis is that institutional differences make too large a difference for meaningful comparisons to be made against data collected and changes within the reach of the e-learning environment to be implemented. This is where a real-world clustered approach comes in: by taking account of context and appreciating the limitation of the scope of our understanding it is possible to develop systems to analyse strategies for learning design.

Ultimately the use of any evaluative system is to aid in the process of teaching and should be designed to help teachers in making and developing their own practice and not in providing off the shelf solutions (Laurillard 2008). Students must still be supported in their learning regardless of the approach. A big part of providing this support in teaching is the scholarship of teaching (Trigwell et al 2000) and the process of understanding what works, taking informed views from pedagogical research, then trying out strategies in a localised context and then using data to evaluate and reflect on the success of the process. This is a process Biggs and Tang (2011) call transformative reflection. There is still an unnatural split between researchers and teachers, why would so many be planning and doing but not evaluating (Laurillard 2008; Beetham and Sharpe 2013). The aim here is to provide a framework that will aid in this process not automate it. Caution should be taken in trying to fully automate any part of the teaching process it is important to look at the internal limitations of data and to ensure that the any presented materials can be fully interpreted (Greller and Drachsler 2012). E-learning is full of examples where attempts to automate the learning/teaching cycle have failed, Intelligent tutoring systems have often failed due to the lack of available soft tacit information (Dyke et al 2007) and adaptive personalised teaching systems often failed because it was unclear how to match learners to alternative approaches. As such, there has been a move away from adaptive approaches towards a blended approach (Beetham 2007).

E-learning has moved towards a pupil-centred pedagogically informed approach, away from the traditional view of a tool for distance learning. It is now a ubiquitous part of teaching. However, its impact is still mainly at the fringes of learning. To have an impact it must help in the full learning and teaching process, by not only having a pedagogically informed approach to its use in learning but by supplying tools that will help to inform teaching and learning. Teachers are time short and under great pressure from budgetary and other factors to deliver quality outcomes for their students, e-learning can help them to achieve this but only if it is involved and aligned with all stages of the teaching cycle from the pedagogical approach to teaching, the collection of data for assessment and the evaluation of the success of teaching. This thesis will present a novel framework to achieve this and will critically evaluate it against its use in a multi-institutional primary setting.

1.2 Aims and Objectives of this Thesis

This thesis reports on a research programme to develop an effective means of integrating evaluation into an e-learning environment. A novel framework for the collection and evaluation of data is proposed and evaluated. The Framework aligns the pedagogical features of the learning environment with formative assessment within the environment. The evaluation approach taken is to use processes integral to learning through the environment without the need to collect additional data. It will look at the specific example of Primary school Science activities with the Bradford Robotic Telescope and its associated e-learning environment, and it will evaluate this example and show how these principles of evaluation and the framework adopted can be applied generally.

Within the schools environment there are natural clusters of students with similar traits, this could be because of the setting of students due to their ability or the social economic factors found in the district where the school is located and these must be taken into consideration in the analysis of data from schools (Taylor et al 2013). In order to look at underlying domain-specific pedagogical considerations, beyond the context of an individual school, class or student within a primary school setting a clustered approach to data collection and analysis will be taken. This approach is tested by modelling data against predicted distributions. The impact of known characteristics of clusters will be examined to look at their effect on the distribution of data. The aim is to show that the data collected within a range of contexts is suitable for statistical analysis without knowledge of all of the contextual issues.

The aim is to provide a tool that allows meaningful comparisons within the classroom situation of a primary school teacher to support the teacher as an *action researcher* (Laurillard, 2008, Biggs & Tang, 2011) involved in a meaningful cycle of investigative practice.

Two approaches to the learning material are then tested: the intervention using an applied approach, in which students use “scientific principles” to solve an authentic problem and the control using a “pure science” approach, where students collect data to provide evidence about natural phenomenon. The data from the two approaches are compared to see the level of impact that changes within the learning environment have. The aim is to allow a judgement on how large a change is required in order to effect meaningful change within students learning.

The results of these experiments are used to critically evaluate the framework and recommendations to aid in the design of future systems will be made.

The next part of this thesis will examine and define the roles of the teacher and the learner as well as the context that this study takes place in, including the use of the Bradford Robotic Telescope.

1.3 Definitions of the Learner and the Teacher their Actions and Roles

There is a need to define the roles of the teacher and the learner and the concept of pedagogy. These definitions help to frame this work as they further explain the approach and lines of examination when trying to understand e-learning and the benefits it could bring. It is important to separate the ideas of learning and teaching, as they are two different processes.

It could be this confusion that leads some to see a gap between the theory and practice. Some would critique constructivism saying that it is not practical in teaching, it is not an instructional design theory but a theory of knowing. Constructivism describes what goes on in a learner's head and so is less directly applied to teaching (Karagiorgi & Symeou 2005). Being a learning theory does not mean that there has not been work to look at the implications of how to teach with it. With a framework of how individuals learn it is possible to formulate how to teach. A clear definition of the two areas of teaching and learning allows a better understanding of the processes involved for both teachers and their students.

In this thesis, *learning* refers solely to the action of the pupil and the process they go through to gain new knowledge skills or ideas.

Teaching refers to the actions taken by a teacher to promote learning in a student.

Others have highlighted this separation as that of the difference between a theory of learning, which is the research-based influences on the learning process and the mechanisms by which these influences act and the pedagogical frameworks: how these theories are applied to learning and teaching practice (Mayes and de Freitas 2004).

Pedagogy sits at the interplay between teaching and learning. The word being derived from Ancient Greek, meaning the one who leads children to their place of learning, this fits with a modern view of aiding pupils to learn. Learners in many pedagogical theories are no longer seen as passive receivers of information but this does not mean that they do not need help with their learning (Beetham & Sharpe 2007). Pupil-centred models of learning do not in any way devalue the role of the teacher (Biggs and Tang 2007). It is only by looking at both teaching and learning together that it is possible to get a clear view of the process.

1.4 The Bradford Robotic Telescope

This thesis focuses on systems being built for the purpose of supporting the use of the Bradford Robotic Telescope and helping, in particular, its schools programme of engagement with the Telescope. The aim was to supply resources that would enable primary school pupils to discover the wonder of the night sky, in an age of light pollution (Baruch et al 2005). The design and principles, however, are proposed as examples to be applied beyond this case.

The telescope's website and control servers supplied both opportunities and restrictions to the implementation of any design. Fundamentally it was the aim of the learning materials to aid users in their use of the robotic telescope, aiding them in the taking of images that would aid their learning of astronomy.

1.4.1 Telescope Outreach Programme

The telescope is one of a number of telescopes available to schools and the general public over the web (Camacho et al 2009). Since its conception the telescope has been designed with public access in mind, the first telescope on Oxenhope Moor coincided with the first years of the internet becoming the World Wide Web and so the two seemed a perfect match, with a publically accessible robot connected through a website. Early versions of the telescope online learning environment on which this work is based were a key part of the outreach in the Tenerife based evolution of the telescope and were aimed at helping primary schools pupils discover for themselves the wonder of space. The aim was to inspire young people into science through access to the wonder of the night sky (Baruch et al 2005).

Modern life and the problems of light pollution means that many are out of touch with astronomy and the fundamental changes to the environment around us (Baruch et al 2004). The phase of the Moon plays only a small role in everyday life. Jaffacake can advertise with a teacher explaining the Moon to school children, with full Moon, half Moon and total eclipse (cultofJaffa 2011). This in no way implies that teachers do not understand the Moon's cycles but rather demonstrates its minimal relevance to everyday life. Amazingly similarly inspired teaching activity with Oreos can be found, as a way of adding some fun to the topic (NASA n.d.). However, this shows the difficulty in teaching a topic that is abstract, and although the use of simulations can help, it is through a robotic telescope that it is possible to make the experience both real and scientific (Baruch et al 2005).

1.4.2 The Telescope Setup

The telescope has a number of different users of the system, public users and the users in schools: teachers and pupils. The system put in place by Tallon (2010) for the interface between the public users, who were thought of as amateur astronomers, and the telescope was based on the idea that these individuals were rich in their astronomy knowledge and only needed a simple to use but highly controllable system for ordering images for their astronomy needs.

The telescope system itself is an autonomous request based telescope (Baruch et al 2004), this is an asynchronous mode of operation where the user places orders onto a central database of requests and the telescope takes the images at the optimum time. The telescope is based at the Mount Teide observatory on Tenerife (Instituto de Astrofisica de Canarias n.d.) The process described by Tallon (2010) uses the position of the celestial object in the night sky and the popularity of the object ordered. The second parameter affects jobs with the same settings which are stacked and the image taken farmed out to all the users that wanted that object.

The ordering of images of a celestial object with the telescope is through a wizard system. Object selection is made either through a name, a catalogue number or right ascension and declination coordinates. Following the selection of the object, an instrument needs to be selected three devices are available:-

- Constellation camera: a 40 degrees wide-field images of the night sky
- Cluster camera: a 4.3 degree image. Good for clusters like the Pleiades
- 24 arc seconds field of view (~0.5 degrees or slightly less than the whole of the Moon)

(Telescope.org n.d.)

Finally, the user must decide on the camera settings, choosing the exposure time and the filters with which to take the image. A range of image processing options are also available which help to improve the quality of the images returned these include, dark frames and flat fielding.

The image taking process is very quick, lasting only a few minutes, however, the time for the image to be returned could be anywhere from a single night to multiple nights, however, most images are returned within a week (Camacho et al 2009).

The asynchronous nature of taking images means that the telescope is best suited to learning activities lasting a number of days and or weeks, thus allowing time for any images taken to be returned to the user.

The setup also massively reduces the complexity of ordering an image with a telescope as no knowledge of the movements and positions of objects in the night sky is required by the user in order to place an observation (Camacho et al 2009).

1.4.3 Website and Associated Technologies

The web interface and the technologies used to implement the resource were developed by Tallon (2010) Fundamentals of the web system could not be changed, neither was it feasible that they could be transitioned to alternative technologies for the benefit of the virtual learning environment to be implemented.

The telescope system was built on a Linux Apache web server stack with PHP and a PostgreSQL database (Tallon 2010). Apache is an open source

HTTP web server software, it is the most popular HTTP server on the web. (Apache n.d.). It is responsible for serving web-based content such as images and text via the HTTP protocol. PHP is a server-side language that allows users to access dynamically generated web pages (PHP n.d.). PostgreSQL is a relational database programme that allows the fast storage and retrieval of information (PostgreSQL n.d.). This combination of software allows the telescope website to dynamically present information to users over the web.

The Bradford Robotic Telescope has a user account system, developed through the above technologies, that allows users to access to the telescope, order pictures and process the images that they received, these could be stored in a central library of processed images for others to view (Tallon 2010).

It was onto this system that the e-learning environment was to be built, due to a large number of systems that it would interact with and the high level of interaction that was required, it was felt that a custom system would allow more control, rather than developing plugins or adapting an already existing system like Moodle (Moodle 2016). Although either approach would have been possible with an open source approach. Without the telescope and its associated functions and users systems, a different approach may have been taken.

1.5 How this Thesis is Structured

Although the narrative of a thesis is linear, the process of evaluation, or design for learning (Beetham and Sharpe 2013) and curriculum alignment

(Biggs and Tang 2007) are all non-linear, iterative processes. An order in which these areas would be approached had to be selected for this thesis. Starting with a pedagogical approach, then a framework for curriculum and assessment design and finally an evaluation strategy. The aim is to show the alignment of design with data collection and evaluation.

Chapter two, following a definition of e-learning, will present a framework of the different areas of research: pedagogy and technology through design and use of technology for learning (Cook et al 2007), management, evaluation and social context of e-learning processes. It is proposed that these need to be looked at together rather than as separate fields as each one affects the others (Conole and Oliver 2007). The adoption of e-learning and the development of technology will be examined. Chapters 3-7 will then develop an aligned framework in which learning, teaching and evaluation of the education process can take place. The third chapter will outline a pedagogical approach in e-learning and the features of an *ideal* learning environment. The fourth chapter will examine the role of assessment in curriculum design and its alignment with the pedagogy of student activity. The aim of these chapters is to outline the framework that the analytical system will use to collect data.

The fifth chapter will look at the approach to the evaluation of educational activities within the system, drawing on the wider community of approaches to educational and pedagogical research.

The rest of the thesis will then focus on the evaluation of the system, looking at its ability to collect and measure data on the users to allow the process of transformational reflection. This is done through chapter six and seven that in detail cover the methodology of the system design, data collection and analysis and its evaluation. Chapter eight will present the data collected to test the system and chapter nine will discuss these results, critically evaluate the system's effectiveness and recommend future work.

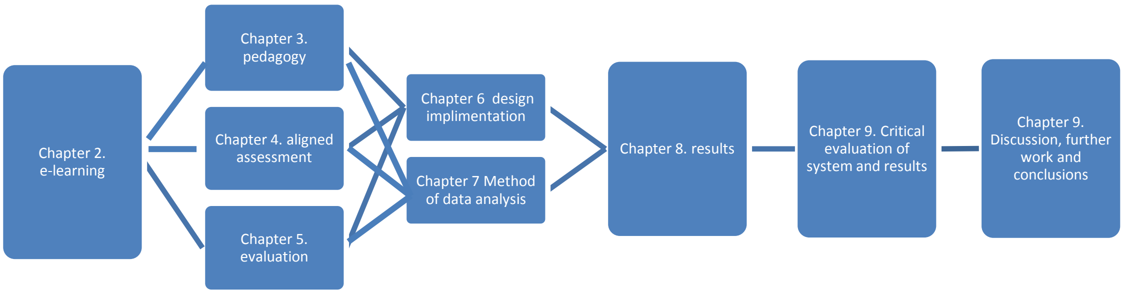

Figure 1.1 Structure of Thesis

Chapter 2: e-learning.

Definition of e-learning and its technologies, review of its adoption across educational levels.

Chapter 3: Pedagogy of e-learning

Impact of pedagogical approaches on e-learning. Framework for e-learning environment.

Chapter 4: Alignment of assessment

Framework for the use of assessment as the tool to help deliver communication and feedback in e-learning.

Chapter 5: Evaluation

Approach to and aims for the use of evaluation within the e-learning environment; alignment of evaluation with pedagogical approaches and assessment.

Chapter 6: Implementation of the e-learning environment

Outlines of the systems and tools developed for the learning environment.

Chapter 7: Method of Analysis and Evaluation

Methods for delivering learning and tools for assessment and evaluation are outlined. Experimental design for testing both the environment and its evaluation tools are described in chapter 6.

Chapter 8: Results

Results from the use of the e-learning system by two groups; the first group of 78 clusters which are used to look at the distributions of a single learning design to determine the impact of contextual factors and the suitability for comparison in the multi-institutional approach. Multiple methods for the three areas of measurement are examined. The second group of 168 clusters are used to examine the effect size of possible approaches to see if approach changes can make meaningful differences compared with contextual changes.

Chapter 9: Discussion and Conclusion

A critical evaluation of the framework is given in light of the results in chapter 8. Recommendations for future work are made.

1.6 Chapter Summary

This chapter has outlined the objectives of the thesis as the presenting a novel approach to the alignment of pedagogical features of e-learning environments with multi-dimensional learning analytics in a cross-contextual primary school situation with the aim of aiding teacher lead transformative reflection and action research in subject-domain-specific approaches.

The chapter has also outlined the structure and approach taken to the thesis and the background in which this study has taken place of the Bradford Robotic Telescope along with the technologies involved.

The next chapter will present a working definition of e-learning through a critical review of the literature.

Chapter 2. E-learning

2.1 The aim of this Chapter

This chapter aims to provide a definition of e-learning that will help to frame the work in future chapters. It will show how e-learning is a multi-dimensional discipline with a range of research areas including pedagogy, learning logistics and evaluation. In order to fulfil these roles a range of technologies have been developed and these will be examined. Finally, this thesis will look at the adoption of these technologies across the different educational areas.

This will set the stage for future chapters that will show how through aligning evaluation with the different aspects of e-learning it is possible to provide an effective framework for developing learning and teaching within these systems.

2.2 Definition of e-learning

The idea of using computers to aid learning is not a new one. The field of computer-assisted instruction can be traced back as far as 1955 when it first appeared as a way of teaching problem solving (Aparicio & Bacao 2013). Educational technology sees its routes going back even further to the audio-visual movement of the 1920s. The development of the definition of instructional technology or educational technology is inclusive of computerised and web-based learning as these technologies develop (Reiser & Ely 1997). The term e-learning is cited as first being used in 1999 at a CBT systems seminar (Al-Saai et al 2011) and was in common use by 2000 by which time the Journal of Network and Computer Applications had dedicated an issue to the emerging field (Ghaoui & Taylor 2000). In 1991 in a paper predicting the future of intelligent tutoring systems, Luckhardt Redfield and Steuck (1991) describe a world which was then 30 years into the future, where learners access a holographic intelligent agent that finds for the user the knowledge they need through a network enabled device that combines both computer and telephone. This Intelligent Tutoring System would be far more integrated into our general use of computers. Despite the probably Back to the Future (Zemeckis 1990) inspired holographic misgiving, the experience described is close to that which many may have using their smartphones to help them find information through Google. It is also clear that Intelligent Tutoring System of the 1990s aspired to a more connected and integrated approach, more akin to what might be seen as e-learning.

The need for clear definitions of e-learning or networked learning lies in the fact that without a clear definition it is hard to outline the areas of development (Bricheno et al 2004). Confusion in the definition often leads to the interchangeability of terms from online learning, or education, technology-based training, web-based or computer-based training, distance learning and education (Kumaran & Nair 2010).

No single definition should be seen as definitive and it is the right of any writer to define their own meaning, and for the reader to agree or disagree with it, however, a clear definition allows us to frame the work and define what is in or out of the field (Reiser & Ely 1997). The definition here although based on others is an attempt to frame this work its limits and the areas it covers.

e-learning is the theory and practice of design, management, utilisation and evaluation of networked learning technologies.

The following sections will examine the definition given above

2.2.1 Theory and Practice

E-learning is a practical applied field and like educational research, it is interested in improvements that can be put to use (Beetham and Sharpe 2013). Often the use of e-learning does not improve learning on its own. There are a range of factors that influence successful e-learning and as such practice must be informed with research (Alexander 2001) and research with best practice.

2.2.2 Design

There has been a shift in approaches from early e-learning which focused on technology to a more complete area of research ranging from technology through content, communication and pedagogical learning strategies involving these tools. (Aparicio & Bacao 2013).

Here Design refers to the advancement and use of technologies with consideration of the pedagogies for their use. Whether this is design for learning (Beethams and Sharpe, 2013) or learning design (Agostinho et al 2013; Lockyer et al 2013; Oliver et al 2013) which are amongst the different approaches within e-learning, it is the consideration of pedagogies and the sharing of practice, to the development of better teaching, that are common in approach.

2.2.3 Management

Although many point out that it is important to focus on the pedagogical changes that e-learning can bring about (Laurillard 2007), the benefits to managing the process of learning through information communication technology include less pedagogically centred benefits of flexibility, coping with increased numbers, sharing and reusing resources and reducing administrations burden (Britain and Liber 1999) and are of great value. It is through this managed process many new opportunities such as learning analytics arise (Lockyer et al 2013).

2.2.4 Utilisation

The field has two influences that of research into education and its enhancement through associated technologies and the social pressures brought about by the rapid adoption of networked technology in wider society. The information age has brought about an epoch change in the way many live their lives and this cannot be ignored in education, even if institutional systems have caused its adoption to lag behind wider social uses (Laurillard 2007).

The Bradford Robotic Telescope school programme first started its development in 2003 a world in which many of the technologies widely available today were being first adopted. Computer rooms were just starting to be introduced to schools and some of the first visits to schools took place in what felt like converted janitors cupboards. Downloading a class load of images brought schools networks to a standstill and many schools still ran internet explorer 5 or even 4, meaning that there were difficulties providing interactive elements. In the years since fast broadband internet has been introduced, mobile phones have become smart with the iPhone appearing in 2007 (Chen 2009), Youtube (Youtube n.d.) has reinvented the way that videos are watched on the web, Myspace has stopped being a popular alternative (Dredge 2015) to Facebook, which was still in its early stages having only been founded in 2004 (FaceBook n.d.) and had nowhere near its two billion users it has today (BBC 2017). It could be argued the progress in e-learning has been far less dramatic held back by the institutional changes rather than progressing at the speed of cultural changes (Laurillard 2007).

Beethams and Sharpe (2007) argues that new technologies offer more than just a few new technologies to help with learning, instead, they are an epoch change in learning, with an information age there are many new opportunities for learning

2.2.5 Evaluation

It is not just the use but the effective use of technology and through evaluation a better understanding and informed field of research. E-learning is a science based on evidence which should provide improvement and understanding, if this is on a local level reporting to teams and projects then many would call it an evaluation. If the evaluation is concerned with theories and a contribution to knowledge then it is research (Oliver et al 2007b)

2.2.6 Networked learning Technology

Key to the technology is its network capability with its ability to join people with each other and information to enable the learning process.

There is some conflict here as to the aim of e-learning. Tavangarian et al (2004) views e-learning as independent learning, rejecting the idea of its use for distance learning. Others, although not seeing it as distance learning, see more than just information being shared. It is the ability to build and sustain community online that is as important as the information, the tacit benefits of learning as well as the implicit. Restricted views of e-learning have been compared with the proverbial elephant with e-learning, only one part of the picture. Just as e-learning is supportive of the traditional classroom-based learning the same is true of distance, or independent learning it is just one of the tools available (Kenney et al 2004).

JISC (2004) takes this extended view of the type and situation of the learning in its definition

"e-Learning can cover a spectrum of activities from supporting learning, to blended learning (the combination of traditional and e-learning practices), to learning that is delivered entirely online. Whatever the technology, however, learning is the vital element."

"E-learning can be seen as network-enabled internet-based learning" (Gunasekaran et al 2002).

The words pedagogical, education, teaching and learning were all considered. Learning technology might make some think that all the focus on the student and so, in turn, might have led some down the path of automation of the learning process. Here the aim is to join student to student, student to teacher and student to information.

"E-learning is defined as pedagogy empowered by digital technology which involves learning done at a computer, usually connected to a network, giving us the opportunity to learn almost anytime, anywhere. “ Kumaran & Nair (2010).

Similarly, the words educational or teaching were not used for their inflexions. Teaching for many would have meant a teacher-centric model,

“The delivery of education (all activities relevant to instructing, teaching and learning) through various electronic media.” (Koohang and Paliszkiewicz 2013).

As can be seen in this definition there is the need to clarify education as a range of activities and approaches. Within the context of this study, the choice of the word pedagogical over learning might have been apt due to the inflexion of the idea that pedagogy is the science of interplay between student and teacher to enable the student to learn (Beetham and Sharpe 2013). It presents a student centred approach but it does not imply that a student is left to discover solely for themselves. It also encompasses a vast variety of approaches to achieve the goal of helping the student to learn. However, the word learning is used over pedagogy as there are examples of e-learning, such as the hole in the wall project (Dangwal et al 2002), where even young students are able to successfully self-organise their education using e-learning in an informal setting. Although the definition is to be used to set the context of this thesis for the reasons outlined above in self-regulated learning it is only part of a bigger picture.

It is not by mistake that this definition draws a comparison to the 1994 definition of educational technology both the history and nature of the fields are closely linked in their development (Reiser & Ely 1997). Although many of the terms might take on a quite different meaning in the definition provided in 1994.

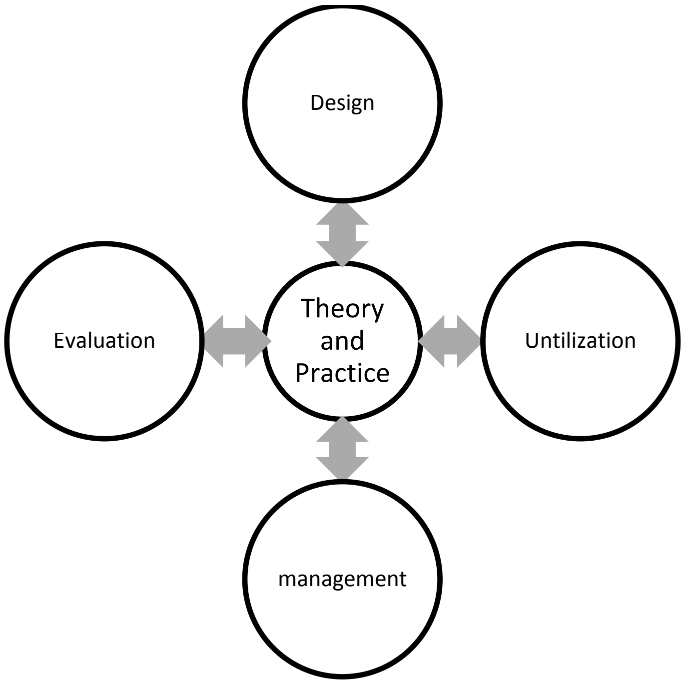

Figure 2.1 Relationship in terms of e-learning definition, adapted from Reiser & Ely (1997) analysis of the definition of educational technology

The importance of a wheel and spoke relationship of key areas (Figure 2.1) is that it shows that it is not a linear process (Reiser & Ely 1997) and that the influences run in both directions both into the field of e-learning and out.

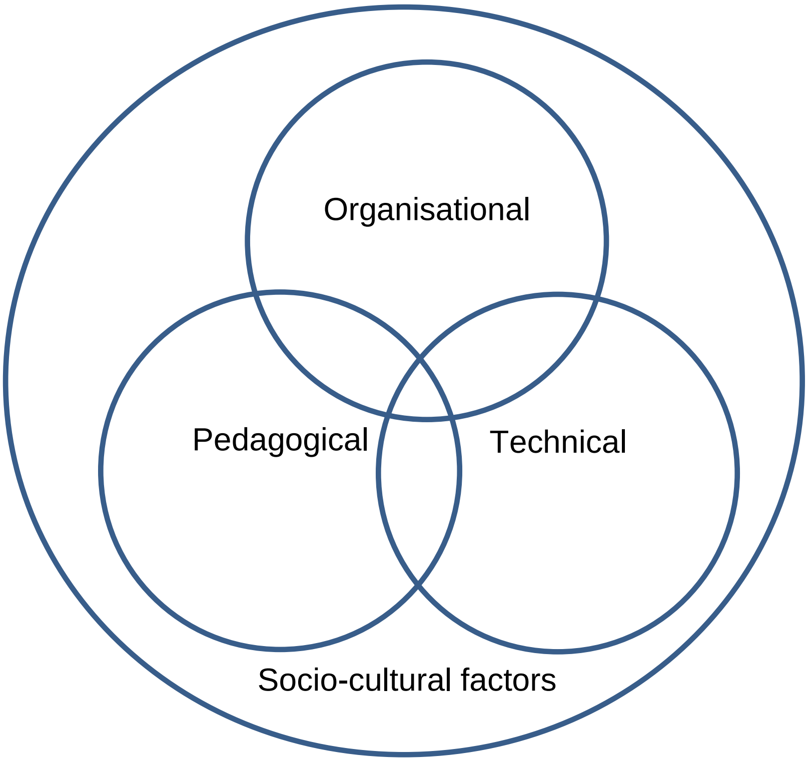

Figure 2.2 Grouping of e-learning research areas, Conole and Oliver (2007)

As can be seen in Figure 2.2 the same themes of e-learning are identified by Conole and Oliver (2007). Here the influence of cultural changes is the sea in which organisational, pedagogical and technical research is framed. A principle of the definition above is that they must be considered together as they all impact on each other.

2.3 Technology in e-learning

There have been many different trends in e-learning technologies over the years. Laurillard categorised technology into five categories in 1993 audiovisual, hypermedia, interactive media, adaptive, discursive (Laurillard 1993) and then re-categorised them in 2002, narrative, interactive, adaptive, communicative and productive. As technologies move forward categorisation will change further and some dominant areas may disappear as they are replaced with new possibilities. Cook et al (2007) take a different approach naming the clustering terms and their associated technology. Kumaran & Nair (2010) also categorise the different areas of e-learning technology by bringing these three models together it is possible to build up a model of the technologies. It is by no means meant to be exhaustive or limiting but is put together to show the foci of e-learning technology.

2.3.1 Asynchronous Technologies and Synchronous Technologies

The largest difference in approach to e-learning is the choice of asynchronous or synchronous technologies. Synchronous technologies aim to either use the technology to bring people together at the same time or to enhance a group of learners learning at the same time. Asynchronous technologies aim to enhance the anytime anywhere nature of e-learning.

The distinction between the two areas is not as clear cut as one might first think. Many institutions choose to take a blended approach incorporating both synchronous and asynchronous as well as e-learning and traditional learning approaches in the same course (Biggs and Tang 2011).

Even technologies that might at first appear to be the definition of one approach have had developments to aid the alternative methods. Lecture board casts and live classroom programs can incorporate capture technology in order to allow users to play back the experience at a later date and are integrated into the environments to promote self-paced learning as well as synchronous learning opportunities (Yang and Liu 2007).

2.3.2 Virtual Learning Environments, Managed Learning Environments, Learning Management Systems

Virtual learning environments are online learning solutions for students, they handle a range of activities including managing students' progress and access to learning materials or activities. Managed learning environments take the features a step beyond managing the learning to managing the full educational process (Britain and Liber 1999).

A wide range of off the shelf solutions are available Moodle (Moodle 2016), and blackboard (Blackboard n.d.) to name but two. Primarily these systems are designed to aid in the logistics of e-learning bringing students, tutors and information together to enable learning. They often include a large variety of learning tools to deliver many of the different areas below.

2.3.3 Narrative and interactive media

Laurillard's (2002) definitions of both narrative and interactive media are separated only by the fact that with the advent of ICT technologies gives access to a much larger non-linear information base. Traditional narrative media, such as a book, provides information and structure but is non-discursive in as much that a book contains the same information each time you read it, a film doesn't have a different ending if you watch it a second time. Interactive media or hypertext-based media allows free navigation through linked topics to provided interactivity through the plethora of media, but revisiting the same page does not change what it said or the content of the video that you watch.

This might be most peoples' idea of e-learning, but the content is by no means the only part of any learning environment (Britain and Liber 1999)

2.3.4 Support and Communication Tools

Online support, forums, chat rooms bulletin boards, email or live instant messaging (Kumaran and Nair 2010), these are often considered an essential part of larger VLE systems (Britain and Liber 1999) but are also found on their own as a tool for learning in the wider web (Kumaran & Nair 2010) where they support a product or are built up around an area of interest. They are listed separately here for this reason.

2.3.5 Adaptive Teaching Systems and Intelligent Tutoring Systems

Adaptive teaching systems cover a broad range of activities in which feedback is given in light of action on the part of the user. This range of tools extends to but is not limited to simulations (Laurillard 2002). The teaching process in the real world is by its very nature adaptive, simple pen and paper activities allow this, discursive methods even more so, by its very nature e-learning needs planning and structure either in the development of the learning materials students are to view or the tools they are to use to communicate (Beetham and Sharpe 2007). Adaptive teaching systems or intelligent tutoring systems are all developed around the principle that the learning and teaching process should have some level of automation in e-learning (Laurillard 1993). For simple narrow domains such as passive voice tutoring systems, it has been possible to develop systems that employ scaffolded methods in order to achieve this (Virvou et al 2000). However, more general approaches that have tried to build on the learning object design and the aim that they could be reused and combined in multiple courses have been less successful. Some have put this down to the idea that this was because they tried to match to defined characteristics rather than use more advanced data mining methods (Atif et al 2003). However, unlike commerce where there is a clearly defined success criterion of the sale education is more complex. Such adaptive systems have found themselves in conflict with the move away from a consumption model towards a constructive learning pedagogy. Intelligent tutoring systems have not been as successful as many hoped they would be failing to provide the level of interaction and intelligence that early work promised (Laurillard 2002).

2.3.6 Learner Authoring Tools and Collaborative Technologies

Collaborative learning is where groups engage in social activities which are mediated through the use of technology. The aim is to use technology to help students learn together (Aparicio & Bacao 2013).

From the mid-2000s web 2.0 was starting to take off and by 2010 had been identified a trend in e-learning. Traditional e-learning follows the approach software that is structured around courses, timetables, and testing this is all too often driven by the needs of the institution rather than the needs of the learner. E-learning 2.0 focuses on tools for users to generate their own content and share it with others. Tools include blogs, wikis, YouTube and can extend to Twitter and even Facebook as it is the users that drive the content of these sites. (Kumaran & Nair 2010).

Specific e-learning examples include e-Portfolios and reflective technologies. e-portfolio tools are specifically designed to support the reflective process in allowing students to gather examples of their work in an online environment which can then be shared with others as evidence of their learning (Welsh 2012).

2.3.7 Massively Open Online Courses

MOOCs, (Massively Open Online Courses) which many believe have the potential to completely change the way learning takes place online with 20,000-180 000 pupils enrolled on a course. In fact, there are much lower numbers taking even the first examination in these courses and so it is important to understand what causes students to enrol but not continue (Aparicio & Bacao 2013).

2.3.8 Mobile Learning

With the increasingly ubiquitous nature of mobile devices, new opportunities for learning are possible particularly in situative learning processes that could include authentic tasks for students. There are also benefits around informal learning, these build on the idea of personalised access by individuals from anywhere at any time (Kukulska-Hulme and Traxler 2013)

2.3.9 Learning Analytics

Learning analytics developed as a trend and by 2011 the first conference on learning analytics is held (Greller and Drachsler 2012). Learning analytics is the use of data from e-learning to better understand/inform the learning and teaching process on a number of different levels including highlighting student behaviour to inform teachers to better understanding learning design. It has become part of many Virtual Learning Environments.

It is a rapidly expanding area of research and this has led to some to call it the new black (Booth 2012). However, a key problem in the field is finding data sets in which research can be carried out. Research often uses sets from Open learning environments supplied by MOOCs or those from Higher Education institutions' own learning management systems (Greller and Drachsler 2012).

2.4 Adoption of e-learning

The approach to adoption and the aims of adoption of e-learning within the different sectors has affected their uptake of different technologies and the speed at which these have been adopted.

Business, Higher Education and School adoption will be looked at separately and then the common themes and lessons will be examined.

2.4.1 E-learning as a Solution to Business Training

Key to the idea of business training has been the goal of anytime anywhere learning (Gunasekaran et al 2002; Kenny et al 2004). Businesses see the advantage of not having to organise their workforce into a single face to face environment as a huge money saving benefit of e-learning. As such many systems for business are geared towards asynchronous managed learning environment solutions. In 2001 there was already a strong belief that education will be the next big thing for the internet (Alexander 2001).

Business is often more interested in the logistical solution rather than the pedagogical excellence of the materials provided so there is only a token amount of evaluation of the quality of training and as with any area of business there have been those less scrupulous and more interested in selling their product than its suitability of it for its context (Gunasekaran et al 2002).

By 2004 it was estimated that over a quarter of European vocational and continuing professional development users' time in training is in blended or pure e-learning (Kenny et al 2004)

The market went through rapid growth in the early part of this century with surveys reporting that 25% of businesses had used e-learning in 2002 (Gunasekaran et al 2002) this had increased to reported levels of 99% in 2004 (Kenny et al 2004).

2.4.2 E-learning Adoption in Higher Education

Adoption of e-learning was taken up as it was believed it would increase access and reduce the cost of Higher Education, which was in a funding crisis. It is hoped that through e-learning an improved quality in teaching can be achieved (Alexander 2001). Virtual learning environments first started to appear in the later part of the 1990s, initially, they were not web-based. Adoption of cross-campus adoptions was common by 2003 (BECTA 2003). Institutional adoption of e-learning was in general to sustain mass learning rather than to develop pedagogical approaches (Bricheno et al 2004).

UK higher education has taken an approach of blending e-learning methods with traditional methods and teaching with the aim of further integrating new technologies such as mobile learning into the model (Kenny et al 2004). Learning management system like Blackboard are used by a large number of universities to facilitate both on campus and distance courses (Gunasekaran et al 2002).

Due to the blended learning approach that many institutions took for the adoption of e-learning, the ability of network learning to be developed has been limited. During the process of adoption little was done to recognise the changing the role of staff and to ensure the quality of teaching or innovate teaching practice (Bricheno et al 2004). As a results e-learning is often used to help support the logistics of learning rather than to tackle the larger pedagogical issues it could help to improve. (Laurillard 2007)

2.4.3 E-learning Adoption in Schools

The adoption of e-learning in schools was slower in schools than Higher Education, with the UK behind examples found in the USA and Europe (BECTA 2003). A key problem in the adoption of e-learning solutions within mainstream education in the UK has been with the haphazard way that institutions have been able to individually invest meaning that systems are not always compatible and/or cost-effective (Department for Education and Skills 2005). Best practice was not shared before mass adoption was undertaken. This has not meant that schools have not had high aims for their adoption with a key benefit being E-delivery which is seen as a more efficient way of communicating to third parties such as parents than traditional print methods. Although they are keen to point out that this is not to replace face to face contact but to aid individuals in being better informed and therefore making these face to face visits more effective. Schools should be able to share everything from the homework that they set to details on the syllabus that the students should be learning. This can go beyond into the world of work with school leavers having an e-portfolio (Department for Education and Skills 2005).

Schools have tended to buy off the shelf products and the pedagogical quality of some of these is questionable. Attempts were made to control the amount of education content that the BBC put together trying to push the idea that not every situation suits e-learning and that it is often better for pupils to learn from concrete experiences such as understanding the push and pull of magnets rather than dealing with the extra level of abstraction that a simulation on a computer might add (BECTA 2005).

The adoption of ICT in primary schools has had a focus on the opportunities that individual tools that software could supply the schools i.e. photo, text and music programs. The adoption of Learning Management Systems is less advanced although the use for the sharing of homework and email is found. In terms of hardware, schools are moving away from fixed equipment to Wi-Fi enabled solutions to better enable the use by pupils. In contrast to HE and business who see potential money-saving opportunities, this is not seen as big an advantaged in a classroom situation. One of the major factors named in the adoption of ICT in schools is the increased motivation of students in use (Kalaš et al 2014).

2.4.4 Common Themes

Although the level of adoption varies across the different sectors and the reason for adoption might not be the same some aspects are common

- Adoption of e-learning is often for logistical reasons but in order for it to be used effectively new pedagogies need to be considered
- There needs to be recognised that the time to adopt new technologies needs time set aside for both students and teachers
- E-learning is adopted as it improves communications and accountability in the learning process.

It could be seen that the speed at which business has adopted e-learning solution is in part due to the fact that it has outsourced the production and delivery of training. However, this has resulted in some questionable practice. In contrast, Higher Education has been slower to adopt and has looked instead for hosting and management systems for which they have developed their own content. This has, in turn, lead to issues around a lack of support and or time allowance for work around developing and supporting e-learning in addition to the traditional teaching and learning time allowance (Blass and Davis 2003). A strong background of research and collaboration in Higher Education means that work has been undertaken to make sure that e-learning tools are developed that are scalable and interoperable (Govindasamy 2002). Of all three schools are the least developed and most disjointed. However, their needs, wants and requirements are no less sophisticated and there are also the demands of third party involvement in the form of parents/guardians which are not be involved in other levels of an educational institution to the same degree.

2.5 Chapter Summary and Conclusions

In this chapter three views of e-learning have been put forward in order to better understand the field and its applications: a definition, the technology areas and its adoption.

E-learning has seen a shift from its traditional view of a consumption-based model towards a more collaborative model of learning. Early systems for e-learning were often more concerned with the logistics and practices of the organisations that they were adapted for rather than the pedagogical benefits that they could bring.

There has been a shift in the definition of e-learning and its aims from enabling anywhere any time learning to technology to enhance pedagogy in both blended and distance situations.

Constructionist approaches to e-learning have redefined the use of technology and the development of web 2.0 technologies have gone hand in hand with this approach, it has caused the decline of ideas around virtual teachers and places the teacher and students as the central parts of an e-learning environment.

The field of e-learning is one that incorporates a wide range of activity in education from the pedagogy of delivery to the management of the logistics of education.

The aim of the next chapter is to look more closely at the range of pedagogical approaches and their implications for technology with the aim of outlining a framework for the telescope and its e-learning environment.

Chapter 3. Pedagogy of e-learning Environments

3.1 The aim of this Chapter

This chapter will consider a range of pedagogies and their impact on the design process, the aim is to create a list of features for an *ideal* learning environment that would facilitate these needs for both the teacher and the learner.

