



GRIFFITH COLLEGE DUBLIN

Student name:	Fergus Toolan		
Student number:	2932232		
Faculty:	Training and Education		
Course:	MA in Training & Education	Stage/Year:	2
Module:	Dissertation		
Study Mode:	Full time	Part-time	X
Lecturer/Supervisor Name:	Angela O'Keefe (Supervisor) – Peter Gillis (Module Lead)		
Dissertation Title:	Strategies for Improving Retention in Online Learning		
No. of pages (excluding this page):	79		
Disk included?	Yes	No	X

(e.g. number of pieces submitted etc.)

Academic Misconduct:

I understand that I will be subject to the penalties imposed for breaches of academic conduct as defined in the College's Academic Misconduct Procedure (QA J6).

I agree to allow Griffith College to deposit this thesis in electronic format in Griffith Online (go.griffith.ie), Griffith College's open access institutional repository on my behalf.

Signature: _____ *F. Toolan* _____ **Date:** _____ 15.04.2020 _____

Please note: Students MUST retain a hard / soft copy of all assignments and must SIGN the Assignment Submission Sheet provided by the lecturer / member of Faculty as proof of submission.

Strategies for Improving Retention in Online Learning

By

Fergus Toolan

Dissertation submitted in partial fulfilment of the requirements for MA in Training and Education (QQI)

Faculty of Teaching and Learning

Griffith College Dublin

April 2020

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of the MA in Training and Education, is my own; based on my personal study and/or research and that I have acknowledged all material and sources used in its preparation. I also certify that I have not copied in part or whole or otherwise plagiarised the work of anyone else, including other learners.

Signed: *F. Toolen*

Dated: 15.04.2020

Abstract

This research seeks to determine if methods exist to identify students in online education who are a retention risk and to develop solutions to help prevent said students from exiting the course prematurely. In order to do this effectively, this study addresses three specific questions:

- What data is provided by VLEs that might help educators to measure student engagement?
- To what extent are educators able to identify those students who are in danger of exiting a course prematurely in the online learning environment?
- What preventative measures are being used by educators to attempt to improve student retention in the online learning environment?

A qualitative approach is used to answer the above questions. Initially the documentation for the most popular VLEs is analysed to identify the information present that would allow educators measure student engagement. Following this online educators are interviewed in order to harness their thoughts and experiences in the identification of students who are a retention risk. The participants are also asked about their preferred strategies for preventing the early drop out of students in online learning.

The result of this primary research is to develop a set of recommendations, both for higher education institutions, and also for educators which aid in the identification of at-risk students. Additionally recommendations are provided for strategies that can be used, both proactively and reactively, in online learning to improve student retention.

Keywords: Online Learning; Virtual Learning Environments; Student Retention.

Table of Contents

CHAPTER 1: INTRODUCTION	8
1.1 MOTIVATION	8
1.2 RESEARCH QUESTIONS	9
1.3 DISSERTATION STRUCTURE	9
1.4 SUMMARY	10
CHAPTER 2: LITERATURE REVIEW.....	11
2.1 WHAT ARE E-LEARNING AND ONLINE LEARNING?.....	11
2.2 THE GROWTH OF ONLINE LEARNING.....	12
2.3 RETENTION IN HIGHER EDUCATION.....	14
2.4 IDENTIFYING AT-RISK CLASSROOM STUDENTS	17
2.5 IDENTIFYING AT-RISK ONLINE STUDENTS	19
2.6 LITERATURE DISCUSSION.....	21
2.7 SUMMARY.....	23
CHAPTER 3: METHODOLOGY	24
3.1 RESEARCH APPROACH.....	24
3.2 RESEARCH PARADIGMS	24
3.3 RESEARCH QUESTIONS	25
3.4 DATA COLLECTION.....	26
3.4.1 <i>Research Participants</i>	28
3.5 ANALYSIS METHODS.....	28
3.6 ETHICAL CONSIDERATIONS.....	29
3.7 SUMMARY.....	30
CHAPTER 4: COMPARISON OF VLE TRACKING DATA	32
4.1 BLACKBOARD	32
4.2 CANVAS.....	34
4.3 MOODLE	35
4.4 COMPARISON OF VLES.....	37
4.5 DISCUSSION	40
4.6 SUMMARY.....	41
CHAPTER 5: THEMATIC ANALYSIS	42
5.1 PARTICIPANTS.....	42
5.2 THEMATIC ANALYSIS	43
5.2.1 <i>Research Question 2: Analysis</i>	44

5.2.2	<i>Research Question 2: Discussion</i>	46
5.2.3	<i>Research Question 3: Analysis</i>	49
5.2.4	<i>Research Question 3: Discussion</i>	50
5.3	SUMMARY.....	52
CHAPTER 6: CONCLUSIONS & RECOMMENDATIONS		54
6.1	CONCLUSIONS.....	54
6.1.1	<i>Research Question 1</i>	54
6.1.2	<i>Research Question 2</i>	54
6.1.3	<i>Research Question 3</i>	55
6.2	RECOMMENDATIONS.....	56
6.3	LIMITATIONS.....	58
6.4	FUTURE WORK	58
6.5	SUMMARY.....	59
REFERENCES		60
APPENDIX A: ETHICAL APPROVAL		70
APPENDIX B: PARTICIPANT INFORMATION SHEET		71
APPENDIX C: CONSENT FORM		72
APPENDIX D: INTERVIEW QUESTION GUIDE		74
APPENDIX E: THEMATIC ANALYSIS RESEARCH QUESTION 2		76
APPENDIX F: THEMATIC ANALYSIS RESEARCH QUESTION 3		77

Acknowledgements

First and foremost let me thank my supervisor, Dr. Angela O’Keefe. Angela provided stimulating discussions and prompt feedback throughout the dissertation process.

Special mention must be made of Dr. Fiona O’Riordan, who sparked my original interest in the scholarship of teaching when we worked together many years ago.

My classmates were a great source of support throughout the tough times during the dissertation process. The WhatsApp group solved lots of problems as the deadline loomed!

To the members of the *Reflective Drinking Society*, which places critical thinking about teaching as one of the aspirations for all – with some drinking also! While there were many members over the years, Henry, Kevin, Ray, and Dave deserve thanks for discussions!

From a career perspective, I wish to thank my employers, Politihøgskolen (The Norwegian Police University College). Particularly my bosses, old and new, Ivar Husby and John Ståle Stannes, who supported me throughout the masters. My direct colleagues, Georgina, Ivar, Kurt, Rune, Ulf, and Yves a fantastic group of teachers always willing to share their expertise.

I must also thank the participants in this research. The insightful points that are found in this document are the result of years of experience and passion for teaching and learning.

Occasionally, the stress of work and study becomes too much for everyone, on those days I would adjourn to Liam O’Riains for beer and banter. Special mention must be made of Ken, Kerry and Liam who would cheer me up by never once speaking of training or education!

To my parents Anna and Donal who have always supported my desire for education, thank you!

Finally to my partner Helen, thanks for the support, encouragement and proof-reading throughout the process. I could not have done this without you!

List of Abbreviations

- **CA:** Continuous Assessment
- **CAO:** Central Applications Office
- **EDM:** Educational Data Mining
- **EQF:** European Qualifications Framework
- **GPA:** Grade Point Average
- **HE:** Higher Education
- **HEA:** Higher Education Authority
- **LMS:** Learning Management System
- **MOOC:** Massively Open Online Course
- **NFQ:** National Framework of Qualifications
- **NLP:** Natural Language Processing
- **T&L:** Teaching and Learning
- **VLE:** Virtual Learning Environment

Chapter 1: Introduction

Online learning, while in academic timeframes a new phenomenon, is growing at a staggering rate. Data from the US shows that more than one in three higher education students are taking one or more courses online as of Autumn 2017 (Ginder et al., 2019). The same study showed that nearly one in five students were taking their entire programme online. Less than one in ten were taking one or more online courses in 2002 (Allen and Seaman, 2010). The benefits of online learning are substantial. For the student, the flexibility inherent in the delivery mechanism, specifically in asynchronous e-learning, is often the only reason that students are able to attend higher education (Jaggers, 2014). For the institution, not requiring physical space for students is a huge advantage in terms of cost of programme delivery and also allows the institution to potentially appeal to a global market (Harris and Martin, 2012).

However, all higher education institutions are concerned with the problem of student retention. Lack of retention has implications for the student, the institution and also society as a whole (O’Keefe, 2013). Patterson and McFadden (2009) discovered that the rate of attrition in online courses is approximately six to seven times greater than that of the traditional university course. Many studies have posited that one of the reasons for this is the lack of a sense of community in online courses (Tinto, 2006; O’Keefe, 2013; Gaytan, 2015). Regardless of the reason, the high attrition rate in online learning, combined with the growth of online learning, is a grave concern to the academic community.

1.1 Motivation

With the growth in online learning, and the move toward life-long learning, it is essential that research focuses, not just on retention as a whole, but more specifically on retention in the online environment. More and more students are joining online education programmes and many are failing to progress. It is necessary for academia to, not only understand the reasons for this attrition, but also to be able to detect students who are at-risk at the earliest possible time. In the traditional classroom this detection could be achieved using simple metrics such as attendance. Numerous studies found a direct link between attendance and performance, for instance Crede et al. (2010) performed a meta-analysis of 68 research papers over an 82 year time frame and showed a direct correlation between attendance and performance. In

online learning there are no simple metrics that reliably capture this correlation. This provides one of the strongest motivations for pursuit of this research.

Additionally, the researcher has been lecturing almost exclusively in the online environment for over ten years. One particular introductory course (at masters level) taught by the researcher has an attrition rate of 20 – 25%. Generally the researcher only identifies those students who are struggling when they make themselves known. It is difficult to predict which students will be at-risk of dropping out. For this reason the researcher is particularly interested in this aspect of educational research.

1.2 Research Questions

The main focus of this research is in the area of retention of online higher education students. The aim of the research is to explore how to identify students who are in danger of prematurely exiting courses in the online learning environment, and identify strategies to help prevent their exit. In order to answer this question three sub-questions are examined:

- **Research Question 1:** What data is provided by Virtual Learning Environments (VLE) that might help educators to measure student engagement?
- **Research Question 2:** To what extent are educators able to identify those students who are in danger of exiting a course prematurely in the online environment?
- **Research Question 3:** What preventative measures are being used by educators to attempt to improve student retention in the online learning environment?

1.3 Dissertation Structure

This dissertation is organised as follows. Chapter 2 describes the literature relevant to this research. Specifically, it defines online learning as used in this research and examines its growth in recent years. It also examines the current retention problem in higher education and governmental strategies to address this issue. It then proceeds to examine techniques that have been used in the traditional classroom system to identify students who are a retention risk. Finally, existing methods, used in the online environment, to identify at-risk students are examined.

Chapter 3 describes the methodologies used in this research. These methodologies are described in terms of the research paradigms that underpin them. Additionally this chapter describes the participants in this study showing their experience in the area of online learning. The chapter also discusses the ethical considerations that were identified in preparation for this research and shows how they have been addressed.

Chapters 4 and 5 present the findings of this research. In Chapter 4 a document analysis related to the first research question is conducted. This document analysis identifies possible data sources in the online environment which might be used to measure student engagement. This chapter discusses the utility of the identified data in predicting at-risk students. Chapter 5 presents the findings relevant to the remaining research questions. This provides a thematic analysis of the interviews conducted with online educators as part of this research and a discussion of those results.

Finally, Chapter 6 provides a summary of the key findings of this research. It also provides recommendations, both for educators and institutions, of strategies for identifying and helping students who are a retention risk in the online environment. Additionally it describes some of the limitations that were inherent in this work, and discusses the possibility for future research in this area.

1.4 Summary

This chapter provided an introduction to the subject area of this research project and also described the motivations for conducting said research. The chapter proceeded to describe the aim of the research and in doing so identified three research questions that will be addressed. Finally the structure of the thesis was described.

Chapter 2: Literature Review

The literature review is a vital piece of all academic research as *'without it you will not acquire an understanding of your topic, of what has already been done, how it has been researched, and what the key issues are'* (Hart, 2018, p. 1). In terms of this study the literature is composed of five distinct areas of research. Firstly the chapter focuses on e-learning and online learning and attempts to define and distinguish between these concepts. Secondly the growth in online learning in recent years and the projections for the future are discussed, additionally the reasons for this growth are examined. Thirdly, the issue of retention in third-level education is examined and some of the factors that are considered to impact retention levels are identified. Following this, previous work in identifying students who are at-risk of dropping-out / failing in the traditional classroom setting is discussed. Some of the metrics that have been used to achieve this are examined to see how successful they have been. Finally the notion of identifying at-risk students in the online learning environment is examined and previous work in this area is discussed. The chapter concludes with a discussion of the literature consulted throughout the research.

2.1 What are E-Learning and Online Learning?

Countless definitions of E-Learning have been proposed over the years. For instance Sun et al. (2008, p. 1183) define e-learning as *'the use of telecommunications technology to deliver information for education and training'*. Welsh et al. (2003, p. 246) define e-learning as *'the use of computer network technology, primarily over an intranet or through the internet, to deliver ... instruction to individuals'*. Sanders (2006, p. 1) defines e-learning as *'the delivery of learning via any form of electronic media'*. Many other definitions have appeared however, all share two criteria, that e-learning always involves the delivery of instructional material and that technical means are used to deliver this material.

Online learning always involves an online component (Moore et al., 2011) thereby distinguishing it from e-learning which merely involves some form of electronic media, although many e-learning definitions do include networks as the form of technology (Welsh et al., 2003; Sun et al., 2008). This implies that with online learning students are remote, and access material through the internet. From the criteria identified previously online learning

can now be described as e-learning in which the internet is the technical means used to deliver instructional material.

Allen and Seaman (2007) classify courses into four distinct types: online; blended / hybrid; web facilitated; and traditional. The classification is based on the amount of content that is delivered via online means. By this classification, in order to be considered online, at least 80% of course material must be delivered via the internet. In order to be considered blended or hybrid, courses would have between 30% and 79% of material delivered online, while those that have 1% to 29% are considered web facilitated. The notion of web facilitated would include traditional courses in which a content management system is used to share course notes. Traditional courses have no online component present.

Online learning can also be classified as synchronous or asynchronous (Sife et al., 2007). In synchronous learning teachers and students *'interact at the same time though they may be dispersed geographically'* (Sife et al., 2007, p.58) whereas asynchronous learning *'supports work relations among learners and with teachers even when participants cannot be online at the same time'* (Hrastinski, 2008, p. 1).

However, there is still some confusion in the literature regarding the definitions of online learning, e-learning and distance learning with these terms *'often interchanged without meaningful definition'* (Moore et al., 2011, p. 129). Due to this fact some of the previous work referred to in this literature review, refers to the concept of e-learning, when in fact online learning is a more appropriate title.

2.2 The Growth of Online Learning

Online learning has grown in recent years to have become one of the main mechanisms of educational delivery presently. Ginder et al. (2019) provide figures for the US market in which they state that in Autumn 2017 17.4% of students in the US were enrolled exclusively in online learning programmes. Furthermore there were another 17.6% of students who were enrolled in some, but not all, online learning courses. Taken together it means that over one in three college students in the US were taking at least one online course as part of their studies in

Autumn 2017. Compare this to Autumn 2008, less than 10 years previously, where 25% of students were taking one or more online courses, while in Autumn 2002, a mere 9.6% of students were taking any online courses (Allen and Seaman, 2010).

One of the many reasons for this growth in online learning is at the policy level. This can be seen at an institutional and national level. For instance in Ireland, at the national level, the Higher Education Authority's strategic plan (HEA, 2018) identifies Digital Transformation as one of its strategic themes. One aspect of digital transformation identified by the HEA is that more courses will be delivered in the online environment. Similarly the National Strategy for Higher Education (Department of Education, 2011, p. 17) states that "*the delivery of higher education in Ireland must be characterised by flexibility and innovation*". One of the delivery mechanisms outlined in this strategy is that of online learning, allowing students to learn in full- or part-time capacities at times and places of their choosing.

Many institutions are also addressing online learning in their strategic plans for instance, Trinity College Dublin aim to '*use flexible and wholly online modes of course delivery to increase the numbers of student online learners*' (Trinity College Dublin, 2014, p. 25). Similarly University College Cork (2017, p. 28) aims to '*strengthen the provision of online programmes and enhance access to lifelong learning*'. However, it is not only institutions and national policy that is driving the growth in online learning. The student is responsible for the growth also.

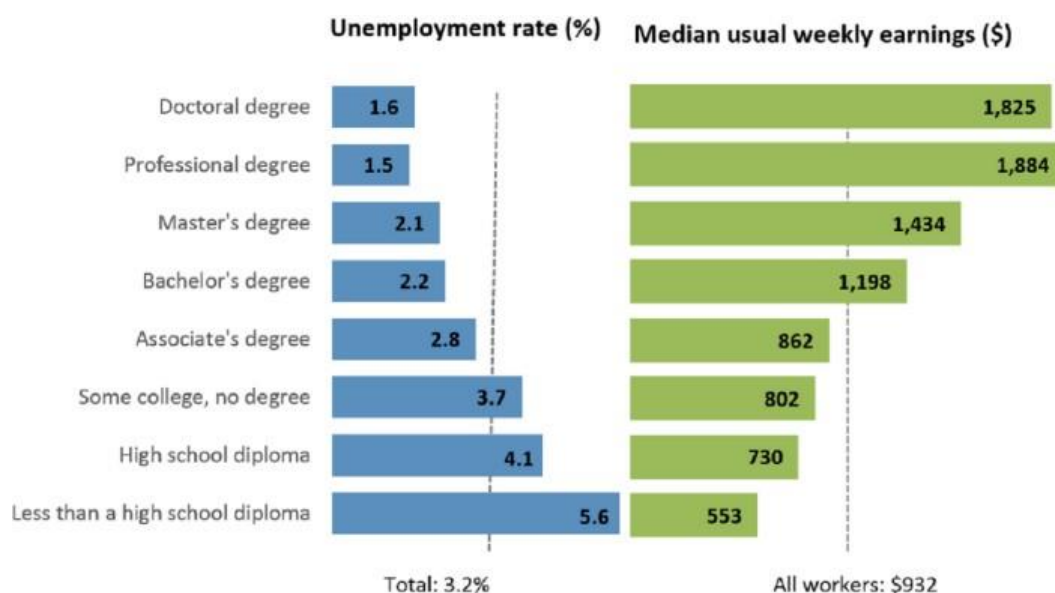
Many studies have examined the reasons that students wish to take online courses over traditional classroom-based courses. Almost all of these have discovered that the flexibility and convenience of online delivery is one of the major factors in their popularity (Harris and Martin, 2012; Jagers, 2014). Many students in the online environment are older than the traditional college student. Harris and Martin (2012) showed that in the 40+ age group in one Canadian College there were more than 4 times the number of online students as classroom students. This is in contrast to the 18 – 22 age group in which 5 of every 6 students were in classroom-based programmes. Harris and Martin (2012) also showed that distance to college was another common reason for the choice of online learning over classroom-based learning. However, only 17% of respondents said that they preferred online learning to classroom

learning, whereas 65% of respondents in classroom-based courses preferred that method of delivery.

2.3 Retention in Higher Education

The word retention comes from the verb to retain, meaning to keep possession of (Dictionary.com, 2020). Swecker et al. (2013, p. 47) define student retention as *‘the continuous enrollment of students from one fall semester to the following fall semester’*. In the higher-education domain it has become an increasingly important topic, with many institutions’ funding dependent on graduating students (Tinto, 2006).