The focus is on the pedagogical features of the learning environments rather than how these functions could be carried out technologically. As discussed in chapter two e-learning tools can be either asynchronous or synchronous, for remote users or users in a blended environment. It is the aim of the tool or the ability it enables in the learning and teaching processes that are of interested, the why not the how. A full description of the implementation of this *ideal* learning environment is given in chapter 6.

This chapter focuses on the process of teaching and learning within the environment and not the logistics of how these processes are managed which are covered in more detail in Chapter 4 through the principles of curriculum alignment and assessment.

3.2 The Impact of Pedagogy on the Design of e-learning Environments

Some would argue that there are no pedagogical models of e-learning only models of learning and methods for how they are enhanced by technology (Mayes and de Freitas 2002) some take an alternative view that the affordances of e-learning technology prescribes pedagogical use (Laurillard et al 2000). It is important to consider the pedagogical approach used, as this enables successfully selection of technological learning tools (Laurillard 1993; Laurillard 2002; Mayes and de Freitas 2013). This is certainly the approach taken by Britain and Liber (1999) in their models for evaluating online learning environments and supported by the process put forward by Alexander (2001), which places the tools used as part of a wider picture in which pedagogy and environment are both considered. This chapter will look into the wider picture of pedagogical approaches and their influences on the development of e-learning environments.

Numerous studies have shown that the use of information technology in and of itself does not improve education (Alexander 2001). An example of this would be the use of PODcasts, few would have argued that the one thing needed in education is more people talking at you (Laurillard 2008). Another example would be the placing of lecture notes online, which does little to promote learning (Cook et al 2007).

An essential part of the development of an e-learning environment is the consideration of pedagogical principles that underlie the system, which is often one of the most neglected factors. Some developers of e-learning even try to distance themselves for the pedagogical implications of their tools as they wish to remain unbiased in their approach or appearing restrictive in the potential uses of their tools (Govindasamy 2002). Others would argue that it is this approach that has meant that e-learning has only been used to change the fringes of educational approaches rather than be used as the transformational tool for learning that it is (Laurillard 2007). Beetham and

Sharpe (2007) supports this idea saying that e-learning should not be thought of in terms of how it can be used in line with current learning, it should be understood that just as the information age has brought about a huge change in our social habits it will bring about a similar change in our learning habits. Despite a plethora of pedagogical models, few are applied to the development of e-learning (Dyke et al 2007). It is important to look at not only what technology is available but how our aims in e-learning affect the pedagogies chosen and in turn the benefits of technology being used. It is important to look at why teaching is happening the way it is and not just if it is possible to use technology to teach.

3.3 Pedagogy in e-learning

There are two common approaches in e-learning which would appear at first to be in opposition to each other. One perspective can be seen as associationist (Mayes and de Frietas 2004), positivists and behaviourist (Oliver et al 2007). An alternative perspective comes from a social viewpoint, which here includes cognitive constructivists, social constructivist and situative (Oliver et al 2007) or cognitive and situative perspective (Mayes and de Frietas 2004).

Pedagogies from the associationist or similar perspective, view knowledge as an object that can be transferred, broken into its constituents and built back up in systematic order by the learner (Mayes and de Frietas 2004). They believe in a positivist view of cause and effect and a behaviourist model of reward and punishment in order to promote learning in students and so success is a result of drill and practice to reinforce concepts through assessment (Oliver et al 2007).

Constructivist aligned pedagogies build on the key concept that learning takes place through our experiences. They place the learner at the centre of a system almost as a scientist testing their concepts in the light of their actions in the world (Dyke et al 2007). Learning and understanding are therefore part of the social construct in which our lives are lived and it is this world that gives our actions meaning (Karagiorgi & Symeou 2005). Vygotsky (1978) uses the example of pointing: at first, it is just trying to reach for something but it soon through the actions of others becomes more than this and so is understood and becomes a tool for learning. This repertoire of tools and symbolisation is built up throughout childhood into more and more abstract processes.

It might be the conclusions of the pupil centred model of learning that students should simply discover for themselves; however, this is could be said to be neglectful on the part of the teacher who should be helping to coach students finding activities in their proximal zones of development (Karagiorgi and Symeou 2005). It is the social aspects of learning that provides motivation and enables progress. Vygotsky's (1978) zones of proximal development are an example of this where the teacher or peer provides the help needed by a student to complete a task that they would otherwise have been unable to complete on their own.

The situative perspective of which Wenger's (Lave & Wenger 2003). communities of practice is an example, is an approach in which it is our interactions with a community that provides value to our actions and so motivates us to learn. These motivations are there even when direct contact with the community might at first not be so obvious, for example, a student writing will do so thinking about what their teacher or peers will think of it (Fowler and Mayes 1999).

In this section, two contrasting philosophies of education have been outlined. This provides a context in which to examine approaches in e-learning and the methods that enable these approaches. However, as the next section will show many institutions now take a mixed approach, viewing the different levels of learning (Mayes and de Freitas 2004) as or institutional influence practice (Alexander 2001). Other post-theoretical approaches claim the difference between the credentials that teacher claim and what they do don't align, combining constructivist ideals of learning through activity with behaviourist reward through assessment (Oliver et al 2007). Good pedagogical research in e-learning must acknowledge its influences and methodology in both the way an environment is built to collect data and the way it is interpreted (Oliver et al 2007). By using a framework for effective practice built around clear pedagogical principles, approaches can be compared and used frame further study (Laurillard 2008).

3.4 Content Centred Learning Environments

The associationist or behaviourist approach is often aligned with ideas in the shareable content object reference model despite the claim that it is pedagogically neutral (Dyke et al 2007). Along with the learning object metadata model (IEEE 2002), they form the standard to which virtual learning environments should be built to enable scalable, interoperable learning content. This, in turn, builds to approaches like instructional design, which is popular with commercial e-learning practices (Oliver et al 2007), where content has in the past been considered king (Waller 2004).

The pedagogical advantage of e-learning from a content centred approach is that compared to traditional learning, which allows only the flow of information from a single teacher to multiple students at a single speed, e-learning can allow students to access this information at their own rate (Carchiolo et al 2002).

Modelling of the interactions in content-driven environments is around 3 principal factors:

- a learner model, what the user already knows, how the user likes to learn and information tracking there use through the system,
- a domain expert: knowledge about the subject matter and how concepts link together
- a pedagogical model: teaching strategies the best order to introduce the topic (Specht and Oppermann 1998).

Alternative models are developed where the role is of the teacher is not entirely replaced by the system (Carchiolo et al 2002).

3.4.1 Student Activity in the Environment

It would be a mistake to think that a content centred approach is not interactive. Use of simulations is just one example of how student activity is supported (Laurillard 2002). Discussion boards are also used to support content and the use of questions to supply feedback to the student and the teacher (Britain and Liber 1999).

It should not be assumed that content-driven learning environments are solely associative or didactic in their approach. A content centred constructivist e-learning environment would allow students to independently explore an information space to obtain content, interaction should be provided by students either being able to manipulate the material or there being multiple paths (Keengwe et al 2014). The aim is to allow students to construct their understanding (Mayes and de Freitas 2002). It is the process of discovery that promotes learning (Laurillard 1993).

3.4.2 Support for Students

There has been a long quest in producing content centred adaptive and therefore supported content centred learning environments (Dyke et al, 2007). There are numerous examples of trying to support learners by providing alternative content and feedback. This can happen through adaptive sequencing (Hockemeyer and Dietrich 1999; Atif et al 2003), navigation of content (Sampson et al 2002) or a combination of these approaches (Specht and Oppermann 1998; Kinshuck and Patel 2001).

Personalisation, in this case, is the adaption of content to fit with a perceived style of learning of the student (Beetham 2013). The idea of learning styles gained traction in the early to mid-2000s. The aim of many was to provide an individualised/personalised approach to learning for students (Coffield et al 2004). It would be easy to associate cognitive principals with learning styles but the approach taken is much more associationist in its application in e-learning environments (Mayes and de Freitas 2002). Amongst many approaches, Howard Gardener's theory of multiple intelligences (Gardener 1984) is one of the more physiological based ideas, as it grew from research into extreme cases that showed the inner workings of the human mind and its capacity to learn skills independently. Other models relied on the use of thematically separating approaches of individuals. Most theories categorise learners based on the skills they have or may prefer to use (Beetham 2013). However, the system needs a learner model in order to do this and which required information to be collected on the user by both initial testing and profiling of users and the results of users whilst using the system (Specht & Oppermann 1998). There are criticisms of learning styles in that it is hard to pin down empirical evidence for the boundaries of style, as Gardener (1984) points out everyone has the ability to enhance these intelligences and so many attributes are dynamic. It is not clear as to the best approach to take with learning styles. Should an individual be matched with their preferred style to aid learning or mismatched in order to challenge or keep learning fresh (Coffield et al 2004; Beetham, 2013)? For these reasons it is recommended that they are approached with an air of caution as they may not provide the results they are intended to have (Mayes and de Freitas 2002).

3.5 User-Generated Content Centred Learning Environments

The alternative view builds from the critique that all too often e-learning means simply putting lecture notes online (Cook et al 2007). Most constructionist teaching approaches advocate the idea of problem-based learning, discursive active activities that provide experiences with active feedback (Keengwe et al 2014). Here two different approaches will be looked at: the first follows the influence of Papert, who placed concepts as tools to be used in the construction of knowledge. The second is the social constructivist view, which sees knowledge being built through the community and it is through discursive actions that knowledge is built and tested.

3.5.1 Learning Through Tools

Constructivists believe in activity driven work, in the example of Papert and LOGO an interactive programming environment was designed to bridge the gap between experience and learning. Papert encouraged students to draw using computers and programming, the link between the environment and the mathematical tools meant that students independently developed (Crawford 1996; Dyke et al 2007). In the learning environment, learning became an integral part of the experience-driven process of using the tool. Further interactions with robots changed the social context of the work and this change promoted more collaborative work (Crawford 1996). Laurillard (2002) holds this up as an example of a microworld in which user action facilitates learning. The idea is learning by making (Cook et al 2007), so students in the online learning environment are not simply consuming content they are productive in the learning environment. The idea is that it is possible to learn through our interactions with the environment and the feedback it gives, as one might when one plays. It is no coincidence then that this work went on to be franchised by Lego in their Mindstorms robotic series, which endeavours to do the same thing (Crawford 1996). Although there is some scepticism about the approach as teachers said that students often needed support through guided discussion and worksheets so it was not pure discovery learning (Dyke et al 2007).

Laurillard (2002) points out that a simulation or a game is not the same approach, although many confuse the two. The difference is that the primary aim of simulations and games is not for the user to create. There is interaction with concepts and feedback but learners cannot produce new concepts only act within the confines of the simulation or game.

3.5.2 Learning Through Communities and Social Interactions

There are many pedagogical bases that result in the same approach of collaboration in learning environments these include and social constructivists theories such as Vygotsky or Piaget and situativists such as Wenger (Jones et al 2007). The aim is to learn collaboratively by sharing and viewing the ideas with others.

Wenger's model of communities of practice supplies us with a parallel approach within e-learning. Wenger's stages of imagination, engagement and alignment, can be seen in an online discussion through conceptualization, as the user comes into contact with others' ideas, construction as they test their knowledge through tasks and dialogue that results in the creation of new concepts. There is the added benefit that as an individual develops in a community it is the community that motivates them as they value the views of others in the community (Fowler and Mayes 1999).

However, there is little agreement in how this model transfers into e-learning. The challenge is to promote social interaction that in turn leads to a community of practice. In many e-learning situations do individuals engage in learning for long enough to form these communities or have enough social interaction with others involved in learning (Dyke et al 2007)?

Another model for social learning is Vygotsky's (1978) zone of proximal development and social development in the meaning of tools and symbols has many applications on the ideas of e-learning as a tool that is given its affordances through our social use (Jones et al 2007)

One model that might help expand on the process is connectivism. Connectivists view learning through the strength of the connections that are made between people, information and digital learning artefacts. cMOOCs are an example of a connectivist environment built on the idea that students should find and build their own notes and discuss ideas with others. Teacher place objectives for students to research and students then build and share their ideas with others in the environment. The connections, interaction and engagement (CIE) framework, is a model of how learning takes place in a cMOOC: it places four levels of interaction at the centre of all the learning processes, the lower of which are simple procedural interactions, and the higher of which facilitate deep learning. There are far fewer of the deeper level creative interactions that act as a catalyst causing a cascade of lower level learning engagement by users. In this relationship, the lower levels help build engagement towards the higher levels of interaction. The levels of interaction are built on the ideas of constructivist, Bloom's revised taxonomy and Laurillard's conversational framework, these principles are then further empirically deductively categorised into the CIE framework. Operational interactions both testing of the tools work and helping others to do this. Wayfinding, actions to signpost information with actions such as liking or linking and forming groups, Sensemaking aggregation/sharing, discussion, reflection and the top level of Artefact creation and remixing (Wang 2017)

What is interesting is the amount of effort that is necessary to engage in higher level activities and it is important that our interactions with learning environments don’t become eddied in the lower levels. This might go some way to explain the high level of fall off in cMOOCs (Aparicio & Bacao 2013). Wang (2017) talked about a large drop after week 6 of a 36-week course and data from the first 6 weeks also show a drop-off, all be it with some spikes in activity in specific weeks, he comments about how procedural problems can have a large knock-on effect on the groups learning.

3.5.3 Reflective Content Systems

A key part of many pedagogies is the idea of reflection (Dyke et al 2007). Studies into the use of Pebblepad, a web portfolio tool that has taken a social constructivist approach, advocate the importance of feedback to students and mechanisms for this from both teachers and peers. Social constructivism is a process by which members of a community work collaboratively sharing ideas and information which are stored for the future community via online means. Three key areas of reflection are highlighted: on your own performance, your performance to set criteria and your performance against others. The ability to work online to share work and comments enables students to engage these 3 levels of reflection (Welsh 2012).

3.5.4 Support for Students

Support for learning is an integral part of constructivist models as outlined above, this support comes either from a more experienced individual like a teacher, tutor or mentor as in Vygotsky’s (1978) zones of proximal development or through peers as in a community of practice. Social constructionists point out the importance of being able to reflect on work by working with others, sharing work with others to enable us to reflect on the work of others and compare our own (Welsh 2012). The tools in microworlds themselves are also a means of support as they can provide feedback to learners (Laurillard 2002).

VSM model by Britain and Liber (2004) outlines key interactions between the group and the teacher. These key actions support the learning process through the negotiation of a learning contract with the teacher where aims and objectives are set, co-ordination, monitoring, individualisation, self-organisation, the adaptation of learning material. If students are to engage in the support then they must be able to organise their learning as part of the group with other group members (Britain and Liber 2004). As such communication and organisational tools become very important in providing this support.

3.6 Motivation of Students

Motivation from a behaviourist point of view uses reward and punishment of tasks (Dyke et al 2007). So the automation of feedback can help to support this. In a purely content driven system improved motivation comes from improved content, better production values on videos, images and look as well as the use of virtual environments and simulations. Simulations or gamification can further enhance the teaching experience.

Maslow's needs hierarchy of Physiological, Safety, Love/belonging, esteem and Self-actualization and Decharms achievement provide a structure to explain student behaviour and motivations in education (Ray 1992) which is contrary to the behaviourist approach. In e-learning systems environments that provide opportunities to fulfil the different levels will provide motivation to students (Mayes and de Freitas 2002).

A key aspect of learning through a user-created content approach is that it is not isolated it is further enhanced by social aspects in both the discovery of examples and the sharing of results, this is just as true in learning with tools as it is in purely social approaches. When Papert's work was extended into robots and a social environment it changed how the students learned (Crawford 1996). Some of the most successful e-learning communities are built around open source communities, these often focus on aspects of learning and sharing around the use of tools and can be seen as social communities of practice (Dyke et al 2007). A key element of the constructivist approach of the use of authentic tasks is also a key form of motivation for students (Mayes and de Freitas 2002).

It should be noted that although many environments and teachers claim a constructivist approach when it comes to motivation they fall back on more behaviourist approaches (Oliver et al 2007). This is seen in the assessment of discussion boards and other social learning activities in order to promote use, which in turn can lead to examples where students set up fake discussions to maximise marks (Oliver et al 2007).

3.7 Adoption of an E-learning Framework

Beetham and Sharpe (2013) outline the idea of design for learning. It is principally concerned with the application of relevant pedagogies for the target audience through a process of iterative development and evaluation. It is the development of a framework for teaching but it is not in so doing meant to be inflexible and should always support a dialogue of change with the students. The aim of this learning environment is to develop a structure that can support such a process by providing a range of tools and a pedagogical framework around which to test new concepts (Laurillard 2008). Any framework developed must be flexible enough for teachers using the system to engage in meaningful change in the development of their own teaching in line with the principals of scholarly teaching as outlined by Biggs and Tang (2013). This will be discussed in more detail in chapter 5.

Two approaches to e-learning environments were examined above with their supporting pedagogies: content centred objectivist/associative/behaviourist and user-generated content centred environments constructivist/situative. These can be viewed as different aspects of the same learning process (Mayes and de Freitas 2004) and this is the approach taken here. Learning design (Lockyer et al 2012; Oliver et al 2013; Agostinho et al 2013), looks for shared practice that can be applied to e-learning through the sharing of successful approaches to learning and teaching the most common approach representing a learning task is as follows:

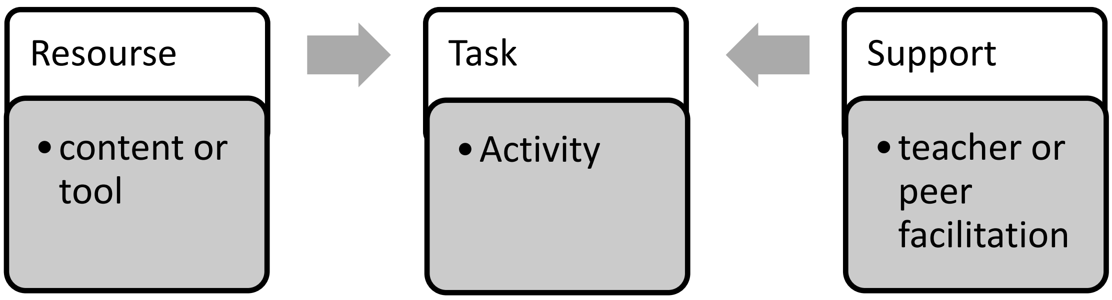

Figure 3.1 Common learning design categories with example adapted from Lockyer et al (2012)

The two approaches are further broken down into three stages these will now be examined to show how they would support learning activity.

1. Content centred
2. User-generated content - tool base
3. User-generated content - social based

Although it is possible to use any one system on its own, the argument here is that by using all three together it enables a more complete learning cycle.

Blooms revised taxonomy (Krathwohl 2002) gives a structure for different levels of activity for students: remember, understand, apply, analyse, evaluate and create. For a user-generated content environment to support the lower levels, they must first read and understand the comments of others. A content centred environment tries to ensure the quality of initial teaching materials in a way so that it deals with misconceptions and they do not persist (Blass and David 2003). Although there may be many criticisms of content centred systems that promote feedback through questioning and drill and practice, this type of learning maps onto the lower levels of Bloom's taxonomy and this can be of particular use in Science teaching (Dyke et al 2007). However, in content centred environments the support and activity around the higher levels of creation is absent and must be provided by supplementary systems or work by the student. By bringing these models together support for multiple levels of learning is possible, and a mastery model can be applied.

The final level looks to promote student learning by providing a social element and so unlocks the motivational benefits of a community of practice to which the student can feel they are adding to (Fowler and Mayes 1999).

Laurillard's (2008) conversational framework shows a model of how this might work with both different levels of resource and support from either peers or a teacher.

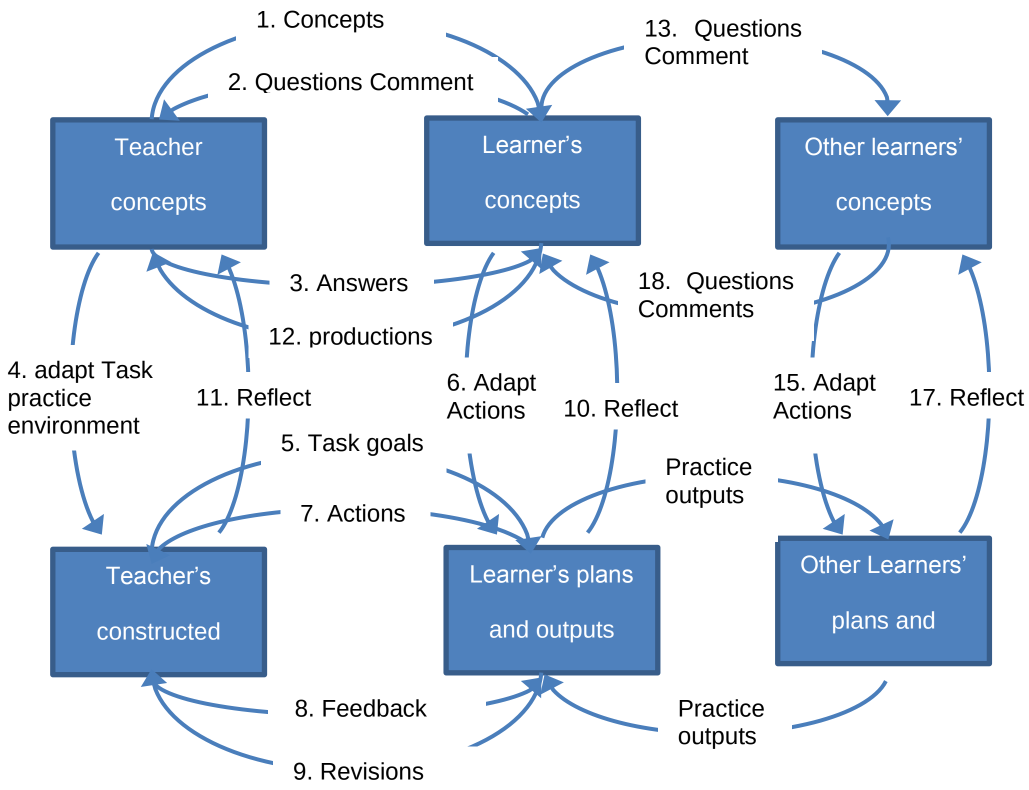

Figure 3.2 Laurillard's conversational framework (2008)

Laurillard's conversational framework has been used as a structure to evaluate existing virtual learning environments (Britain and Liber 1999). By assuming that all learning takes place through the online learning environment then it can be evaluated against the conversational framework by asking how the environment supports each part of the processes. So one might ask how does the environment present concepts to the student, how are the teachers able to adapt constructed environment tasks for the student or how is the student able to reflect on concepts and present them back to the teacher? This might place a heavy emphasis on communication tools within the learning environment that allows this conversation to happen directly. However, this is not the only possibility for the conversation, simulations allow this conversation to happen. In this case the teacher is replaced by the computers feedback, however, what is important is that the teacher is still able to assess the progress of the student as they would be able to do in a direct conversation if they are to be able to continue. Black et al (2007) outline how assessment in both formal and informal forms can be used to illuminate understanding between student and teacher in order to promote adaptation of teaching. Methods by which this can be undertaken will be examined in chapter 4.

Criticism of the conversational framework argues that it focuses on “articulated” knowledge, which might be conceived as too narrow a view of the aim of education (Goodyear 2001). However, this view is not held by all (Mayes and de Freitas 2002). The second phase of the learning cycle can be experiential in nature if the activity is set up in that way. Where this is not possible in online tools Laurillard (2008) talks about blending e-learning with a traditional approach so that it maximises the resources available, so for example in a Chemistry course the constructed environment might revolve around a real laboratory experiment.

Earlier versions were also criticized for not connecting beyond the teacher-student relationship to the role of peers or the group (Britain and Liber 1999) and although the model is built so that it could be viewed as the fully supported one to one relationship, Laurillard (2008) added a peer element and by drawing parallels to the social structures found in the HE teaching: the lecture moving into the tutorial or laboratory experience, showed how the model represents structures and social practices found in traditional teaching. Govindasamy (2002) takes the opinion that much of the teacher's role within the framework can be automated although this requires prediction of possible pitfalls of students in order to introduce features for performance support. Systems also need to be built to monitor rates of access in order to monitor student and data collected can be used both to target support and provide motivation to students.

An important point of Laurillard's learning cycle is that is splits learning and teaching into two phases the first conceptual and the second experiential. This reflects the interplay of the two approaches discussed so far in this chapter and so the conversational framework provides a basis for our learning environment.

It is partially noteworthy here as a theory as it tries to deal directly with the use of technology-enhanced learning tools and as such is one of the few pedagogies developed with this aim in mind. Laurillard's Rethinking University Teaching (2002) elucidates the importance of contemplating the pedagogy of teaching in a new learning environment rather than focusing on the logistics of teaching to mass audiences. Laurillard's framework is used here as a reference in order to evaluate the three-stage framework suggested below and parallels can be drawn.

3.8 Implementation of Ideal Learning Environment

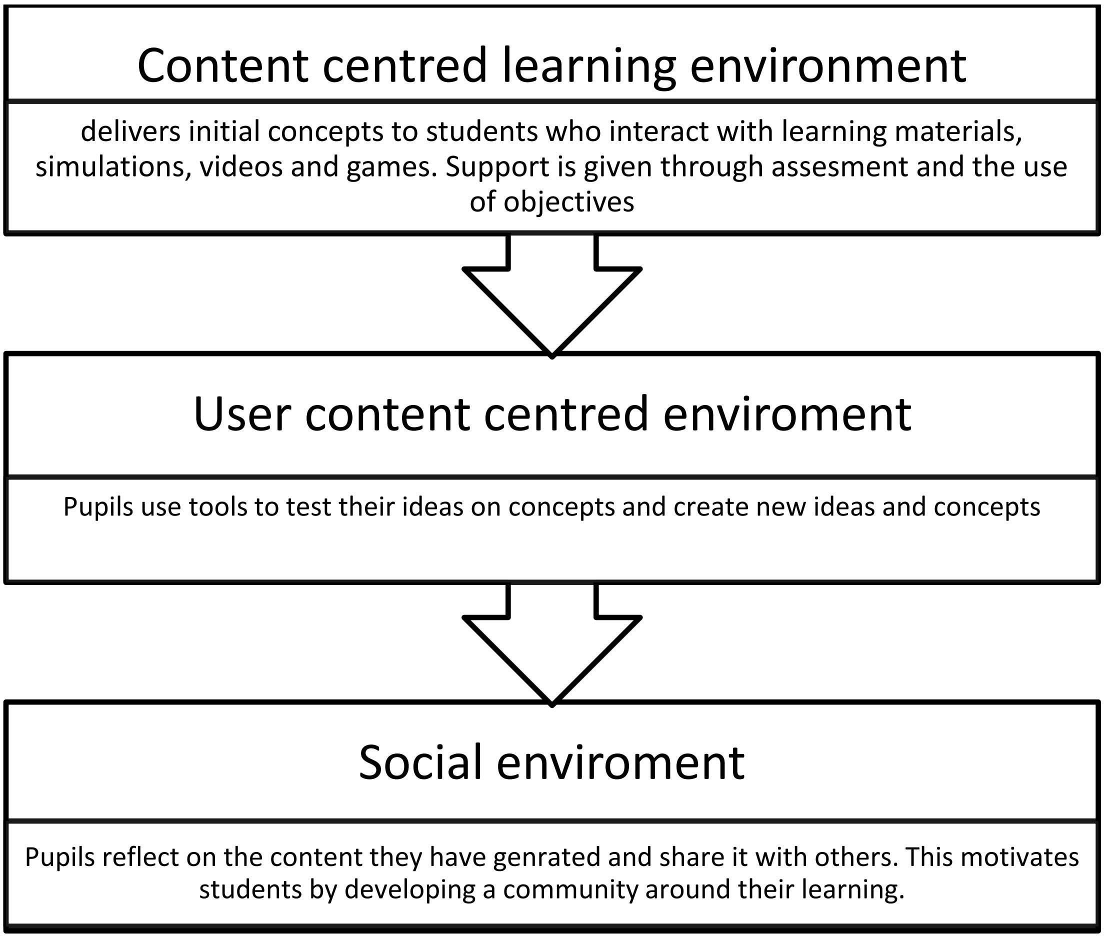

Figure 3.3 Framework for e-learning environment

Although the environment is shown in a linear fashion (Figure 3.3), this is not how it would be used after the first iteration of learning. Once students had moved through the system to produce their first shared output, they would start to look at the work of others, reflect on this, revisit initial concepts and generate further content of their own to share with others.

Each stage will be looked at in more details for the pedagogical process involved.

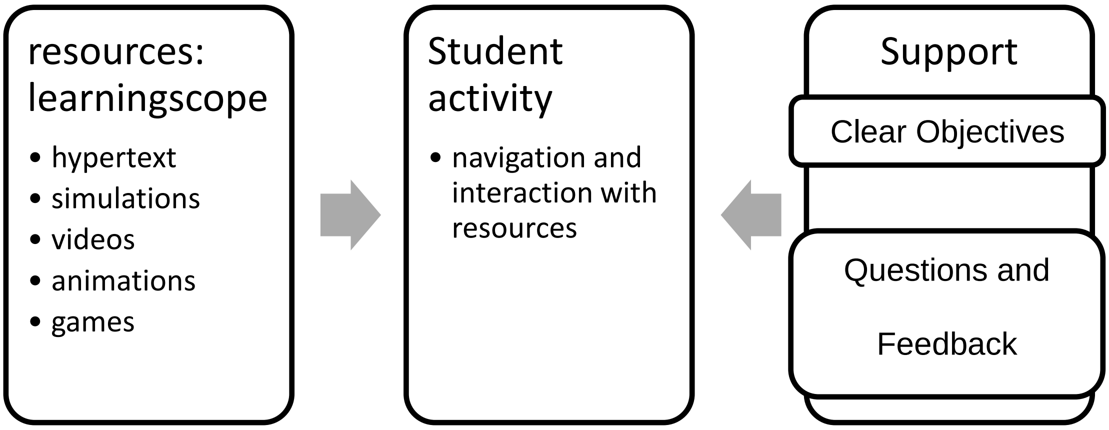

Figure 3.4 Model of the First level e-learning Environment: Content Centred Environment. Learningscope

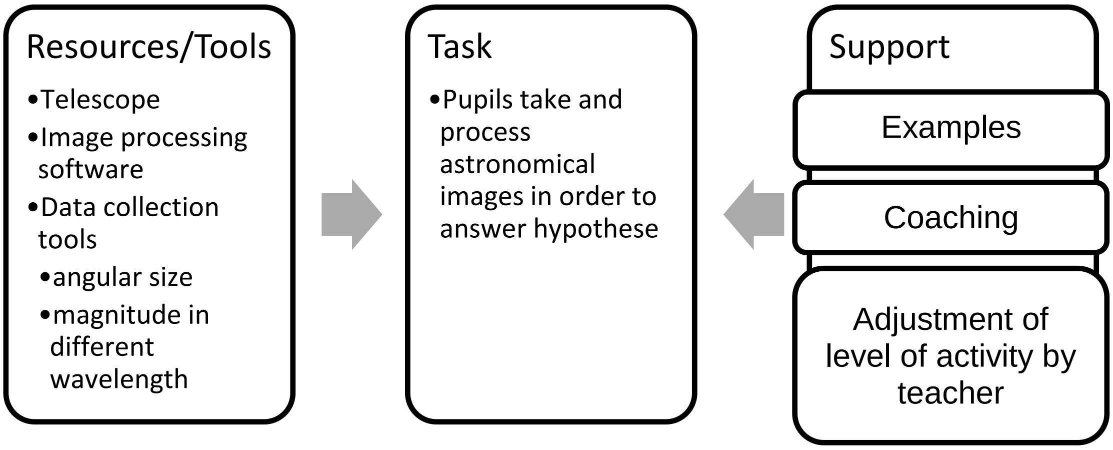

Figure 3.5 Second level of e-learning environment: tools level

In this first level (Figure 3.4) students are free to explore content and interact with resources. Concepts are presented to users through a range of resources. Support is given both directly through the use of simulations and from the use of questions and teacher feedback. Objectives are negotiated with students to give clear aims to learning (Britain and Liber 2004) and in doing so this helps to focus the learning of pupils in the environment (Laurillard 2000).

The aim of the second level of the e-learning environment (Figure 3.5) is to engage students in learning by making (Cook et al 2007). The activities are supported through adjustment of the activity by the teacher as outlined in the conversational framework (Laurillard 2002) and the provision of examples and coaching in line with Vygotskian zones of proximal development (1978).

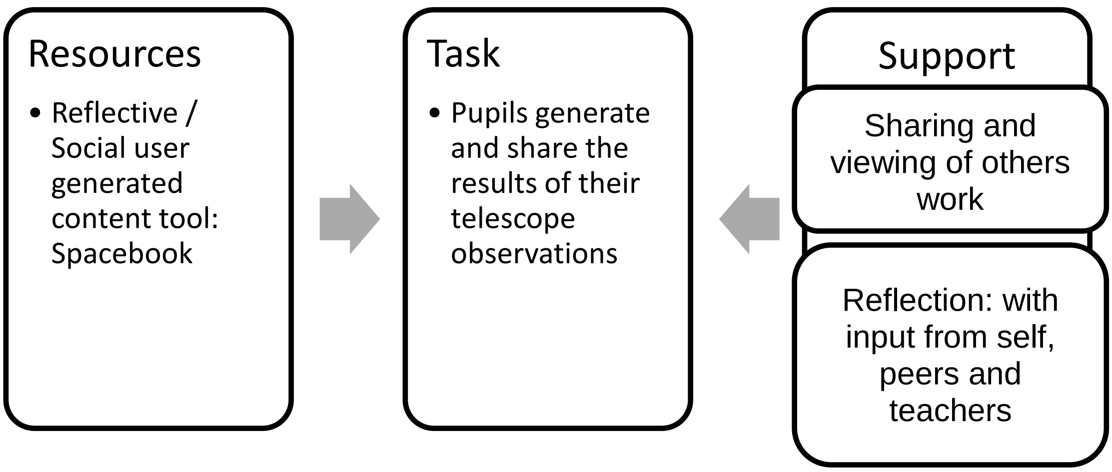

Figure 3.6 Social level of the learning environment: Spacebook

The social level of the learning environment (Figure 3.6) provides a creative output to the learning taking place and in doing so it provides motivational benefits as learners are able to share within a community. This might be a long-standing community such as their class or a new community of the wider telescope user base. This level also provides an important reflective phase to the learning as seen in an e-portfolio (Welsh 2012), with the learner able to reflect on their own and others work and on feedback from their teacher. It also supports the peer-based reflection shown in the conversational framework.

Although this model has been put together using the telescope as an example it is by no means unique to its application, the framework could be used to promote any learning activity. For example, the tools stage could be an online programming environment with user making and sharing games or a comic book drawing app.

3.9 Chapter Summary and Conclusions

A three-stage framework for an e-learning environment has been outlined (Figure 3.3) that provides a mastery model of learning through two levels first discursive (Figure 3.4) and then experiential (Figure 3.5). These levels enable it to fulfil Laurillard's (2008) conversational framework as elucidated as a framework for evaluation by Britain and Liber (1999). A third stage (Figure 3.6) is added to provide both a reflective (Welsh 2012) and social aspect to the learning. The aim of the social aspect is to provide motivation to learners through a sense of belonging. This stage also supports the conversation between students and teachers and students and their peers in the experimental level of learning.

This model highlights the need for lines by which teacher and student communicate and this will be examined in more detail in the following chapter, which will look at assessment in its broadest sense as a means of collecting and acting on data from students (Black et al 2007).

Chapter 4. Alignment of Pedagogy and Assessment

4.1 Chapter Aims and Objectives

In the previous chapter, a structure for teaching and learning was laid out for an e-learning environment. An outline was given of the tools that this system required in order to fulfil these requirements. The previous chapter focused on teaching and learning rather than the larger structure of designing a curriculum. This is because the same objectives and assessments could be achieved via a variety of approaches both by student and teacher. There are a number of commonly accepted features within the good design of curricula, students need clear expectations and engaging activities, feedback and time for consolidation, these should be aligned with learning outcomes (Beetham and Sharpe 2007).

The aim of this chapter is to define the methods around which activity would function, how the students will know what they are to do and how their teacher will know that they have succeeded in these tasks. These include the setting of objectives for learning and the assessment within the e-learning environment and by extension the data that is available for both students and teachers to measure their progress against. This chapter will look at the role that assessment plays in providing feedback to both student and teachers in order to adapt and meet their aims and objectives (Black and Wiliam 1998).

The separation of pedagogy and assessment as it has broadly been termed here is that learning and teaching are possible without this structure. It is not uncommon for adults to learn about subjects without an aim to their learning beyond their simple enjoyment of it and although informal learning can employ the strategies outlined in this chapter it does not need to. However, this is not a luxury afforded to teaching in the UK system at present and so this chapter will outline systems to help guide learning and measure progress.

Assessment does not stand on its own and so this chapter will also be looking at how it integrates with teaching and learning through learning objectives in what Biggs and Tang (2011) call an aligned curriculum.

4.2 The Aim of Assessment

Black and William's seminal paper "Inside the black box" (1998) powered a revolution in assessment and how it is used in teaching. "Assessment for learning" has been widely adopted in schools across England (The Department for Children, Schools and Families 2008), its adoption in HE has been quite different, following a more constructivist approach in line with Vygotsky and Wenger's communities of practice (Sambell et al 2013).

Key to the original idea is that it is not possible to improve education if it is not known what is happening and the best way to do this is through assessment. It is through changing assessment that the biggest impact on teaching and learning can be made (Black & Wiliam 1998).

There are three purposes for assessment in schools in the UK: assessment for league tables for schools in order to provide accountability for providers of education, assessment involved in the certification and grading of individuals using high stakes activity, which inform progression of pupils into higher levels of education or employment. These must be robust enough to draw comparisons between students across different schools. Finally, there is assessment for learning. Assessment for learning is formative assessment designed to promote the learning of pupils and extends to be informal in nature. Assessment for learning can inform learning and provide feedback for both students and teachers. This is not restricted to the activity of the teacher and can be self-assessment and reflection from the pupil (Black et al 2007).

The later of the three views are taken here as assessment and learning are very much intertwined, with an understanding of where a pupil is at and what they might be capable of with help, it is possible to design an intervention to help the individual. This is the principal behind Vygotsky's (1978) zones of proximal development and is further illustrated in the adaptive task setting presented in Laurillard's conversational framework (2008). The key here is adaptation, it is not simply a case of trying to pre-assess a student and then use this one off assessment to place a student on a set path, as this has been shown to be unsuccessful (Black et al 2007) instead the aim is to alter our teaching in line with continual feedback from the student. An aim of assessment for learning can be to draw out the misconceptions of students, and this can be successful when students are given options to discuss and agree or disagree with, this is part of the process of making students feel comfortable to get things wrong as without this they will not engage with the process (Hodgson and Pyle 2010).

The approach taken in HE to assessment for learning has been very different from that in schools. The development of assessment techniques to promote learning has focused on six key factors: authentic assessment, balancing assessment, practice and rehearsal, formal feedback, Informal feedback and students as self-assessors (Sambell et al 2013). This is perhaps more in line with a wider constructivist view that tasks should be authentic tasks and that students are responsible for their own learning. However the principals at its core are the same: get assessment right it enables teaching and learning, get it wrong and it demotivates learners or worse still promotes surface learning rather than a deeper understanding of the subject.

Biggs (2003) points out that it is important for us to design our curricula around objectives and to align our activities and assessments through a cycle of alignment. In a well-designed curriculum, one should be indistinguishable from the other.

To summarise: assessment is the tool that provides data to aid the loop of feedback between the student and the teacher (Black et al 2007). Using this data the teacher can understand how the student is progressing and give them feedback on their progress, but the loop is complete when the teacher also uses this data to adjust their future teaching, so it is both parties that should adjust as a result of available data (Black et al 2007; Laurillard 2003; Biggs and Tang 2011). This is not a one-off, it must be continual if our aim is to understand the best possible route to help the student learn (Black et al 2007; Vygotsky 1978). This assessment does not need to be formal (Black et al 2007) and can be reflective (Sambell et al 2013), led by peers (Black et al 2007; Sambell et al 2013) and aided by computers.