Poor student retention affects many stake holders in society. These include the actual student, the higher education institutions and society as a whole. The effect on the student can be catastrophic. Numerous studies have shown a link between educational attainment and career prospects. For instance Figure 1 (US Bureau of Labour Statistics, 2019) shows the link between educational attainment and salary (right) and unemployment (left). It can be seen that those with the highest level of educational attainment are more likely to be employed and also will earn over three times the salary of those with little education.



Note: Data are for persons age 25 and over. Earnings are for full-time wage and salary workers.
 Source: U.S. Bureau of Labor Statistics, Current Population Survey.

Figure 1: Unemployment and earnings by educational attainment in the US in 2018.

From Figure 1 we can also extrapolate one of the effects that student attrition has on society as a whole. Higher unemployment requires that a greater percentage of the tax base be spent on social welfare. Additionally, socioeconomic status regardless of measurement '*by income, education or occupational status, is amongst the most robust determinants of variations in health outcomes*' (Williams et al., 2016, p. 407). The causal linkage between educational attainment and health adds further cost to the taxpayer based on the levels of retention, or lack thereof, in higher education.

Finally higher-level institutions are also affected through the inability to retain third-level students. In certain cases the funding for the institution is directly dependent on the numbers of graduating students (Tinto, 2006). In other cases low retention rates lead to reputational damage to the institution and, as a consequence, less students applying to that institution in the future (Lau, 2003).

But how large is the problem? It is acknowledged that the United States has one of the lowest retention rates in the developed world, with O'Keefe (2013) suggesting this may be as high as 30% - 50% when all factors are considered. In Ireland the Higher Education Authority commissioned a study on 41,441 first year students entering the third level education system between March 2015 and March 2016 (Liston et al., 2018). This study discovered that 14% of these students did not progress beyond first year. However, the full rate of attrition might be higher than this, further studies are required to determine how many more dropped out in subsequent years of their education.

Liston et al. (2018) discovered that retention rates in certain disciplines were above 20%, in particular construction and related topics. They also found significant differences between the level of course. Higher attrition rates were discovered in courses at NFQ levels 6 and 7 rather than at level 8. They conclude that this is most likely related to past educational attainment, as the level 6 and 7 courses generally require less points to enter. To continue this they discovered that the lowest rate of attrition was found in medicine, in which the majority of students had near maximum CAO points.

The trend of past educational attainment being a good predictor of student retention (Liston et al., 2018), is not only a factor in Ireland. Internationally past educational attainment has been linked with retention in numerous studies. For instance Fike and Fike (2008), in a study of 9,200 college students in the US, discovered a correlation between those that pass basic academic skills tests (reading, writing and mathematics) and those that complete their college education.

In addition to past educational attainment other factors have been linked to poor retention rates. These include: personal and financial circumstances (Lau, 2003; O’Keefe, 2013; Lisciandro and Gibbs, 2016); motivation (Lau, 2003; Friedman and Mandel, 2011); and the sense of community or belonging (Lau, 2003; O’Keefe, 2013; Swecker et al., 2013). Personal and financial strains are commonly cited as causes of student attrition. These include family issues / responsibilities, financial difficulties, medical issues, and emotional issues (Lau, 2003; Lisciandro and Gibbs, 2016). Lisciandro and Gibbs’ (2016) study found that these were the reason for over 70% of drop-outs. However, of these personal circumstances, financial was the least important (only 7%).

Motivation, or lack thereof, is another commonly cited reason for attrition in higher education. Lau (2003) claims that the lack of motivation is linked to an inability on the student’s part to understand the importance of education or to apply classroom learning in the real-world. Motivation has been cited as one of the main reasons for student attrition. Tinto (2006) stated that in the early years of research in the area of retention, many institutions adopted a ‘blame the victim’ approach in which lack of motivation was the key factor. Friedman and Mandel (2011) examined performance in terms of motivation and discovered motivation at the start of college predicted student GPA in first year.

Another category commonly identified is that of the sense of community or belonging. According to O’Keefe (2013, p. 607) *‘developing a sense of belonging is critical to the success of college students’*, however, many factors are at odds with this sense of belonging. For instance O’Keefe (2013) mentions that part-time students and in particular online students feel more removed from the college community. Additionally diversity of the student body means more students have personal and financial commitments beyond those traditionally

associated with college students, leading to a sense of disconnect from the wider college population. Tinto (2006) argues that the student's involvement in the classroom is important for retention. In other words the student must be engaged in the classroom.

The majority of the previously cited work refers to traditional classroom settings, but the online environment is different. Patterson and McFadden (2009) found that attrition in US college degree programs was six to seven times higher in the online environment than in the traditional classroom. Gaytan (2015) identified a number of factors that led to this increase in attrition rates. This study found that the student's motivation was considered of vital importance by faculty members. Yet students themselves viewed the quality of student / faculty interactions and the institutional support provided to students as being the most important factors influencing attrition. This finding links strongly with Tinto's (2006) belief that the sense of belonging or community is vital for student retention.

2.4 Identifying At-Risk Classroom Students

Numerous means of identifying at-risk students in the classroom have been proposed over the years. These include attendance (Crede et al., 2010; Schneider and Preckel, 2017), previous academic performance, student motivation, and socioeconomic status (McKenzie and Schweitzer, 2001).

Attendance has often been seen as a means of predicting both student retention and achievement in traditional classroom-based teaching. Crede et al. (2010) performed a meta-analysis of published research on the question of classroom attendance and its correlation with student achievement. The scope of the meta-analysis was literature published between 1927 and 2009. They discovered that attendance and achievement (and hence retention) were directly correlated, and that this correlation had not changed over the study's duration. In total their study analysed 68 papers over the 82 year period. Based on their analysis they concluded that "*class attendance appears to be a better predictor of college grades than any other known predictor*" (Crede et al., 2010, p. 288).

Previous academic performance has often been linked with academic achievement in higher education (McKenzie and Schweitzer, 2001). Most countries have a competitive examination-based system for admission to higher-level education (e.g. the Irish Leaving Certificate, the UK A-Levels etc). Performance in these end of school examinations has been shown to have a direct correlation with academic performance (McKenzie et al., 2004; Day et al., 2010). While McKenzie et al. (2004) used previous academic performance to predict achievement at first year level, Day et al. (2010, p. 552) showed that A-Level scores *'shared a significant positive correlation with final degree mark'*. There is, however, one caveat with this result, for certain groups of students it does not have the same predictive ability (McKenzie and Schweitzer, 2001). McKenzie and Schweitzer (2001, p. 4) state that *'secondary school grades are not as good predictors for mature student's performance as they are for school leaver's performance'*.

Another commonly cited means of predicting students who are in danger of dropping out / failing is through the analysis of student motivation (Pintrich et al., 1993; Busato et al., 2000). Pintrich et al. (1993) examined the predictive power of motivation using the Motivated Strategies for Learning Questionnaire (MSLQ) with 380 college students across a range of disciplines and showed that *'motivational sub-scales showed significant correlations with final grade'* (Pintrich et al., 1993, p. 810). Busato et al. (2000) studied 409 psychology students and found that *'achievement motivation was associated positively with academic success'* (Busato et al., 2000, p. 1064).

Socioeconomic status is generally agreed to incorporate parental income, parental education, and parental occupation (Park and Bauer, 2002). A longitudinal study involving over 100,000 students Sirin (2005, p. 438) found that socioeconomic status is *'one of the strongest correlates of academic performance'*. However, this result was determined to be most applicable to school leavers entering higher education, other groups of students (e.g. mature students) did not show as strong a correlation between socioeconomic status and retention rates.

2.5 Identifying At-Risk Online Students

As the popularity of online learning increases (Ginder et al., 2019) predicting achievement in the online learning environment has become more important. However, it presents challenges that are not present in the traditional classroom environment. According to Baker and Inventado (2016) the online educator has more challenges than the local educator in identifying at-risk students. This is due to the lack of direct interaction between student and educator. As a result, educational data mining (also known as learning analytics) has become a large area of research in the past decade (Minaei-Bidgoli et al., 2003; Hämäläinen and Vinni, 2006; Romero et al., 2008; Romero et al., 2013; Jiang et al., 2014; Shahiri and Husain, 2015; Baker and Inventado, 2016; Sclater et al., 2016). Educational data mining is *'an emerging discipline, concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings and using those methods to better understand students'* (EDM Society, 2020, p. 1). This has been enabled by the growth in data that is being stored about interactions in the online learning environment.

VLEs store a vast amount of data about each student interaction with the course material. This is contained in web logs, semi-structured textual data, which contain all requests made to the web server. This includes every student login, page view, discussion posting, etc. There are a variety of VLEs in use today such as: Moodle (Moodle, 2020a); Canvas (Instructure, 2020a); Brightspace (D2L, 2020a); Blackboard (Blackboard, 2020a); etc. All of these contain logging data which can be used for the purpose of educational data mining. Sclater et al. (2016) view the VLE as being the main source of data for educational data mining and learning analytics.

While educational data mining has many uses (Baker and Inventado, 2016) the one of most interest in this research is that of prediction. In prediction a model is developed for inferring a particular aspect of the data (Baker and Inventado, 2016), in this case student retention and achievement are the aspects to be inferred.

Educational data mining has been successful. Hämäläinen and Vinni (2006) achieved over 80% accuracy in predicting if students would pass or fail a course. They were also able to predict

this half way through the course allowing sufficient time for beneficial intervention. Minaei-Bidgoli et al. (2003) also achieved over 80% accuracy in predicting whether students pass or fail a course. Additionally they had a 33% accuracy in predicting a student's grade. Romero et al. (2008) achieved accuracies of over 65% in predicting students' grades and final marks.