4.3 Alignment of Assessment

It is the dichotomy of the logistics of running formal assessment and the requirements of learning that leads many involved in the educational field to shy away from the idea of assessment. The everyday realities of having to mark and process formal assessment are somewhat at odds from what many teachers might want to implement with the individual (Stawser 2009). In the UK education system teacher may find themselves under the pressure of formal external exams and the desire to work to test (Black et al 2007).

Up until this point this thesis has been talking about assessment as a standalone concept, this thesis has mentioned that it is intertwined with learning but it has not talked about how these two actions are linked together. It is through the setting of objectives for pupils that it is possible to define the outcome of learning and it is thus through outcomes that the method and aim of assessment are set. Regardless of whether this assessment is to be summative or simply a formative tool to help the individual reflect on their current progress, the assessment must be developed so that it meets the needs of our intended objectives (Biggs & Tang 2011).

Further to this, it is important to be clear in our understanding of knowledge and the levels of understanding that students need to gain if clear objectives are to be written and matched with the actions that students will undertake (Biggs & Tang 2011).

Bloom's revised taxonomy describes the different levels of assessment or learning, from simple low psychological load tasks like remembering to higher level order task involving creative actions (Krathwohl 2002). Other taxonomies like SOLO also explain how students understanding progresses from disconnected domain-specific concepts to connected and abstracted understanding that can be applied in new contexts. SOLO taxonomy describes the levels of understanding that scaffolds a user to form *declarative* to *functional* understanding. Through a stage of linking concepts and moving through threshold concepts to build a complete and connected network of ideas which become *functional* understanding with a level of abstraction of the knowledge which allows it to be applied in new situations. Objectives must be set aligned to assessment so that they encourage students to adopt a deeper approach to obtain *functional* understanding. It is not enough to ask students to describe students must hypothesise and reflect using the concepts of the domain that they need to learn about (Biggs & Tang 2011).

However, the link between types of assessment and the promotion of deep learning within students is not as clear cut as one might first think. This could be put down to not being able to develop suitable learning. However, others think that deep learning strategies, although universally seen as better, are not per se more effective in the short term (Gijbels et al 2005).

As Biggs and Tang (2011) point out it is essential that assessment and learning objectives are aligned. Alignment is achieved if the verb that is used in the teaching objective is also used in the learning task and the assessment activity. Multiple learning objectives can be used to define the different levels of understanding that you wish a student to move through and with the help of rubrics to students can make judgements on standard grading criteria. It should also be noted that the approach taken by students to their learning is not fixed. Students will change their approach between deep and surface learning based on the course style and assessment (Yonker 2011).

Given that the SOLO taxonomy model is a mastery model (Biggs & tang 2011) it can be combined with the ideas of Vygotsky (1978) proximal zones of learning and concepts in assessment for learning (Black et al 2007) to scaffold the learning process for the individual. In order to do this, different objectives need to set for pupils at different points in their learning.

Using the Laurillard's conversational models (2003) which has two levels of interaction between the student and the teacher: the first is on the understanding of initial concepts and the second is on the application of these concepts in a higher order task, this is comparable to SOLOs stages of declarative and functional (Biggs & Tang 2011). Two very different assessment techniques can be employed to aid these two processes. Multiple choice questions are suitable for the demonstration of understanding of concepts and for the constructed task of the experimental phase should employ longer answer project based questions that help develop functional understanding.

4.4 The Use of Multiple Choice Questions

Some point out that the use of MCQ is an inevitability of working with larger numbers (Yonker 2011). It can be difficult to develop MCQ in line with the higher order thinking in line with Bloom's revised taxonomy. Some of the difficulty might be in the fact that it is hard for domain experts to estimate the difficulty of questions Tractenberg et al (2013) suggest a matrix for developing questions for different levels of difficulty. Test banks have been developed to help measure both shallow outcomes knowledge or factual (shallow cognitive) or applied (deeper cognitive processing) MCQ can be used to measure understanding and application of knowledge and others can assess the ability to analyse situations and solve problems (higher level thinking). It should be possible to distinguish between surface and deep learners in order to reward the latter. Assignment essays rewarded deep learners, formal exams did not. Studies have shown that there was not a correlation between deep learning and results of MCQ designed either to test factual or applied knowledge although there was a detrimental correlation between the adoption of a surface strategy approach and test scores. However, they concluded that this could be through the inability to develop complex enough questions to elicit a positive correlation between deeper learning and results. This supports the idea that surface approach inhibits results rather than a deeper approach to learning improving them (Yonker 2011).

4.5 Measurement of Confidence

One criticism of multiple choice questions is that students can, unlike long answer questions, “lucky guess” answers, to combat this a method to measure the students' confidence in their answers will be used. Confidence based testing (CBT) as it is commonly known uses a variety of techniques to measure students’ confidence; however, most want some sort of scale of confidence as the outcome. It is then possible to categorise answers into four different areas that allow for further analysis

Table 4.1 Confidence categories from Kampmeyer et al (2015)

Wrong		Right	
High confidence	Low confidence	Low confidence	High confidence
Misinformed	Uninformed	Partially informed	Well informed

The difference here is that the uniformed points to an area of the curriculum that would need covering in more detail in the future whereas misinformed required remedial actions and possible changes to the way concepts have been covered.

In an example of confidence based testing given by Steward et al (2013) students voted on their confidence in a class voting system: students were shown a question, then students discussed the question before the results were shown to the group, they are asked to rate their confidence. Either using set statements or saying either high or low confidence. This study showed a strong relationship between those that are confident and get the question right; however, the other 3 groups are of about equal size. The system was used to identify areas in which the pupils might need help to improve. The responses used to identify possible areas of improvement were correct and low confidence and incorrect with either level of confidence. So in a question relating to a given subject, they looked at how many got it right and then if there was a higher than usual amount of any of these categories, as it indicated an area for intervention (Stewart et al 2013).

Confidence testing has been successfully used in studies in UK secondary schools where pupils were asked to mark their confidence on a 1-10 scale with marks given to pupils for correct answers and marks lost for incorrect answers. A 1-10 scale was found to be easily understandable by a pupil. A strong relationship was found between pupils' facility and confidence in Maths. Pupils also commented that they understood and enjoyed the process and that it encouraged them to be more reflective in their answer and encouraged them to take a deeper approach than they would normally. The study also found that there was no significant difference in gender to the tools ability to measure confidence (Foster 2015).

Confidence based testing can be used to add an extra dimension to multiple choice questions, giving information on the understanding of pupils beyond their ability to get the answer right. If combined with questions that allow students to reflect on misconceptions as elucidated by Hodgson and Pyle (2010), then it should be possible to build a more comprehensive understanding in the same way that face to face interactions contain tacit information about pupils understanding beyond the verbal answer they give. By causing pupils to be reflective it also promotes deeper learning strategies by making pupils consider how they know something not just on the right answer.

4.6 Assessment to Promote Deep Learning Outcomes

The higher levels of Bloom's involve the creative process which by its very essence involves using knowledge to make something new (Krathwohl 2002). In order to promote deep learning in students a rich project or inquiry-based task should be used for assessment. The assessment task should be rich, authentic, they should involve activities that are meaningful to the student and will be seen by more than use the teacher but a wider audience

(Sambell et al 2013) and are aligned with the aims and objectives of a course (Biggs & Tang 2011). Within this authentic approach students should be given multiple chances to attempt and modify in light of feedback (Sambell et al 2013), this might be conceived of as a workload stumbling block (Yonker 2011), but it suggested that this process of reflection can be aided by access to ideal answers in an e-learning environment that can be opened by the students after they have attempted questions (Laurillard et al 2000). Reflection as an essential part of the feedback process can come in a number of forms: self-reflection, peer reflection and teacher lead reflection and online systems can aid the links required for all of these types of reflection (Welsh 2012). Peer feedback is key to the approach taken in assessment for learning (Hodgson and Pyle 2010), and a "communities of practice" model that links learning outputs to social benefits can be integrated into the approach of a learning environment as part of the assessment. Although it is important to be careful in our interpretation and assessment of social connections as they can easily be manipulated if these interactions themselves are being marked (Oliver et al 2007)

4.7 Chapter Summary and Conclusions

For a learning environment to be effective it must integrate effective assessment tools, using taxonomy such as Bloom's (Krathwohl 2002) and SOLO (Biggs and Tang 2011) along with pedagogies such as Laurillard's (2002) conversational framework it is possible to align different types of assessment with different stages of the learning cycle.

Assessment is only useful for the learning cycle if it evokes an understanding of the pupil understanding for the teacher and then the teacher acts on this feedback to change the learning and teaching that is taking place (Black et al 2007). Vygotsky's (1978) theory of a zone of proximal development elucidates that it isn't enough to know what a student knows at the moment, this should be combined with what they are able to achieve in an assisted situation, to develop a strategy for learning and teaching.

Assessment can be provided by a range of activities from simple gestures to more formal examinations Black et al (2007). The pedagogical framework described in chapter 3 can be aligned with two levels of assessment the first on the declarative level and the second for the applied level.

Multiple choice questions can be an effective way of testing both understanding and application, when combined with confidence testing can promote deeper reflective thinking in pupils (Foster 2015) as well as providing an increased level of insight into the learning and understanding of students to teachers. This type of practice and drill assessment aligns well with the lower levels of Bloom's revised taxonomy and the behaviourist principals that provide motivation and reinforcement through reward and punishment (Oliver et al 2007). Although it should be noted here that the approach is aligned with assessment for learning principals that assessment should provide an environment in which students feel safe to get answers wrong and see this as part of the learning process (Hodgson and Pyle 2010)

Longer more complex tasks can form part of the assessment cycle. Activity and assessment must be aligned the objectives of these longer tasks with the learning and teaching activity and the assessment method, only together do they then provide an effective assessment strategy for the applied levels of understanding or functional understanding (Biggs and Tang 2011). Examples include portfolio tools, but the aim is to provide a system that allows us to not only set objectives but a space in which the student can reflect on their own answers. These tasks should be authentic in that they are meaningful both in that they reflect real-world tasks but that they also have a use beyond assessment and so are meaningful to the student in the larger social context (Sambell et al 2013).

This chapter has outlined and aligned the methods of assessment that aids learning in the different stages of the e-learning environment. The next chapter will build on the processes of assessment as the key method of communication between teacher and student to show how it can be further aligned with a process of evaluation.

Chapter 5. Evaluation

5.1 Aims and Objectives

This chapter aims to outline both the purpose of the evaluation process and the methods that it will use to evaluate the learning and teaching taking place. This chapter will include:

- An outline of the need for evaluation in e-learning systems
- Possible approaches to evaluation will be examined.
- The reasons for integrating evaluation will be outlined.
- Finally, this chapter will outline the process by which these tools are aligned with the educational environment and the points at which tools will evaluate the educational process.

However, this chapter will stop short of detailed explanations of data collection and interpretation which are detailed in the methodology chapter 6 describing the data collection tools and 7 the methods by which this data is to be analysed. Although this is a specific example of the Bradford Robotic Telescope the approach is taken in the chapter will also discuss the general applications of these tools.

5.2 The Need for Evaluation Processes in e-learning

It is increasingly important that effective teaching practices are shared and re-used as class size and other factors have applied pressure to the teaching situation (Beetham, & Sharpe 2007). Pressures are not only pedagogical, but there is also a fiscal drive to improve education which has increased the need to provide quality control mechanisms in teaching and learning (Lockyer et al 2013).

Evaluation may make many think of a summative process. Projects are often evaluated in their final stages to measure their success or impact. However, this is not the aim of the feedback system presented here. The aim is to close the loop of feedback between teacher and student by allowing the teacher to change their actions based on the data given by the student. A key part of “assessment for learning” is that it is important to close the feedback loop by acting on information (Black et al 2007). Informed practice should be shared between teachers to improve teaching practice (Biggs and Tang 2011; Laurillard 2008; Beetham and Sharpe 2007; Trigwell et al 2000). Data collected on students through their use of online learning environments has the potential to inform at many different levels that of the student in self-reflection, the teacher in a reflection of the learning of students and so, in turn, it can inform the institution and government policy. This process is not simply hierarchical it has the potential to be collaborative, with students supporting students at the bottom level and teachers supporting teachers in their practice (Greller & Drachsler 2012).

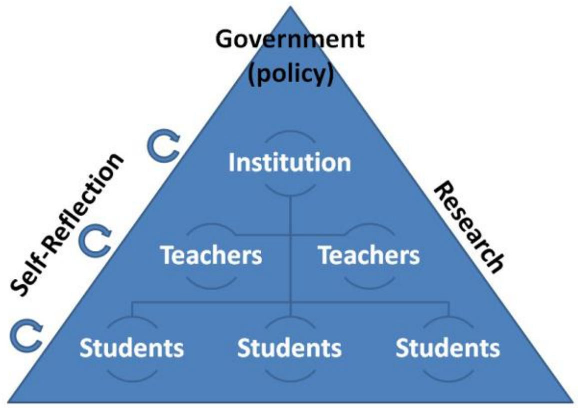

Figure 5.1 Data use and applications from Greller & Drachsler (2012)

Beetham, & Sharpe (2007) outline the need for design for learning, where the concept of design is the iterative process of investigation of principals and relevant theories, application, representation of these ideas to the developer and users and then testing and evaluation in order to decide on improvements. It is suggested that research can help to uncover similar key techniques for subject-specific pedagogies that help to move these areas forward (Beetham & Sharpe 2007; Laurillard, 2008). An example has been the use of phonics in teaching children to read, a now commonly adopted technique in the UK educational system (DfE 2013). It is still an open question if these universal patterns exist and the best that can be hoped for is to design for learning or plan for learning within the contexts there will always be a need for a contingency plan (Beetham and Sharpe 2007).

To do this it must be possible to evaluate through a system that will allow us to test ideas, find and reveal patterns and compare approaches across the multiple contexts in which learning might take place in the e-learning environment.

5.3 Who Should be Doing the Research?

It has been noted by some researchers that evaluation of learning technology is not a robust or cohesive area of research, partly due to the speed of development of technologies (Selwyn 2010).

This is further worsened by the wide range of those that claim the field

"a loose assortment of technologically minded psychologists, pedagogy experts, maths and science educators, computer scientists, systems developers, and the like." (Selwyn 2010: 65)

Beetham & Sharpe (2007) outline a division between those doing teaching and those researching teaching. This is contrary to the discipline-based nature of the field of pedagogy. Why should practice and research take place in isolation? Although there are some within teaching who present models to help narrow this gap: Biggs and Tang (2011) transformative reflection, Schon (1991) reflective practice, Scholarship of teaching (Trigwell et al 2000).

One aim of this thesis is the development and evaluation of an analytical framework that aids the statistical evaluation process that allows judgements on approaches to learning to be quicker and easier to make for teachers. This must be done with caution as some level of interpretation is always required (Greller & Drachsler 2012). Before it is possible to engage in this process of evaluation there must be a framework on which to base our changes and this must facilitate the teacher, a time poor researcher, in engaging in the process (Laurillard 2008). This framework was defined in chapters 3 and 4. Using this framework the process of aligning analytical tools for evaluation can take place.

5.4 How can Teachers Engage in the Research Process?

There are a number of different models by which teachers can engage in a research-informed and informing practice.

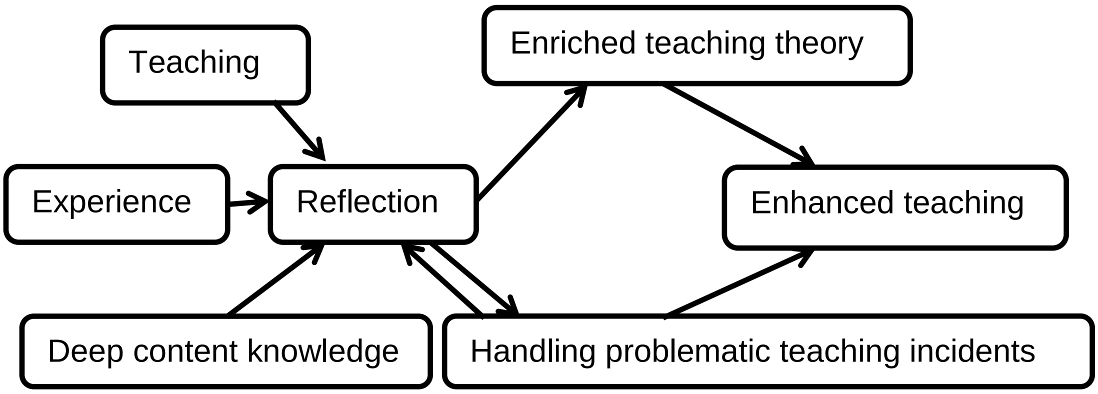

Figure 5.2 Model of transformative reflection presented by Biggs and Tang (2007) p49

Under a model of transformative reflection, a teacher uses their own experiences combined with a deep understanding of the knowledge domain they are teaching informed by teaching theory to make a judgement on the best way to deliver a topic using the resources currently available to them. They then reflect on their teaching using data that tells them how their students are learning to deal with problems within teaching with the hope that this will enhance both their own teaching and the theory of teaching within their subject. This process of reflective practice could also take an action research methodology, principally employed with doing as an action to cause change (Biggs and Tang 2007).

The action research methodology is a practitioner-led format that takes the approach of action, evaluation and then informed further action, this process normally takes a minimum of two cycles (Oliver et al 2007). This differs from the more traditional views of a researcher as an objective independent observer (Scott & Usher 2011).

Teachers can and often are willing to undertake the role of the action researcher reflecting and building on their practice, however, teachers are not well supported in undertaking a more formal approach to learning and teaching. They are time poor and so need a framework on which to build their research. Teachers, who want to innovate, want control over the process they need tools to adapt, reflect, collaborate and share their finding with others (Laurillard 2008). At its core is the idea that it is a process to make informed improvements within the context of the study that might then be applied to those in a similar situation. The criticism of this approach to educational research is that it can never truly be separated from the original context and it is also difficult to attribute improvements to any single factor as there is implicitly no control in the study as the aim is to affect change for as many as possible. It is difficult to separate or prioritise the ethical and the epistemological aspects of the research as could be claimed by more pure experimental methods (Scott & Usher 2011). Slavin et al (2014) point out the importance of achievement measures which are independent of the experimental treatment and as in action research, the aim is to affect change in a particular measure their independence could be brought into question. This is where technology can help too by providing a strong evidence base for the data collected for the studies and supplying a quasi-experimental backing to the practitioner-led framework. By modelling data against normal distribution then it is possible to look at the impact of known factors (Heron

2012) allowing the individual action researcher to take a more objective approach.

5.5 Evaluation Methodology

There are a wide number of influences on the field of educational research from the great philosophical aspects which are often overlooked in general educational research (Scott & Usher 2011) to the influences from technologists and their desired pedagogical approach or the subject in which the technology is to be applied (Alexander 2001). The success of a virtual learning environment for English and grammar (Virvou et al 2000) might be very different from that of science activities as these will all have domain-specific pedagogies as well.

To compound this problem most literature describes what the teacher has done and does not report on the student experience, often if this is reported then there is a significant difference between this and the expected. If the aim is to improve the results of students then data on what the students do and achieve as a result of the changes made needs to be collected (Alexander 2001). A feature of good e-learning research is that it is patent in its approach, this allows for a critique of assumptions. It is impossible to carry out e-learning research without reference to a methodology and a pedagogical approach since any claims will build on these assertions (Oliver et al 2007).

Greller and Drachsler (2012) outline a framework for learning analytics with 6 key factors to consider: stakeholders' objectives, what data is available, instruments, external limitations, internal limitations, although they state that factors are not limited to these 6. The objective of this evaluation process has already been outlined above, which in this case is to allow teachers to reflect on the teaching method. The subject of this thesis is to evaluate different approaches to teaching so the main stakeholders are the students who generate the information with the clients being teachers and designers and schools who would benefit from an improved approach. External limitations: as this is data to be shared among users it must be fully anonymous data that is untraceable to either individuals or institutions in line with the user agreement of the telescope site. This leaves three areas to examine: data, instruments and internal limitations. By extending the Greller and Drachsler (2012) framework there are the following three areas within data and instruments that must be considered before an evaluation strategy can be adopted.

First, a philosophical approach to evaluation must be considered. Evaluation of educational approaches normally falls into two areas informative and comparative. The different types of educational research are induction, deduction, retroduction and abduction (Scott & Usher 2011). A definition of successful teaching and learning is required, i.e. the outcomes that have been defined in order to call the intervention a success (RCUK 2014). The pedagogy and learning design for this framework have been described in chapters 3 and 4 this chapter will place the definitions within the context of evaluation.

The framework needs to ask what is knowledge, which is seen as often ignored the initial cornerstone for our basis for educational research (Scott & Usher 2011). This is more than just simple learning outcomes but overarching principals. With an answer to these questions, the tools for evaluation can be aligned with the teaching and learning strategy (Lockyer et al 2013).

Once all these factors have been considered it is then possible to look at the tools for data collection and analytical approach that will allow us to collect the data required in an online learning environment and interpret it in the way that was decided in the first stage.

Although this is presented as a linear process, it is not: the approach chosen will be affected by our definition of successful outcomes of education and this will, in turn, affect the applicability of the research tools available to use.

In addition this study is interested in evaluating the tools provided to teachers through the system for their validly, teaching and learning is taking place in an online environment with a wide range of users from a variety of backgrounds and in order to draw and apply any conclusions there is a need to be able to understand this range of contexts. The methodology used must be robust enough to work across these contexts and still show meaningful results. This is covered in more detail in chapter 7 rather than here as this chapter is principally aimed at a framework in which to develop tools for teachers to use in transformative reflection through an action research approach.

5.5.1 Selection of a Research Approach

The different types of educational research are induction, deduction, retroduction and abduction (Scott & Usher 2011). At an early stage choices need to be made about the aim of the approach taken: is the research trying to uncover behaviour, does the research wish to find evidence to support an approach or develop our understanding of a particular situation?

Inductive studies look to generalisation, studying a small sample that might then allow more generally applied principals. Deduction draws logical conclusions from generally accepted facts and statements. Studies might, therefore, try to draw a conclusion from comparisons about the best approach for a set situation. Abduction means away and here the aim is to take away the best explanation for a set of data, in many ways, it is similar to retroduction but the direction of inference is different. Studies might use grounded theory to categorise phenomenon to create theoretical frameworks or empirically find evidence of an approach or theory within a dataset.

A big factor on the approach taken will be the competences of the users (Greller and Drachsler 2012) and at this point, it is believed that as simple an approach as possible is required. The reflective practice model described in section 5.4 fits well within a comparative model, so a deductive study with an experimental structure will be used as the frame for our tools. Teachers can test a simple question of the system "is my new method of teaching an improvement on the normal performance of a group?"

Often an approach might be mixed in its methods in order to provide both empirical evidence for the theory and at the same time inform the theory going forward. One example of this is Heron (2012) where students were not only compared in their performance of multiple choice questions at different points to measure the effect of teaching but were also asked why they had given the answer they had given in order to uncover the misconceptions that persisted in spite of teaching. As such the study was not only able to deduce that lectures were having little or no effect on students' performance but to infer what changes needed to be made to teaching to make a real difference.

Although data sets collected in this study could support a mixed method or other approaches they will be left to future work in the area. Once a method for measuring success in e-learning has been chosen, it is then possible to look at the factors that cause or prevent it.

5.5.2 Developing an Experimental Method

The scientific method is often applied to educational research, it places at its core the idea of cause and effect and that these can be observed in an objective way by the researcher. Some would say that this is not the case in an educational research environment any hermeneutics carried out would be done within the context both of the researcher's views and the environment for learning. However, this instead of a problem is a consideration of the research itself (Scott & Usher 2011). This has already been laid out in chapters 2-4 for the environment and the research basis is covered here.

Between-subject methodologies look at two groups: one a control study and the other with the intervention, often with pre and post-study data which can be used to look at the effect the study has in comparison to the control (Churches and McAleavy 2015).

Black et al (2007) in their work to show improvements in assessment for learning in UK schools, used a simple model of finding a control group, that undertook the area of study in question under matched contexts. To do this they looked within the school for either a similar class in that year that had not had the treatment or a group from a different year. However they noted that there were problems with this method in finding matches, and the elimination of factors such as the quality of the teaching particularly in cases with new teachers or effects of accidental change in practice on the control groups once the teacher changed their practice and noticed an improvement (Black et al 2007).

Small between-subject interventions are often criticised as the measured effect of the study could be down to factors outside of the study's control. Were the students simply more talented, was the context a greater effect than the intervention? Studies get around this by trying to increase the study size as this means that they are able to get a better picture of the background level. However, this style can raise issues as a study could have mathematically significant effects that are not useful in the real world, a drug that consistently removed 10% of an infection would not be used as it doesn't have the desired results, the same can be true for the impact of an intervention. It is essential to consider the effect size of a study, not just the significance (Coe 2002).

Within-subject design seemingly offers a solution to this problem of context as all groups are exposed to all the intervention and the control to compare which has a greater effect (Churches and McAleavy 2015). However, although a successful application of this methodology can be made in some teaching situations where a technique is being learnt and practised, it is harder to apply in an area such as science were a topic might be covered only once. Here the intervention goes against the nature of the teaching required and so becomes prohibitive.

In the United States of America there has been a movement towards large randomised experimental design studies (Taylor et al 2013), although this comes with its difficulties not only in the convincing of schools to take part in such studies with the pressures of exam results for their pupils but also in the designing of such studies. Clustering of pupils in the school setting is the first problem, both the school and the class present their own context for the learning taking place it is therefore essential that a large number of clusters are used in order to show that a single intervention has an impact. Secondly, experimental design often requires some level of randomisation as to which clusters undertake the intervention. Randomisation or matching can add further pressures to the recruitment of schools and if schools are not active in the intervention they might drop out. Ideally, these studies should employ between 25-50 clusters (Taylor et al 2013).

Slavin et al in their review of experimental studies into elementary education discusses the lack of studies that meet their requirements, where the criteria included that the intervention was of at least 4 weeks, had pre and post measurements tools that were not focused too closely on the changes that the study wished to make. Less than 10% of over 300 studies initially identified as experimental in their approach met their requirement. Although these are strict guides this low level of studies shows the difficulty involved in trying to develop interventions with such attributes Clearly, there are substantial boundaries to conducting such studies.

Measurement tools present many problems. Use of national tests would at first seem ideal but they are out of the control of the researcher and so subject to change. Too narrow an instrument might also result in unwanted outcomes within learning as the intervention might be prone to pushing or developing aspects of learning that are otherwise not needed (Taylor et al 2013). Heron outlines a method of evaluation which uses clustering to examine between groups comparing their effectiveness at answering a quiz as part of their studies. She is able to show how teaching had little effect on the assessment and views of pupils. The study method relies on a large number of clusters undertaking the test and the modelling of clusters against a binomial distribution, further testing then allows comparison against factors that could affect the outcome of learning. However, this study took a large number of years to put together it is the hope that by combining this method with online learning environments it is possible to speed up the process and by increasing the number of groups undertaking the activity beyond what a single institution could enable.

An objective of this framework is to aid a practitioner-led cycle of reflection through evaluation as outlined in section 5.3. As Greller and Drachsler (2012) advocate, the limitations of the end user must be considered when data is presented for interpretation, primary school teachers in England, who are the target user of the system, are not required to have a qualification beyond a grade C in GCSE mathematics (Department for Education n.d.). As a result of this, the interpretation of data should be through as simple a form as possible, using statistics that are easily understood. With this in mind, it is important that within this methodology that the effect and model remains a normal distribution against analysis in order for further comparisons to be made. This is one critique of e-learning research in that it attributes normal distributions to data that does not necessarily follow this distribution and so it is important there is evidence this is the case before moving on to more complex comparisons (Mitchell 2000). Although other distributions for parametric data could be used or non-parametric analysis they could require knowledge of distributions that is beyond the end user. This is discussed in more detail in section 7.4.

5.6 Alignment of Learning Analytics with Pedagogy and Learning Design

Lockyer et al (2013) point out that there is a need to understand the pedagogical aims and learning design of our system before looking at the data collected. Data from a discussion group would be very different in a peer discussion and a tutor-led situation, and without this knowledge of approach, one might comment that there was a failure in the tutor-led session through the dominance of a single person (Lockyer et al 2013). The pedagogy and Design for learning for the website have already been defined in chapters 3 and 4. This will be built on through alignment with the data collection and evaluation strategies. Using the principals put forward by Lockyer et al (2013) it is possible to look at our learning design and infer the data points that would be useful. There are two types of data that can be assessed in most learning analytics systems these are *checkpoint* and *process* analytics. Checkpoint analytics looks to see if a task has been attempted or completed but not at the learning process, this might be the viewing of content or the handing in of coursework, whereas process analytics looks more at the learning cycle with data revealing the process (Lockyer et al 2013). Within the learning environment presented in chapters 3 and 4, there are two stages the first is the *declarative* phase and the second is the *functional* stage. This follows a mastery model of Bloom's taxonomy in terms of levels and transformative levels of SOLO. Under this learning design, the aim is to promote learners into a deeper model of understanding of learning and engagement.

This Thesis will take the novel approach of extending and aligning the Lockyer et al (2013) framework with the principals of Kirkpatrick in the evaluation. The Kirkpatrick model (Kirkpatrick & Kirkpatrick 2008) takes evaluation from the lower levels of reaction as a measure of satisfaction, through knowledge gained, behavioural changes into improvement in results. The approach was recommended by research councils UK (2005). Key to the model is that it links through the lower levels into the upper levels, so if someone enjoys an experience they are more likely to learn something. If someone learns something then it is more likely that they will change their behaviour and if they change their behaviour then they are more likely to see a change in their results. Learning on this level links to *declarative* understanding whereas behavioural changes are brought about by a functional understanding.

The framework for this combined model will be used here to align evaluation with learning analytics and learning design, considered across the first 3 levels of Kirkpatrick, listed below as satisfaction, learning and behavioural change.

5.7 Measuring Satisfaction

It has been shown that in the use of discussion boards and other web 2.0 technology that there is a relationship between the quality of interactions and communications and the satisfaction of learners which, in turn, lead to an increase in the reuse of the site. Quality->Satisfaction->Loyalty (Wang & Chiu 2011). Delon and McClean (2003) Information Systems success model shows how this is part of a bigger model in which there is clear feedback between use, benefits and satisfaction with a positive relationship. Using this inferred linkage this study will look at satisfaction through the use of the website and learners returning to the experiences. The concept here is simple: if the learner enjoys an experience then they are more likely to do it again. If a learner does not enjoy an experience then why would they repeat it? Delon and McClean (2003) suggest several methods for measuring use: frequency of use, time of use, number of accesses, usage pattern, and dependency. The first three will all be used here as they can be simply calculated from the history that is provided to each learner and teacher as an aid for tracking their own progress. Assuming the positive causal relationship they also give a simple figure to interpret: for frequency the higher the figure the higher the satisfaction, for the time of use the longer they use it for the more satisfied and for the number of accesses, the higher the number of accesses the higher the satisfaction. There is a question of whether this is employing *checkpoint* or *process* analytics as defined by Lockyer et al (2013) as this study is attempting to use *checkpoint* data to tell us about the *process* of learning via examining access patterns not simply access on its own.

5.8 Measuring Learning

The gaining of knowledge by the pupils involved would appear to be the simplest to measure, but as outlined in chapter 4 assessment, there are different levels of understanding and ways of using knowledge. On top of this, it is important to consider the difference between a level of understanding and the progress an individual has made when looking into the impact of the intervention.

The previous chapter on assessment discussed the different levels of understanding from both SOLO taxonomy (Biggs and Tang 2011) and Bloom's revised taxonomy (Krathwohl 2002). Both grapple with the issue of knowledge from different angles. Using Laurillard's conversational (2008) framework this study concluded that it is interested in two stages discursive and experimental, which here are mapped to understanding first being declarative and then being functional as defined by Biggs and Tang (2011).

As mentioned in the methodology outlined by Heron (2012) in section 5.5.2 it is possible to mitigate the need for pre and post-tests if measurements of the distribution of the population are made. As such the answers given to multiple choice questions in the content part of the website will be used to measure learning within the site using the above-stated methods. This sort of analysis should be *process* analytics as it not only shows us that they have engaged but students understanding as well.

5.9 Behavioural Change a Move Towards Deep Learning

It is commonly agreed that it is beneficial to promote deeper learning in pupils and that this can be achieved through pupils attempting more complex tasks that allow them to apply the knowledge that they have learnt (Biggs and Tang 2011). It should be possible to use pupils' engagement with these more complex tasks as a measure of behavioural change in a beneficial direction will enable pupils to go to achieve better results. Using Bloom's taxonomy as a model for higher levels of achievement (Krathwohl 2002) and the notion of functional understanding it is possible to identify the actions in learning that requires the student to analysis, evaluate and be creative. In order to use spacebook, the students' reflective space, students must have engaged in many of these aspects. They must have planned a set of images to order in order to fulfil their experimental requirements, they will have had to process and analyse the data that came back and finally they will have written something about it in their spacebook entry. At this point, no judgement will be made on the quality of learning that has taken place, it is the action of the student engaging in these higher order acts that are of interest. An increase in the number of spacebook interactions is seen as an indicator that pupils are more interested in engaging in higher cognitive load activities and as such have changed their behaviour towards activities that promote deeper learning.

5.10 Evaluation Conclusions

The split that has developed in the educational community between researchers and practitioners is an unnatural split (Beetham and Sharpe 2007), the addition of understanding and utilizing new technologies can further compound this but teachers want tools to help them not off the shelf solutions to their teaching problems (Laurillard 2008). Practitioners use their tacit understanding to make judgement calls on a day to day basis in reflective practice (Schon 1991), so by developing tools and frameworks (Laurillard 2008), it can aid teachers to take part in a formalised data-driven transformative reflective practice (Biggs and Tang 2011).

Alexander (2001) points out that all too often research ends up focusing on what the teacher has done and not the outcomes for the students, Learning analytics of virtual learning environments offer a potential solution to this problem (Lockyer et al 2013; Greller & Drachsler 2012) with access to large amounts of user-generated data. Heron suggests a methodology for comparative studies, that presents a solution to the difficulties found in other cluster study methods (Taylor at al 2013; Slavin et al 2014). By combining the Lockyers et al (2013) principals of aligned learning analytics with the Kirkpatrick model of evaluation (Kirkpatrick & Kirkpatrick 2008) a framework can be built for evaluation that allows the research to remain objective in their approach. Removing the difficulty of many studies of actively changing what is measured within the study with the result of leading to narrow or unwarranted learning improvement (Slavin et al 2014). This multi-dimensional framework means that a successful intervention must affect change in 3 areas: satisfaction, learning and behavioural changes, not just one area to the detriment of the other two. So it should not be possible to focus on satisfaction by just making something fun if the students then fail to learn and it should not be possible to promote shallow learning or memorisation of facts to the detriment of students undertaking deeper learning behaviours.

This Chapter has shown how the processes of assessment in an e-learning environment can be aligned with 3 levels of evaluation to enable a practitioner to develop their teaching methods through a reflective cycle. The next chapter will outline the implementation of this within the Bradford Robotic Telescope and the tools used to collect data for evaluation.

Chapter 6. Implementation of the e-Learning Environment

6.1 Outline of this Chapter

This chapter will outline the learning environment that was developed. The structures used to deliver the system and the tools developed as a result of the pedagogical concepts laid out in chapters 3 and 4.

The processes outlined in chapter 3 will be developed using theories of Virtual Learning Environment design in order to build a robust flexible environment. Within this environment, the tools for evaluation are placed and this chapter describes these tools in detail including their structures and main features. The base learning environment, developed for this project, will be explained, this will include descriptions of the tools developed and used to make and display content.

Testing of this environment and the tools for data collection and analysis are covered in chapters 7, 8 and 9.

6.2 Implementation of The Telescope e-learning System

Chapters 3 and 4 outlined the systems required for an online learning environment. In this section, the implementations of these systems are described.

The e-learning system was built on an existing framework for the telescope and its associated interfaces and user account system, this system was developed by Tallon (2010) and the telescope team. This is described in section1.4.

The system is broken down into 4 areas of action in which users can generate and view data

- **Learningscope:** The Content Management system

  Allows for the creation and display of content including, text, multiple choice questions, video, simulations, games
- **The Telescope Interface:** Allows for the taking and processing of images, stores images for users
- **Space Book:** The reflective area for user-generated content

  Personal web space for users to comment on and share their telescope observations
- **The Objective System:** Allows the setting of objectives for pupils, and for pupils to respond to objectives. This sits outside the three levels of the system described in section 3.9 which relate to the 3 previous bullet points as the setting of objectives could be for any of these three systems.

On top of this, there are 3 levels of user interaction defined in the system. Each builds on the last in terms of its capabilities and any account can be upgraded to the next level if required.

- **The student level interactions:** Stores information about students, their interactions with content, their answers to questions and comments on objectives, the group that they are in i.e. their class

- **The teacher level interactions:** stores information on the teacher and their relationships to students within the system. This section will also cover the way simple actions that can be completed by the teacher.
- **The moderator and developer level interactions:** Can approve and develop webpages or comment and suggest changes to be made

6.2.1 Learning Objects and Content Centred Virtual Learning Environments

In the late 1990s and the early to mid-2000s, there was much discussion on the best way to structure content in the virtual learning environment and meta tag its content so that it could be shared between learning environments and connected together in order to create a course of work. This need primarily arrived because of the time it takes to produce online learning resources and the potential savings made by being able to share this content.

In 2002 the learning object metadata standard was developed as an extensive format to help describe learning objects. Unlike the principal of the single argument lesson as a learning object or an atomic structure that supports a single learning concept the protocol is cable of supporting many different structures below it from a single image through to a course of work (IEEE 2002).

However, the Learning Object Metadata standard describes many of the features that you would hope to know about a learning object, divided into the categories

- general, description of the object as a whole
- lifecycle
- rights
- technical, information needed to run the object
- educational, pedagogical description of the objects approach
- relational how this learning object is linked to others
- classification, for use with classification systems

(IEEE 2002)

as well as pointers towards the address that the final information the learning object contains. This metadata system leaves the structure of the learning object very open and as such, it is important to describe how it is to be used within this telescope system.

Key to the concept was the learning object, this is a single concept that can be presented to the user without the need for any other information. Designed well a learner should be able to view the object independently of any other information and still be able to understand the information contained within it. Three possible models for the storage of learning objects were possible. The first would involve a system where they are separate files uploaded into the system, this is very much the approach taken in systems such as Blackboard (Blackboard n.d.) where the power points presentations and other file formats can be uploaded into the system for students to view. The second is the use of an interpretive language to build pages much like that found on Wikipedia (Wikipedia 2017) or other wiki systems and the third was a database driven system which separated structure from content in the system or any sort of mark-up language.

One aim was to separate completely the content from the structure of the webpage. This is similar in a way to CSS (W3C 2017) but taking this a step further, as the amount of content as well as the look of the content could be easily controlled. For example, the length of a learning object in terms of content could be independent of the number of pages that it appeared on. Sections where defined to stop splitting on information where it would be impractical (Figure 6.1) but otherwise this allowed content to be fitted to a screen or split over various pages on the fly rather than being hard-coded into the page.

A page builder was developed, which allowed administrational users of the system to design learning objects by uploading content such as images, videos, animations and games and text editor that included the use of HTML content and links both external pages and internal learning objects. Multiple choice questions could also be built with a question and branch options with responses for students if an option was chosen by the user. The user side of the question system will be explained in more depth in a section later in the thesis.

6.2.1.1 Implementation of Learning Object System

Rational database tables are used to describe the learning objects so that they can be built dynamically, and alterations tracked within the system.

A key feature of the storing and generating of learning objects on the fly was the idea that they could contain different approaches to the same subject, which could be organised on the page as required at the time. Similar to the approach suggested by Atif et al (2003). Each learning object is subdivided into sections to allow alternatives to sections such as language or to allow the same content to be used in multiple places a lookup table is added to allow a many to many relationship.

Table 6.1 Database Table for Sections in Learning Objects Implementation

section_id	unique identifier
LO_id	identifier of the object which this section is part of
Position	position in the learning object, given as a number. i.e this is the first part to appear in the learning object. Note sections with the same position are allowed as these represent alternatives, at this higher level, this was handled in the system as drop down content.

Table 6.2 Database Table for Layout in Learning Object Implementation

Layout ID	unique identifier
Section ID	identifier of the section which this item is part of
Position	Numeric
Type	text, question, animation tells the system which sub table to find information about this section in

Tables are then used for each different type of data to be stored with key values stored in the table but files stored under a directory structure table 6.3 shows an example for text.

Table 6.3 Example database table for text content

look up id	used to find the content for a section of a learning object
Content	the text to appear in the final Learning Object

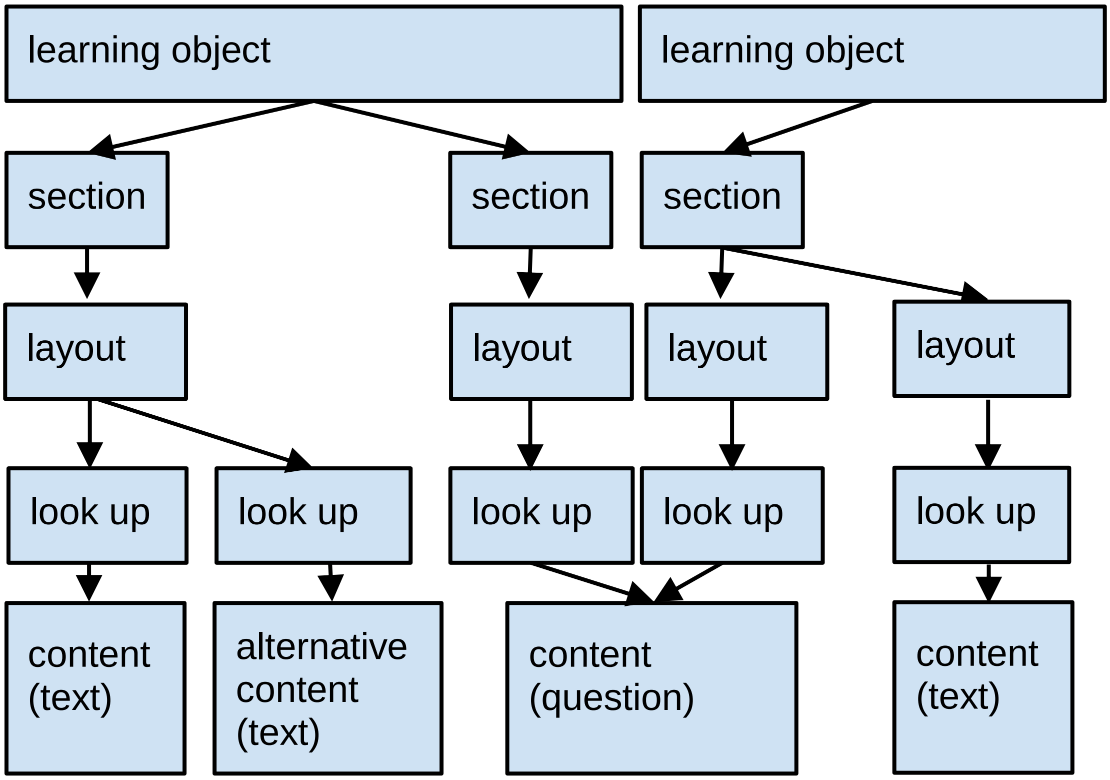

Figure 6.1: Example of learning object data structure

Figure 6.1 shows one possible structure available in two learning objects. The first learning object contains two sections of content the first a block of text, for which two alternative versions are available and the second is a question. Possible alternative versions of text could be different languages or different levels of complexity, they could even be revisions of the text made over time, all are possible after this level of dissociation. In the second learning object shown there is one section with two parts of content. The question is connected via the lookup table to the same question in the first learning object and this is followed by some text. A many to many structure allows functionality such as only having to update information once on the system if it is used in multiple places and allowing for alternatives such as different resolutions or formats of image files or videos.

6.2.2 Lesson Plan Route System

Learning objects could be collected together in a structure called a route. This allowed lesson plans to be designed around a number of different learning objects or activities. Routing allowed for both sequential and flat structures with alternatives. Learning objects are numbered to give them a position in the route and if given the same position number as another learning object they would appear as a flat structure offering alternatives within the route which could be explored via the navigation menu. Each route was given its own ID and a meta-data table could be used to describe amongst other things its learning objectives. Links could also be made in the learning object to other learning objects in the system. This is slightly outside the normal use of learning objects as they are meant to be self-contained but is very much a key affordance of web-based learning as people expect the web to be a network of interlinked pages (Laurillard 2000).

6.2.3 Multiple Choice Question System

Multiple choice questions were added to the system in order to provide feedback to both teachers and students on the student's current understanding as outlined in chapter 3 and 4.

All questions were accompanied by a star points game, students could risk up to twenty stars on getting a question right. If they got the question right they won the stars and if they didn't they lost the stars. All students started with 100 points and negative points were not possible, if and when a student reached 0 they were allowed to risk stars that they didn't have, in the same way, they had to up to that point.

All questions were set up so that they could only be answered once by any student. Questions were placed in line with activities, to reduce the separation between engaging with the activities and answering the questions on them.

During initial testing it was found that if students had multiple attempts they would gameplay retaking questions to find the right answer, so a single attempt was introduced to reduce this behaviour. This did however illicit team game playing where one student would attempt a question before their friend in order to find the right answer. To minimise the negative effects of this, the positions of possible answers were randomised, this way discussions around this sort of game playing took a content led rather than position led dialogue.

Typical conversations before the change.

> “I’ll try the first option to see if it’s right”

Typical conversation after the change.

> “I’ll try half-moon to see if it is the right answer”

Teachers had the ability to reset any questions as unanswered to students and to give points to any student that they wanted to. The main reason for this was during testing it was found that a small percentage of students became upset if they didn’t initially understand the game whilst answering the first question, these two options gave teachers ways of mitigating these circumstances.

A high score table was available so that students could see how many points they had in relation to the rest of their group.

6.2.4 Storage of User Response to Multiple Choice Question

Information was stored on the user response to each question

Table 6.4 Database Table for Storage of Question Response

User ID	QuestionID	Response ID	Mark	Stars	Time

From this data, it was possible to highlight to teachers students that were struggling with questions by showing a percentage of correct answers, which could be further broken down to show which responses were most popular.

6.2.5 Telescope Systems

Systems for the taking and processing of images were developed by Tallon (2010) and the telescope team. Taking images was via a wizard system that completed an order template. Links could be provided to examples of pre-filled templates for observations relevant to the activity i.e. a template was supplied that allowed students to take an image of the whole moon each night for a week. These templates could be further edited if the student wanted to. So in the above example, the telescope optical device could be changed to alter the level of zoom. See 1.4.2 for a more detailed description of telescope hardware.

Images returned could be manipulated to bring out the features students wanted to highlight and then saved as images in their processed forms.

6.2.6 Social and Reflective System: Spacebook

The telescope social and reflect environment was called spacebook. It allowed users to place processed images produced through the process described in section 6.2.4 into their own personal web page. They could then annotate them both with text and drawing over the top of these images and text and basic line graphs in line with the images.

Through a menu system, users could add content to support images including text and basic line graphs from data sets, these could then be arranged by users.

Users were also able to create new pages so that they could sort images and other elements into topics of interest.

Teachers were able to approve pages for sharing with others on either a school or class level.

Class level approval could be turned on during a session to save teachers having to approve every update. School level required individual approval of posts. Interschool approval was available by system admins after school level approval had gone ahead. Due to the nature of the approval system, multiple versions of all components were stored in the system so that only the correct level of approved content was shared at the approved level.

Teachers were able to edit and comment on spacebook pages and had the same level of access as the student that had developed the page.

It should be noted out of these tools that the image annotation overlay system was developed by another member of the telescope team to the specification outline for this thesis and all other parts of the system were developed as part of space book system were developed as part of this thesis.

6.2.7 Objective System

Independent of these systems an objective system was developed that allowed teachers to set tasks for the users and provided a text-based space for students to respond if needed to the instructions. Objectives were visible on both the student's home page and from the dynamic drop-down menu where they could alter their response or notes and mark the task as complete.

6.2.8 User Interface

The user interface was split into 3 areas

- Navigation
- Activity interface
- Menu interface

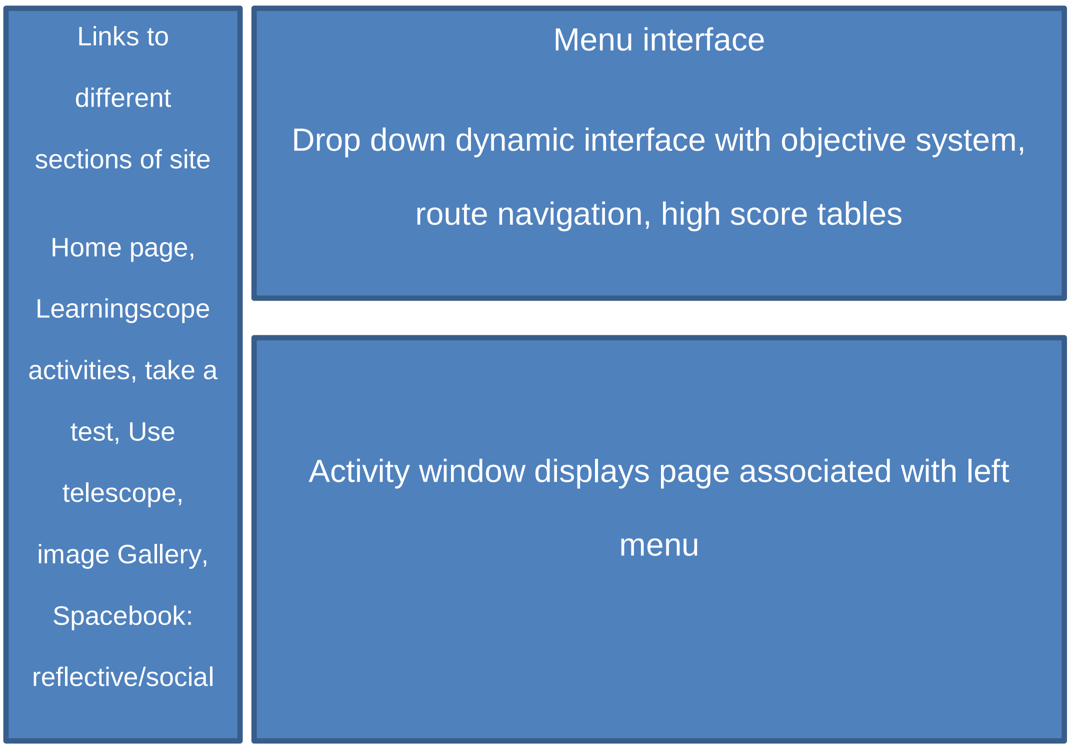

Figure 6.2 Website interface

6.2.8.1 Home page

A home page was provided to users with links to the last page they looked at information on their progress through the material and the percentage of questions they had answered.

6.2.8.2 Route Navigation Menu

A map of the current route was supplied to pupils with information on which pages they had visited and the questions on each page and if they had been answered. This was situated in the drop-down menu at the top of the page and so could be viewed at any time by the pupils.

6.2.8.3 Objective menu

The objective system was located in the pull-down menu at the top of the page and so could be accessed at any time without the need to disengage from the current activity. i.e. using spacebook or viewing content in the learningscope system.

6.2.9 Implementation of Student Model

Outlined below are the database structures used to achieve a profile of the student base user information table.

Table 6.5 Database Table for Student Core Information

User_ID	unique main telescope account user ID links functionality of the system to the learner
LSUser_ID	unique ID of the learning system. kept as separate from the main system ID as this would allow a single account to have multiple learning accounts if required. As such a single telescope user could take part in two courses and progress independently through both.
LSposition	current learning object of the user, allows the user to access the last page of the learning scope system that they view between sessions.
Route	the lesson plan that the user is currently following

A history of the pupil's progress is also kept

Table 6.6 Database Table for Student Access

ID	unique identifier
Lsuser_id	used to link to the lsuser table
lo id	used to link to the learning object table
Time	the timestamp of when the page was visited

This information was used to generate reports for the pupil and teacher showing the pages that they had visited that related to the topic that they were studying at the moment. They also produced a history for users and teachers in the system and were used to enhance navigation through the system.

Information on the questions the user has answered is also stored. This is discussed in more detail in 6.2.3.

6.2.10 Teacher Model

A teacher needs to function as the owner of groups of pupils, they need to be able to set work for pupils and access the results of pupils, mark work and give feedback.

The following functions were available to teachers

- General function
  - Select which class they want to view the information of
  - Add/remove or relocate students in groups
- Objective functions
  - Set an objective
  - View student responses to objectives
- Content system function
  - Place students on a route
  - Lock pages from a group
  - Hide pages from a group
  - Position student to a page in the learning material
  - View current position of students in the system
  - View a history of student access to the system
  - View answers of students to questions on route grouped either by a student or by a question
- Social / reflective system functions
  - View students' spacebook pages comment on and edit
  - Approve content sharing
  - Set auto approval of class level sharing

6.3 Summary of Chapter

The development of the learning environments was based on the existing system for the telescope systems and user accounts as developed by Tallon (2010). This chapter has outlined the four systems used by the website

- The content management system called learningscope

  Allowed the creation and display of content. Including a multiple choice question system.

- The telescope tools system

  Allows pupils to take and process images with the telescope for constructed environment tasks.

- The social/reflective environment called spacebook

  Allows users to comment on and share images that they have taken with the telescope.

- The objectives system

  Allowed teachers to set objectives and students to respond to these. This system was external to the above three systems to allow the setting of objectives across one or more of the above areas within the website.

Three levels of use were possible in the website: editor, teacher and pupil. Pupils generated information and content in the system through their interactions with the above four systems which were associated with their individual user account. Teachers could carry out a wide range of activities to allow them to control and monitor access and content of pupils in the system. Finally, editors could develop content in the system for others to view and approve the posts made by others.

This chapter has outlined the different areas of the telescope website, how they were intended to be used by the different levels of user and the data that they were enabled to collect on these interactions. The next chapter will look at the methods by which this data should be processed and analysed in order to enable simple interpretation for the end user in their evaluation of the learning process.

Chapter 7. Method of Data Analysis and Evaluation

7.1 Evaluation of the delivery methods

This chapter will outline the experimental design to test the systems described in this thesis up to this point, it will outline the approach taken to analyse the data the system collects and as a result highlight the strength and shortcomings in the approach presented in this thesis. The system will have to be able to see past the context of learning of individual schools and classes if meaningful comparisons are to be made about approaches to learning. Within the schools environment there are natural clusters of students with similar traits, which could be because of the setting of students due to their ability or the social economic factors found in the district where the school is located and these must be taken into consideration in the analysis of data from schools (Taylor et al). It is the clusters themselves that will allow us to look at the effects of these measurable factors (Slavin et al 2014) and model these factors against predicted normal distributions (Heron 2012). However, this study itself is contextual in that it is looking at primary school science in Key Stage 2, 7-11 year old in the UK, with the target age group was upper key stage 2, 9-11-year-olds, and so has its limitation. Detail of the context of the schools that took part in the study will be given. A better understanding of the sorts of data sets produced by virtual learning environments is required if they are to be of use.

7.2 The Context of the Study

The following section will outline the context of the study including, teaching requirements of the national curriculum in the area of astronomy for Key stage 2 students, the background of teachers to this area, the classroom setting for teaching and the profile of schools in the area where the study took place.

7.2.1 National Curriculum Requirements for Teaching

In primary schools, the UK national curriculum covers the fundamentals of space has been part of the National Curriculum for the duration of this study. The 1999 national curriculum asks students to understand that the earth sun and moon are spherical, that the position of the sun changes during the day and this causes shadows to change, that the rotation of the earth causes night and day and that the Earth takes one year to orbit the sun and the moon takes approximately 28 days to orbit the Earth. (National Archive 2010). These requirements remain unchanged in curriculum revisions until 2013 (National Archive 2014). In the 2013 revision, the national curriculum is rearranged from a purely key stage structure, with guidance for teaching order, to a more structured school year based program in an attempt to give clearer information about how pupils should be progressing. This change brings the topics relating to the Earth and space to year 5 (9-10-year-olds) teaching (DfE 2013). The solar system is part of the national curriculum for the first time (DfE 2013), although many schools had previously covered the subject. All of these topics allow for simple observations that needed the minimum of manipulation and only simple data collection and so are suitable for teaching with the telescope.

Although there have been a number of changes revisions to the National Curriculum for primary schools, their effects on the teaching of astronomy have been limited.

The majority of the data collection for this study took place between 2010 until 2013, meaning that changes to the curriculum made in 2013 were just coming in and being implemented by forward-looking schools.

7.2.2 Profile of UK Primary School Teacher

In comparison to the amateur astronomer the telescope system was designed for (Tallon 2010), the number of teachers that are specialised in astronomy is limited. In 2013 it was reported that only 55% of secondary schools physics teachers had a degree in that subject and that most primary schools had no teachers who had studied science at degree level (Thomas 2013).

However, the requirements of the UK national curriculum for this level are, as outlined above, basic concepts, all of which are based around observable changes that could be observed in everyday life such as the phases of the Moon.

7.2.3 The Setting of Learning and Teaching in most UK Schools

This study took place in the formal teaching environment within schools. In this setting, a teacher delivers a class in a single room to a group of pupils.

As such the system is built around a similar formal learning setting, with the aim that the website is an aid to this teaching.

This carries a very important impact on any systems designed, i.e. that there is always an alternative system of direct communication for the teachers to use in the classroom. Any system to be successfully used must provide an improved service. If a system doesn't help a teacher in their teaching then it will not be used.

7.2.4 Bradford and the Surrounding Area

The majority of this study took place in Bradford and locations in the surrounding area, with additional schools from Leeds, Calderdale, North Yorkshire, Rochdale. A small number of schools came from Reading and Wokingham.

Bradford is part of the West Yorkshire conurbation that makes up one of the most densely populated areas of the UK outside of greater London. This region contains 17% of the UK's 1% most deprived areas (Index of Multiple Deprivation 2010). However, in contrast to other similar urban areas in the UK Bradford is growing with an 11% increase in the 2010 census. 22% of the population is aged 0-14 years old, only two other authorities have higher percentages: Slough and the London Borough of Barking & Dagenham (City of Bradford Metropolitan District Council 2012).

Bradford has around 150 primary schools in the local educational authority. This number fluctuated year to year with 155 in 2010 and 157 recorded in 2009 (Ofsted, n.d.). The district has consistently underperformed compared with the national average, and although there has been a consistent closing of the gap the area is still below the national average. 2011 saw the number of pupils achieving level 4 or above in English and Maths at Key Stage 2 increasing nationally to 74% but in it dropped to 71% this drop continues in the District to 70% in 2012 and 69% in 2013 where the national level stays at 75%. In Bradford, this level increased in 2014 to 73% (national 78%). Finally, this continues to 76% in Bradford against a national average of 80% (Department for education 2015).

The picture in Science is similar, with Bradford primary schools performing behind national average with nationally 89% of pupils achieving level 4 or above in Science and the Bradford District achieving 85%

Table 7.1 Percentage of student achieving English and Maths level 4+ at key stage 2 Department for Education (2015)

Year	2009	2010	2011	2012
National average (state funded schools only)	72	73	74	79
Bradford (excluding independent)	68	73	71	75
North Yorkshire	74	78	77	80
Leeds	72	74	73	77
Rochdale	73	75	72	77
Calderdale	76	75	77	82
Reading	69	72	69	77
Wokingham	79	80	79	82

All areas show an improvement. Bradford and Reading preforming similarly but slightly under the national average. Leeds and Rochdale both perform approximately in line with the national average falling slightly behind in 2012. North Yorkshire, Calderdale, and Wokingham all outperform the national average year on year. This means that there is a good range in areas from which the study can be drawn.

Table 7.2 Science performance by district 2010 – 2013, Department for Education (2015).

year	2010	2011	2012	2013
National	85%	85%	86%	88%
Bradford	81%	81%	82%	82%
North Yorkshire	87%	87%	87%	89%
Leeds	83%	82%	84%	86%
Rochdale	83%	83%	82%	86%
Calderdale	85%	85%	88%	89%
Reading	83%	81%	83%	83%
Wokingham	90%	90%	90%	93%

Performance in science is more consistent in the Bradford district with a slow increase between 2009 and 2012

Table 7.3 Percentage of schools seen as good or outstanding by district (Ofsted n.d.)

Year	2009	2010	2011
Bradford	57% (of 157)	61% (of 155)	68% (of 155)
North Yorkshire	72 % (of 324)	76% (of 323)	76% (of 322)
Leeds	66% (of 219)	66% (of 219)	68% (of 217)
Rochdale	71% (of 68)	71% (of 69)	74% (of 69)
Calderdale	64% (of 85)	66% (of 83)	68% (of 83)
Reading	46% (of 37)	46% (of 37)	50% (of 34)
Wokingham	72% (of 47)	71% (of 48)	71% (of 49)

7.3 Experimental Design

This section of the thesis examines the methods to answer three key questions:

1. Is it possible to compare results between the clusters, is there any sort of consistent performance in the 3 identified factors for enjoyment, learning and behavioural changes, that could then compared to a different approach? Or is the null hypothesis true that the context of the individual learners has too large an effect on the performance of students within the environment and it is not possible to make any sort of judgement about the quality of material and methods used in the teaching environment?
2. In terms of learning can confidence data be used as a more sensitive tool than just the number of right answers, is there a positive relationship between the two and what extra information could it provide in the learning stage of the programme?
3. What is the effect size of changing approach within the study?

The approach of this study is to take a quasi-experimental basis and is therefore comparative in nature, with the aim to see if alternative subject-specific pedagogical approaches can be seen to make a difference when delivered through an online blended environment. This data will then be used for a critical evaluation of how effective various tools used in evaluation have been in helping us to make this evaluation of taught content.

In order to provide a measurable effect of the change in teaching strategy a clustered analysis model as presented by Heron (2012) was used. A clusters approach is supported by Slavin et al (2014).

Through the use of clustering, the effects of factors beyond the control of the study can be taken into account. This is of particular interest in this study as it is interested in seeing the difference between the effect of the changing strategy for teaching rather than the context of teaching.

7.3.1 Delivery of Introduction to the Bradford Robotic Telescope

The telescope system was introduced to schools using a two lesson model. Each lesson was a week apart. Five members of the telescope team were involved with visits to schools during the period of the study. Deliverers were free to introduce the telescope using either of the routes and where allowed to choose their preferred model. It was felt that the deliverer had to be comfortable or this would have an undue effect on the results.

Due to the large numbers involved with the trial a fair spread of schools would appear in both cohorts. However, this was also felt to be consistent with the future use of the system in which it would be used.

7.3.2 Development of Learning Materials for Comparison

For this study, two routes are compared to analyse the impact of changes in learning designs. The telescope delivered activities over a wide range of topics for teaching astronomy to 8-year-olds to 18-year-olds. Topics covered everything from the basics on the cause of night and day or the phase of the Moon through to more advanced topics on Hubble's constant and Cepheid variables. Each of these learning designs was called a route in the Bradford Robotic Telescope system and had its own unique ID, given to it sequentially on creation, there are over 100 learning designs covering a range of topics and pedagogical approaches. Each route also had a title and metadata to describe it, this structure is described in more detail in section 6.2.2.

The two routes used in this study had IDs within the system of 65 and 84. These routes were used as the learning designs covered the same area of the national curriculum, were aimed at the same age group, had a similar pedagogical structure, learning objectives and learning outcomes, thus limiting the number of factors that might affect site use outside of the proposed change in pedagogical approach. Both routes covered the national curriculum requirements for the phase of the Moon. This was one of the more established areas of curriculum on the site, as it was the focus of early redevelopment of the telescope (Baruch et al 2005). Both routes had a similar approach to pedagogical structure, which followed that outlined in 3.8: first teaching an understanding of the phase of the Moon through materials in the learningscope system, before pupils experimented using observations of the Moon to predict how the phases would change, finally in both routes pupils were encouraged to share their results with others through their spacebook pages. As such the pedagogical change looked at is a minor change rather than a complete redesign in approach.

The routes differed in their approach in the context of the experimental phase: route 65 was a *pure scientific* study, pupils learned about the phases of the Moon and took images to observe the changes. Route 84 was contextualised with an authentic context of planning a Moon landing. Pupils used the phase of the Moon to choose a landing site with the maximum number of hours of sunlight. Observations would be needed in order to check the phase of the Moon and predict future phases for both routes.

More detail on the routes can be seen in tables 7.4 and 7.5

Table 7.4 Learning Design Metadata for Route 65

ID	65
Title	The Phase of the Moon
Summary	Pupils are introduced to the concepts of the phases of the Moon and their progression. Students use this information to plan a set of observations of the Moon observe its phases.
Topics	The Phase of the Moon
Learning outcome	• To understand the cause of the phases of the Moon and how they change over time. • Using a moon phase diary, Use the current phase of the moon to predict future phases over time (approx. 28 days for a full cycle)
National curriculum links	**The Sun, Earth and Moon** that the Sun, Earth and Moon are approximately spherical Periodic changes d. that the Earth orbits the Sun once each year, and that the Moon takes approximately 28 days to orbit the Earth **Everyday effects of light** that light travels from a source b. that light cannot pass through some materials, and how this leads to the formation of shadows c. that light is reflected from surfaces [for example, mirrors, polished metals]
Rationale	Assimilation phase using the learningscope system, followed by an experimental phase using the telescope and then the sharing of results through spacebook
Duration	2 x 2-3 hour slots a week apart
Learners	Key stage 2 students

Table 7.5 Learning Design Metadata for Route 84

ID	84
Title	Planning a Moon landing
Summary	Pupils are introduced to the concepts of the phases of the Moon and their progression. Students use this information to plan a Moon landing. They then take observations and adjust their plans in light of their observations
Topics	The Phase of the Moon
Learning outcome	• To understand the cause of the phases of the Moon and how they change over time. • Plan a trip to the Moon, Use the current phase of the moon to predict future phases over time (approx. 28 days for a full cycle)
National curriculum links	**The Sun, Earth and Moon** that the Sun, Earth and Moon are approximately spherical Periodic changes d. that the Earth orbits the Sun once each year, and that the Moon takes approximately 28 days to orbit the Earth **Everyday effects of light** that light travels from a source b. that light cannot pass through some materials, and how this leads to the formation of shadows c. that light is reflected from surfaces [for example, mirrors, polished metals]
Rationale	Assimilation phase using the learningscope system, followed by an experimental phase using the telescope and then the sharing of results through spacebook. Authentic task
duration	2 x 2-3 hour slots a week apart
learners	Key stage 2 students

Learning materials were designed by telescope team members separately, so as not to bias the result of a preferred model during the design phase. This was felt to be in line with the future use of such systems were teachers or in this case outreach officers, would develop materials themselves and test to see if their approach was an improvement on previous approaches.

At this point, an attempt is not being made to attribute a reason why one approach might be better than another. This study is also not trying to match an approach to a particular type of user.

7.3.3 Pilot Study

A pilot study was presented as a paper at ICEILT in 2015 (Machell and Baruch 2015) that looked at both the consistency in questions answered and the engagement with the site. This pilot looked at the three areas: satisfaction using the number of page views, learning through clustering of question results and confidence based testing system.

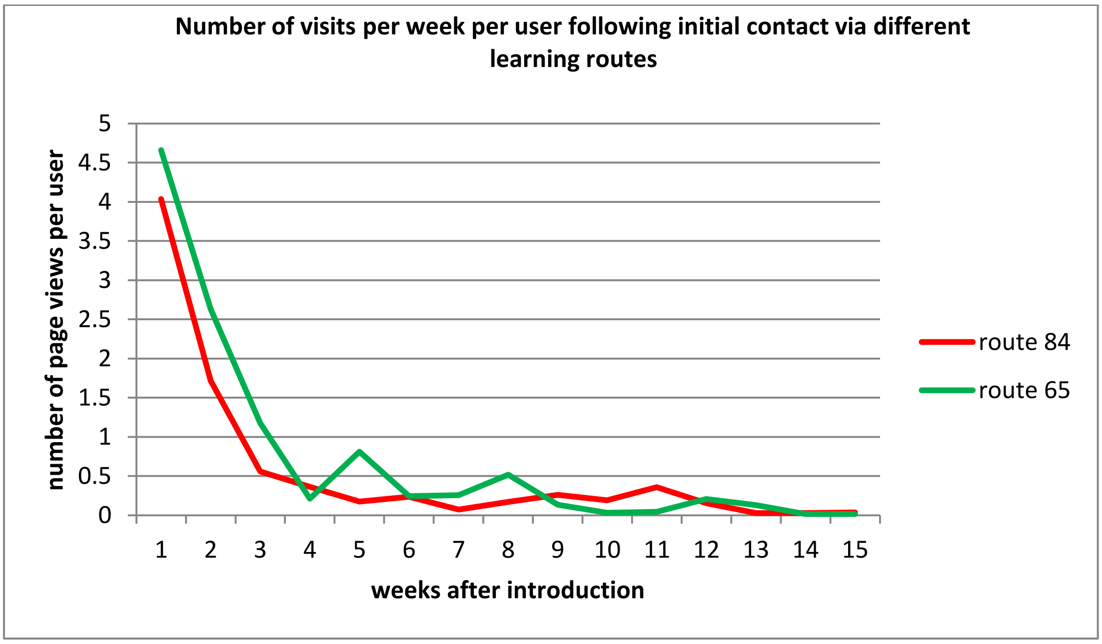

Figure 7.1 Number of visits per week per user following initial contact via different learning routes (Machell and Baruch, 2015)

Both routes appeared to have a similar visits pattern. Route 65 appeared to have slightly higher weekly visits, although tailed off around the 4-week point.

A more detailed breakdown using clustering was needed to better understand the context implications.

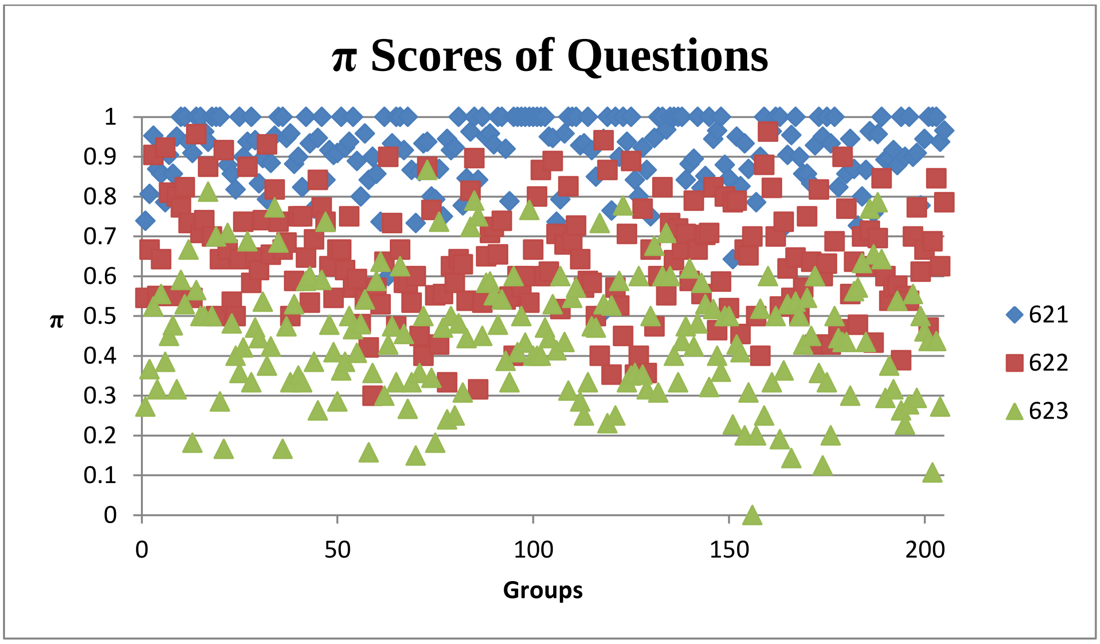

Figure 7.2 Spread of groups answer questions in the pilot study (Machell and Baruch, 2015)

Figure 7.2 shows individual cluster's responses to questions, the grouping of difficulties is possible to see. However, there was not enough data at this point to see if this was constant across a route as well as single questions.

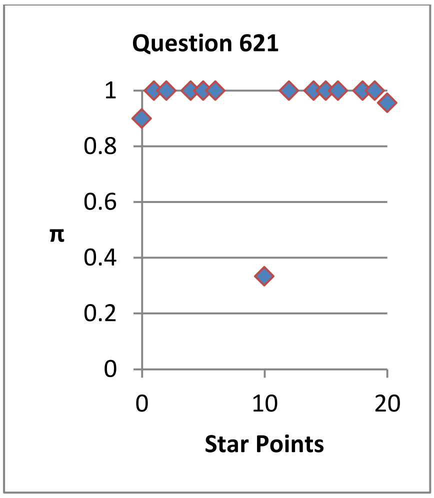

Figure 7.3 Confidence question 621 (Machell and Baruch 2015)

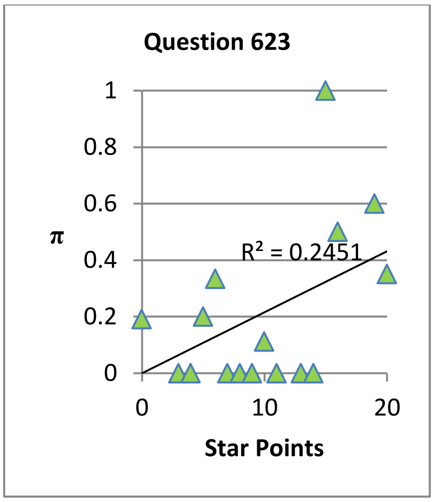

Figure 7.4 Confidence question 623 (Machell and Baruch, 2015)

In the above graphs (Figure 7.3 and Figure 7.4) the percentage of correct answers π, calculated as the number of correct answers/number of attempts, was plotted against the number of stars the individuals choose to risk. This shows the relationship between confidence and correctness. Figure 7.3. Shows the relationship for a question with an overall high item facility of ~0.9 (item facility is the percentage of correct answers/number of attempts) and Figure 7.4 of a much lower overall item facility ~0.6. As can be seen, there is almost no relationship between correctness and confidence in Figure 7.3. But there is a relationship in Figure 7.4. As such it was concluded that the relationship between confidence and correctness was not as simple as the linear relationship suggested in other papers (Foster 2015) and that learner had a base level of confidence that was in turn affected by the difficulty of the question. This relationship is further investigated in the method listed below.

7.4 Clustered Method of Analysis

In this study analysis through the use of clustering provided by the groups that data was originally collected through has three aims:-

1. to establish the suitability of data for interpretation by normal distributions
2. to look at the impact of known factors
3. to see if a meaningful impact can be made form a change in approach

A cluster in this study refers to the natural grouping of students that happens in the classroom environment for education. It is a feature of the data as it was collected and should not be confused with the clustering approach found in data mining, which is an analytical approach applied to data after it has been collected. Each cluster is therefore collected with its own unique set of contexts; these might be the results of efforts made in the school i.e. setting due to ability or the impacts of socioeconomic factors of the wider environment.

The number of users in a cluster is given by the number of accounts created for a group that have at least one successful login to the telescope system. Accounts created but never used will not be counted as part of the cluster.

Using a modified version of Heron's (2012) method for comparing the results of first-year physics students, which examined the effects of teaching in a lecture course, it is possible to examine the impacts of learning designs. Once the variance in a single learning design's distribution has been examined, the effects of known factors can be examined to look at their impacts; finally, a comparison can be made to an alternative method of delivery to examine its impact.

The distribution of clusters is compared to Gaussian or Binomial distributions as they offer simple interpretation via the use of a mean and the standard deviation. A key factor in design that must be considered are the competences of the end user (Greller and Drachsler 2012), which in this case are primary school teachers. In England, primary school teachers are only required to have a GCSE grade C in Mathematics (Department for Education n.d.), so a normal distribution model has been used in order to support simple statistical analysis by the end user. Other distributions such as the Poisson distribution might be more suitable for time-based events (Thompson 2001), however, the added benefits of these distributions allowing for skew in data might cause misinterpretation by the end user of the significance of the impact of their changes (Cohen et al 2011). Non-parametric analysis of data sets would also allow for a wider range of distributions to be examined (Ibid), however again more sophisticated analysis methods would be required by the end user.

A chi-squared comparison to a normal distribution modelled using the samples mean and standard deviation will be carried out, this will show if a sample is suitable for simple interpretation using of both mean and standard deviation by the end user.

Effects of known contextual factors will then examined for correlations in results to see if any of these factors bias the system.

The aim is to give the end user confidence in their interpretation rather than assuming normal distributions of alternative parametric or non-parametric data sets, thus causing problems in the analysis (Mitchell 2000). The method does not hide or automate (Siemens 2013) or over generalise the interpretation process (Gašević et al 2016), but aids the end user in their own interpretation of a context-driven analysis (Laurillard 2009; Greller & Drachsler 2012). By automating the more complicated statistical elements, it allows for simple interpretation by the end user through simple statistical techniques as a starting point, rather than hidden statistical methods to produce a single number that might have no real meaning to the end user, thus removing their ability to meaningfully interpret it in a given context.

7.5 Measuring satisfaction

Satisfaction will be measured through reuse (Wang & Chiu 2011) and will be measured through 3 methods: number of visits, the frequency of visits and length of use (Delon and McClean 2003). The system stores the time each time a student accesses a page for their personal records and for their teacher. All methods will use data collected by the access log database table outlined in 6.2.8.

- The number of visits will be measured from the number of different views of content pages from the student.
- Length of use will be calculated from the first logged use of the learningscope pages to the last logged visit for an individual. This will then be averaged for the cluster by dividing it by the number of users in the group.

$$mean\ number\ of\ days\ use\ for\ cluster$$

$$= \frac{\sum x_{time\ of\ last\ use} - x_{time\ of\ first\ use}}{number\ of\ active\ accounts\ in\ group}$$

- The frequency of use will be calculated by dividing the number of visits in the first 30 days by an individual by 30. The mean frequency of visits will be calculated for each cluster. Frequency varied as the period of use increased. As 30 days was the point at which the majority users stopped using the telescope system in the pilot study (Machell & Baruch 2015) it was used to avoid distortions which occurred if the use was averaged over the total time of use for any user.

7.6 Measuring learning

Questions will be analysed to work out the number of correct answer against the number of users in the cluster.

This will produce an item facility for the question which is the probability of a correct answer. Using the following formula

$$item\ facility = \frac{number\ of\ correct\ answers}{number\ of\ learner\ in\ group}$$

A learner will only be counted as in the group if their account has been logged into a least once. In this way, any account that was made for a user that never used the system will not distort the result.

This will be compared for all the questions on a route, individual questions and 3 key questions for comparison with an alternative method.

$$route\ facility = \frac{\sum item\ facilities\ by\ group}{number\ of\ questions\ attempted\ by\ that\ group}$$

Clusters that do not attempt at least half the questions on the route will be rejected as not completing the route.

Groups where less than half the users in the group at most attempt any of the questions will be rejected. As it is deemed they did not engage with the route as a group.

Results will be modelled against the predicted results using a binomial distribution where the probability is given by the route facility. For simplicity of modelling the average group size will be used following the method presented by Heron (2012). This will be visually compared and a goodness to fit chi-squared calculation carried out.

Cronbach's alpha will be used to see if this is a suitable tool for measuring their understanding.

$$\alpha = \frac{k}{k-1} \times \left(1 - \frac{\sum_{i=0}^{k} \sigma_{y_i}^2}{\sigma_x^2}\right)$$

Where $\sigma_{y_i}^2$ is the variance on the individual item and σ_x^2 is the variance of observed total test scores.

As well as this, three key questions that are found in both approaches to the route will be compared to see if they are consistent.

7.6.1 Confidence Measurement

The confidence of students answering questions will be examined. More detail on the collection of this data can be found in 6.2.3 this section looks at the methods for analysing the data collected.