The issue with educational data mining is the difficulty of the approach for non-technical instructors. For instance Hämmäläinen and Vinni (2006) used a Naïve Bayesian approach to classification which is based on the Bayesian theorem related to prior probabilities.

$$P(\#_s | \tau) = \frac{P(\#_s)P(\tau | \#_s)}{P(\tau)}$$

In order to implement Naïve Bayes classification one must understand, and implement, the above equation. Minaei-Bidgoli et al. (2003) used genetic algorithms in their system while Romero et al. (2008) used multiple algorithms such as: Neural Networks; Decision Trees; Nearest Neighbour; and Linear Regression. All of these algorithms are highly mathematical in nature and very difficult to implement. Also, not only does the algorithm need to be implemented, but generally data has to be pre-processed (Dutt et al., 2017), another non-trivial task. Hence these techniques are not directly usable by non-technical educators.

A limited amount of work has been performed on some means of predicting student retention and achievement in the online environment using less technical means. Robinson et al. (2016) analysed student survey data prior to a Massively Open Online Course (MOOC) and ran Natural Language Processing (NLP) algorithms over this data. This was able to generate sets of key words that students used and correlate these keywords with success or failure¹. While the NLP algorithm is complex it only needs to be run once (unless a course changes), non-technical educators could look for the presence / absence of these keywords in survey forms in the future in order to attempt to predict achievement. However, it should be noted that there were over 40,000 students in the MOOC meaning analysis would very time consuming but still technically feasible.

¹ Some of the words that corresponded highly with success were 'innovate', 'engage', and 'impact'.

A study conducted in Nottingham Trent University (NTU) (JISC, 2019) found that engagement with the VLE was one of the most important factors in predicting achievement. They found that those with consistently 'good' engagement throughout the semester were much less likely to fail than those students who were judged to have had poor engagement. NTU uses commercial software to rate student engagement and as such the algorithms are unknown and hence engagement is not sufficiently well defined.

2.6 Literature Discussion

This section has described some of the literature relevant to the area of student retention and achievement and has helped to motivate the specific research questions addressed in this research. Retention has long been considered an issue in higher education, however, with the unprecedented demand for online education in recent years (Ginder et al., 2019), both from students and government, it has become an even larger issue. Patterson and McFadden (2009) identified that the attrition rate in online learning is six to seven times that of traditional learning.

All of the factors identified for traditional classroom-based learning apply in the online domain, for instance student motivation is still considered important (Gaytan, 2015). However, many of the factors that lead to attrition in the classroom are magnified in the online domain. For instance online learning often involves a more diverse group of students, often older students with family commitments (O'Keefe, 2013). The flexibility of online learning means that they are able to study, but their increased personal responsibilities can make this more difficult. Hence, this group are more likely to fail to complete a course of study. Similarly, the sense of community identified by many researchers (Tinto, 2006; Gaytan, 2015) is more difficult to achieve. Gaytan's (2015) study showed that students perceive there to be very little contact between faculty and students in the online environment compared to the traditional classroom. Classroom-based teaching provides a basic sense of community through physical presence at the same location. The same sense is not present in the online environment.

Current research in predicting those who will succeed in online learning requires high technical knowledge, and generally an advanced understanding of complex mathematical topics such as machine learning and data mining (Baker and Inventado, 2016). For non-technical lecturers this is potentially a step too far, lecturers are unable to generate the prediction rules and most likely would not understand any of the generated rules.

Many, simpler, success predictors have been proposed for the traditional classroom. The aim of this research is to determine if similar predictors can be defined for online learning. It is expected many of the same predictors would be valid. For instance student motivation will still play a part in success in the online learning environment, although it is often considered more difficult to motivate students in online learning (Muilenburg and Berge, 2005). However, student motivation levels are gathered through use of complex questionnaires which are time-consuming and often have a poor response rate (Glynn et al., 2011). Hence, it is difficult to use this predictor as we would generally have incomplete knowledge of our students and may not have the time to implement the survey.

Other predictors that were discussed include socio-economic status and previous academic performance (McKenzie and Schweitzer, 2001; Lau, 2003; Tinto, 2006). Both of these have shown success in the traditional classroom setting, however, they may not be beneficial in terms of online learning. Online learning can be utilised in terms of higher education, where this information is most likely available to the educators. However, it can also be utilised as part of a Massive Open Online Course (MOOC), in which case many courses will not record any of this information about a student, and hence predictors based on this information cannot be used.

These predictors become less useful with certain groups of mature students in particular (McKenzie and Schweitzer, 2001; Sirin, 2005) as the desire and motivation for education may have changed based on the student's lifetime and experience. However, this group of students are more heavily represented in the online learning environment than the traditional 18 – 22 year old student (Harris and Martin, 2012), meaning it is unlikely that these predictors will be of great use in the online environment.

The final predictor that was examined was that of attendance. Attendance has long been linked with academic success (Schneider and Preckel, 2017) and this information can easily be gathered by a lecturer in the traditional classroom setting. However, it is not easy to develop a simple metric such as attendance for the online environment. Most of the literature relating to prediction in the online environment is based on educational data mining which is inherently difficult to utilise and even to understand (Minaei-Bidgoli et al., 2003; Hämmäläinen and Vinni, 2006; Romero et al., 2008; Dutt et al., 2017). However, some work in the area has shown that some simple metrics may be the most beneficial, for instance student engagement with the VLE (JISC, 2019). This leads to the overall purpose of this research, to determine if strategies exist which will prevent students from prematurely exiting an online course of study.

2.7 Summary

This chapter provided the setting for the research that follows. Initially e-learning and online learning were described and an attempt to distinguish between them was made. The lack of agreement on definitions means that the terms are often used interchangeably. Secondly the current state of the online learning environment, in terms of the recent growth in that market and the predictions for the future were discussed. It appears that online learning is still growing in popularity and will be present for many years to come. Subsequently the issue of retention in third-level education was examined along with the reasons for the current student attrition rates. It was also discovered that the attrition rates in online learning are higher than those in traditional classroom learning. Finally some of the means that have been considered for predicting students who are unlikely to succeed in education, both in the traditional classroom environment and also in the online environment, were examined.

Chapter 3: Methodology

This chapter examines the methodological paradigms that underpin this research and the actual data collection and analysis methods that will be used throughout. Before proceeding, it is essential to distinguish between methods and methodology. Gabriel (2011, p. 1) describes a method as *'simply a research tool'*. This would include a qualitative interview or a quantitative statistical analysis technique. Methodology, on the other hand, is the *'justification for using a particular method'* (Gabriel, 2011, p. 1).

3.1 Research Approach

Kothari (2004) considers that there are two main approaches to research, quantitative and qualitative. He defines the quantitative approach as *'the generation of data in quantitative form which can be subjected to rigorous quantitative analysis in a formal and rigid fashion'* (Kothari, 2004, p. 5). Conversely he defines the qualitative approach as being concerned with *'subjective assessment of attitudes, opinions and behaviour'* (Kothari, 2004, p. 5). Mixed methods research is a third approach which aims to draw on the strengths and minimise the weaknesses of both quantitative and qualitative approaches (Johnson and Onwuegbuzie, 2004). This research comprises two distinct threads. Firstly relevant documentation is analysed to gather information about the various VLEs in use in higher education, and in particular the information that is available in these systems that might help to measure engagement. Secondly this research utilised semi-structured interviewing of a number of experienced educators in the online learning environment, and as such it is a qualitative approach that best suits this research.

3.2 Research Paradigms

Hammersley (2012) describes four methodological paradigms of research: positivism; interpretivism; critical research; and constructivism. Positivism is based on the belief that science is the only source of knowledge and that all social and educational research should use the models of research proposed in physics and chemistry (Hammersley, 2012). The positivist paradigm generally involves empirical data upon which laws are logically inferred. The interpretivist paradigms believes that there is *'a fundamental difference between the nature of the phenomena investigated by the natural sciences and those studied by ...*

educational researchers' (Hammersley, 2012, p. 22). One of these differences is that it is difficult to fully understand what people do or why they do it. Interviews are one means of conducting interpretivist research. Critical research focuses on perspectives of society and whether the research can provide a practical guide for transforming society (Hammersley, 2012). Finally constructionism can be viewed as a radical form of interpretivism which believes that knowledge is not insight but instead is constructed through the social interactions of people (Hammersley, 2012).

The main methods used throughout this research involve document analysis and interviews with experienced online educators, to gain their insight into the issues of retention and achievement, and as such the interpretivist paradigm appears to be the most suitable for this work. However, it could also be argued that the initial document analysis could be considered positivist in nature, the data is available, and an attempt is made to draw conclusions from that data.

Having now identified that the interpretivist paradigm is most appropriate for this research, the next step is to identify the means of logical inference to be utilised. Generally there are two methods of logical reasoning in scientific research: inductive and deductive. Inductive reasoning is the '*process of reasoning from a part to a whole*' (Kothari, 2004, p. 20). Deductive reasoning, on the other hand, is '*the process of reasoning from some premise to a conclusion which follows from that premise*' (Kothari, 2004, p. 20). The choice of reasoning is therefore based on whether we begin with data (inductive) or with a theory (deductive). In this case we will begin with the interview transcriptions as the data set, and hence inductive reasoning will be the most applicable.

To summarise, the methodology that will be used in this research is a qualitative approach to research, involving the interpretivist paradigm and utilising inductive reasoning.

3.3 Research Questions

As stated in Chapter 1, the aim of this research is to discover how to identify students who are in danger of dropping out / failing in the online learning environment, and after identifying

said students develop strategies to help them. The literature review showed some of the techniques that have been used in the classroom and also some of the, admittedly very complex, techniques that have been used to identify these students in the online environment. Based on the diversity of students who attend online learning programmes some of the traditional predictors such as socioeconomic status may not be as effective. Previous research has shown that the concept of 'engagement' is linked with retention (JISC, 2019), which leads to the first research question:

Research Question 1: What data is provided by VLEs that might help educators to measure student engagement?

Studies have shown that faculty can struggle to identify students who are in danger of dropping out / failing in the online environment. It would be desirable to determine what techniques are being used in the online domain in order to identify these students. This leads to the second research question:

Research Question 2: To what extent are educators able to identify those students who are in danger of exiting a course prematurely in the online environment?

Additionally the aim of this research is to determine what can be done to improve the retention rate in online learning. This leads to the final research question:

Research Question 3: What preventative measures are being used by educators to attempt to improve student retention in the online learning environment?

It is hoped that in answering these research questions it will then be possible to describe what is currently being done, and also to suggest best-practice to other online educators.

3.4 Data Collection

This section looks at the methods of data collection that will be employed throughout this research. The first research question asks what data is provided by VLEs that would aid

educators measure student engagement in the online learning environment. The method of data collection for this, involves reading a number of manuals for various online learning platforms. The platforms chosen for this part of the research are: Canvas (Instructure, 2020a); Blackboard (Blackboard, 2020a); and Moodle (Moodle, 2020a). These three platforms were chosen due to their popularity. In Autumn 2019, they represented the three most common Learning Management Systems (LMS) in use across the US higher education system, accounting for almost 75% of LMS usage (Edutechnica, 2020).

Research questions two and three aim to determine what efforts are being made by online educators to identify students who are at-risk and also what preventative measures are being used to aid student retention. In order to answer these questions interviews are conducted with a number of experienced online educators from the third-level sector.

Generally, interviews are divided into three categories based on the level of structuring that is present (Zhang and Wildemuth, 2009a). These categories are: structured; semi-structured; and unstructured. The structured interview is more akin to a spoken survey than an actual interview. The interviewer has a set of pre-defined questions which are asked in the same order to all participants (Zhang and Wildemuth, 2009a). Semi-structured interviews are more flexible. With this technique the interviewer has an interview guide, generally containing open and closed questions. The interviewer is free to deviate from the order of questions in this guide and also to ask other questions based on circumstance (Lao and Wildemuth, 2009). Unstructured interviews rely on '*social interaction between the researcher and the informant*' (Zhang and Wildemuth, 2009a, p. 222), meaning that no questions are specified in advance.

For this research the method of semi-structured interviews will be used. There are two reasons for this. Firstly, a certain level of structure is required so that answers can be compared between participants and themes can be discovered. An unstructured approach would make the identification of themes more difficult, if not impossible. Secondly, the participants in this study have a wealth of experience which can be drawn upon. A structured interview would limit access to this knowledge.

All of the interviews will be conducted using Skype (or similar video conferencing software) and will be recorded. These interviews will then be transcribed and analysed. The interview guide is presented in Appendix D.

3.4.1 Research Participants

Participants in this research will be faculty members of academic institutions, all of whom are lecturing courses that involve a strong online learning element. These courses may be fully online or blended. Academics who teach solely classroom-based material will not be interviewed as part of this study. This includes academics who use a VLE as a content management system from which students can access notes or submit assignments. The working definition that will be used is that participants will teach new material to students through online means. Chapter 5 provides an overview of the participants and their experience.

3.5 Analysis Methods

Document analysis will be performed on the data collected from the VLE documentation in order to answer research question one. Bowen (2009) identifies five specific uses of document analysis in research. These are: to provide contextual data; to suggest interview questions; to provide supplemental research data; to track change and development; and to verify and corroborate evidence from other sources. This research uses document analysis both to provide contextual data and to suggest interview questions. The contextual data identified during the document analysis will be used to describe the methods presently implemented in VLEs which will allow for the measurement of student engagement. Additionally the results of document analysis will be used to inform the suggested questions for the interviews.

Regarding the analysis of the interview data the method of thematic analysis will be used. Qualitative thematic analysis is used to explore the meanings of physical messages (Zhang and Wildemuth, 2009b). For the purposes of this research these physical messages are the interviews. The aim of thematic analysis is to condense raw data into themes or categories,

differing from traditional content analysis which aims to merely count textual elements (Spurgin and Wildemuth, 2009).

Zhang and Wildemuth (2009b) proposed a method for performing qualitative thematic analysis. The proposed steps include:

1. Prepare the data: This step involves the transcription of the interviews.
2. Define the analysis unit: Qualitative thematic analysis uses themes as the coding unit. This is different from content analysis which uses physical linguistic units (words, sentences, etc).
3. Develop categories: Categories will be developed inductively from the data as the aim is not to test a particular theory, but to explore. It is advised when identifying categories to use the constant comparative method. This involves comparing potential category members to all current members to ensure that it belongs to the same category.
4. Assess categorisation consistency: All categorisation needs to be checked for consistency. There are a number of reasons that the categorisation can become inconsistent over time. These include reviewer fatigue, evolution of the category, and the addition of new categories during the process.
5. Draw conclusions from categorised data: This stage involves the development of theory supported by the identified categories.
6. Report methods and findings: In order to allow for repeatability the entire method must be documented along with the results of the analysis.

This means of data analysis will allow for the development of new theories and models and also provide detailed descriptions for the phenomenon that is online student retention.

3.6 Ethical Considerations

Ethics are closely associated with morals and involve embracing moral issues in the context of working with humans (Gregory, 2003, cited in Ramrathan et al., 2017). Ethics have become a central issue in educational research and we must conduct research with due regard to

ethics (Ramrathan et al., 2017). To that end, an ethical approval request was made to Griffith College Dublin in January 2020. Approval for this research was received on January 18th 2020. A copy of the approval notice is contained in Appendix A.

The interviews that are conducted as part of data collection involve human subjects and are therefore subject to ethical considerations (Walliman, 2017). According to Walliman (2017, p. 47) when dealing with interview subjects they will '*decide whether to take part according to the information they receive about the research*'. This information will allow the participant to provide informed consent to the research. In this research all participants will be provided with the participant information sheet (Appendix B). Following this they will be asked to sign a consent form (Appendix C). Additionally Walliman (2017) states that participants should have the right to withdraw from the research at any time. In this research participants are informed, both in the participant information sheet and also in the consent form, that they have the right to withdraw from the research up to the point of data analysis. At this stage all information will be anonymised and it will be impossible to identify individual participants in the data.

Another ethical issue is that of General Data Protection Regulation, more commonly known as GDPR, which governs any institution or person handling personal data. Personal data is '*data relating to a person who is or can be identified ... from the data itself*' (Department of Business, Enterprise and Innovation, 2020, p. 1). Voice recordings are regarded as personal data and as such are subject to GDPR. Recordings will be stored in a secure encrypted container on the researcher's computer. The recording will be transcribed and any identifying comments will be removed from the transcription. Once the transcription is complete the recording will be destroyed following best practice for data disposal.

3.7 Summary

This chapter introduced the methodologies that underpin this particular research and then looked in detail at the methods that will be used in the research. Firstly the research approach was identified as being qualitative within the interpretivist paradigm using inductive inference. The methods of data collection to be used throughout the research were then

described for each individual research question. Subsequently a brief description of the type of participant that was sought during this research was provided. Following this the methods of data analysis that will be utilised throughout this research were described. These can be summarised as: 1) a document analysis in order to identify the features of the VLE that might allow the measurement of student engagement; and 2) a qualitative thematic analysis to draw conclusions from the interview data relating to the identification of at-risk students and possible preventative measures to aid student retention. Finally the ethical issues that are likely to affect this research were described.

Chapter 4: Comparison of VLE Tracking Data

In an effort to answer research question 1, what data is provided by VLEs that might help educators to measure student engagement, a review of documentation for the three leading VLEs was conducted. Edutechnica (2020) provided the data on VLE popularity showing that the three most popular in Autumn 2019 were: Blackboard (Blackboard, 2020a); Canvas (Instructure, 2020a); and Moodle (Moodle, 2020a). The aim of this aspect of the research is to identify student tracking data that is available to all lecturers in the online environment and to compare the available data across the most popular platforms.

4.1 Blackboard

The Blackboard Learn system was first released in 1997 (Blackboard, 2020a) and between 2016 and early 2019 was the most common VLE used in the US higher education sector (Edutechnica, 2020). Blackboard Inc. state that the system has over 100 million users around the globe (Blackboard, 2020a), while Edutechnica (2020) stated that Blackboard had over 5.5 million users in US higher education institutions in Autumn 2019.

Blackboard Learn provides direct functionality to detect at-risk students, through the retention centre. Figure 2 (Blackboard, 2020b) shows a screen shot from Blackboard which is alerting the lecturer to the fact that there are multiple at-risk students detected in the current course. This table lists missed deadlines, grades, course activity and course access. Additionally instructors can create rules in this retention centre in order to further enhance the functionality. The retention centre will also allow the instructor to track any interventions that they themselves have made such as communication, deadline changes, and so forth.

In terms of basic measurements of student progress, Blackboard Learn records a number of these including: number of logins; number of submissions; number of discussion posts; time-in-course; and number of interactions (Blackboard, 2020b). The number of logins, submissions and discussion posts are self-explanatory. The time-in-course measures the total amount of time that the student has spent accessing the course material. The number of interactions is not clearly defined in the documentation. It is a measure of the number of interactions that the student has had with the course material, but it is uncertain as to the

exact actions that are being counted. All that can be interpreted from this metric is that a higher number implies that the student had more interaction with the VLE than a lower number.