Students marked themselves on a scale of 0-20, with a correct answer gaining them the points and an incorrect answer losing them then the points. This draws comparisons with the methods presented by Foster (2015) in which secondary students marked their confidence in a maths test. The study showed that students understood the concept of a confidence scale around a win-lose model and it showed a high degree of correlation between item facility and confidence. Analysis by Cronbach's alpha to showed that it was a better measure than answers alone.

A key difference to this study is the number of stars that can be risked is 20 rather than 10 and the fact that instant feedback is given to the students after each question rather than the questions being marked as a whole.

Although the system was explained to students before they started the route no set explanation was given. This again was done as it was felt to be in line with future use where such tight control over language would not be possible.

7.6.1.1 Analysis of confidence based testing

The relationship between item facility and confidence will be examined to see if there is any sort of correlation**.** Harder questions should make pupils less confident in the right answer (Foster 2015).

Based on the pilot study that had shown that users in low difficulty questions did not show a relationship between stars risked and correctness (Machell and Baruch 2015), it was believed that users in the system had a set confidence level that was then affected by each question. So, for example, a confident user would risk high on easy questions and then still risk high on questions they were unsure about but might reduce this slightly. However, an unconfident user might decide they only want to risk stars on questions they know they will get right.

This proposed hypothesis in behaviour will be tested by looking at the variance in the number of stars risked from the mean of each user. The aim is to predict the overall confidence of the user and then to compare this to the confidence of the user on the given question to measure the impact the question has.

Due to the instant feedback of the system, it was felt that the progression of confidence within users might also have an effect. More complex models such as moving averages were considered but are discussed as further work.

If either of the above methods shows a relationship between confidence and correctness then Cronbach's alpha can then be calculated for the method. This will be compared to Cronbach's alpha for test scores to compare their ability to measure learning.

7.7 Measuring Behavioural Changes

Here the aim is to see how the initial learning experience encourages higher levels of achievement using a Bloom's taxonomy model. To do this, this study will look at the amount of creative work that they do in the system. If this study is trying to promote a behavioural change to deep learning then it will want the students to start engaging with these higher levels.

The mean number of spacebook entries for a student in a group will be calculated by finding the total number of entries for a group and dividing it the number of pupils in the group.

This will then be compared to a predicted model using normal distribution and the goodness to fit calculated using a chi-squared calculation.

7.8 Analysis Against Known Characteristics of Subject Clusters

In order to further examine the effects of context, this study will look at factors that could affect the performance of groups

- Group size
- School performance
    - Science
        - Percentage level 4 or above in school
    - Math and English
        - Percentage achieving level 4 or above in school
- Social/economic factors
    - Percentage of disadvantaged students

Group size will be taken from the number of active accounts in the cluster. School performance and social/economic factors will be taken from the national statics for the school in the years that the study took place.

Although in some cases these will not be the results of the students that have undertaken the study they are a key league table statistic for school and therefore representative of the school as a whole and so will be used as a guide to the context for the school's performance in that year.

7.9 Analysis of Correlations Between Evaluation Tools

Each of the following relationships will be examined by plotting them against each other and calculating the R^2 value to see if there is a relationship

- Enjoyment with learning
- Learning with behavioural changes
- Behavioural changes with enjoyment

as suggested in the Kirkpatrick model for evaluation (Kirkpatrick & Kirkpatrick 2008).

7.10 Comparison of Different Approaches

Two different approaches to the curriculum will be tested and compared using a T-test of the clustered results to show the significance. However, due to the large sample size, this thesis will also look at the effect size as this will allow comparison to other interventions (Coe 2002).

$$Effect\ size = \frac{(Mean_{intervention} - Mean_{control})}{\sqrt{\frac{\text{Standard Deviation}^2_{intervention} + \text{Standard Deviation}^2_{control}}{2}}}$$

Chapter 8. Results

The following results were obtained in 2010-13 during which period the telescope system was used with schools with two approaches to teaching the phases of the Moon. These are referred to as route 65 and route 84 in the literature below. Route 65 took a "pure science" approach and route 84 used a "real world" context of planning a Moon landing, these are described in more detail in section 7.3.2.

For route 65, 78 clusters of school children in Key stage 2 (9 to 11 years of age) used the telescope system. The average cluster size 25.6, with a total of 1997 students.

For route 84, 168 clusters of school children in Key Stage 2 (9 to 11 years of age) used the telescope system. The average cluster size 25.9, with a total of 4351 students.

Confidence testing was carried out on the third set of data with 29 clusters of school children key Stage 2 (9 to 11 years of age). Average cluster size 24.8, with a total of 719 students

8.1 Modelling Against Predicted Distributions

Route 65 was used as the control data set by which to evaluate the effectiveness of the system. In the following section, data is modelled against normal distributions in order to evaluate its suitability for statistical analysis.

Comparisons were made between known factors that might have been expected to have an impact on the data to see if any of these factors had an impact on the data collected.

1. Number of pupils in the cluster
2. Achievement in the school via
    a. Percentage of students achieving 4+ in English
    b. Percentage of students achieving 4+ in science
3. Percentage of disadvantaged students in the school.

For comparison factors 2 and 3 the number of clusters = 69, average cluster size = 26.4 total number of students = 1768. Due to unavailable data 7 potential clusters could not be part of the comparison to known factors. All statistics were compiled from national league tables on the schools that took part (DfES 2015) which are national indicators of school performance.

8.2 Satisfaction

Satisfaction was measured by the students' reuse of the learningscope website material. Using access logs of student ID, page ID loaded and time of access. Data is looked at per cluster with the mean calculated for each cluster. Three different methods are examined:

- Method 1. Mean number of visits per user
- Method 2. Total time use
- Method 3. The frequency of use in the first 30 days

Full details of the different methods are available in section 7.5.

8.2.1 Satisfaction Method 1: Number of Visits

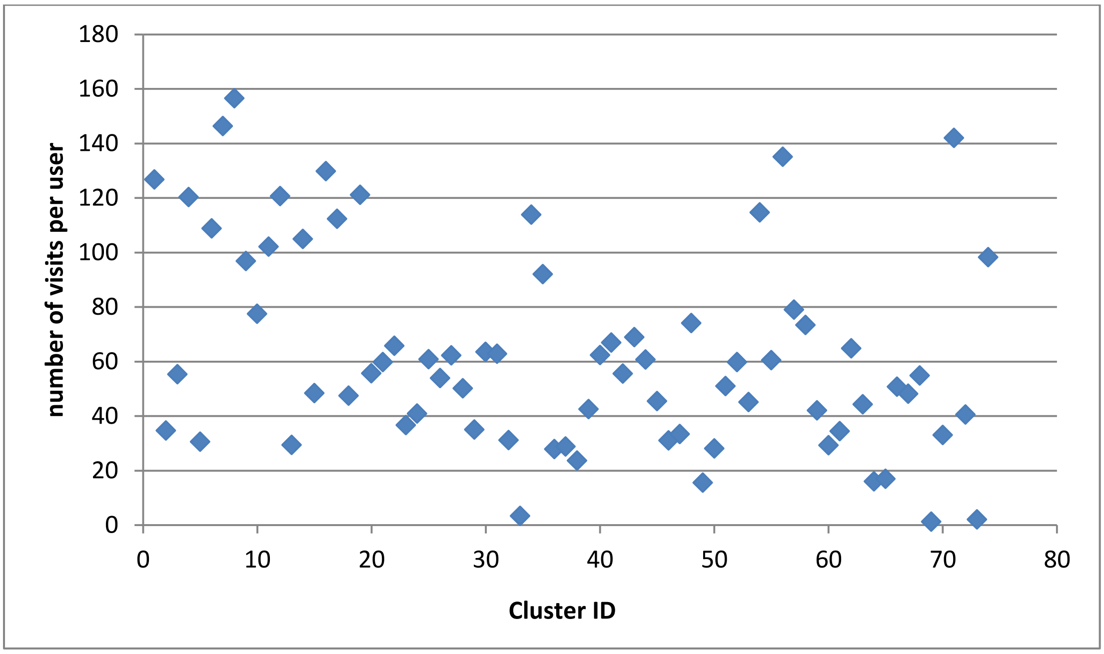

Figure 8.1 Plot of the mean number of visits per user grouped by clusters

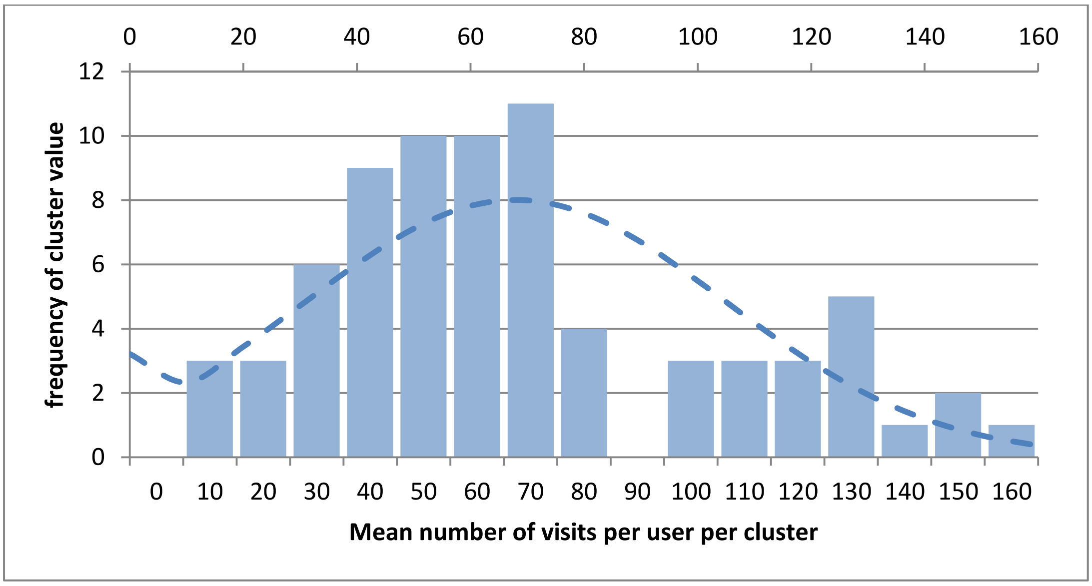

Figure 8.2 Histogram of the frequency of the mean number of logins per user grouped by cluster

Visual inspection of this Figure 8.2 shows a less consistent spread of data, with almost a bimodal spread of data around 50 and 130.

A goodness to fit Chi-squared calculation showed a significant statistical difference from the predicted normal distribution ($P < 0.05$).

8.2.2 Satisfaction Method 2: Period of Use

For each cluster the mean number of days use per member was calculated by working out the period of time each user had used the system by taking the time of the first visit to the learning scope educational pages from the time of the last visit of a unique user ID. These were summed for the cluster and then divided by the number of active accounts with at least one successful login to the telescope site in a group.

$$mean\ number\ of\ days\ use\ for\ cluster = \frac{\sum x_{time\ of\ last\ use} - x_{time\ of\ first\ use}}{number\ of\ active\ accounts\ in\ group}$$

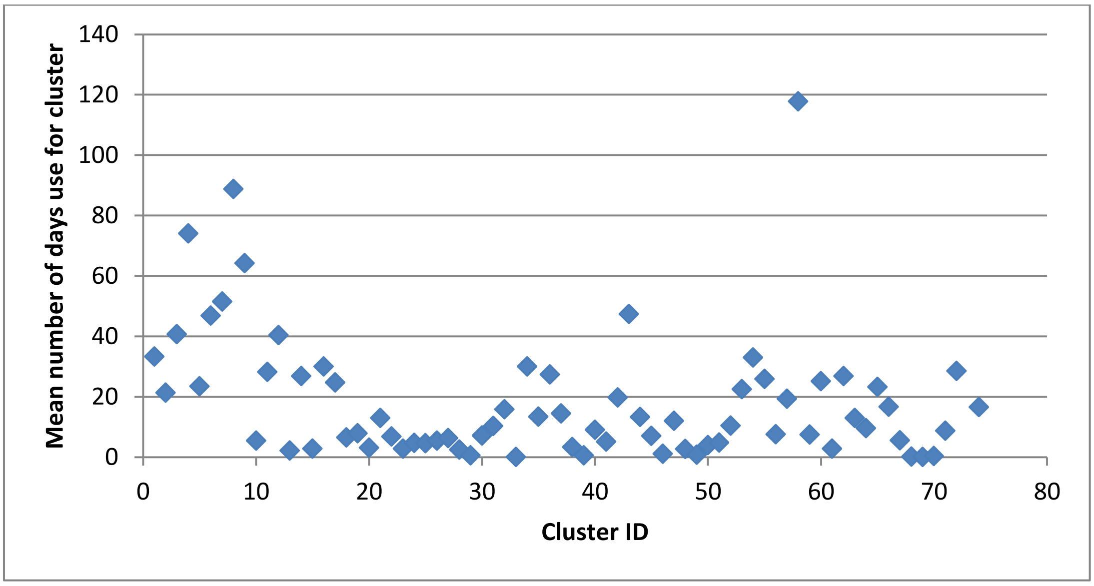

Figure 8.3 Mean number of days use for clusters

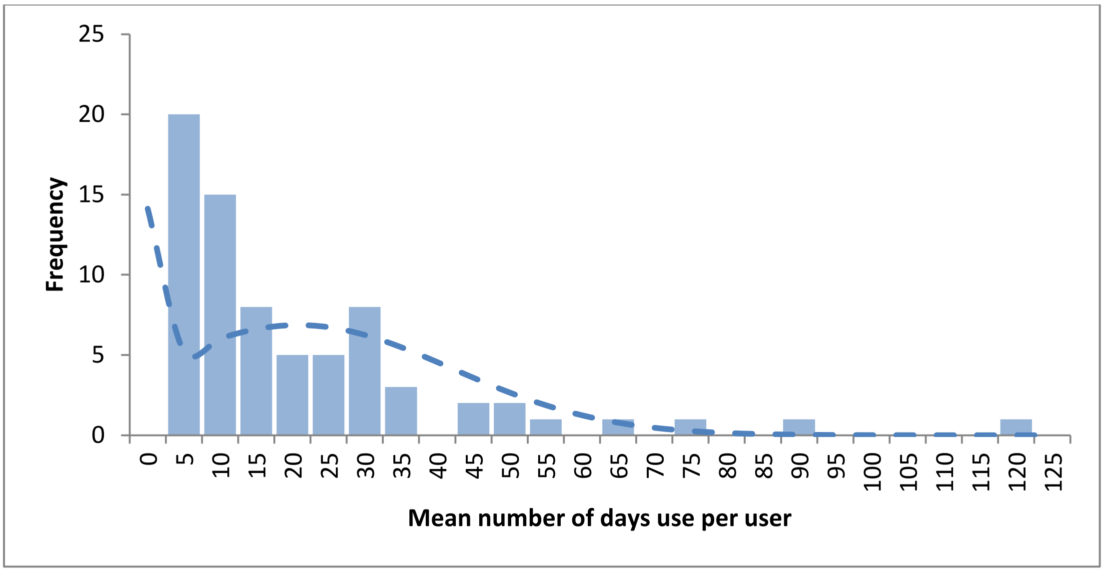

Figure 8.4 Histogram of the mean number of days use per user in a group

Mean = 18.3, Standard Deviation = 21.2. In Figure 8.4 a clear drop of use from 5 days day is visible, most use of the site has finished by the 35 days of the website. However, 5 groups used the website for more than 50 days with a maximum period of use of 120 days for one cluster's mean period of user per user. This is not long enough to have been caused by reuse of accounts after a year. This shows a large difference from a normal distribution around the mean period with the modal period being just 2 days. Comparison to normal distribution shows a large variation. This is due nature of the data following a decay pattern, rather than a normal distribution about the mean.

8.2.3 Satisfaction Method 3: Frequency of Use

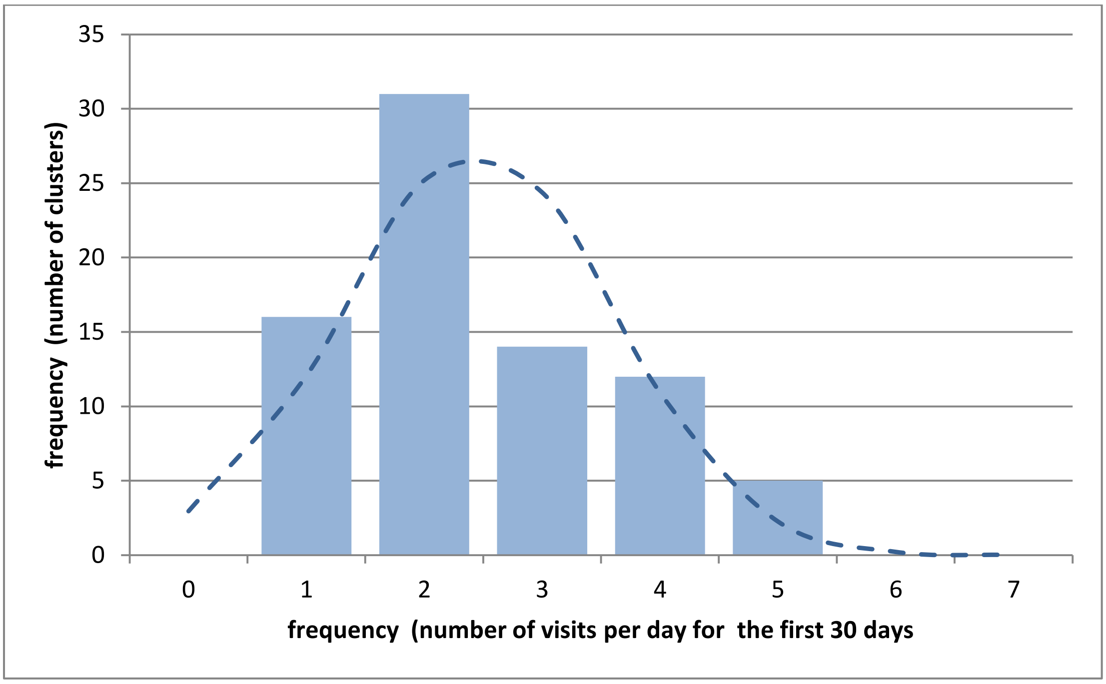

Figure 8.5 Mean frequency of visits per person for clusters

Mean = 2.0. A visual inspection of Figure 8.1Figure 8.5 against the predicted shows that the results are skewed to the lower side of the range. A chi-squared goodness to fit test shows this a borderline significant variance (P = 0.058).

8.2.4 Analysis of Known Factors for satisfaction measurements

8.2.4.1 Cluster size

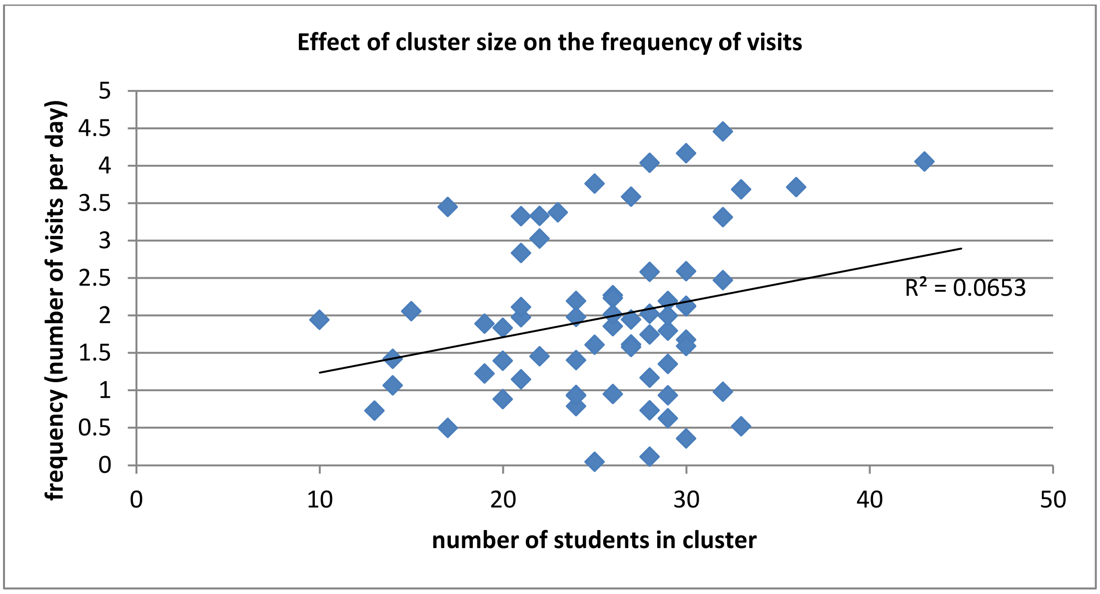

Figure 8.6 Effect of cluster size on the frequency of visits

Figure 8.6 shows there is no strong correlation between the number in the group and the frequency of use. Groups ranged from 10 to 43 in size. There is a slight upward trend in the regression however this is probably caused by the low-frequency items of high leverage.

8.2.4.2 Maths and English Attainment in School

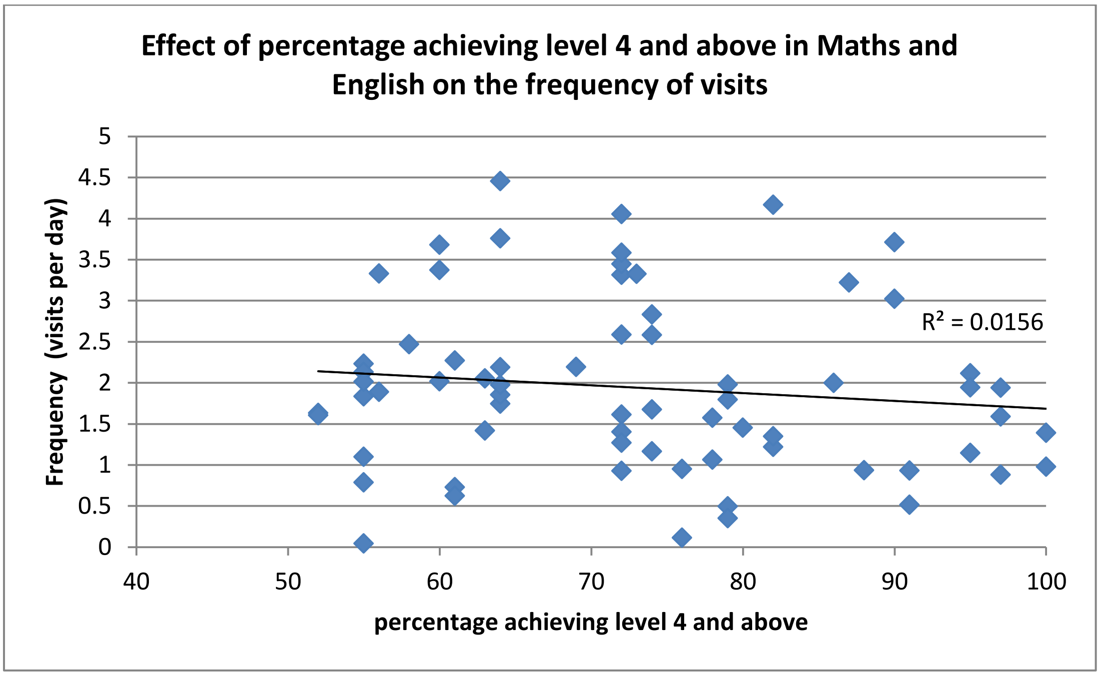

Figure 8.7 Effect of percentage of students achieving level 4 and above in maths and English on the frequency of visits

Figure 8.7 shows there is no strong correlation between the percentage of students achieving level 4 and above in Maths and English on the frequency of visits.

8.2.4.3 Science attainment in school

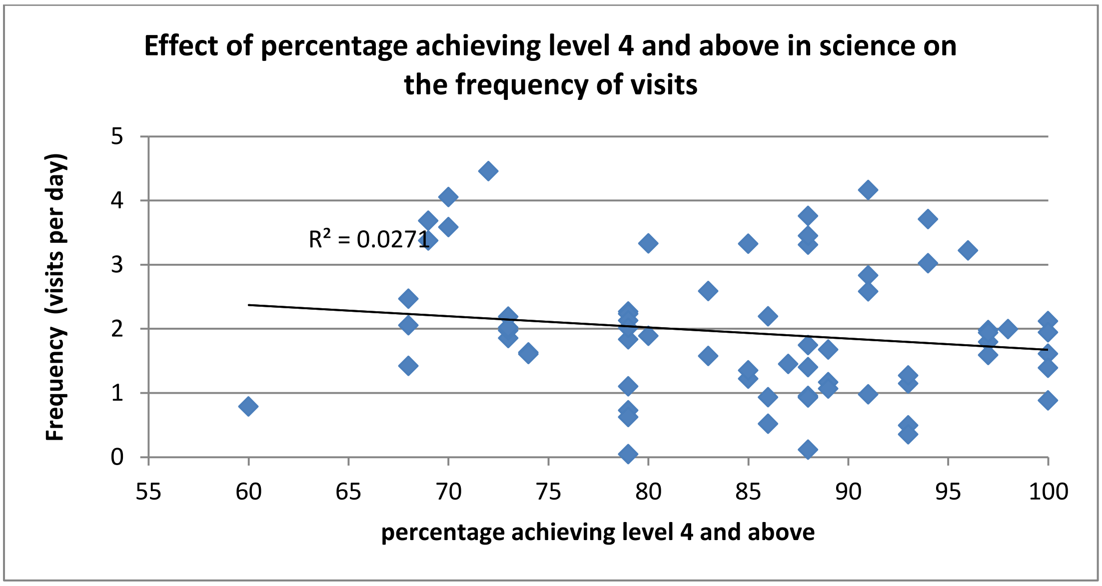

Figure 8.8 Effect of percentage of students achieving level 4 and above in Science on the frequency of visits

Figure 8.8 shows that there is no strong correlation between the percentage of students achieving level 4 and above in the school in Science and the frequency of visits.

8.2.4.4 Percentage of disadvantaged students in school

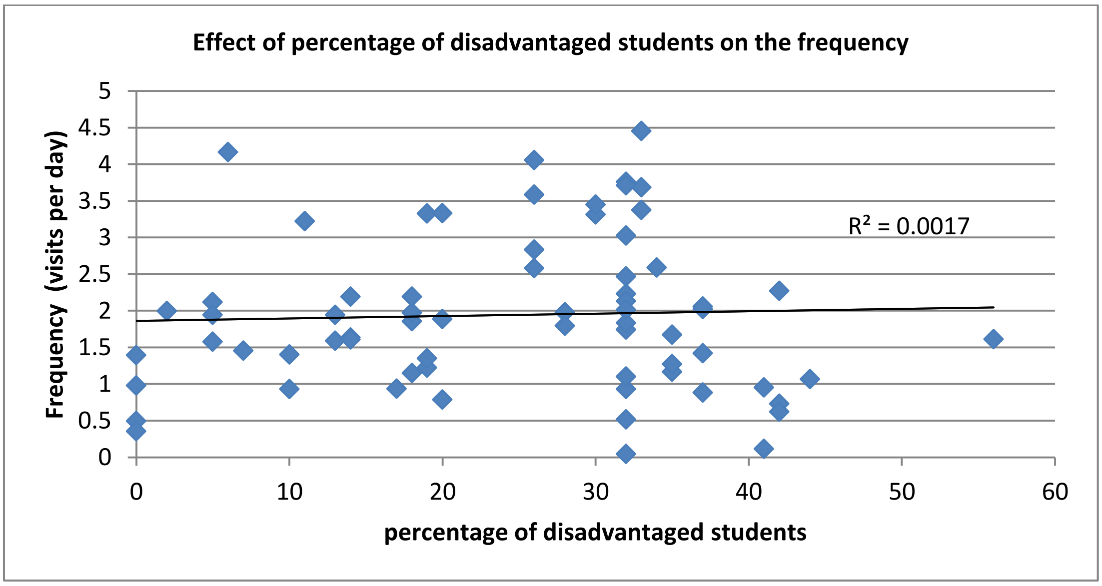

Figure 8.9: Effects of percentage disadvantaged pupils on the frequency of visits

Figure 8.9 shows that there is no strong correlation between the percentage of disadvantaged students in the school and the frequency with which students visit the site. Four of the clusters from schools with the lowest percentage of disadvantaged students can be seen to perform comparably to some of the higher percentages.

8.3 Analysis of Learning Measurement for Clusters

Percentage of correct answers for all questions on a route was worked as detailed in 7.6

Shown below are the raw results and the results after filtering was applied as outlined in 7.6

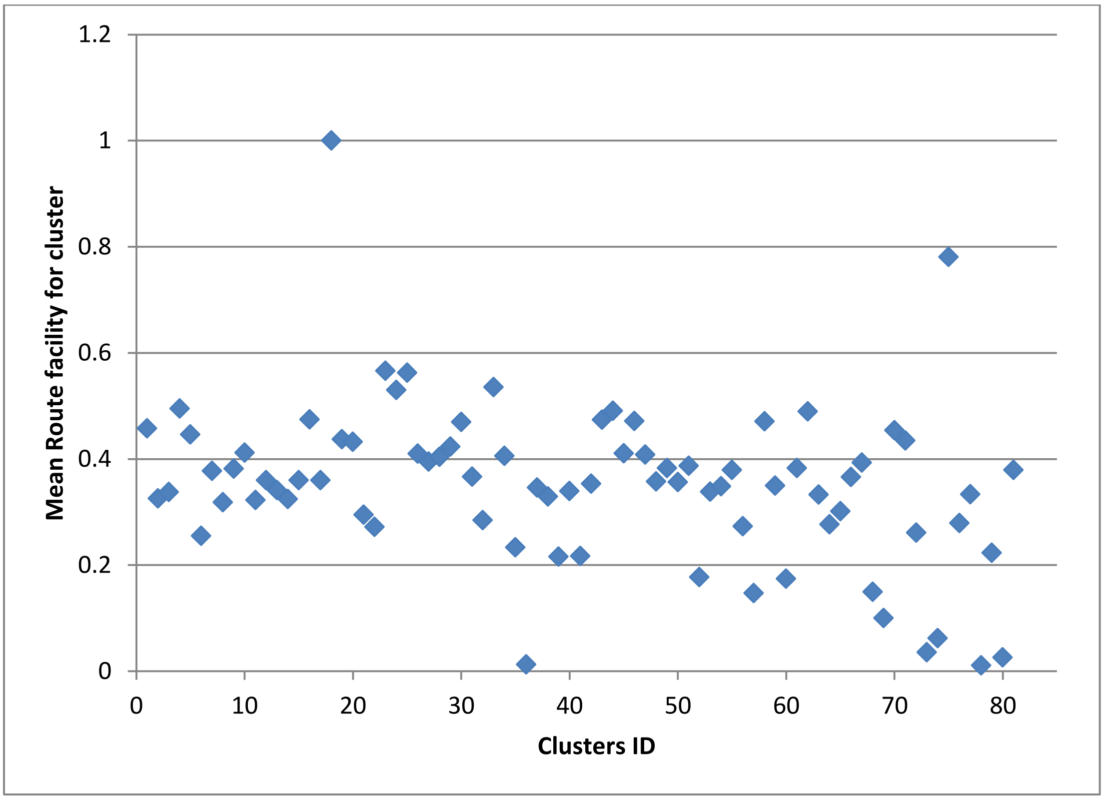

Figure 8.10 Raw data for clusters mean route facility

In Figure 8.10 Raw data all clusters are shown, before any cleaning of data based incomplete data collection from either cluster size or percentage of questions attempted by the user.

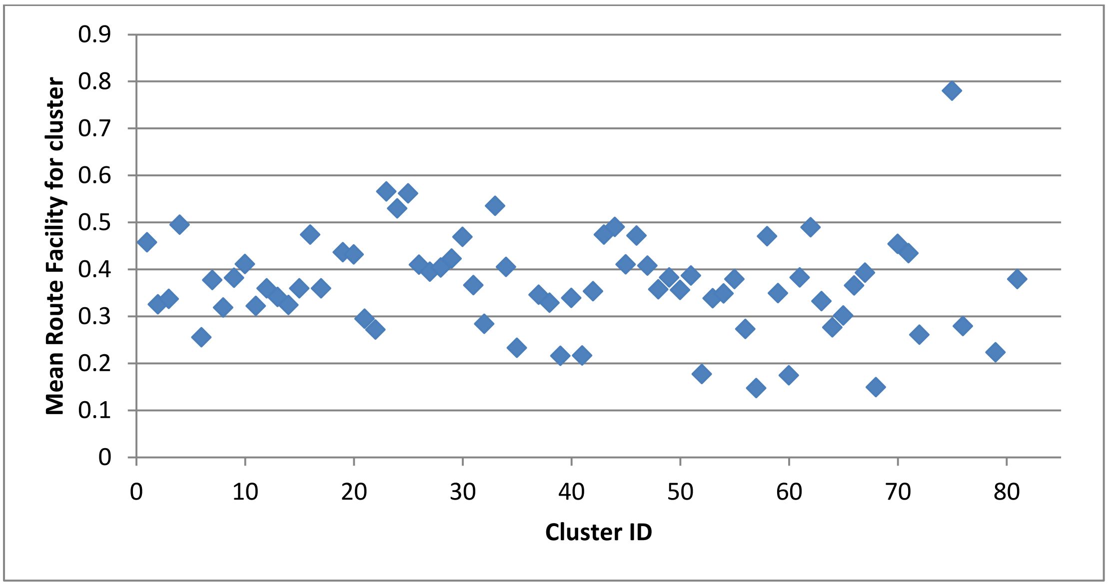

Figure 8.11 Results for cluster mean route facility.

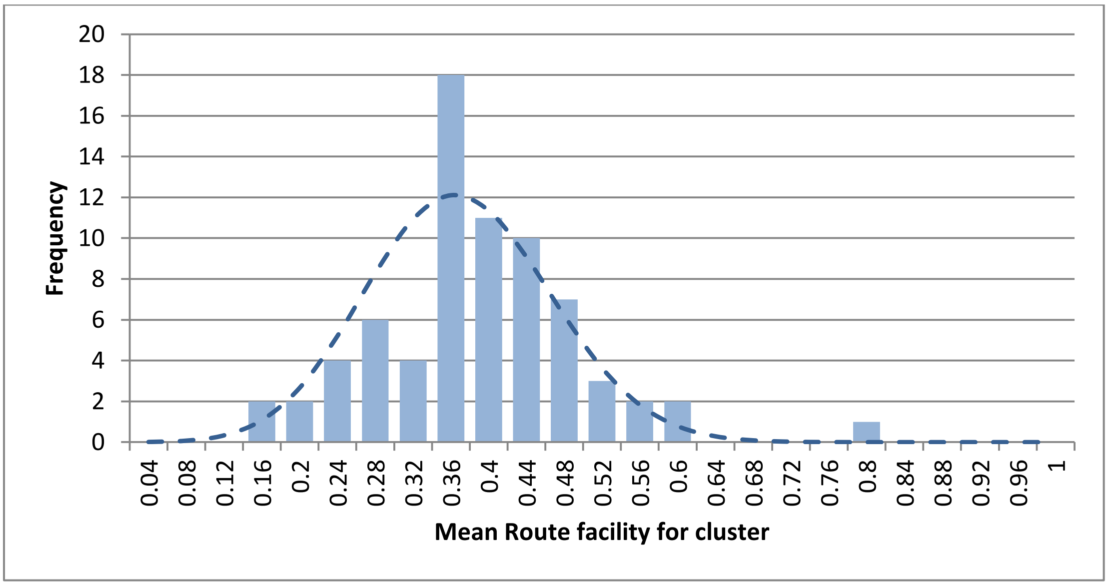

Figure 8.12 Histogram frequency of groups achieving a given route facility

Data was cleaned of incomplete data sets (Figure 8.11). Clusters that answer less than 9 questions out of those found on the route (questions n =

19), were removed as they were deemed to have complete an insufficient proportion of the test. They had not reached the halfway point. This removed 6 groups from the results. On average the percentage of questions answered per cluster was 18.6 out of 19. Two further groups were removed due to less than half the group attempting any of the questions on the route. Both of these scenarios were ruled as incomplete as too high a proportion >50% would have to be assumed wrong through non-response.

Predicted results were plotted using a binomial distribution with a probability set to the mean value of 0.369 and a group size of 25.

Visual inspection of Figure 8.12 shows a good level of fit. More pupils achieve the mean score than predicted n= 18 actual against a predicted level of n =12. One group achieve 78%, this was investigated due to its high level, but the group was found to be of standard size and a high percentage of all questions had been answered by all its users.

Comparison by chi-squared goodness to fit test showed P = 0.37. As such, there is a good level of fit to the predicted.

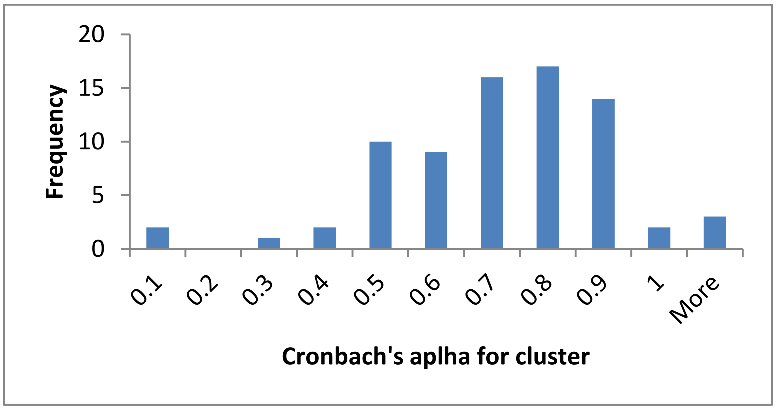

Figure 8.13 Cronbach's alpha for learning measurement

Figure 8.13 shows the Cronbach's alpha for each cluster, the vast majority are of an acceptable level of 0.7 – 0.9, however, scores of greater than 0.9 meaning that there might be redundancy in the assessment and lower the 0.7 meaning that it might not be a good tool for measurement are also achieved (Tavakol & Dennick, 2011). This large range of results means that the result is inconclusive overall.

8.3.1 Analysis of Known Factors for Learning Measurements

8.3.1.1 Cluster size

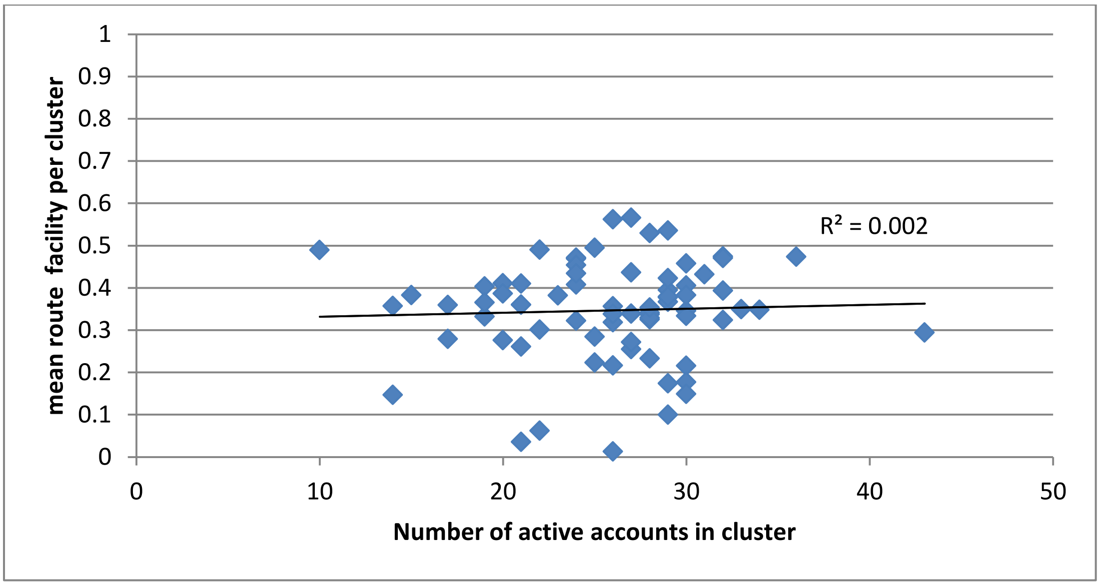

Figure 8.14 Route facility against the number in a cluster

Figure 8.14 shows no strong correlation between the number of students in a cluster and the percentage of right answers on average for that cluster (R^2= 0.002).

8.3.1.2 Maths and English achievement in school

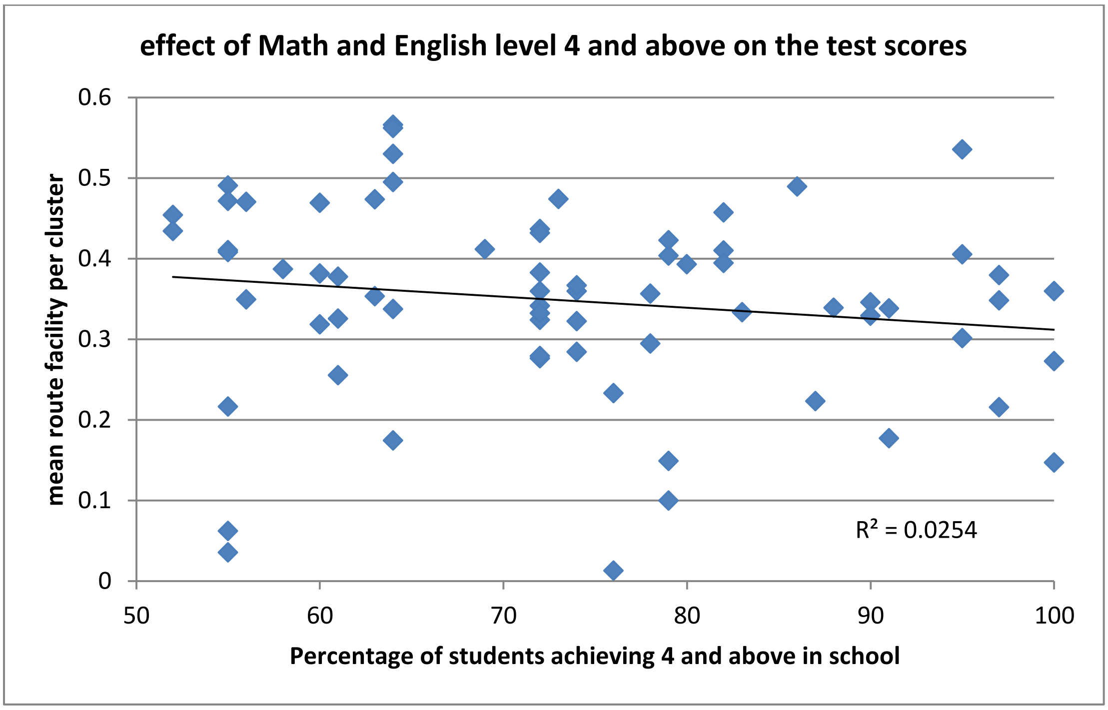

Figure 8.15 Effect of Math and English level 4 and above on the route facility

Figure 8.15 shows no strong correlation between the percentage of pupils achieving level 4 and above in the school in English and Maths and the percentage of correct answers for the average pupil in that cluster (R^2 = 0.03).

8.3.1.3 Science achievement in school

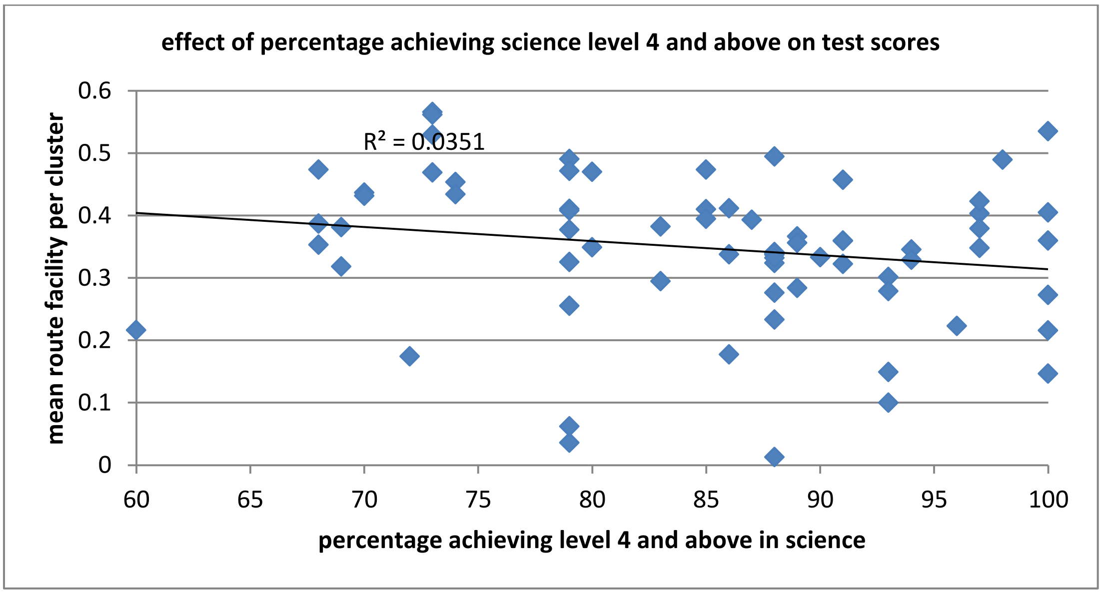

Figure 8.16 Effect of percentage achieving science level 4 and above on the test scores

Figure 8.16 shows that there is no strong correlation between the percentage achieving level 5 or above in science and the test scores of the students in the system ($R^2 = 0.04$).

8.3.1.4 Percentage of disadvantaged students in school

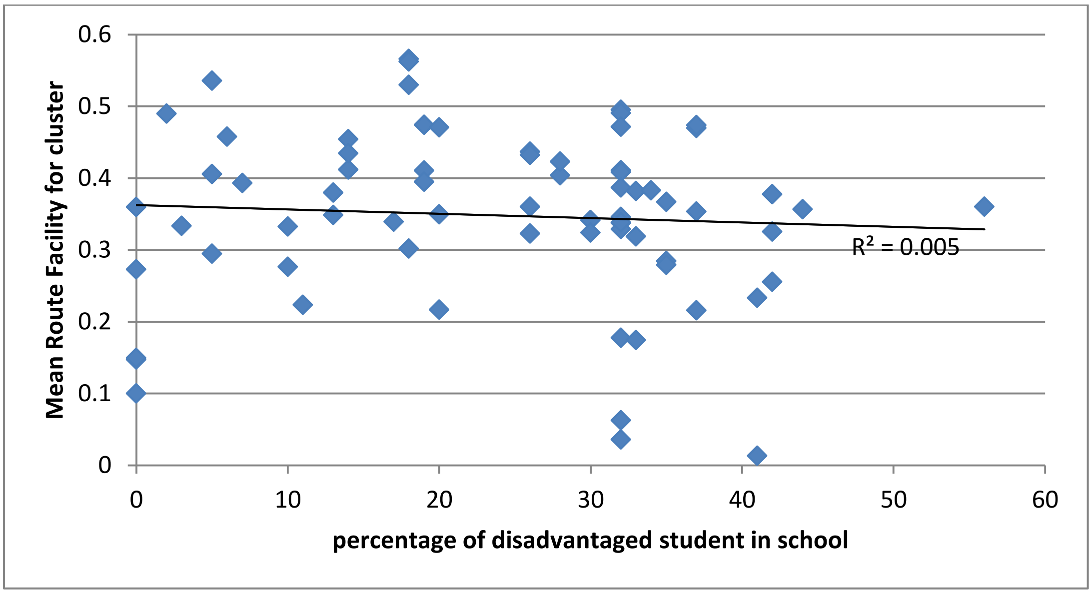

Figure 8.17 Effect of percentage disadvantaged students in school on route facility

Figure 8.17 shows that there is not a strong correlation between the percentage of disadvantaged students in the school and the test scores of individuals using the system ($R^2 = 0.005$).

8.3.2 Confidence-Based Testing

Students were asked to risk between 0-20 stars on getting a question right. If answered correctly they won the stars, whereas an incorrect response lost them the same number of stars. Results here examine if there is any sort of useful relationship that could be used as a measure for learning.

Figure 8.18shows the effect of item facility per questions on confidence, there is no strong correlation between item facility (number of users getting the question right) and confidence. As such it is not possible to say that as confidence increases so does the chance of getting the question right (R^2 = 0.02).

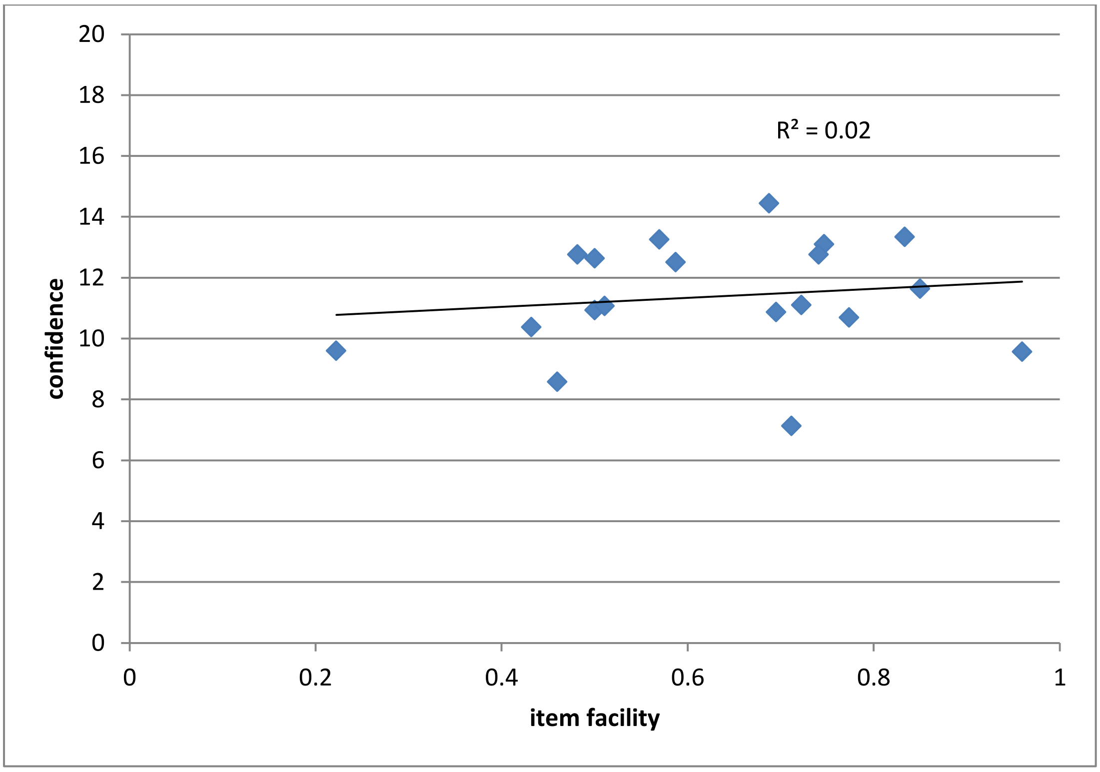

Figure 8.18 Effect of item facility on confidence grouped per question

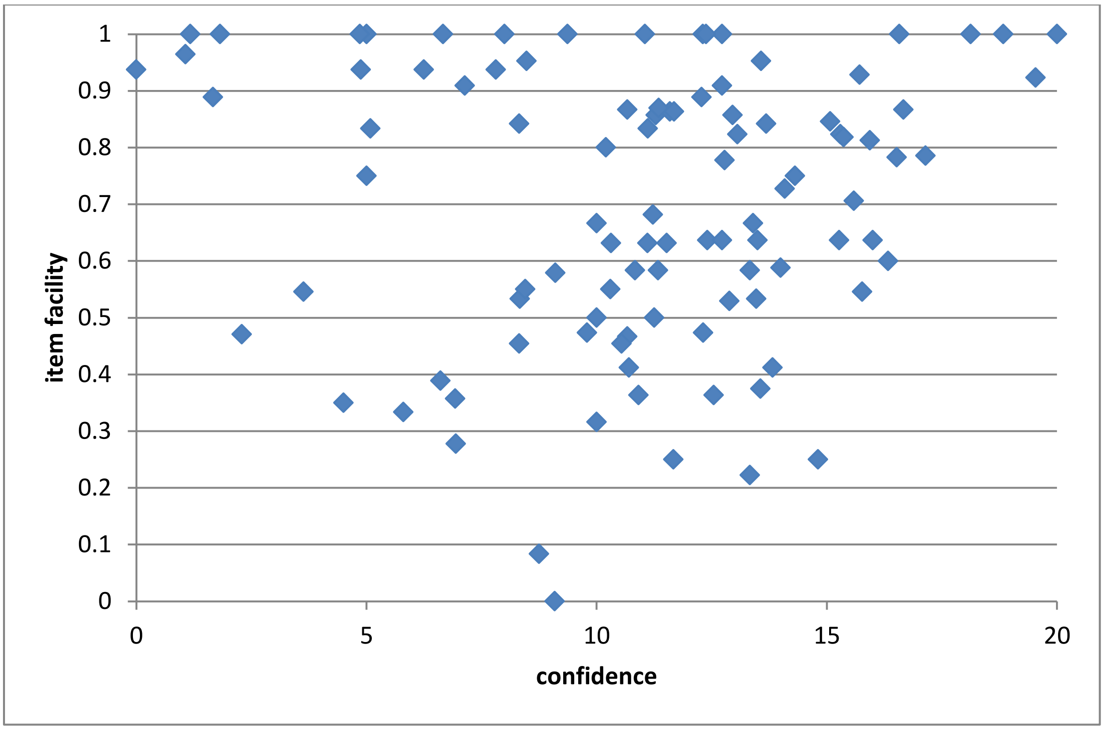

Figure 8.19 Effect of confidence on item facility

Clustered results for each question per group show a similar loose association (Figure 8.19), however, this graph more clearly shows that when a group finds a question easy this relationship breaks down, as a full range of confidences are found for questions with an item facility of greater than 0.9.

8.3.2.1 Confidence Method 2:

The mean confidence level was calculated for each student and the difference between this and the question is used to give a measure of confidence for each student.

Despite this adjustment, the data in Figure 8.20 of the residual confidence against the percentage of correct answers for a group show no strong correlation.

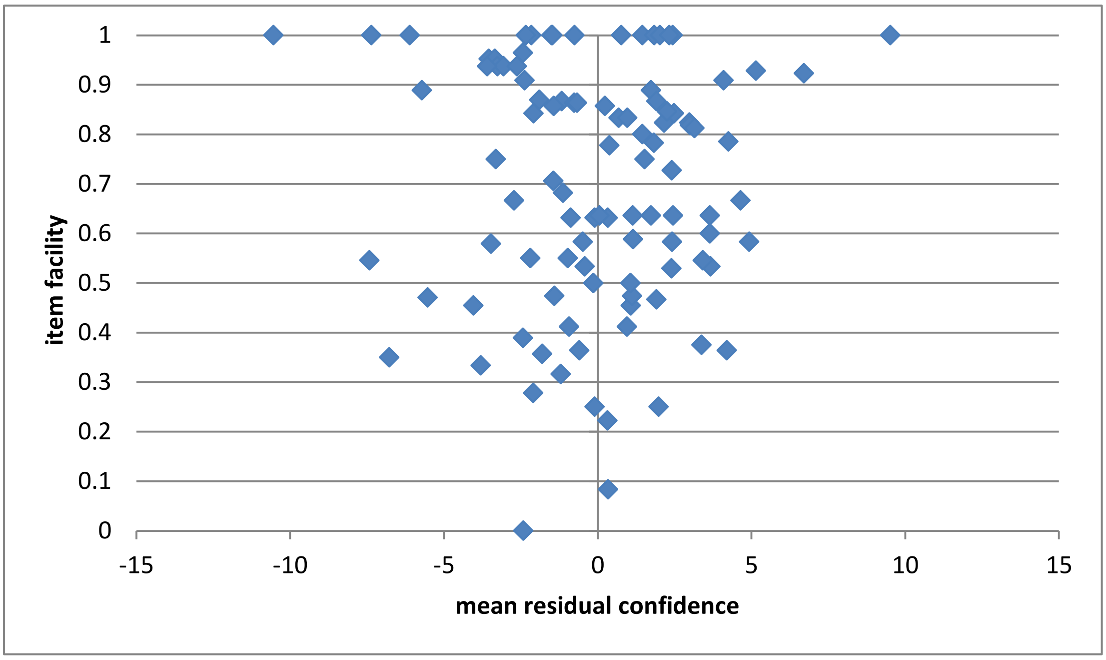

Figure 8.20 Effect of mean residual confidence on item facility

Due to the lack of any correlation found by either of the above methods, a comparison of the Cronbach's alpha was not carried out to see if there was any improvement on the route facility method for learning. A full discussion of this can be found in the discussion in chapter 9.

8.4 For Behaviour

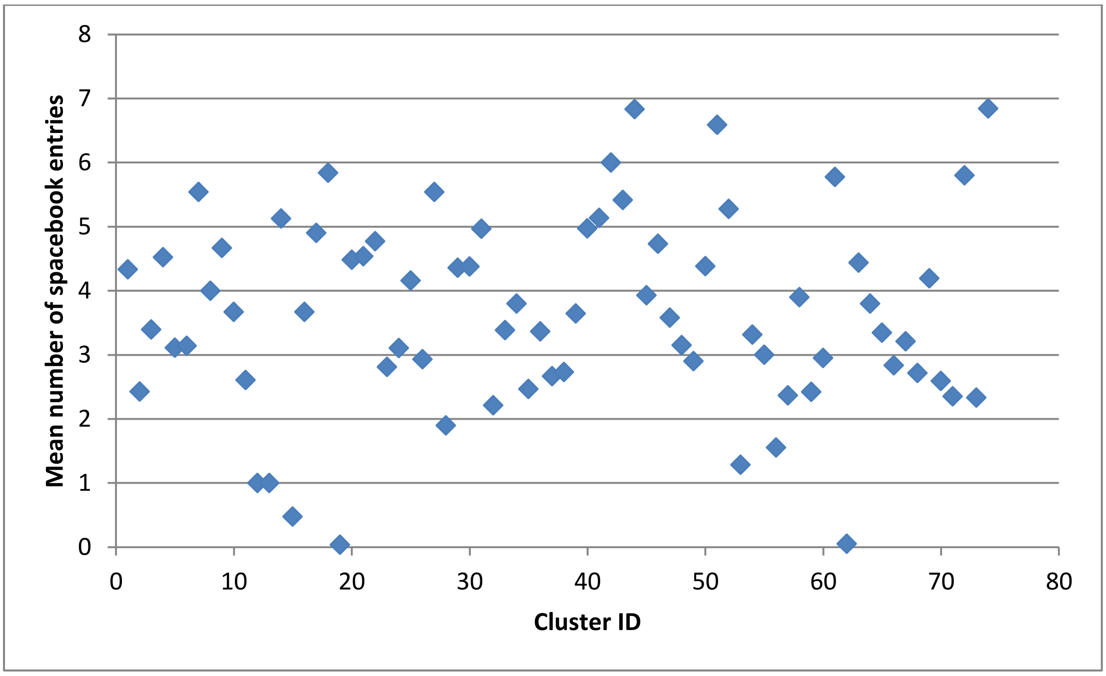

Figure 8.21 Mean Number of Spacebook Entries for Class Clusters

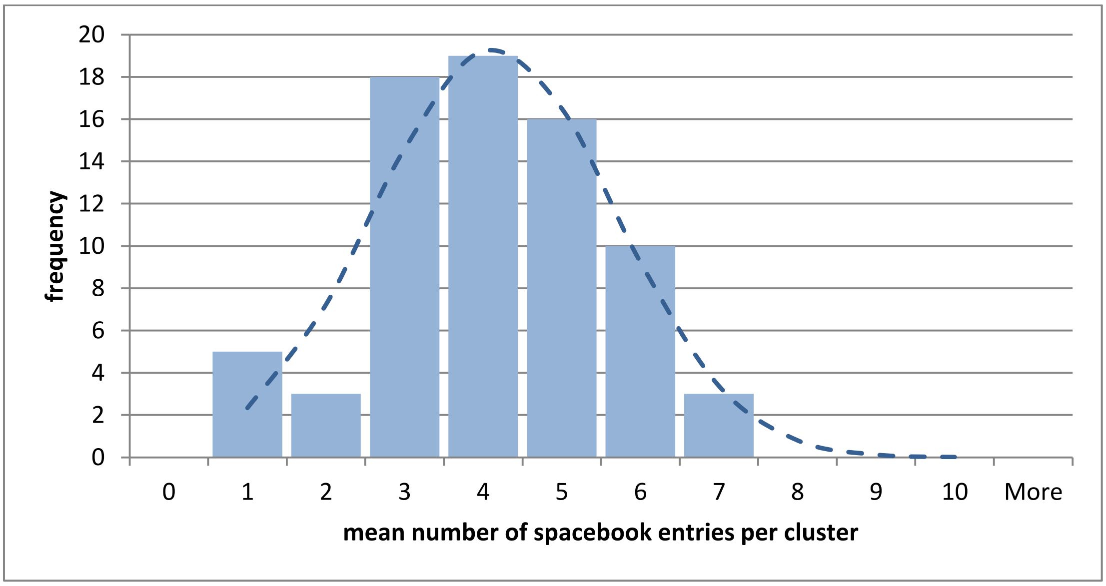

Figure 8.22 Frequency of clustered mean space book entries per pupils

The mean number of clustered spacebook entries per user = 3.6. Visual inspection of the relationship against the predicted model shows a high level of fit (Figure 8.22). The most frequent number of entries is found at 4 entries per user. With a full range of just one entry in a class of 27 users to 6.84

entries per user. There is a slightly higher than expected number of classes that had an average of just one entry per page, a slightly lower number that had 2 and again a slightly higher than predicted number for 3.

A goodness to fit chi-squared calculation against a normal distribution shows a probability of P>0.20 (P = 0.24). As such this is a good model for comparison as the behaviour of students is consistent and can be modelled using normal distributions.

8.4.1 Analysis of known factors for behavioural Measurements

8.4.1.1 Cluster size

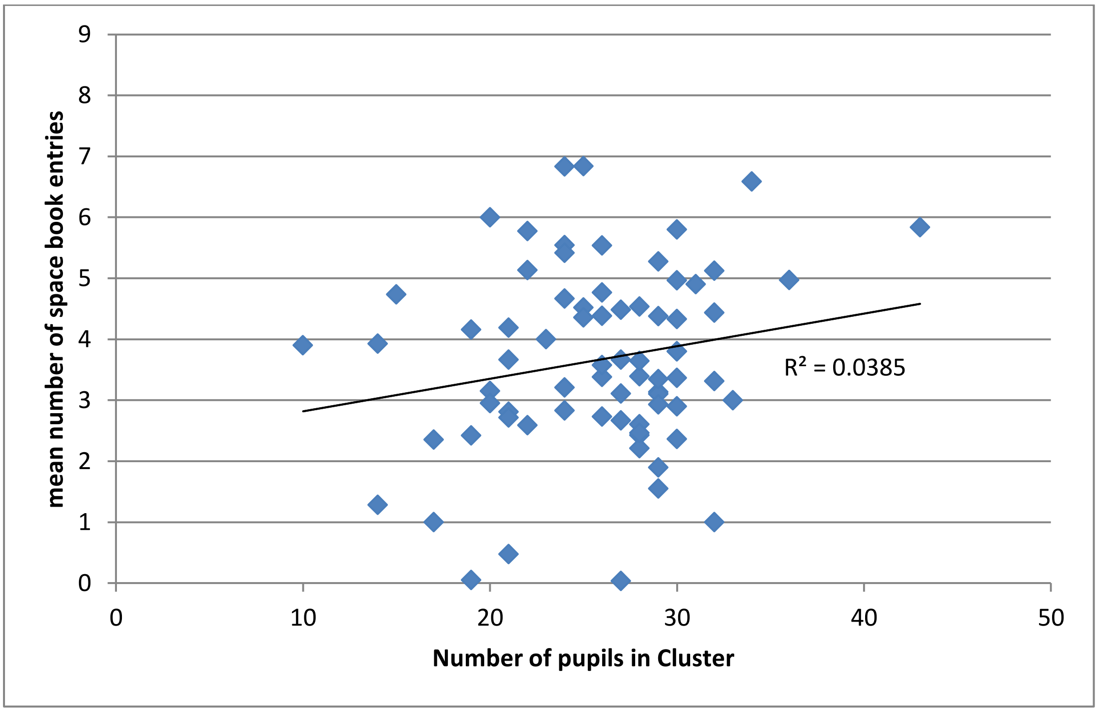

Figure 8.23 Number of pupils in a cluster against the number of space book entries

There is little correlation between the number of students in a group and the mean number of space book entries for that group (Figure 8.23) (R^2 = 0.04).

8.4.1.2 Maths and English achievement in school

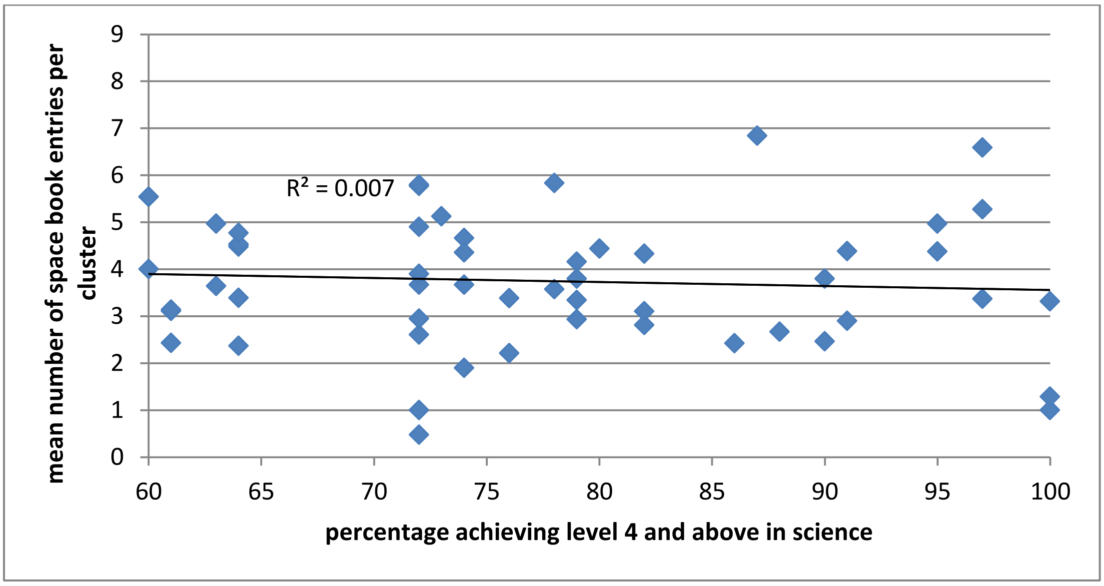

Figure 8.24 Effect of percentage achieving Maths and English level 4 and above on behavioural changes

8.4.1.3 Science achievement in school

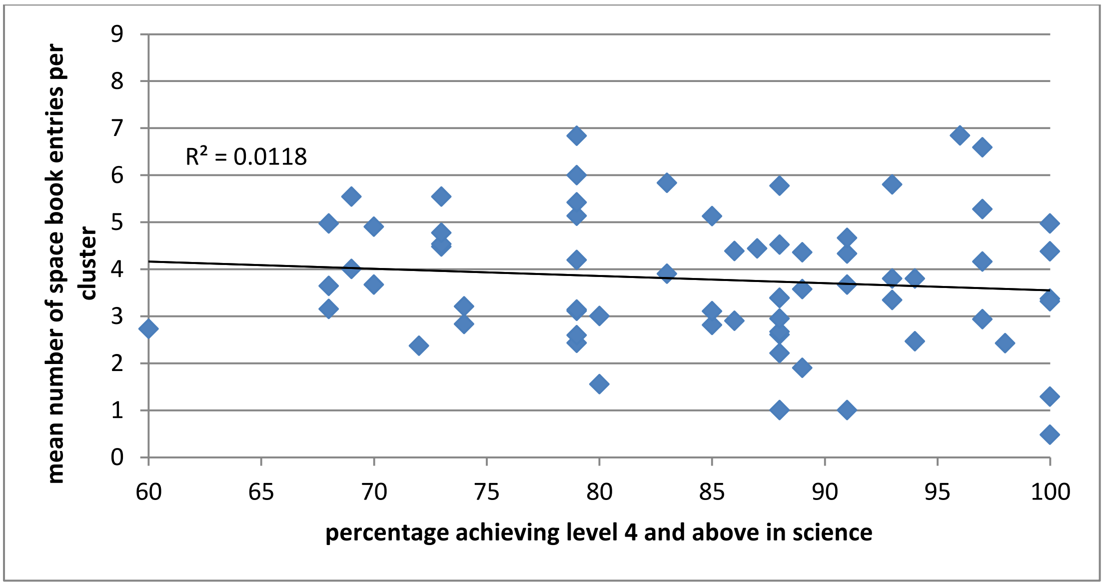

Figure 8.25 Effect of percentage achieving science level 4 and above on behavioural changes

Figure 8.25 shows that no strong correlation between the percentage achieving level 4 and above in science in the school and the mean number space book pages per cluster which is the measure for behavioural change (R^2= 0.01).

8.4.1.4 Percentage of disadvantaged students in school

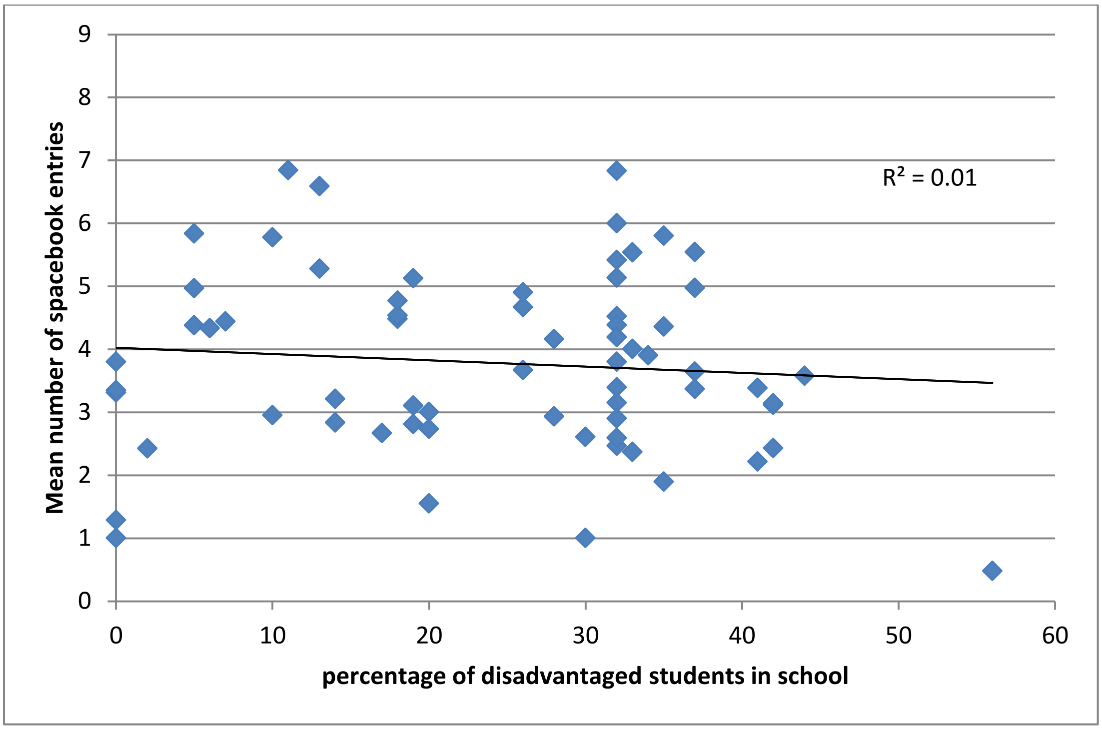

Figure 8.26 Effect of Percentage of Disadvantage Students on Behavioural Measures

Figure 8.26 shows that there is not a strong correlation between the percentage of disadvantaged students at a school and the mean number of space book entries they make (R^2= 0.01).

8.5 Summary of Validation of Measures for Evaluation

Satisfaction

Method 3: frequency showed the best level of fit to a normal distribution of results (Figure 8.5), although slightly skewed to it was not significant (p> 0.05) and a mean value of 2.0. Method 1: number of visits showed an almost bimodal split (Figure 8.2) and method 2: period saw a large difference between the mean and the modal values (Figure 8.4).

Learning

The mean test results for the clusters showed a good fit to a binomial distribution (P> 0.37) and a mean value of 0.4 (Figure 8.12). Cronbach's alpha values (Figure 8.13) were inconclusive with a range of values above 0.5, with a modal value of 0.8

Confidence based testing was inconclusive with no clear correlation for either method with the number of correct answers for questions.

Behaviour

Showed a good level of fit to the normal distribution (P> 0.2) and a mean value of 3.6 (Figure 8.22)

Effect of known factors

All of the measures showed no strong correlations with known factors

8.6 Comparison of Approaches

Comparisons of the success of learning were made on the two approaches to learning route 65 a pure scientific approach which encouraged students to take images of the moon to observe the natural phenomenon of its phases and route 84 an applied approach that saw pupils planning a trip to the moon and had to make sure that they would land in the correct phase.

8.6.1 Satisfaction Test Comparison

Of the 3 methods outlined for satisfaction in section 7.5, the third method was used for comparison. This is because it was not skewed by single user high use in the same way the other two methods were. Method 1 which used the number of visits could be affected by a single user in the cluster using the site a lot. It also did not show a good level of fit to the predicted models. Method 2 which used period of use could again be distorted by a single user continuing with the site for significantly longer than the other in the group. Method 3 that looked only at the first 30 days of use did not suffer from these distortions in the same way as only visits in the first 30 days counted.

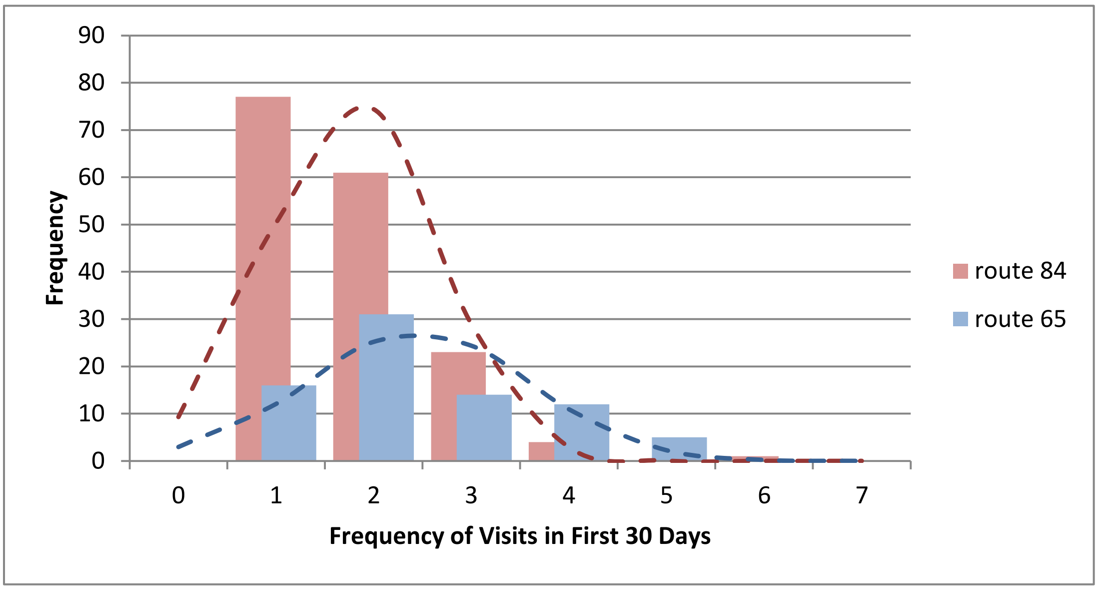

Figure 8**.27 Comparison on satisfaction Method 3 Frequency of use Table 8.1 Comparison of Frequency of Visits in the First 30 Days of Use**

	r65	*r84*
Mean	2.0	1.3
Variance	1.2	0.7
Number of clusters	78	168

Figure **8**.27 shows there was a significant (P<0.01) negative impact on the frequency of visiting the educational website in the first 30 days for route 84.

With an effect size of -0.69. The approach clearly discouraged users from returning to the educational material. This effect is larger as there is skew of results about the mean. However, this does bring into question the goodness to fit of the data for a normal distribution.

8.6.2 Learning Test Comparisons

A reduced selection of questions were used for the comparison as only four questions were asked in both routes and so allowed direct comparison.

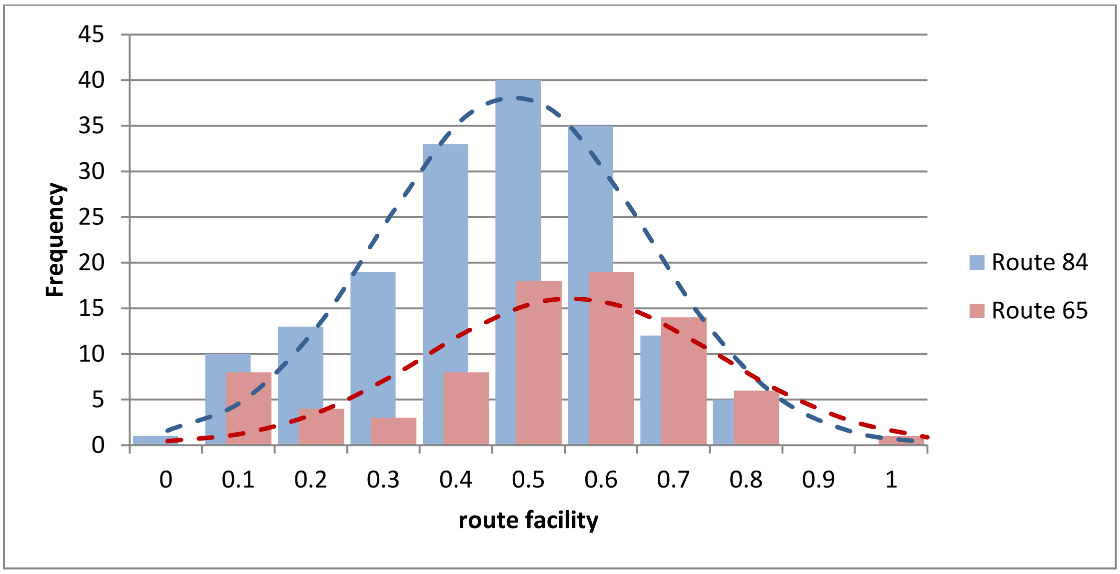

Figure 8.28 Comparison of Learning Measure

Table 8.2 Comparison of Learning Measure

	R65	*R84*
Mean	0.5	0.4
Variance	0.04	0.03
Number of clusters	78	168

Figure 8.28 shows a significant negative impact (P < 0.01) effect size -0.43

8.6.3 Behavioural Test Comparison

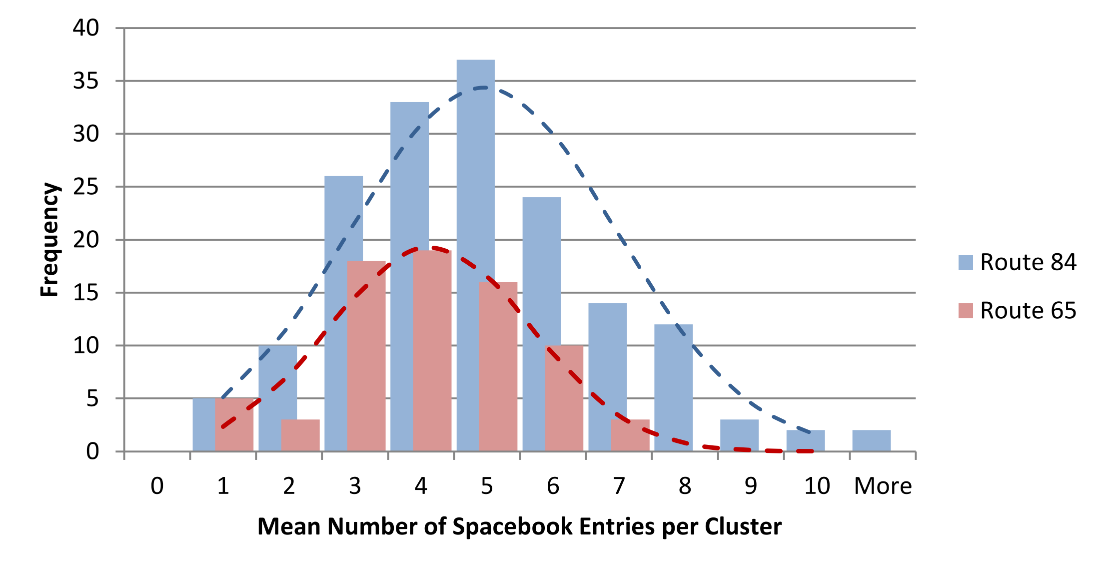

Figure 8.29 Comparison of Behavioural Changes

Table 8.3 Results from Behavioural Comparison

	Mean number of space book entries per person clustered for Route 65	Mean number of space book entries per person clustered for Route 84
Mean	3.6	4.4
Variance	2.26	4.00
Number of clusters	74	168

Figure 8.29 shows there was a significant positive effect on the number of space book entries for clusters. The mean number of entries per person in a group of 4.4 for route 84 compared to 3.6 as the cluster mean for route 65 ($P<0.01$). This related to an effect size of 0.40. So route 84 clearly promoted an approach in which more creative work was carried out by the users of the system.