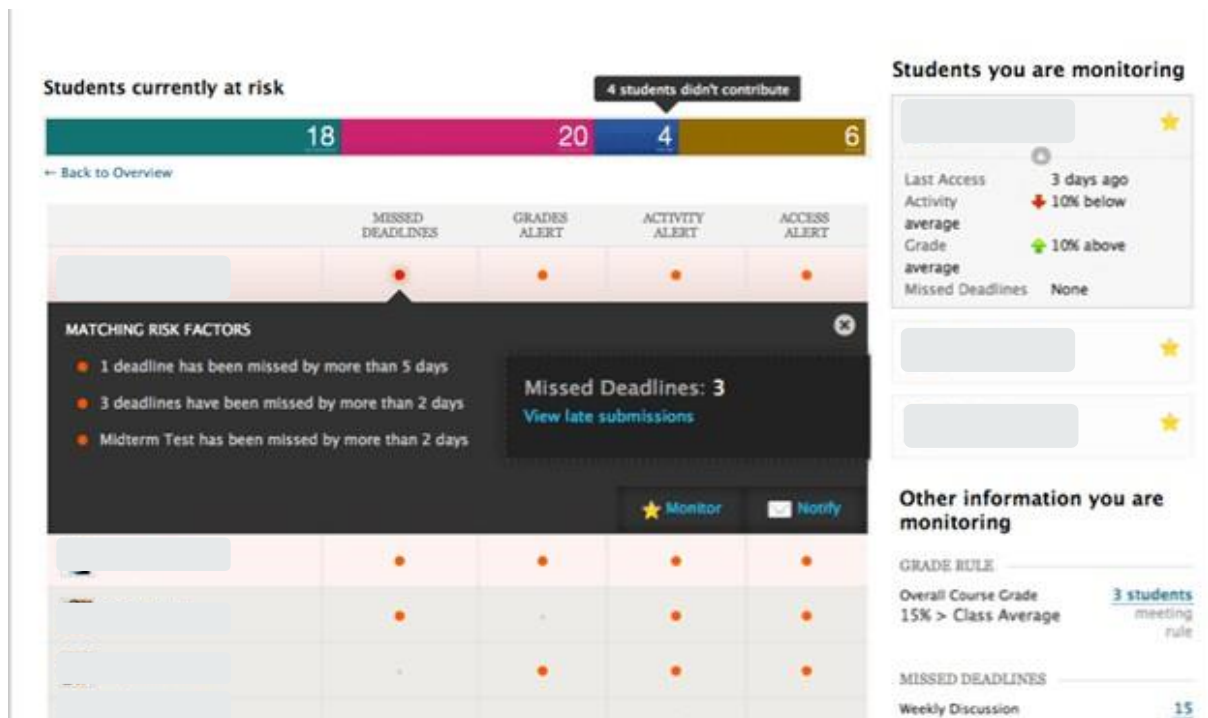


Figure 2: At-risk student detection in the Blackboard Learn system. Source (Blackboard, 2020b).

Blackboard Learn also provides a means of determining the number of times a student viewed each individual page and the time spent on each page (Blackboard, 2020b). It also provides an activity matrix which allows for student activity to be compared across the entire class using scatter plots, providing a visual means of identifying at risk students. The activity matrix can also show students whose grades differ from the expected grade. This would include students whose performance drops in a particular test.

Finally, Blackboard Learn provides achievements as a form of motivational tool. Achievements take the form of a certificate or a badge which reward students for accomplishing a certain task. However, they can also be used to monitor progress by checking how many achievements have been earned by each student. The instructor can customise these achievements for each course.

4.2 Canvas

The Canvas LMS was first released in 2010 (Instructure, 2020a). In Autumn 2019 Canvas overtook Blackboard as the most popular LMS in the US higher education market with 1,094 institutions and over 6.8 million users in total (Edutechnica, 2020). According to Instructure, Canvas now has over 30 million users globally (Instructure, 2020a).

In terms of analytics Canvas provides much information for instructors. Figure 3 (Instructure, 2020b) shows sample data from the Canvas analytics module. The system provides the number of page views for a user and can also be used to gather the number of page views by a user for each individual page in the course material (Instructure, 2020b). Canvas also provides the cumulative scores of all students enrolled in a course. Additionally Canvas provides the total number of submissions (either quizzes or assignments) from each student. The number of submissions is further broken down into three sub-fields: the number of on-time submissions; the number of late submissions; and the number of missing submissions.



The screenshot displays a table with 8 columns and 6 rows of student data. The columns are labeled with circled numbers 1 through 8. The data is as follows:

1 Student	2 Page Views	3 Participations	4 Submissions	5 On Time	6 Late	7 Missing	8 Current Score
[Redacted]	72	8	30	12	3	3	86.21%
[Redacted]	13	4	30	4	1	6	98.67%
[Redacted]	29	6	30	6	1	5	94%
[Redacted]			30	0	0	8	45%
[Redacted]	2	0	30	0	0	8	60.63%
[Redacted]			30	2	0	8	56.25%

Figure 3: Canvas analytics module output. Source (Instructure, 2020b).

As can be seen in Figure 3 (Instructure, 2020b) Canvas also provides a participations metric which appears to be similar to the number of interactions in Blackboard (Blackboard, 2020b). However, the Canvas documentation does explain how this metric is calculated from the following parameters: the number of announcement comments; the number of assignment submissions; the number of discussion posts; the number of wiki pages created; the number of quizzes submitted; and the number of quizzes started (Instructure, 2020b). It would appear

that the only place the number of discussion posts is recorded is as part of the number of participations metric, and as such is not always recoverable. Canvas does however, record the number of messages sent / received between the student and the instructor (Instructure, 2020b).

Canvas is currently in the process of moving from the traditional Analytics module to the New Analytics module (Instructure, 2020b). The metrics mentioned above are all still present, however, they have added a means of viewing data on a week by week basis, instead of just at the entire course level. These metrics could be useful to show a drop off in student interactions with the course material.

4.3 Moodle

The Moodle LMS was first released in 2001 (Moodle, 2020a), and is the only open-source solution of the four most popular learning environments analysed as part of this research. Open source software refers to the availability of the actual computer code used to create the system. In the case of a truly open source solution, such as Moodle, this code is freely available to anyone. Moodle is also free, although if you require support there is an associated fee (Moodle, 2020a). This means that anyone can install and test Moodle, without the need to enter a service contract with a commercial LMS vendor. Edutechnica (2020) considered Moodle to be the third most popular LMS in the US higher education market, with over 17% market share, consisting of 588 institutions with over 2 million users.

Like Blackboard, Moodle are also introducing a *Students At-Risk of Dropping Out* module which aims to determine which students are likely to drop out of a course (Moodle, 2020c). The module is based on the work of Garrison et al. (2010) on communities of inquiry. The Moodle module will measure the *cognitive presence* of a student on a scale of 0–5, with zero meaning no interaction with the material, 1 meaning viewed the material, up to 5 which means resubmitting after feedback. This measure is combined with the social presence measure which identifies the number of opportunities that a participant took to communicate with others. Again this is measured on a scale of 0 – 5. Combining these metrics, and using a machine learning approach, allows Moodle to predict at-risk students automatically.

However, it should be noted that, as this is a machine learning approach, the system does require training (Moodle, 2020c). The module could not be used on a newly installed system, or even a newly created course, it must have historical course data in order to make predictions for the current student group.

Moodle also provides some of the simpler metrics such as grades / quiz scores and page views. However, Moodle does this in an unusual way, through the actual system logs (Moodle, 2020b). The user of a Moodle system is able to filter the logs based on a particular information need. For instance, the instructor can look at one single student, or one single resource. These can be done over a user specified time frame. This system can be used to gather statistics such as the number of page views and submissions.

Figure 4 (Moodle 2020b) shows the Moodle report relating to forum (discussion) access. Moodle provides more information than most of the systems analysed about forum activity. From Figure 4 we see that the forum activity report is broken into the number of discussions started, the number of replies made and the number of attachments included with posts. It also provides the total number of words / characters that the student has posted to the system.

Summary report - First (feline) impressions

First name ^ / Surname	Number of discussions posted	Number of replies posted	Number of attachments	Number of views	Word count	Character count	Earliest post	Most recent post
[Redacted]	2	1	1	15	304	1677	Wednesday, 23 October 2019, 5:31 PM	Thursday, 24 October 2019, 9:09 AM
[Redacted]	2	4	0	28	269	1483	Wednesday, 23 October 2019, 5:43 PM	Wednesday, 23 October 2019, 5:45 PM
[Redacted]	0	2	0	10	169	841	Thursday, 24 October 2019, 9:10 AM	Thursday, 24 October 2019, 9:10 AM
[Redacted]	2	4	1	31	544	3000	Wednesday, 23 October 2019, 5:39 PM	Thursday, 24 October 2019, 9:06 AM

Figure 4: Detailed report of forum activity in Moodle. Source (Moodle, 2020b).

Activity completion and the badge concept also appear in Moodle (Moodle, 2020b). With this the instructor may define certain rewards that are automatically provided to students after successfully completing certain activities. These badges are used as a motivating factor for students, but can also be used to predict under-performing students. Those students who are not attaining the badges are most likely not participating fully in the course.

4.4 Comparison of VLEs

The aim of this chapter is to determine what data is provided by VLEs that might help educators to measure student engagement? To achieve this the documentation relating to student progress tracking in the three main VLE's as identified by the Edutechnica (2020) report has been analysed. Table 1 provides a comparison of these three learning management systems based on their functionality.

	Blackboard	Canvas	Moodle
Logins	Yes	No	No
Page Views	Yes	Yes	Yes
Time in Course	Yes	Yes	No
Submissions	Yes	Yes	Yes
On-time Submissions	Yes	Yes	Not directly – inferred from deadlines and submission dates.
Interactions / Participations	Yes	Yes	No
Discussion Activity	Yes	Not directly – inferred from participations	Yes – including entries for views and posts.
Achievements / Badges	Yes	Yes	Yes
Activity Matrix	Yes	No	No
Quiz Scores / Grades	Yes	Yes	Yes
Number of messages sent	No	Yes	No
Time-based views²	No	Yes	Yes
At-risk detection	Yes	No	Yes

Table 1: Comparison of potential engagement metrics available in the four most popular virtual learning environments.

²This refers to the ability to generate statistics for a set period of time, for instance for the past week.