The reduced number of clusters in this test results comes from the fact that there were 4 groups that did not add a single entry to space book. If these were added to the group it would increase the difference between the two groups by further lowering mean for route 65. However, the reason that there are no entries at all cannot be known and so they have been excluded as incomplete rather than a group that attempted and failed to make any entries.

8.7 Evidence of Relationship: Satisfaction -> Learning -> Behaviour

The following section presents results looking for relationships in the data that follow the relationship put forward in the framework presented in chapters 3-5.

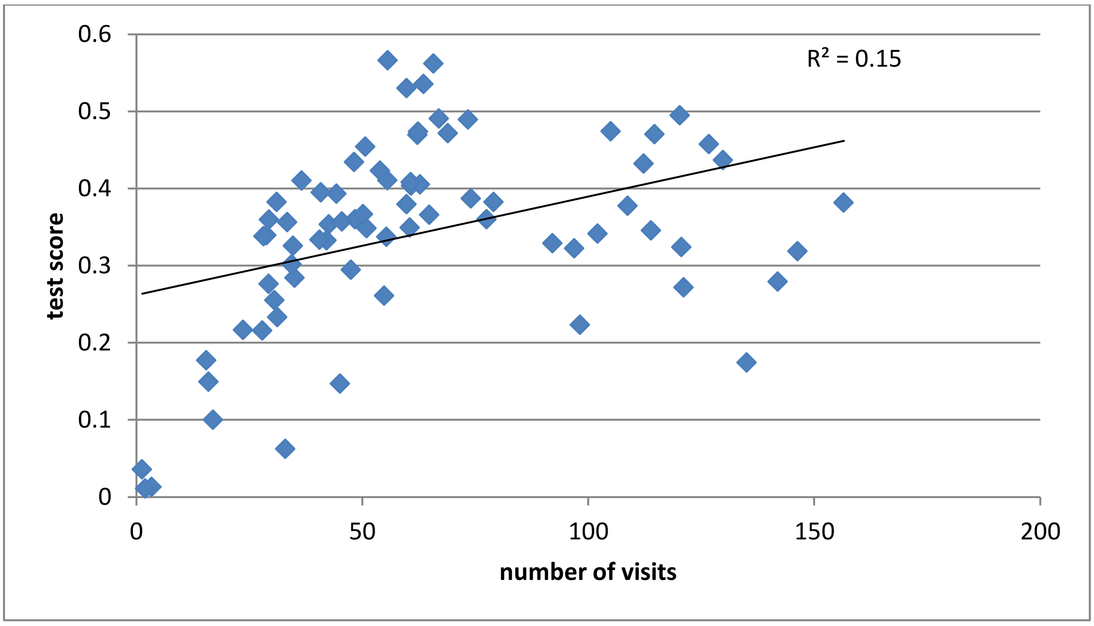

Figure 8.30 Graph of the Correlation of Satisfaction Method 1 and Learning

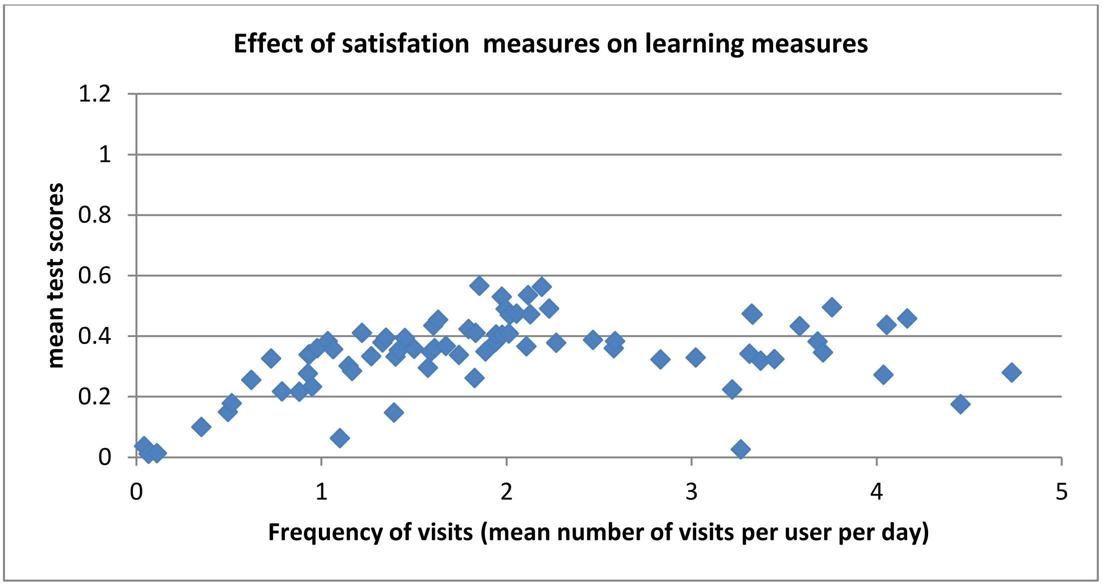

Figure 8.31 Correlation of Satisfaction Method 3 and Learning Measures.

Figure 8.30 shows that from 0-70 average visits there appears to be a connection between the number of visits and the scores for students in the system. However, above 70 visits the correlation is less strong meaning that overall there is a weak association between the two overall. This is in line with the bimodal nature of the result. A similar result can be seen in Figure 8.31 where there is a trailing of the impact of the increase in the frequency of visits or satisfaction on the learning of the students. Below 2.2 visits per day $R^2 = 0.68$ but over the data as a whole this is much lower $R^2 = 0.12$

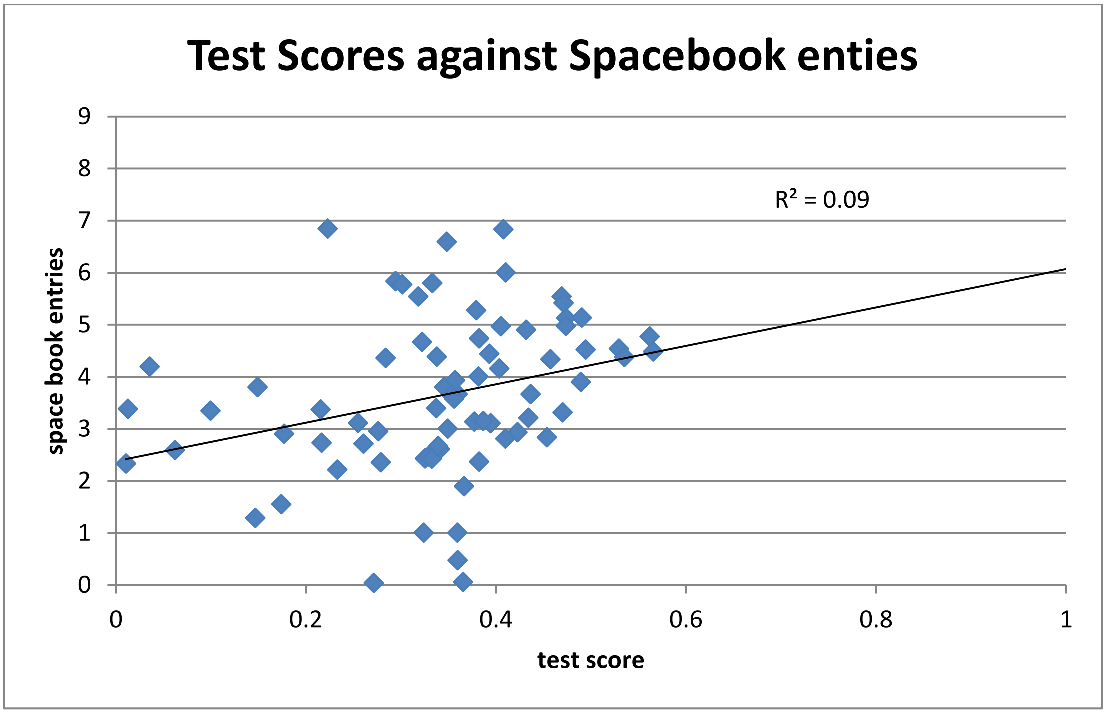

Figure 8.32 Graph of the correlation between learning and behavioural changes

There is no strong association between test scores and students engaging in developing and writing about their images in their space book reflective area (R^2= 0.09).

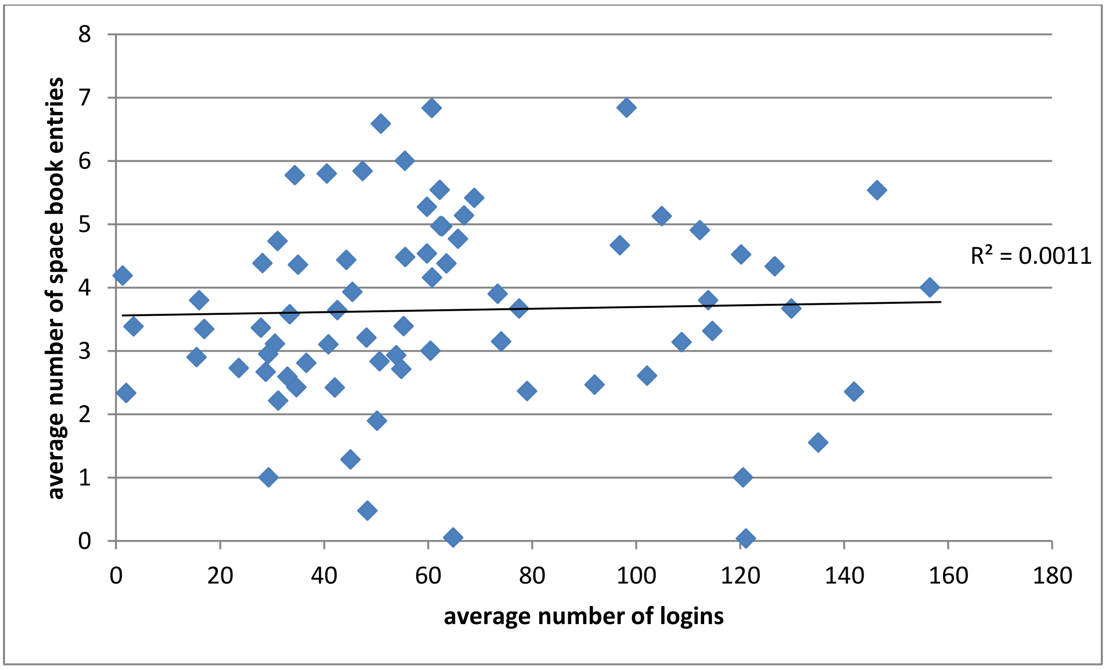

Figure 8.33 Graph of the correlation between satisfaction and behavioural changes

There is almost no correlation ($R^2 = 0.001$) between the number of visits to the learning material and students engaging in higher-order activities such as producing their own space book entries.

Chapter 9. Discussion and Conclusion

There were three aims to this thesis.

- The first was to present a novel framework for the collection and analysis of data in a multi-institutional primary school setting.
- The second was to evaluate the effectiveness of this framework by analysing the sample of data collected. In this aspect, there were three key questions.
    1. was the data collected fit for analysis, did it follow normal distributions?
    2. what was the level of influences of external factors, could they be compensated for through the clustered methods of analysis or where their influences too large.
    3. was the framework for the e-learning environment flexible enough to allow changes within it that would have an impact on learning or did the framework only allow superficial changes with too small an impact to be measurable against the influences of background effects.
- The third aim was to critically evaluate the framework in light of the data it was able to collect.

This section will try to answer these three questions, and discuss future work as a result of findings in these areas. The first aim was answered in chapters

1 – 6 of this thesis. The second aim will be discussed in sections 9.1, 9.2 and 9.3 below and the third aim will be discussed in 9.3, 9.4 and 9.5 below.

9.1 Discussion of Tools for Measuring Key Factors in Educational Interventions

In this section the suitability of data collection methods will be examined. One criticism of educational evaluation is that it often assumes normal distributions for measures that are in fact not normally distributed (Mitchell 2000). By analysing data against normal and binomial distributions as elucidated by Heron (2012) it is possible to draw conclusions about the impact of other factors on the data. The following section will show how this was successful in all three areas of evaluation satisfaction (9.1.1), learning (9.1.2 and 9.1.3) and behavioural changes (9.1.4). External factors were examined (9.1.5) and no strong correlations or impacts were found beyond that which might be expected in a normal distribution. As such it is concluded that the data is suitable to use and the tools are successful in the collection of data for evaluation purposes.

9.1.1 The Measure of Satisfaction

All the measures of satisfaction used the connection between satisfaction and reuse of the website adapted from methods outlined in models by Wang and Chiu (2011), Wang (2008) and Delon and McClean (2003). The first measure of satisfaction used an average number of page visits or uses of the website as a measure for satisfaction. This did not display a good fit to a normal distribution ($P < 0.05$). As such this was rejected as a tool for comparison with other groups. In this case, the distribution of access appears to be almost bimodal. This is on the face of it quite surprising, the use of access logs is a common evaluative technique within e-learning, and there is assumption more visits shows an increased value on the resource (Elliott & Neal 2016; Wang and Chiu 2011). However, looking at the relationship between learning and satisfaction in Figure 8.30 there is a weak association. However, the bimodal model continues to show what initially looks like a strong correlation between satisfaction and learning below 90 visits per user breaks down. Laurillard (2008) offers a possible explanation for this when she talks about individuals eddying in systems, this type of behaviour would produce a higher than the average number of page visits and would also lead to poor results in learning.

The second method used looked at the length of time in days that groups used the site. Again there was a large variation with a mean of 18.3 days and Standard Deviation = 21.2 days. As expected the mean is low and there is a drop off over time. However, the modal periods are even lower and as such, it was not a good example for modelling against normal distributions and comparison with statistical methods built on this assumption of a spread around the mean.

The third measure of satisfaction the frequency of visits in the first 30 days showed a much better fit to a normal distribution ($P > 0.05$). Mean = 1.96. Although it was still skewed to the lower side and so was borderline in its acceptability. This method was chosen as the best method for future comparisons. The relationship between the frequency of visits and learning again show a strong correlation with the initial increase in visits that tails off at around 2 visits per day (Figure 8.31). A similar relationship is seen in the frequency of visits as the number of visits in their correlation with learning. An initially strong relationship breaks down at a higher frequency. Below 2.2

visits per day $R^2 = 0.68$ but over the data as a whole, this is much lower $R^2 = 0.12$. As such it is concluded that during comparison with alternative teaching methods an increase in this value should be viewed with caution, as higher values than the control group may be connected to non-beneficial behaviours, future developments due to this implication are discussed in section 9.4.1.

9.1.2 Measure for Learning

Performance within the learning material was consistent across the groups and when compared to a binomial model the data was found to be a good fit ($p = 0.37$). It should be noted that one way in which it deviated from the predicted model is that there was a larger number of students achieving the arithmetic mean than predicted (observed = 15, predicted =11) so the data is more concentrated about the mean.

There were some issues around incomplete data sets from groups that had to be handled. If the number of questions attempted by the group or the percentage of the group answering a question was low it distorted the results. This is in line with the binomial modelling the larger the group size better the data should be. Small groups have a greater chance of fluctuation from the mean and so incomplete sets have the same issues. This is in line with Greller and Drachsler (2012), who points out that one of the biggest problems in the analysis of data sets are clean datasets free from dummy accounts, and some of these smaller groups might be the results of teacher tests. Similarly, larger mostly unused accounts in groups might be the result of a mismatch between accounts being created for groups that had access to fewer computers and time than was expected. Although filtering was put into place so that only accounts with at least one successful log in were counted as active and so part of the group. Although a comparison to group size was made and found to have little impact on the learning results this is within the 10 to 43 students range (Figure 8.14).

The initial method compared the right answers against the number of students in a group. Not answering a question was seen to be closely aligned with a wrong answer, as both demonstrate a student which is not confident in their understanding. This assumption becomes less plausible if only a small percentage from a group attempt the question. It is recommended that by combining checkpoint analytics for page views it would be possible to improve the filtering put in place by the active account filtration in cluster size. A clear idea of those that never saw a question could be achieved rather than those that decided not to answer a question.

This would not remove the statistical problem of small groups but would remove statistical doubts of the low percentage of answers given.

It is felt that removing incompletes is fair and in line with other studies (Greller and Drachsler 2012; Foster 2015), although it should be noted that with these results are included it moves the model away from the normal distribution.

Analysis for internal validity was carried out by Cronbach's alpha, this was clusters by the group and showed a range of values with the majority between 0.5 and 1, modal 0.8. (Figure 8.13). Scores above 0.7 are seen as an acceptable level of measuring a single variable, in this case, the level of understanding of the phase of the Moon, scores above 0.9 showing evidence of redundancy (Tavakol and Dennick 2011). Despite the narrow range in learning results that can cause problems with Cronbach's values, these values are seen as generally acceptable. It is recommended that this test is run for each iteration and that stock values for known tests should not be used (Tavakol and Dennick 2011). As such it could be used as a filtering factor for groups to eliminate groups that do not show good levels of internal validity. However, this method was not used in the comparison, in this case, see section 8.2 due to the low number of items in the comparison test which can cause large impacts on the Cronbach's alpha figure calculated (Tavakol et al 2011). Countermeasures for this are discussed in future work (section 9.4.2)

9.1.3 Analysis of Confidence-Based Questions

Both methods for confidence based testing methods failed to show any relationship against the percentage of students giving the correct answer. This is contrary to the finding of Foster (2015), whose work with secondary students in mathematics found that students both understood the process of risking points for answers and answering questions in relation to how likely they thought they would get it right. An alternative method of measuring confidence based on residual levels of confidence. Residual confidence was calculated by taking the mean confidence for a user from the value risked for a question. This hoped to compensate for users that showed a constant confidence level, for example always risking the maximum number of stars regardless of the difficulty of the question, however, it did not improve the results. Both of these results are interesting as they imply that users were neither consistent in their confidence thus showing an inherent confidence level from which they might deviate if confronted with a question they thought more or less difficult. Nor was the relation based on a linear relationship to the difficulty of the question. One possible conclusion is that users were just uninformed or misinformed (Kampmeyer et al 2015) causing them to guess and there were questions that users did better at guessing. However, from the levels seen in the Cronbach's analysis, this seems unlikely. Although for the sake of this study use of confidence was rejected it requires further work to better understand what has the potential for a very useful tool. This is looked at in more detail in section 8.4.2 of further work. But without a causal relationship, it is impossible to start analysis on a deeper categorised level.

9.1.4 Measures of Behavioural Change

Data collected for behaviour changes were successful. The tool was used consistently by clusters to varying levels and so produced data that could be further examined. The mean 3.64 entries per user for the average cluster.

Despite a slight peak in the number of users producing just a single entry into their space book which could be explained by minimal teaching requirements placed on students by teachers and a dip in the number producing two entries, there is a strong correlation to a normal distribution. Goodness to fit chi-squared calculation against a normal distribution shows a probability of $P>0.20$ and so the null hypothesis can be rejected. As such this is a good model for comparison as the behaviour of students is consistent and can be modelled using normal distributions. This allows for comparison against other samples.

Again there is a problem with data collection in null results. If a student fails to make any spacebook entries then the system does not know if this is because they never attempted to make any or if they failed to make any. There were 4 clusters where no spacebook entries were made by any members of the cluster. It is proposed that analytics systems could be improved if visits to this area of the site were also monitored. It would then be possible to get a measure of how many visits it takes to produce or not produce an entry.

9.1.5 Analysis of Known Factors

Clustered results in the above three areas where compared to known factors that could have influenced results in order to evaluate the effects of these factors.

The number of students in a cluster was analysed against satisfaction, learning and behavioural measure, all three showed very little correlation with the number in the group. Although it should be noted that there was generally a very small range in the number in the groups mean = 25.77 standard deviation = 5.48. It can be concluded that this does not have an effect on any of the indicators sampled within this given range. Both measures for levels of achievement in schools showed little correlation between the clusters and their satisfaction, achievement in learning or behavioural changes. It should be noted that there is a very loose negative relationship with the percentage of students achieving level 4+ in English and Maths and in Science and the measurements for learning and satisfaction. It is as such a possibility that the system was better at helping lower achievers than higher achievers. This is in line with the finding of lecture capture access which was shown to help lower achiever more than higher achievers who only accessed the resource a limited number of times (Owston et al

2011). However, the association is weak R^2 = 0.02 for achievement in English and Maths and R^2=0.035 in Science. Satisfaction saw a correlation of R^2 =0.02 with Maths and English and R^2 =0.03 with Science. This is normally too low a value to be significant but it is mentioned here as the same negative relationship was seen across all comparisons. It is felt that the impact of this was small, as clusters with low scores were in line with clusters achieving high scores.

Comparison against the percentage of disadvantaged students in schools showed only weak correlations with small effects on any of the evaluation methods.

In summary, in all areas results were in line with the spread of results found in the general population when compared to the known factors. However, this should be used with caution. This thesis is not saying that this result means that context can be ignored only that within the contextual changes a range of results is still achieved in line with the general population. As such it is possible to confidently compare subject domain-specific methods knowing that outside factors should not adversely compromise the results and as such going forward there is not the need for the system to require such information to be added to the system for groups in order to make comparisons.

9.2 Comparison with Alternative Delivery Method

Two approaches to learning were examined. The first (route 65) used an approach that saw pupils learning about the science behind the natural phenomenon of the phases of the Moon and then encouraged them to take their own observations and keep a Moon diary. The second approach (route

84) took a more applied task as the basis for the study. Students learnt about planning a mission to the Moon and then used the telescope to take images to plan their trip to the Moon.

Students undertaking route 84 were found to access the site significantly less frequently ($P < 0.01$), with a negative effect size of -0.69. They also achieved significantly worse ($P < 0.01$) in the questions with a negative effect size of -0.43. However, they performed significantly better in the behavioural changes ($P < 0.01$) with a positive effect size of 0.40.

The first thing that can be concluded before looking more deeply at the results is that is possible to make changes in the system that have a meaningful effect on the students using the system. It is the effect size that this study is most interested in as this allows comparison across different studies (Black et al 2007 and Coe 2002). The effect size of the study being in line with the effect size of the taking of practice tests (effect size = 0.32) and the difference between inquiry and traditional science curriculum approaches (effect size = 0.30) (Coe 2002). The second one of which this study draws comparisons to, as route 65 could be seen as a traditional approach and route 84 an inquiry or problem-based approach. For a grade at GCSE, a study would need an effect size of 0.5 or larger (Black et al 2007) and it is suggested that research should focus on activities that cause an effect size changes of at least 0.4 (Biggs and Tang 2011). This is seen as achievable within the system as it has caused a similar if at times negative effect. It should be noted that the aim of this study was never to look at the cause of the changes or to at this point find methods for improving teaching and learning, but only to see if changes could be made that would cause impacts on learning of the scale stated above and this has been achieved.

This moves us on to the second point which is the multi-dimensional nature of the analysis. Route 84 although negative in two of the measures was positive in the third. Examination against Bloom's (Krathwohl 2002) would suggest that the first two measures are of lower levels of Bloom's and the third is on higher level actions. As such they are designed to measure different feature of learning and so performance at different levels is to be expected. It could also be said that the more traditional method might promote lower levels of learning (Biggs and Tang 2011). However, it would be hoped that in promoting deeper learning approaches which are supported by the high levels of taxonomies such as Bloom's or SOLO which an inquiry or problem-based approach is meant to encourage (Biggs and Tang 2011) that the approach would enable the lower levels as well. This result is in line with findings of other studies where deep learning has not resulted in better test results (Yonker 2011). This could be put down to the tests not being able to make suitable measurements and a clear benefit was seen in the approach in the amount of creative work that the groups completed. It is a strength of the system that it is able to measure the impact of the change of approach in this way and can be used to inform future design of curriculum tasks as in an ideal world the practitioner would develop an approach that causes increases in all areas.

The system could be further strengthened by improving and combining data collection techniques a clearer picture of the process of learning. Whereas the learning tool used *process analytics*, behaviour changes and satisfaction were only a *checkpoint analytic*, they simply measure of if the student engaged in the task not how well or what they did (Lockyer et al 2013). The satisfaction measure should be extended to cover a larger part of the site as a checkpoint for engagement as the system only measured access to educational materials. Behavioural changes that caused pupils to use these areas of the site more might have caused the satisfaction measure to drop in the way it did, however, without data on the frequency of use in this area of the site, it is not possible to confirm if this was the cause of the reduction.

The analysis of learning across the routes had to use a reduced number of questions as only 4 questions were used in both routes. This low number of questions meant that it was not suitable for validation via Cronbach's alpha (Tavakol & Dennick 2011), which could be used as a way of filtering groups in the comparison. Although both routes use the same objectives questions were contextualised to the approach meaning that direct comparisons were not possible. Recommendations for alterations and further work are made in 9.4.2 so that these problems can be avoided in the future.

9.3 Comparison to Other Studies

In recent years there has been an increase in the number of studies looking at the impact of learning designs on learning analytics and behaviours in e-learning environments to promote learning (Mangaroska & Giannakos 2018).

Most studies fall into three broad categories those involving a single group, larger single institution studies involving multiple groups and studies involving large numbers of students in a single MOOC. Studies are more commonly of a smaller size with less than 40 learners (Mangaroska & Giannakos 2018).

Analytical techniques typically reflect the environment of their collection with smaller groups using the subjective use of analytics to aid orchestration and revision of learning design, this can be supported through analysis via expert opinion surveys or critical analysis of use. Examples of this can be seen in Melero et al (2015), Majumdar and Iyer (2016) and Florian-Gaviria et al (2013), comparisons are not made here to this scale of study as they differ in size, the use of aggregation and analysis methods.

Larger studies commonly try to correlate behaviours in learning environments with success of students or correlate thematic traits within learning designs to behaviours in the learning environment (Nguyen et al 2017; Rienties & Toetenel 2016; Toetenel & Rienties 2016; Pursel et al 2016). It is this area of study that will be looked at in more detail here as they are the closest to the methods used in this thesis. Although there is the key difference that this thesis was set in multi-institutional environment and looked at the impacts on enactment of contextual factors of a single learning design and a comparator, whereas the comparator studies either use a large range of learning designs across subject and discipline areas (Nyugen et al 2017; Rienties & Toetenel 2016; Toetenel & Rienties 2016) or the single enactment of a model within a MOOC (Pursel et al 2016).

Table 9.1 shows a list of similar studies and their results for comparison. As Coe (2000) recommends this thesis presents the effect size for a change in approach to learning by a given learning design as a statistic which can be used for comparison. Here comparisons are made with correlation data, although not the same statistic, correlation statistics are also a measure of the effect size of a variable, with both measures showing moderate effect size at >0.5, however, any comparison should not be as simple categorisation: small, medium and large and should be a comparison to previous studies (Cohen et al 2011). Caution should be used here as both the data processing and methodology of designs differ in the comparator studies. This thesis reported effect sizes of Satisfaction (-0.69), learning (-0.43) and behavioural changes (0.40).

Heron's (2012) study was used as the basis for statistical methods in this thesis and so is the most simplistic comparison. Similar results were found in terms of distributions, despite the difference in studies context; Heron looked at a single institution with face to face teaching administered online, this study looked at a blended method of teaching across multiple teaching establishments. Although different versions of Chi-squared significance tests are used, comparisons of probability can be made. In this study for learning, distributions showed a good level of fit (P=0.37). Behavioural change measure showed a good fit to the predicted (P=0.24). In Heron's study, the null hypothesis that distributions are normal was not rejected to two questions at P=0.05 and one question at P=0.01. Both studies showed statistics that were well represented by a normal distribution.

However, in this thesis less reliable fits were found in satisfaction with a significant differentiation from the predicted model at P<0.05 for both the number of days use and the number of visits with only the frequency of visits showing a borderline level of fit (P>0.05).

The conclusion from the Heron study is that the teaching or the approach to teaching taken in the environment of a lecture had no significant impact on pre or post teaching groups on their level of learning. By comparison, the results of different approaches from within the e-learning environment and its different approaches to teaching in this thesis showed a significant difference to the method of instruction ($P<0.01$) in all three areas, satisfaction, learning and behavioural changes.

Similar work which looks at the impact of learning design on satisfaction, engagement and learning has also been undertaken through a range of studies at the Open University (Nguyen et al 2017; Rienties & Toetenel 2016; Toetenel & Rienties 2016). These studies use a wide range of modules taught at the Open University and aggregated data from the use of online systems, their satisfaction surveys and the results of students to look at the impact of learning designs. The group sizes from which data is aggregated in all of the Open University's studies are far bigger than in this thesis but each design is only run a limited number of times. Learning designs are mapped against 7 categories: Assimilative, Finding and handling information, Communication, Productive, Experiential, Interactive/adaptive and Assessment with the percentage of each course given to each type of activity. Thus the Open University's studies rely on the strength of the categorisation and the generalisation of what an assimilation activity might involve. Laurillard (2012) shows how it is possible to redesign an assimilative activity to have an improved teaching design with greater amounts of feedback whilst still remaining an assimilation design at its core, as such the granularity of mapping becomes an important factor in the interpretation. The key difference is that in this thesis a single comparison of multiple interactions of just two learning designs were used to compare the repeatability of measurement within the environment and estimate impacts of other factors before looking at the difference that a single change would make, thus negating the need to thematically categorise the designs. A range of computational techniques are used to compensate for possible variance in data in the Open University's studies.

Comparable impacts were found to the changes in learning design. Nguyen et al (2017) report that a change in learning design accounts for 69% of the time spent on the e-learning environment. This is similar to change in satisfaction found in this study which was measured through site use. Smaller impacts were found on learning (0.11) in the Nguyen et al study compared to the changes made in this thesis (-0.43), one explanation to this could be that more mature students have developed their own meta-cognitive practices for learning and so are able to succeed despite design, these same strategies might not be available to the younger pupils in this study, meaning that the given method had a bigger effect. However, without an in-depth case study of student interactions, it is not possible to confirm the cause of behaviour in either case. Learning design changes were also shown to have a possible negative effect in learning with Toetenel and Rienties (2016) showing a significant negative effect on assimilative activities (-0.326, $p<0.05$). This study showed a similar negative effect on learning when changing context to promote a real-world example with the aim of promoting deeper learning. Both of these changes are comparable in size, however, this result is somewhat contrary to the changes measured in this study as the alternative learning design promotes learning activities more in line with a non-assimilative nature. This could be because the assessment in question was more aligned as elucidated by Biggs (2003) in each case with the outcomes associated with the type of learning activity that saw positive gains in learning results.

Rienties & Toetenel (2016) also demonstrated a correlation between satisfaction and engagement that would support the methods for satisfaction measurement used in this study that relied on engagement.

Correlations in behaviour in MOOCs can also be compared to the correlations between learning and satisfaction (through the frequency of use) in this thesis. Pursel et al (2016) found a correlation between the viewing of online materials and success (0.16). Pursel et al (2016) also found separation in online activities that support the multidimensional approach taken here, as the pedagogical structure for analysis was designed around the idea that there were different levels of interactions within the environment that were independent points of failure within the design.

Table 9.1 Results and Methods of Analysis of Comparable Studies

	Method of Collection	**Method of Analysis**	**Key Results**
Heron (2012)	Online administered undergraduate module test results collected between 2000 and 2012 average group size 100	Normal probability plot correlation coefficient test on 3 questions within the test designated CAP1, CAP 2 and DRB1	Using a "normal probability plot correlation coefficient test" and using p<= 0.05 the null hypothesis (that the distributions are normal) is not rejected for DRB1 or CAP2, while using p<=0.01 it is not rejected for CAP1 (The distribution for CAP1 is skewed, expected given the high mean).
		T-tests and (non-parametric) permutation tests on results for groups to tests given before, during and after instruction	CAP1, the "during instruction" average is lower than the "before instruction" average (p<=0.05). CAP2, "during instruction" average is higher than the "before instruction" average (p<=0.05). No statistically significant difference between the "before instruction" and "after instruction" results.
Toetenel & Rienties (2016)	60,000+ students, 157 learning designs. Aggregated module data for students and thematic mapping of % type of activity in learning design.	Correlation (Person r Coefficient) of the percentage of learning design designated to one of 7 thematically mapped categories of activity to pass rate of students who completed the course.	A negative correlation of assimilation activities in design and pass rate of completed students (-0.326, p<0.05).

	Method of Collection	Method of Analysis	Key Results
Rienties & Toetenel (2016)	151 modules with an average of 753.15 students per module. Aggregated module data for students and thematic mapping of % type of activity in learning design.	Multiple regression analyses, between thematic categorisation of pedagogical approaches in teaching, the satisfaction of learners, retention and engagement	A significant negative correlation was found between assimilative activities and academic retention (r = 0.268, p<0.01). A significant positive correlation was found in terms of communication (r=0.269, p<0.01). Satisfaction: positive correlation between assimilative activities and Average SEAM satisfaction data (r= 0.333, p<0.01) Negative correlation to finding information (r=0.258, p<0.01) Communication activities and satisfaction (r=0.224, p < 0.01) Average time spent in the VLE correlated positively with finding and handling information (r = 0.318, p<0.01), communication activities (r = 0.471, p < 0.01), experiential (r = 0.376, p<0.01) and total workload (r = 0.456, p<0.01), whilst, a negative relation was found with assimilative activities (r = 0.300, p<0.01). VLE engagement per session positively predicted learner satisfaction. 0.192 (p<0.05)

	Method of Collection	Method of Analysis	Key Results
Nguyen et al (2017)	74 undergraduate modules and their 72,377 students. Aggregated module data for students and thematic mapping of % type of activity in learning design.	Thematic learning design categorisation and mapping are analysed using a Fixed effect model with assessment as the dependent variable was conducted. A mixed method of case studies with cases selected from across the range of results is used	Adjusted R-squared for fixed effect model 0.69 for the effect of learning design on engagement Linear regression adjusted R squared 0.29 for Satisfaction, and 0.11 for pass rate against learning design
Pursel et al (2016)	User survey and data gathered from MOOC use 90000 + in MOOC 9000 in survey Known bais in the completion rate of survey responders again full MOOC sample.	Correlation analysis	Key indicators of success in MOOC are the use of forums (0.18)and the viewing of online materials (0.16) However low correlation found between previous attainment (0.05) Good independence of variables videos views and posts per week (0.07)

	Method of Collection	Method of Analysis	Key Results
This thesis	Online data generated by use of the virtual learning environment. Aggregated data from 78 groups used to analyse the nature of distributions And a further 168 groups used for a comparison of the effects of a change in learning design. Three measures satisfaction through the frequency of use, learning through formative multiple choice questions and behavioural changes through engagement in activities that promote deeper learning approaches.	Suitable measurements for interpretation by simple Gaussian distributions are identified using a chi-squared goodness to fit modelled against the mean and SD of the sample for behavioural changes and satisfaction and binomial distributions using the mean value as the probability for learning. R^2 analysis of known factors to look at their impact on distributions Comparison of learning designs using T-tests and cohen's d for effect size	Pupils responded consistently to the same learning design across a range of the contexts. satisfaction through the frequency of use with a probability of ($P>0.05$), learning through formative assessment ($P>0.3$) and behavioural changes through engagement with higher order activities ($P>0.2$). National indicators for social/economic, academic achievement and group size are examined for bias with no significant factors found. $R^2 < 0.07$ in all areas. Change in learning design within the e-learning system satisfaction (-0.69, $P<0.01$), learning (-0.43, $P<0.01$) behavioural changes (0.40, $P<0.01$).

9.4 Conclusions from Multi-dimensional Evaluation Framework Tools

This study was able to show that the performance of users within clusters, in this case, real-world class groups, that are exposed to the same learning material perform in a way that means they can be compared to others in different schools in different cities from different backgrounds to draw meaningful conclusions about pedagogical approaches.

This shows great potential in the semi-automation of evaluative techniques in order to aid the science of pedagogy and in aiding teachers to engage in evaluation of their teaching through a process of transformation reflection.

One key advantage of the framework developed is that they are multi-dimensional, as such it avoids the temptation to alter an intervention to increase the measurement of any one single factor within the study (Slavin et al 2014). Although not done in this study for a school teacher the final level of Kirkpatrick attainment (Kirkpatrick and Kirkpatrick 2008) can easily be linked as they will have access to this for their own students.

Of the three measures, satisfaction showed the greatest variation in all the measures and differed from the normal distribution in all but the frequency measure, which would make comparison hard with other groups. Even the frequency of visits presented skewed results although not significantly, most noticeable in use by those who took part in route 84. Individuals had the greatest effect on satisfaction measure of both visit count and period of use. A single child that either visits the site an enormous amount or continued to use the site for a number of years had large effects on their group. These sorts of users were not rare in the data with a small percentage of users continuing to use the telescope site after 15 weeks, and the majority of traffic ending after just 4 weeks (Machell and Baruch 2015).

The relationship between achievement in learning and reuse showed that although initially, more use is equitable to better learning this relationship trails off as use increases beyond an average of 2.2 visits per day. This result is contrary to the assumptions made by many researchers who view access as a measure of success for a project and models of use presented by Wang (2017) who links lower level activities which they call wayfinding activities with a higher level activity of creating new content. It is better explained by the process of eddying (Laurillard et al 2000). Alternative methods of analysis are suggested in further work in line with this.

Route facility appears to be a good measure for learning with users performing consistently as do the behavioural changes. The two factors showed little association with each other with users who showed higher levels of learning producing more entries in spacebook ($R^2 = 0.09$), however, this could be put down to the narrow range in space book entries for a single route. It might also show that they are measuring two different aspects of the learning cycle with behavioural measures focused on higher level creative activities and the learning measure looking at the lower level discursive or declarative understanding.

Due to the nature of questions being part of the learning process when this was changed so did the phrasing of many of the questions asked, this made comparison across approaches difficult and only a limited number of questions were asked in the same way to both groups.

Behaviour changes again modelled well against normal distributions allowing comparison with statistical methods. However, there is a need to expand the checkpoint analytics to process-driven analytics as was shown in the comparison of designs, one might ask the question if they did not understand the basics what could they have achieved in the higher levels. However, this was not the initial aim of this tool as this study wanted to see if pupils were encouraged to engage in higher level and social practices within learning rather than their achievement in this area.

The following framework is proposed for the evaluation process (figure 9.1). The relationship shows how the lower levels of learning and satisfaction support behavioural changes, and although the areas are interrelated and there is a spot in which all can be achieved it is the case that achievement can happen in one or more of the areas without success in the others. It is also possible that too much of a focus on any one of the areas will have a detrimental effect on the other two. This is held up by the relationships seen between data which showed a loose association (section 8.7) and the results of the comparison of approaches to learning which showed a method that enhanced one area to the detriment of the other two. Each of the areas of evaluation is aligned with an activity in the e-learning environment and each stage of the learning process helps to build on the last and so enables achievement in the next.