The following list presents all of the comparison headings discovered and describes the findings related to each:

1. **Logins:** This metric provides a measure of the number of times that a student logged into a system and often the distribution of said logins. This property is present in one of the three LMS analysed (Blackboard).
2. **Page Views:** This metric provides the number of times a student accessed pages. This data is generally presented in two ways, either as a total number of page views across the entire course, or as the number of page views for each content item in the course. All three LMS provide this metric.
3. **Time in course:** This measures the amount of time that the user has spent in the course material. In some cases this measure is broken down to time at each individual content item. Two of the three LMS record this information (Blackboard and Canvas), however, it should be noted that any attempt to measure the time a user spends on a page is an estimate. This is due to the inability to determine when someone leaves the site (if they fail to logout) and also the fact that the user might have multiple browser tabs open, but unused (Analytics Edge, 2020).
4. **Submissions:** The number of submitted pieces of course work which includes assignments and quizzes. This information is recorded by all LMSs.
5. **On-time submissions:** The number of submitted pieces of course work, including assignments and quizzes, which were submitted on or before the deadline. While this can be inferred through the submissions and the knowledge of deadlines for all of the LMS surveyed, only Canvas and Blackboard records this information explicitly.
6. **Interactions / Participations:** This metric measures the level of interaction a student has with the system, however, it is defined differently in the two systems that use the metric (Blackboard and Canvas). Generally it includes discussion posts, submissions,

announcement comments, communications, etc. However, it is dependent on the implementation in each LMS that utilises it.

7. **Discussion Activity:** This, generally, measures the number of posts that the user has made to the discussion forum. However, the Canvas LMS does not record this directly, although it is sometimes possible to infer it from the number of participations. Blackboard records the number of discussion posts made by a student during the course. However, Moodle provides much more detail, not only does it provide the number of posts, it also provides the number of new discussions created, the number of replies on existing threads and the number of discussion posts that have been read by the student.
8. **Achievements / Badges:** While Badges are mainly seen as a form of student motivation (Broer and Breiter, 2015), they can also be used as a means of predicting those students who are at-risk. This can be achieved through the identification of those students who have not obtained badges. As gamification of education has become more popular achievements and badges are now found in all major LMS.
9. **Activity Matrix:** The activity matrix is a visual representation of other activity scores such as interaction or participation. It is included as a separate section due to the visual nature of this data, and provides a simple means by which lecturers can identify students who are not interacting as often as others, and therefore potentially at-risk. This is only provided in Blackboard.
10. **Quiz Scores / Grades:** All of the LMS provide some means of determining how students are performing in any graded element of the course. Poor performance on this metric may indicate students who are at-risk of failure.
11. **Number of messages sent:** This measures the number of messages that the student has sent through the internal communication system. Only Canvas allows the number of messages sent by a user to be monitored.

12. **Time-based views:** This refers to the ability to limit the period for particular metrics. Consider the page view metric, one student might have 1,000 page views, which would be well above the average, during the entire course. However, that student may not have accessed any pages in the last two weeks, indicating a potential problem. If there is only a single page view metric the instructor may identify this student as performing well, where in fact they might be struggling. Two of the three LMS (Canvas and Moodle) provide this facility.
13. **At-risk Detection:** The final category are specific at-risk detection modules provided by the LMS. Two of the three LMS studied in this research (all except Canvas) have recently implemented functionality to identify at-risk students. These algorithms appear to be based on either machine learning (Moodle) or heuristic methods (Blackboard).

In the next section the utility of some of these metrics / techniques in helping to identify the struggling student is discussed.

4.5 Discussion

This chapter aims to answer the first research question, namely what data is provided by VLEs that might help educators to measure student engagement? In answering this question the researcher analysed the documentation for the student tracking aspects of the three most popular LMS in the US higher education sphere (Edutechnica, 2020). In performing this analysis some general tracking data was identified that might be beneficial in identifying at-risk students, and is also available in many LMS.

The first key point to note is that two of the three LMS analysed have, in recent versions, introduced modules that aim to identify at-risk students. The first conclusion to be drawn from this is to the importance of the topic. The vendors are developing this functionality based on requests from the user base. Secondly, the methods that are being used by the LMS vendors are shown. One approach uses machine learning techniques (Moodle) which require a large training set (past course data) and time to train. The remaining platform (Blackboard)

uses heuristic techniques in which they try to predict a student grade and look for students who may differ from that. The approaches that lecturers could manually implement would be based on the heuristic rather than the machine learning style.

All LMS record simple numeric data such as the number of logins and / or the number of page views for each student. Additionally some LMS (Blackboard and Canvas) create a summary score for the number of interactions / participations in the course. All of these simple metrics could be an online substitute for attendance. For instance the student who has only viewed 4 pages from a course with 100 pages could be considered to have *poor attendance* and the inherent risk factors that brings. Blackboard and Canvas also record the amount of time that students spend on each course resource, however, this may not be as accurate a measure as people assume (Analytics Edge, 2020).

In addition to the course wide summary statistics mentioned in the previous paragraph, some of the LMS vendors (Canvas and Moodle) are allowing access to this for a specific time interval. This will allow for changes in engagement to be measured throughout the course. For instance some students may put in large amounts of effort (reflected in the number of page views and logins, etc) in the early part of a course and then fall behind later. This can only be detected with time-based analysis of this data.

Finally all of the LMS provide some idea of the interaction that students have with the discussion board, although in the case of Canvas this metric is not direct. This may in turn allow for the community aspects of learning to be evaluated. Those students who are not partaking in the community might be considered to be more at-risk of failure than those who are.

4.6 Summary

This chapter examined the data that is recorded about students in various learning management systems, specifically Blackboard, Canvas, and Moodle. It then provided a comparison of these and finally identified those that might be beneficial in predicting student engagement with the course material.

Chapter 5: Thematic Analysis

For the purpose of answering research questions two and three, a series of interviews were conducted with teachers in higher educational institutions who are currently involved in online teaching. This chapter presents the results of the qualitative analysis of the interview data. The chapter commences with a description of the participants, this focuses on their teaching experience, and in particular their teaching experience in the online environment. It also describes their educational attainment levels in the area of teaching and learning (all participants had at least a masters level qualification in their chosen academic discipline). The chapter then proceeds to perform a thematic analysis of the interview responses received from each of the participants and to use these themes to address the research questions. The chapter concludes with a discussion of the results obtained.

5.1 Participants

A total of 5 participants agreed to participate in the interview process. These interviews were conducted between February 11th and February 18th 2020. Each interview lasted between 25 and 35 minutes. Table 2 summarises the experience of each of the participants who took part in the study.

	P1	P2	P3	P4	P5
<i>How many years are you teaching in HE?</i>	6	12	10	7	18
<i>How many years are you teaching online in HE?</i>	1	6	10	7	14
<i>What percentage of your current teaching is online?</i>	100%	50%	100%	100%	20%
<i>Are you full- / part-time?</i>	Full-Time	Full-Time	Part-Time	Full-Time	Full-Time
<i>Have you studied T&L?</i>	Yes (7.5 ECTS)	No	No	No	Yes (30 ECTS)
<i>Which VLE are you using most frequently?</i>	Canvas	Moodle	Moodle & Canvas	Canvas	Blackboard & Moodle
<i>Have you received training in the use of the VLE?</i>	No	No	No	No	Yes

Table 2: General information about study participants.

The participants represented four distinct institutions in three countries across Europe. Additionally one participant also provided some continuous professional development training for law enforcement through a European wide training network. Table 2 shows the range of experience of the participants. All are experienced third level teachers with a minimum of 6 years' experience, however, their experience in the online environment has a larger range, between one and 14 years. Four of the participants were full time and one was part-time, working in two institutions. Participants used three different VLEs (Moodle, Blackboard, and Canvas), with two participants using multiple VLEs.

For the purposes of this interview, each of the participants was asked to focus specifically on one particular course when considering the questions during the interview. Initially participants were asked about the courses that they were focusing on. The responses to these questions are contained in Table 3.

	P1	P2	P3	P4	P5
EQF / NFQ Level	7 / 9	7 / 9	7 / 9	7 / 9	7 / 9
Full- / Part-Time	Part-time	Part-time	Part-time	Part-time	Both
Mature Students	Yes	Yes	Yes	Yes	Both
Approx # Students	<20	c. 200	c. 50	<20	c. 100

Table 3: Course profiles for interview participants.

From this we can see that all of these lecturers are teaching at masters level, with the majority teaching part-time, mature students. This must be considered when analysing the participant responses.

5.2 Thematic Analysis

Data collected from the interviews was categorised according to a three-step process which involves: data processing leading to data condensation; data display; and drawing conclusions from the data (Miles et al., 2019).

Processing involves continually re-reading the transcripts of the interviews and condensing this data into a set of themes which reflect the general views of the participants as a group. Data collection gathered almost 140 minutes of audio recording which, when transcribed,

resulted in 44 pages of single-spaced text for analysis. The following sub-sections provide analysis and discussion of this data in terms of each of the research questions.

5.2.1 Research Question 2: Analysis

The purpose of this research question is to discover the extent to which educators are able to identify those students who are in danger of dropping out / failing in the online environment. Generally participants felt that there were difficulties inherent in identifying at-risk students in the online environment, and that this process was more difficult to perform in the online environment than in the classroom environment. However, participants disagreed on whether students were more likely to drop out of the online environment than the classroom environment. Two participants considered it much easier to drop out in the online environment, with one saying that online the lecturer '*might not notice if they are not there, or not engaging*'. However, two others contradicted this statement saying that having access to all material gives the student a chance to catch up near exam time. While the student may not realise their full potential, they still have a chance to succeed in the exam, something not available if they have missed a purely classroom-based course.

Before examining the actual question in detail we need to see how proactive the participants are in trying to identify those students who are in danger of dropping out. The interviews showed that the participants were highly involved in this process. Participants were asked how often they consulted the VLE data to attempt to predict if students were struggling. It was heartening to see that the *worst* of these was still once per month. The others in the study were either weekly or every three weeks. It is also interesting to note that the person who consulted this data once per month was the only part-time participant in this study.

Appendix E presents the results of the thematic analysis for this question. There were five themes identified from the data condensation process. These are:

- 1 **Lack of Training:** Only one participant in the study has received any training in the use of the VLE (Blackboard). The remaining participants have all learned how to use the systems by '*blunder[ing] around until you know how to do something*'. Additionally only two participants have received any training in teaching and learning for either

classroom or online environments. However, one participant, while never receiving training in teaching and learning or VLE usage has developed and delivered courses in both areas. The part-time participant mentions that some of the full-time staff members receive training in VLE usage, but the external lecturers are generally not included in this, or are unable to attend due to other work commitments.