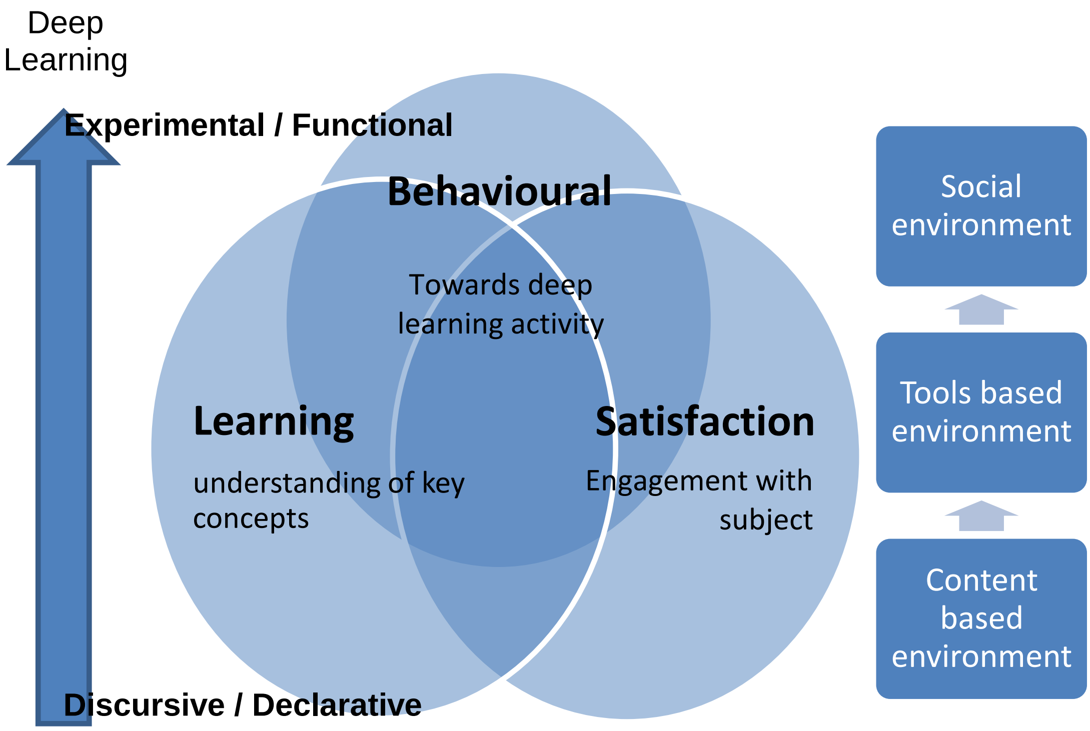

Figure 9.1 Framework for evaluation and learning relationship.

9.5 Further work

There are two directions for future work the first is an extended trial with teachers carrying out their own research using these tools to look at their impact and usability, of particular interest would be the representation of data that involved both tools for presenting designs for learning (Oliver et al 2013, Agostinho et al 2013) with learning analytics (Persico and Pozzi 2015) so that both the approaches and success of these approaches could be shared within the system. In particular, the builder tools used to develop content could be combined with the learning design file format IMS-LD standard (Agostinho et al 2013) to allow sharing beyond the site with others.

The second is to further develop the data collection tools themselves to improve the measurements that they take and the conclusions it is possible to gain from them. This will be examined in more detail here as it is the most relevant to the work of this thesis.

9.5.1 Measures of Satisfaction.

At the time of design of this system in 2009/10, the infiltration of Facebook into everyday life was still in its infancy with around 500 million users (statista n.d.), the idea of the like was not as ubiquitous as it is today, with Facebook at two billion users (BBC 2017). As such visits were used as the key indicator for satisfaction. It was felt that the use of a Likert scale would go unused or would be overly biased as to whether they liked the teacher rather than the activity. The like is now the undeniable system for satisfaction and so would be the obvious choice. With the diversification of likes into a range of responses, it is not inconceivable to design a traffic light system that would not only provide information on satisfaction but could provide useful information to the teacher on the current status of students and how they feel in their understanding of learning.

This could be combined with statistics on page visits to tell a more complete view of satisfaction. Similar approaches have been used but to substantiate the perceived values of lecture chapter, where use logs were used to support the perception of value from student surveys (Elliott & Neal 2016). As discussed in section 9.1.1 a simple measures such number of visits did not present the relationships that might have been expected, More complex studies of navigation were not suitable for this study, however, a mixed methods suggested above would supply information on not only the amount of use but could help to link more complex navigational behaviours to satisfaction.

It is also suggested that the satisfaction tool is extended to other areas of the site, as it is possible that users might enjoy other areas and so be reuse might not be fully tracked, all possible areas should be considered in order to give a complete picture if this method is used.

9.5.2 Measures of Learning

It is recommended that studies should use a pre and post test to measure difference made (Kirkpatrick and Kirkpatrick 2008), however, it was felt that this might be prohibitive as it might be in conflict with the natural flow of the lesson. This study used clustering or normal distributions instead of pre and post. However, the method could be further enhanced by incorporating pedagogies that have come to the foreground since the development of the system in 2010. Flipped lectures have become increasingly popular as a method of teaching in universities in the UK in recent years. Many that use the approach also include some sort of class voting system in order to help reinforce the learning in the lecture part of the flipped activity (Lancaster 2013). Students first answer a question via a voting system, percentage results are then shown to the class but the correct answer not revealed to the group, students are then asked to find someone in the room that has given a different answer to them and discuss the question before re-voting. Round of this may be repeated until the teacher is happy that the class has reached a common and correct, understanding. This may require intervention by the teacher to help students to this point. Finally, the teacher brings the group together to reveal the correct answer to the group (Majumdar and Iyer 2016). This process could be automated with an app for students that would then allow the student to chat using their devices, and could even auto match students with other of different answers, teachers would have an overview of the process and could highlight key comments to the group if they wished.

This would create two lots of useful data, the first is an impact, it would be possible to see the number of people that changed their understanding based on the session allowing us to measure impact and making comparison between groups easier as would have a baseline for each group secondly it would provide a reach stream of data that could be mined for misconceptions. Studies into the use of a clicker and with a similar pedagogical approach and learning analytics have recently been carried out on a class level of feedback (Majumdar and Iyer 2016) although the application of data analytics and the stakeholders would be quite different.

An alternative approach that would allow a similar analysis is the team-based learning method, here an individual readiness test and group readiness test are used (Tweddell et al 2016). Students take the same test twice once on their own and once as a group where they must all agree on the answer they are going to give for a question. Again this could be combined with messaging tools to allow the entire process to happen online. This again would provide a baseline and final level of understanding along with a rich stream to mine for misconceptions.

One key consideration in the adoption of either method is the providing of content, both method separate presentation of initial concepts from the testing cycle this is different from the current parallel testing approach, the impacts of changes to the approach would have to be considered before either methods were adopted.

In section 9.2 the problem of comparing questions across teaching approaches was covered, as the approach to the curriculum framed the wording of questions. There are two possible solutions to this the first is the addition of end of topic tests which could be constant as not displayed against subject material. Tests were available in the system presented here but questions in the test could be attempted multiple times by students and so did not supply meaningful data. The tests were implemented to promote reflection and revision in line with an assessment for learning ethos in the site rather than a summative judgement as so did not align well with evaluation methods. An alternative would be to extend the concept of reusable content in the site to encourage the use of common questions this would increase the number of similar questions in future work. The approach of using question banks is popular in the United States where they allow teachers and lecturers to build quizzes with the confidence of item difficulty (Yonker 2011). Similar services such as exam pro (AQA n.d.) are available for secondary schools in the UK to help teachers develop practice examinations in line with national tests. However, a key aim of the assessment in the site is to facilitate the conversation between the teacher and the student and their understanding, as such some level of questioning to enable this will always be required, meaning a full comparison of all questions isn’t possible.

It was also suggested in 9.2 that the reflective spaces presented in this study could be enhanced by including process analytics, as well as checkpoint analytics currently used. The choice of checkpoint analytics was aligned with the pedagogical approach of a free reflective space that was not intended to be marked with a number, instead, teachers were able to edit the pages and then teachers to allow the sharing of work as a reward for good work. However, there are some features that could be added to the environment that would allow a simpler quantitative measurement of achievement.

Self-marking or peer marking would be possible if an ideal answer was provided (Laurillard et al 2000) However, it was found in the system that the objective tool was underused by most teachers. One reason for this could be the time it takes to prepare this sort of activity online rather than verbally ask a group in a class. So would a teacher take the time to prepare ideal answers for their students if a system was available? The reflective area of space book was very well used by learners as such it is suggested that the two features are combined in future versions of the website to increase the use of objectives in the system.

One solution to ideal answers would be to provide templates for reflective areas and objectives, however, this might be restrictive for teachers and learners. A practical solution is to provide both the tools to develop your own or edit or use existing templates to help teachers get started. An alternative method would be to allow teachers to highlight the good work of students to other students, thus removing the need for the teacher to prepare ideal answers.

9.5.3 Behavioural Measures

At the current time tracking of access to the different pages of the social parts of the website, space book, are not stored. However, there is the possibility to extend this area of the website so that it could produce a similar network analysis that currently used on discussion boards (Lockyer et al 2013). As well as the more formal marking tools suggested in future work for learning these tools would help to model the highly creative aspects of the framework and would enable modelling of student actions over time to better understand the processes which inspire them to create and share. It would also help to give a more complete view of the satisfaction measures.

9.6 Conclusions

Mor et al (2015) in a call for action in the British Journal of Educational Technology asked for more work to be done in the development of learning analytics for the use and development of the design for learning and the making of tools available for teachers to engage in reflective practice and educational research. This shows that these issues are still current and areas of research that need contributions. When this call was made initial results from this thesis were being presented examining how learning analytics could be used to evaluate approaches to learning and so inform practice (Machell & Baruch 2015).

In a recent speech, the Rt Hon Nick Gibb MP (Minister of State for school Standards) states the importance of an evidence-informed profession (Gibb 2017). This is a sentiment echoed here. This thesis has examined a novel framework for the collection and interpretation of data from an e-learning environment. It has shown how evaluation aligned with the pedagogical principals of the learning environment can help develop a balanced method for evaluation.

Persico and Francesca Pozzi (2015) reiterate the need for teachers to be engaged in a reflective practice that they can use to improve the design of their learning activities and share with other teachers. There is increasingly a need for this to a scientific enquiry based approach.

This thesis has examined some of the basic assumptions in the nature of data collected within a multi-institutional Primary school environment. Analysis of distributions of clustered data were compared using statistical methods across the range of contexts, which would be common to e-learning environments in a primary school setting, and has shown that it is possible to move beyond the single classroom and the use of analytics to help a teacher's understanding of their group, to inform the wider principals of design of educational tasks within a structure of scientific enquiry.

Comparisons between alternative teaching methods have shown how changes to tasks within the system can have a measurable effect on the outcomes of learning. The framework for the learning environment presented, which is built on a framework to sound pedagogical principals, is flexible enough to allow a range of approaches within learning, rather than limiting changes to superficial changes. Data analysed from the two alternative methods of teaching highlighted the benefits of a multi-dimensional approach in the evaluation. Conclusions from any single dimension of the learning process could be misleading and might ignore benefits in other areas. This was the case in the approaches evaluated where the alternative method improved marker for behavioural changes towards deep learning but had a negative effect on others.

This emphasises the importance of many of the basic questions that need to be answered for a meaningful evaluation to happen. The philosophical basis for learning and teaching must be in place if it is to be possible to evaluate outcomes and align analytical tools with the activities of students. Without these steps and a framework on which to evaluate, meaningful comparisons are not possible.

Greller and Drachsler (2012) point out that access to datasets to study is a priority for the understanding of research in this area but also a potential problem. Primary educational sets for large samples rather than localised data sets are less common than datasets for Open and Higher education. This thesis has allowed the examination of data from a multi-institutional primary school setting and has shown the potential for data to be collected and analysed across the contexts that the data has been collected from. It has shown that the use of natural real-world clusters is the key to understanding data. The dataset used will be cleaned of all personal data and fully anonymised data will be released in line with the user agreement for further study.

References

Agostinho, S., Bennett, S., Lockyer, L. Jones, J., and Harper, B. (2013) Learning Designs as a Stimulus and Support for Teachers' Design Practices. in Beetham and Sharpe (editors) *Rethinking pedagogy in a digital age, designing for 21st century learning* 2nd edition, 2013 Routledge 119-132

Alexander, S. (2001) E-learning developments and experiences. *Education and training* volume 43, number ⅘ 240-248

Al-Saai, A., Al-Kaabi, A. and Al-Muftah, S, (2011) Effect of a Blended e-Learning Environment on Students' Achievement and Attitudes toward Using E-Learning in Teaching and Learning at the University Level. *International Journal for Research in Education* (IJRE) No. 29,

Amazon (n.d.) *Amazon Echo - Alexa Voice Service - Amazon.co.uk* [online] https://www.amazon.co.uk/Amazon-SK705DI-Echo-Black/dp/B01GAGVIE4 accessed 30th June 2017

Apache (n.d.) [online] available https://httpd.apache.org/ accessed 6th July 2017

Aparicio, M., & Bacao, F. (2013). e-learning concept trends. ISDOC '13 Proceedings of the 2013 *International Conference on Information Systems and Design of Communication.* 81-86

AQA (2017) *exampro* [online] http://www.exampro.co.uk/sec/index.asp accessed 9th of July 2017

Atif, Y., Benlamri, R., and Berri, J. (2003) Dynamic Learning Modeller. *Educational technology and society* 6(4) P60-72

Baruch, J.E.F., Hedges, D.G, Machell, J. and Tallon, C.J. (2004) Robotic telescopes and the public. Astron. Nachr. / AN 325, No. 6–8, 457–461

Baruch, J.E.F. Machell, J. and Norris, K. (2005) Watching the Moon from Tenerife. *Primary Science Review,* 88, 7-9

Beetham, H., & Sharpe, R. (2007). An Introduction to Rethinking Pedagogy. In Beetham and Sharpe (editors) *Rethinking Pedagogy in a Digital Age, Designing for 21st Century Learning* 1st edition, Routledge p1-10

Beetham, H. and Sharpe, R. (2013) An Introduction to Rethinking Pedagogy. In Beetham and Sharpe (editors) *Rethinking Pedagogy in a Digital Age, Designing for 21st Century Learning* 2nd edition, Routledge p1-12

Beetham, H (2013) designing for active learning in technology-Rich Contexts. In Beetham and Sharpe (editors) *Rethinking Pedagogy in a Digital Age, Designing for 21st Century Learning* 2nd edition, Routledge p31-48

Black P, Harrison C, and Lee C, William, D. (2007) *Assessment for Learning*. McGraw-Hill Education.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *The Phi Delta Kappan*, 80(2), 139-148

Blackboard (n.d.) [online] http://uk.blackboard.com/ accessed 5th of July 2017

Blass, E. and Davis, A. (2003) Building on solid foundations: establishing criteria for e-learning development. *Journal of Further and Higher Education* Vol. 27 (3) 227-245

Biggs, J. (2003) Aligning teaching and assessment to curriculum objectives [online] available at https://www.heacademy.ac.uk/system/files/biggs-aligning-teaching-and-assessment.pdf accessed 6th July 2017

Biggs, J. and Tang (2011) *Teaching for Quality Learning at University.* 4th Edition. SRHE and Open University Press Imprint.

Booth M, (2012) Learning Analytics: The New Black. *Educause Review* July/August 2012 p 52-53

Bricheno, P., Higgison, C. and Weedon, E. (2004) The Impact of networked learning on Education institutions. The JISC INLEI Project

Britain, S and Liber, O. (1999) *A framework for the pedagogical evaluation of virtual learning Environments.* JTAP JISC technology application programme [online] http://www.leeds.ac.uk/educol/documents/00001237.htm accessed 9th of July 2017

Britain, S and Liber, O. (2004) *A framework for the pedagogical evaluation of eLearning Environments.* Educational Cybernetics: Reports. Paper 2. [online] http://digitalcommons.bolton.ac.uk/iec_reports/2 accessed 9th of July 2017

British Broadcasting Corporation - BBC (2017) *Facebook hits two billion users* http://www.bbc.co.uk/news/business-40424769 accessed 10th July 2017

British educational communications and technology agency - BECTA (2005) *ICT, amenability and the BBC digital curriculum service in England - Supplement 1* [online] http://archive.teachfind.com/becta/foi.becta.org.uk/display1b9c.html?resID=35715 accessed 9th of July 2017

British educational communications and technology agency - BECTA (2003) *A Review of the Research Literature on the use of Managed Learning Environments and Virtual Learning Environments in Education, and a Consideration of the Implications for Schools in the United Kingdom* [online] available http://archive.teachfind.com/becta/research.becta.org.uk/index928e.html accessed on the 2nd of July 2017

Camacho A, Holt R, Marley S, Roberts S. (2009) Robotic telescopes- classroom activities at the cutting edge of science, Secondary Science Review, June 2009 90(333) 73-82

Carchiolo V, Longheu A, Malgeri M, (2002) Adaptive formative paths in web-based learning environment, *educational technology and society* 5(4) 2002 P64-P75

Chen B.X (2009) *June 29, 2007: iPhone, You Phone, We All Wanna iPhone.* Wired [online] https://www.wired.com/2009/06/dayintech_0629/ accessed 9th of July 2017

Churches R. and McAleavy T. (2015) *Evidence that counts- what happens when teachers apply scientific methods to their practice.* CfBT Education Trust [online] available at https://eric.ed.gov/?id=ED562564 accessed 8th July 2017

City of Bradford Metropolitan District Council (2012) *2011 Census Release 2.1: Key Statistics for Local Authorities Bradford District* [online] available https://ubd.bradford.gov.uk/media/1275/2011-census-second-release-11-december-summary-note.pdf accessed 10th July 2017

Coe R (2002) *It's the Effect Size, Stupid. What effect size is and why it is important. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England, 12-14 September 2002* [online] https://www.leeds.ac.uk/educol/documents/00002182.htm accessed 10th of July 2017

Coffield, F., Moseley, D., Hall, E. and Ecclestone, K., (2004), Should we be using learning styles? What research has to say. *Learning and skills research centre.* ISBN 1 85338 914 5

Cohen L., Manion L. Morrison K. (2011) *Research Methods in Education*. 7th Edition, Routledge

Cook, J., White, S., Sharples, M., Sclater N, Davis, H. (2007) The design of learning technologies, In Conole, G. and Oliver, M. (editors) *Contemporary Perspectives in e-learning research themes methods and impact practice.* Routledge. 55-68

Conole G. and Oliver (2007), introduction. In Conole G. and Oliver (editors) *Contemporary perspectives in e-learning research themes methods and impact practice.* Routledge p. 3 -20

Crawford K (1996) Vygotskian Approaches in Human Development in the Information Era. *Educational Studies in Mathematics*, Vol. 31, No. 1/2 pp. 43-62

Cultofjaffa (2011) [online], Jaffa Cakes Half Moon Full Moon Total Eclipse Advert, https://www.youtube.com/watch?v=-dtmRmOHK78 accessed 31st June 2017

Dangwal R and Gope S (2011) Indian adaptation of Motivated Strategy Learning Questionnaire in the context of Hole-in-the-wall. *International Journal of Education and Development using Information and Communication Technology* (IJEDICT), Vol. 7, Issue 3, 74-95.

Deci E. L., Koestner R., Ryan R. M. (2001) Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again. *Review of Educational Research* Vol. 71, No. 1, pp. 1–27

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*,19(4), 9–30.

The Department for Children, Schools and Families (2008) The Assessment for Learning Strategy [online] available http://webarchive.nationalarchives.gov.uk/20130322234706/https://www.education.gov.uk/publications/eOrderingDownload/DCSF-00341-2008.pdf accessed 6th July 2017

Department for Education (n.d.) Eligibility for teacher training, Getintoteaching.education.gov.uk accessed 9th February 2019

Department for Education (2015) *Schools performance tables* [online] available at http://webarchive.nationalarchives.gov.uk/20150812112357/http://www.education.gov.uk/schools/performance/archive/index.shtml accessed 8th July 2017

Department for Education (DFE) (2013) Phonics check leaflet [online] https://www.gov.uk/government/uploads/system/uploads/attachment_da ta/file/194057/phonics_check_leaflet_2013_.pdf accessed 10th June 2017

Department of Education and Skills (DfES) (2005), Harnessing technology transforming learning and children's services: summary version. [online] available http://webarchive.nationalarchives.gov.uk/20060213221855/http://www.dfes.gov.uk/publications/e-strategy/ accessed 9th of July 2017

Dredge, S (2015) *MySpace – what went wrong: 'The site was a massive spaghetti-ball mess'. The Guardian* [online] available at https://www.theguardian.com/technology/2015/mar/06/myspace-what-went-wrong-sean-percival-spotify accessed on 9th of July 2017

Dyke M., Conole G., Ravenscroft A., de Freitas S. (2007) learning theory and its application to e-learning. In Conole G. and Oliver (editors) *Contemporary Perspectives in e-learning Research Themes Methods and Impact practice* Routledge. 82 -97

Elliott, C. Neal, D. (2016). Evaluating the use of lecture capture using a revealed preference approach. *Active Learning in Higher Education,* Vol. 17(2), p. 153–167.

Facebook (n.d.) *About* [online] available https://www.facebook.com/pg/facebook/about accessed 8th July 2017

Florian-Gaviria B., Glahn C., and Gesa R. F. (2013) A software suite for efficient use of the European qualifications framework in online and blended courses. *IEEE Transactions on Learning Technologies*, vol. 6, no. 3, pp. 283–296

Foster C (2015), Confidence and competence with mathematical procedures, *Educ Stud Math* (2016) 91:271–288 DOI 10.1007/s10649-015-9660-9

Fowler C.J.H. and Mayes J.T. (1999) Learning relationships from theory to design In Squires D., Conole G., Jacobs G., (editors) *The Changing Face of Learning technology.* University of Wales Press. ALT. P.39 – 50.

Gardner, H. (1984). *Frames of mind: The theory of multiple intelligences*. London: Heinemann

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. The Internet and Higher Education, 28, 68-84

Ghaoui C., Taylor M.J. (2000) Support for Flexible e-Learning on the WWW: A Special Issue, *Journal of Network and Computer Applications* (2000) 23, 337-338

Gibb, N., (2017) *The importance of an evidence-informed profession.* Gov.uk [online] available https://www.gov.uk/government/speeches/nick-gibb-the-importance-of-an-evidence-informed-profession accessed 10th July 2017

Gijbels, D., van de Watering, G., Dochy, F. and Van den Bossche, P. (2005) The relationship between students' approaches to learning and the assessment of learning outcomes *European Journal of Psychology of Education*, Vol. 20, No. 4 (December 2005), pp.327-341

Goodyear (2001), *Effective networked learning in higher education: notes and guidelines Networked Learning in Higher Education Project*, Networked Learning in Higher Education Project (JCALT). Centre for Studies in Advanced Learning Technology, Lancaster University 2001.

Google (n.d.) Google Home – made by Google [online] available https://madeby.google.com/home/ accessed 30th June 2017

Govindasamy, T. (2002) successful implementation of e-learning pedagogical considerations. *Internet and Higher Education* 4 (2002) 287 - 299

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, *15* (3), 42–57.

Gunasekaran, A., McNeil, R.D., and Shaul, D. (2002) E-learning: research and applications, *Industrial and commercial training*, volume 34, number 2, 2002 44-53

Hao Yang H (2013) New World, New Learning: Trends and Issues of E-Learning, *Selected Papers of Beijing Forum 2010, Procedia - Social and Behavioral Sciences* 77 (2013) 429 – 442

Heron, P.R.L. (2012) Student Performance on Conceptual Questions: Does Instruction Matter? *Physics Education Research Conference. AIP Conference Proceedings*, Volume 1513. AIP Conference Proceedings, Volume 1513, Issue 1, p.174-177

Hockemeyer, C. and Dietrich A. (1999). The adaptive tutoring system RATH. *ICL99 Workshop Interactive Computer aided Learning: Tools and Applications.* Villach, Austria: Carinthia Tech Institute. 1999.

Hodgson, C and Pyle, K. (2010) *A literature review of Assessment for Learning in science. National Foundation for Educational Research* [online] https://www.nfer.ac.uk/publications/aas01/ accessed 5th of July 2017

IEEE (2002) *Draft Standard for Learning Object Metadata,* [online] available from http://129.115.100.158/txlor/docs/IEEE_LOM_1484_12_1_v1_Final_Draft.pdf accessed on 15 March 2015

Index of Multiple Deprivation -IMD (2010) *Official Statistics English indices of deprivation 2010* [online] available https://www.gov.uk/government/statistics/english-indices-of-deprivation-2010 10th July 2017

Instituto de Astrofísica de Canarias - IAC, (n.d.) *Observatorio del Teide, telescope* [online] http://www.iac.es/eno.php?op1=3&op2=6&lang=en 10th of July 2017

JISC (2004). Effective Practice with e-Learning [online]. Available from: http://www.elearning.ac.uk/effprac/index.htm accessed 8th June 2017

Jones C., Cook j., Jones A., De Laat M., (2007) collaboration. Learning Theory and its Application to e-learning. In Conole G. and Oliver (editors) *Contemporary Perspectives in E-learning Research Themes Methods and Impact Practice.* Routledge. 175 -189

Kalaš, I., Laval, E., Laurillard, D, Lim C.P., Meyer F, Musgrave, S., Senteni, A., Tokareva, N., Turcsányi-Szabó, M. (2014) *ICT in Primary Education Analytical Survey. Volume 2 Policy, Practices, and Recommendations.* UNESCO Institute for Information Technologies in Education [online] available http://unesdoc.unesco.org/images/0023/002319/231933e.pdf accessed 8th July 2017

Kampmeyer, D., Matthes, J., Herzig, S., (2015) Lucky guess or knowledge: a cross-sectional study using the Bland and Altman analysis to compare confidence based testing of pharmacological knowledge in 3rd and 5th year medical students. *Adv in Health Sci Educ* 20:431–440

Karagiorgi Y., Symeou L., (2005) translating constructivism into instructional design: potential and limitations. *Education Technology & Society*, 8 (1), 17 – 27

Keengwe J, Onchwari G, Agamba J (2014) Promoting effective e-learning practices through the constructivist pedagogy *Educ Inf Technol* (2014) 19:887–898

Kenney. J., Hermens a., and Clarke, T., (2004) the political economy of e-learning educational development strategies, standardisation and scalability. *Education + training* volume 46 - number 6/7 2004 pp370-379

Kinshuck, N.A. and Patel, A. (2001) Adaptive Tutoring in Business Education Using Fuzzy Backpropagation approach, In Smith M.J. Salvendy G. Harris D and Koubek *(editors) Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent and Virtual Reality, Proceedings of the 9th International Conference on Human-Computer Interaction.* New Orleans USA, 5th-10th August 2001, Lawrence Erlbaum 465-468

Kirkpatrick, D. L., Kirkpatrick, J. D. (2008*). Transferring learning to behaviour: Using the four levels to improve performance.* 3rd edition. San Francisco, CA: Berrett-Koehler Publishers

Koohang, A. and Paliszkiewicz, J. (2013) Knowledge construction in e-learning: an empirical validation of an active learning model*. The journal of computer information systems*, spring 2013, 53, 3. P. 109

Krathwohl, D. R, (2002) A Revision of Bloom's Taxonomy: An Overview, Theory Into Practice, Vol. 41, No. 4*, Revising Bloom's Taxonomy* (Autumn, 2002), pp. 212-218

Kukelska-Hulme A, Traxler J (2013) Design principles for mobile learning. In Beetham and Sharpe (editors) *Rethinking Pedagogy in a Digital Age, Designing for 21st Century Learning* 2nd edition, Routledge 244-257

Kumaran, K. S., & Nair, V. M. (2010). Future trends in E-learning. *4th International Conference on Distance Learning and Education* 170-173.

Lancaster, S.J. (2013) The Flipped Lecture NDIR, Vol 9, Issue 1 28-32

Laurillard, D. (1993) *Rethinking University Teaching: a framework for the effective use of educational technology.* 1st edition Routledge

Laurillard, D., Stratfold, M., Luckin, R., Plowman and L. & Taylor, J (2000) Affordances for Learning in a Non-Linear Narrative Medium*. Journal of Interactive Media in Education*, 2000 (2) [www-jime.open.ac.uk/00/2]

Laurillard, D. (2002) Rethinking University Teaching: a framework for the effective use of learning technology. 2nd edition Routledge/Falmer

Laurillard, D. (2007) foreword, In Beetham, H., & Sharpe, R. (2007). *Rethinking pedagogy for a digital age: Designing and delivering e-learning*. London: 1st edition, Routledge.

Laurillard, D (2008) The teacher as action researcher: Using technology to capture pedagogic form, *Studies in Higher Education*, 33(2), 139-154.

Laurillard (2012) *Teaching as a Design Science Building Pedagogical patterns for learning and Technology.* 1st Edition, Routledge

Laurillard, D. (2013) foreword, In Beetham, H., & Sharpe, R. (2013). *Rethinking pedagogy for a digital age: Designing and delivering e-learning*. London: 2nd edition, Routledge.

Lave, J., & Wenger, E. (2003). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lockyer L, Heathcote E, and Dawson S, (2013) Informing Pedagogical Action: Aligning Learning Analytics With Learning Design*. American Behavioural Scientist* 57(10) 1439–1459

Luckhardt Redfield, C. and Steuck, K. (1991) The Future of Intelligent Tutoring Systems. In Burns, H., Parlett, J.W. and Luckhardt Redfield, C. *Intelligent Tutoring Systems Evolutions in Design.* Laurence Erlbaum Associates 265 -284

Machell, J.P. and Baruch, J.E.F. (2015) Evaluation Tools for Online Learning for the Bradford Robotic Telescope In *The International Congress on Education, Innovation and Learning Technologies,* Granada, 21 to 23 September 2015. http://www.scienceknowconferences.com/files/extended_abstracts/iceilt2015/?dir=Innovation%20and%20Learning%20Technologies accessed 9th July 2017

Majumdar, R. and Iyer, S. (2016) iSAT: a visual learning analytics tool for instructors*. Research and Practice in Technology Enhanced Learning* (2016) 11:16

Mangaroska, K., & Giannakos, M. N. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies* 1-1

Mayes, T. and de Freitas, S. (2004) Review of e-learning theories, frameworks and models. London: Joint Information Systems Committee. [online] available http://www.jisc.ac.uk/whatwedo/programmes/elearningpedagogy/outcomes.aspx accessed 9th July 2017

Mayes, T. and de Freitas, S. (2013) Technology Enhanced Learning the Role of Theory. In Beetham and Sharpe (editors) *rethinking pedagogy in a digital age, designing for 21st century learning* 2nd edition, Routledge. p17-30

Melero J., Hernández-Leo D., Sun J., Santos P. and Blat J., (2015) How was the activity? a visualization support for a case of location-based learning design. *British Journal of Educational Technology*, vol. 46, no. 2, pp. 317–329

Moodle (2016) About Moodle – MoodleDocs [online] available from https://docs.moodle.org/33/en/About_Moodle accessed 30 June 2017

Mor Y, Ferguson R and Wasson B (2015) Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action. *British Journal of Educational Technology* Vol 46, No 2, 2015, 221–229

National Archives (2010) *National Curriculum*[online] available http://webarchive.nationalarchives.gov.uk/20100202100434/http://curriculum.qcda.gov.uk/key-stages-1-and-2/subjects/science/keystage2/index.aspx access 8th July 2017

National Archives (2014) *Primary National Curriculum until 2014*. Department for Education http://webarchive.nationalarchives.gov.uk/20131202172639/http://www.education.gov.uk/schools/teachingandlearning/curriculum/primary

Nguyen Q., Rienties B., Toetenel L., Ferguson R., and Whitelock D., (2017) Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates. *Computers in Human Behavior*, vol. 76, pp. 703–714

Ofsted (n.d.) *Ofsted Reports* [online] available https://reports.ofsted.gov.uk/ accessed 8th of July 2017

Oliver, M., Roberts, G., Beetham H., Ingraham, B., Dyke M., Levy P. (2007) knowledge, society and perspectives on learning in Technology, In Conole, G. and Oliver, M. (editors) *Contemporary Perspectives in E-learning Research Themes Methods and Impact Practice* Routledge. 3-20

Oliver, M., Harvey, J., Conole, G. and Jones, A. (2007 b). Evaluation. In Conole, G. and Oliver, M. (editors) *Contemporary Perspectives in E-learning Research Themes Methods and Impact Practice* Routledge. 203-216

Oliver R., Harper B., Wills, S., Agostinho, S., and Hedbreg, J.G. (2013) Describing ICT-Based Learning Designs that promote Quality Learning Outcomes. In Beetham and Sharpe (editors) *Rethinking pedagogy in a digital age, designing for 21st century learning.* 2nd edition, 2013 Routledge p102-118

Owston, R. Lupshenyuk, D . and Wideman, H. (2011) Lecture capture in large undergraduate classes: Student perceptions and academic performance. *Internet and Higher Education* 14 (2011) 262–268

Persico, D. and Pozzi, F. Informing learning design with learning analytics to improve teacher inquiry *British Journal of Educational Technology* Vol 46, No 2, 2015, 230–248

PHP (n.d.) intro what is php [online] http://php.net/manual/en/intro-whatis.php accessed 6th July 2017

Postgresql (n.d.) about [online] available https://www.postgresql.org/about/ accessed 6th July 2017

Pursel B. K., Zhang L., Jablokow K. W., Choi G., and Velegol D. (2016) Understanding MOOC students: motivations and behaviours indicative of MOOC completion. *Journal of Computer Assisted Learning*, vol. 32, no. 3, pp. 202–217

Ray, N.L. (1992) Motivation in Education. Accessed ERIC Institute of Science [online] available https://eric.ed.gov/?id=ED349298 accessed 20th April 2017

Reiser R.A. and Ely D. P. (1997)The Field of Educational Technology as Reflected Through Its Definitions, *ETR&D*, Vol, 45, No. 3, 1997, 63-72

Research Council UK - RCUK (2005) *Evaluation: Practical guidelines* The Research Councils UK and The Office of Science and Technology. [online] available http://www.perfeval.pol.ulaval.ca/en/node/1405 accessed 3rd July 2017

Research Councils UK (2014) *Evaluation: Practical Guidelines. A guide for evaluating public engagement activities.* [online] available http://www.rcuk.ac.uk/Publications/policy/Evaluation/ accessed 8th July 2017

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behaviour*, 60, 333-341

Sambell, K., McDowell, L., and Montgomemery, C. (2013) *Assessment for learning in higher education,* Routledge

Sampson, D. Karagiannidis, C. and Cardinali, F. (2002) An Architecture for Web-based e-learning Promoting re-usable Adaptive Educational e-content. *Education, Technology and Society* 5(4) 2002 27-37

Schon, D. A. (1991). *The reflective practitioner: How professionals think in action*. Aldershot: Ashgate Publishing.

Scott, D. and Usher, R. (2011) *Researching Education Data, Methods and Theory in Educational Enquiry* 2nd edition Continuum International publishing group

Selwyn N. (2010) Looking beyond learning: notes towards the critical study of educational technology, *Journal of Computer Assisted Learning* (2010), 26, 65–73

Slavin, R.E. Cynthia, Hanley L.P., and Thurston A., (2014) Experimental Evaluations of Elementary Science Programs: A Best-Evidence Synthesis, *JOURNAL OF RESEARCH IN SCIENCE TEACHING* VOL. 51, NO. 7, PP. 870–901

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380-1400

Specht, M. and Oppermann, R. (1998) ATS _ Adaptive Teaching systems a WWW based ITS, IEE Colloquium on Web-Based Knowledge Servers, 1998 page 6

Statista (n.d.) Number of monthly active Facebook users worldwide as of 1st quarter 2017 (in millions) [online] available at https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ accessed 10th of July 2017

Stawser, M. (2009) Assessing Assessment: Toward a Hermeneutic-Phenomenological Perspective. *InSight: A Journal of Scholarly Teaching*. Volume 4, 56-68

Stewart, A., Storm, C. & VonEpps, L. (2013) Analyzing Student Confidence in Classroom Voting With Multiple Choice Questions, *PRIMUS*, 23:8, 718-732

Tallon, C.J. (2010) *A Dynamic Web Interface to a Remote Robot Evaluated with a Robotic Telescope.* Ph.D Thesis University of Bradford

Tavakol, M., Dennick, R. (2011) Making sense of Cronbach's alpha. *International Journal of Medical Education*. 2011; 2 53-55

Tavangarian D, Markus E. Leypold, Kristin Nölting, Marc Röser and Denny Voig (2004) Is e-Learning the Solution for Individual Learning? *Electronic Journal of e-Learning* Volume 2, Issue 2, 2004, 273-280

Taylor, J. Kowalski, S. Wilson, C. Getty, S and Carlson, J (2013) Conducting Causal Effects Studies in Science Education: Considering Methodological Trade-offs in the Context of Policies Affecting Research in Schools. *Journal of Research in Science Teaching* VOL. 50, NO. 9, PP. 1127–1141

Telescope.org (n.d.) [online] http://www.telescope.org/infopage.php?title=BRT_Tenerife_Hardware accessed 30th June 2017

Thomas, K. (2013) *Science teaching under the microscope. The Guardian* [online] https://www.theguardian.com/teacher-network/2013/may/21/science-teaching-professional-development accessed 20th June 2017

Thompson, W. J. (2001). Poisson distributions. Computing in Science & Engineering, 3(3), 78-82

Toetenel, L., & Rienties, B. (2016). Analysing 157 learning designs using learning analytic approaches as a means to evaluate the impact of pedagogical decision making: Analysing learning designs through learning analytic methods. British Journal of Educational Technology, 47(5), 981-992

Tractenberg, R.E., Gushta M. M., Mulroney, S.E. and Weissinger, P.A. (2013) Multiple choice questions can be designed or revised to challenge learners' critical thinking. *Adv in Health Sci Educ* 18: 945–961

TRIGWELL, K., MARTIN, E., BENJAMIN, J. and PROSSER, M. (2000) Scholarship of Teaching: a model. *Higher Education Research & Development*, Vol. 19, No. 2, 2000

Tweddell S. Clark D., Nelson, M. (2016) Team-based learning in pharmacy: The faculty experience *Currents in Pharmacy Teaching and Learning*. 8 (2016)7–17

Virvou, M, Maras D, Tsiriga V (2000) Student Modelling in an Intelligent Tutoring System for the Passive Voice of English Language Educational

Technology & Society 3(4) [online] available http://www.ifets.info/journals/3_4/virvou.html accessed 8th July 2017

Vygotsky, L.S. (1978) *Mind in Society the development of higher psychological processes*, Cambridge MA: Harvard University Press.

Waller, V (2004). The State of the Nation: e-learning in the 21st century**,** *Training Journal;* Apr 2004; Business Premium Collection pg. 30

Wang, H. C., & Chiu, Y. F. (2011). Assessing e-learning 2.0 system success. *Computers & Education.* 57(2), 1790-1800. doi:10.1016/j.compedu.2011.03.009

Wang, Y (2008), Assessing e-commerce systems success: a respecification and validation of the DeLone and McLean model of IS success. *Info Systems J* (2008) 18, 529–557

Wang Z, Anderson T, Chen L and Barbera E (2017) Interaction pattern analysis in cMOOCs based on the connectivist interaction and engagement framework. *British Journal of Educational Technology* Vol 48 No 2 2017 683–699 doi:10.1111/bjet.12433

Welsh M (2012) Student perceptions of using the PebblePad e-portfolio system to support self- and peer-based formative assessment. *Technology, Pedagogy and Education*, 21:1, 57-83, DOI: 10.1080/1475939X.2012.659884

Wikipedia (2017) *Wikipedia: Tutorial/Editing* [online] available https://en.wikipedia.org/wiki/Wikipedia:Tutorial/Editing accessed 3rd July 2017

W3C (2017) *Cascading Style Sheets* [online] available https://www.w3.org/Style/CSS/Overview.en.html accessed 3rd July 2017

Yang, Z. and Liu, Q. (2007) Research and development of web-based virtual online classroom. *Computers and Education* Volume 48, Issue 2, 171-184

Yonker J.E. (2011) The relationship of deep and surface study approaches on factual and applied test-bank multiple-choice question performance. *Assessment & Evaluation in Higher Education* Vol. 36, No. 6, 673–686

Youtube (n.d.) *about press* [online] available https://www.youtube.com/yt/about/press/ accessed 9th of July 2017

Zemeckis, R. (1990) *Back to the future part III* [film] Universal Pictures