- 2 **Time Pressures:** Most of the participants cited time pressures, or lack of time, as being an issue in allowing them the opportunity to identify struggling students. Again the part-time lecturer assumes that full-time staff have more time for this, however, this is not supported in their responses. In particular the lack of time becomes more of an issue as people become more senior in the institution.
- 3 **Data Knowledge:** While all participants identified the number of logins and page views as being sources of data in the VLE, there were very little other data sources identified. The first research question identified the number of possible metrics that are in use in each of the VLE's most participants only identified a small subset of these metrics. In addition to the number of logins and page views some participants identified the use of student submissions and quiz scores as a means of identifying at-risk students. Others described the use of discussion and communication data as a possible means of measuring engagement.
- 4 **Quizzes / Continuous Assessment:** All participants created quizzes (or exercises) inside the lecture content as a means of identifying students who were not engaging. Most participants felt that this was the best means of identifying struggling students. Participant opinion differed as to whether quizzes were more effective if part of continuous assessment grades, or was the simple fact of having a quiz / submission sufficient to help identify the struggling student.
- 5 **Institutional Knowledge:** One participant (part-time) wished to have access to student's progress in other modules. As a part-time lecturer they had difficulty in accessing this, and therefore could not say if the student was a retention risk, or if the student disliked the particular module. Another participant felt that institutions had

fallen behind in online learning and criteria for the identification of struggling students. That participant felt that without these criteria, that often exist for classroom courses (for instance recording attendance), that they were *'just doing it on [their] own gut feeling.'*

5.2.2 Research Question 2: Discussion

Research question 2 asked to what extent participants are able to identify students who might prove to be a retention risk. Firstly, to be able to do something participants must be willing to do it. Hence, it was heartening to see that all five participants were actively attempting to monitor students with the aim of identifying any who might be struggling during their studies. Figure 5 compares the number of times that participants checked on student progress during a semester with the participant's teaching experience.

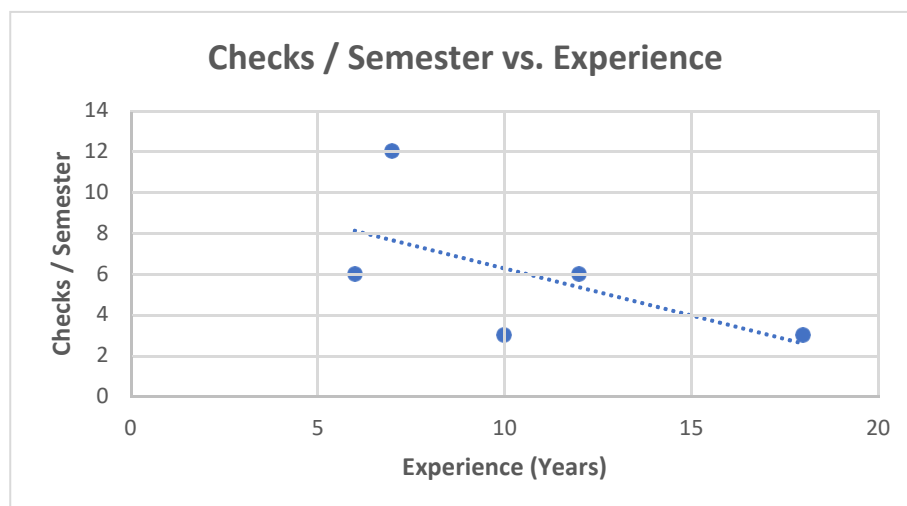


Figure 5: Checks / semester for participants as a function of experience.

From this we see a decrease in the number of checks with increasing experience. This might be explained by the increasing work-load and responsibility as people progress in an academic career. Studies have shown that non-academic work is taking up more time of senior academics, for instance Houston et al. (2006, p. 19) states that Universities have developed *'increasingly complex and time-consuming, control, audit and assurance mechanisms'* to meet external demands. This extra work-load is taking away from time spent teaching, and interacting with the student body. This leads to the first issue that was identified by participants, that of time pressure. All participants, even those new to teaching, identified

this as a problem, however, due to the additional work load related to career progression it becomes more difficult for educators to allocate time to checking student data and looking for those who are struggling.

In order to be able to identify students who might be at risk, the academic must understand the data that is available to them. The first research question examined the four most popular VLEs and identified all of the data that was available in each. Each VLE had between 7 and 10 (Table 1) metrics available for tracking progress, however, only two of these metrics were mentioned by all participants in this study. These metrics were the number of page views and the number of logins. This shows that participants are not necessarily aware of all of the information that can be found in the VLE.

To study this further, the number of individual metrics mentioned by each participant during the interviews was compared with the total number of available metrics for the VLE used by the participant which were identified in the previous chapter (Table 1). In the case of participants who used more than one VLE the total number of metrics was the combination of all metrics provided by at least one of the VLEs in question. Figure 6 shows the results of this analysis.

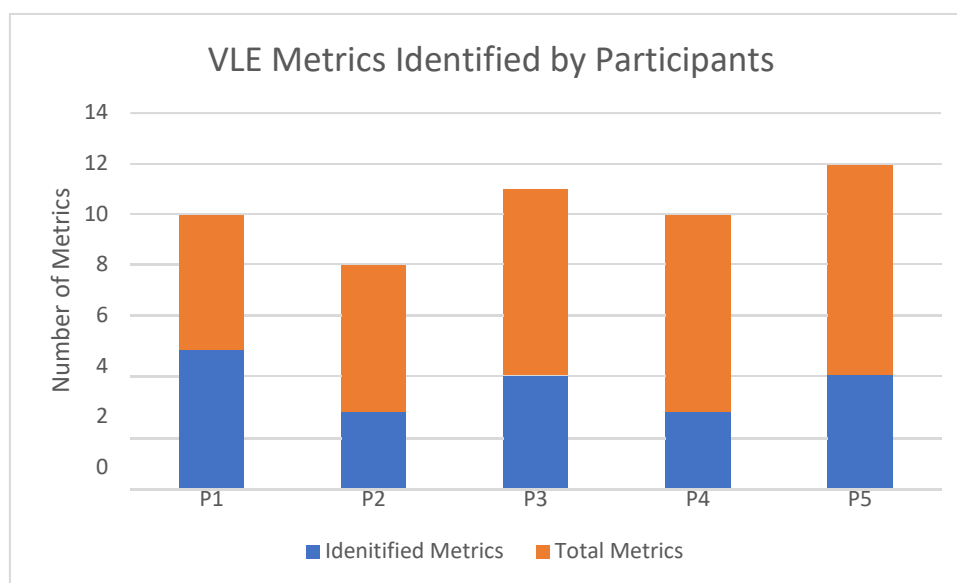


Figure 6: VLE metrics identified by participants as a function of the total number of metrics available.

From this it can be seen that the most knowledgeable participant identified 50% of the available metrics in the VLE that they were using. This lack of knowledge of available metrics in the various VLEs makes it more difficult for educators to identify those students who are struggling.

A potential explanation for this is the apparent lack of training received by participants. All participants were asked two questions related to education or training that they had received. Firstly they were asked if they had any qualification in teaching and learning. Only two participants had received any accredited training in the scholarship of teaching, one with 30 ECTS and another with 7.5 ECTS. In both cases this was a requirement of the institution in which they worked. In terms of training in the use of the VLE, only one participant had received any training in effective use of the VLE in question. The lack of knowledge of the available data for measuring student progress in the VLE might be explained by this lack of training in the use of the VLE. It should also be noted that the participants in this study are all from a technical background, and should find it easier than most to navigate through the VLE and discover the information that is present.

All participants in this study have adopted some form of regular knowledge check exercise in their courses. This can take the form of exercises that require submission, quizzes, or even continuous assessment components. In certain cases the next lesson is not available to students until they have obtained a pre-determined standard in the knowledge check exercise. Figure 7 shows an example of a course in Canvas in which the quiz is used in every lesson. Examining the output quickly shows that one student (highlighted) may be struggling with the course material.

Q1: Introduction to P... Out of 10	Q2: Basic ... Out of 10	Q3: Flow of C... Out of 10	Q4: Data Stru... Out of 10	Q5: Functions & Mod... Out of 10	Q6: Algori... Out of 10
10	10	10	10	10	10
10	10	9	10	8	8
10	10	8	-	-	-
10	10	10	10	10	10
10	10	8	10	9	10
9	10	8	9	9	-

Figure 7: Quiz grades in Canvas with potential at-risk student highlighted. This student is three weeks behind at this stage in the course. Note names have been removed.

5.2.3 Research Question 3: Analysis

The purpose of this research question was to discover what preventative measures are being used by educators to improve student retention in the online learning environment. The thematic analysis identified five distinct themes which can be further classified into two overall response styles, proactive and reactive responses. Proactive responses involve decisions that can be made during course development and management that attempt to reduce the difficulties students have with the material and thereby improve retention. Contrary to this, a reactive response is something that the educator must do when they have identified a student who is struggling and deals with the individual rather than the course as a whole.

The thematic analysis (Appendix F) identified three proactive and two reactive prevention strategies that are being used by the participants. The proactive strategies are:

- 1 **Building a sense of community:** The sense of community is considered difficult to create, or even missing, in online education. Many of the participants described ways in which they were attempting to develop a sense of community through discussion boards, synchronous content, and social media. Many of the participants discussed the use of discussion forums for student introductions '*as a platform to say, hi my name is ...*'. To encourage this, two participants felt that marks should be awarded for participation in the discussion boards. Another '*injects into threads ... adjacent [topics] and see if [they] get picked up*'. Participants generally felt that synchronous content delivery helps build the community in online learning as people '*prefer seeing faces*' and as it provides for '*dialogue rather than monologue*'.
- 2 **Quality of the teaching material:** The participants attempt to create material that is both relevant and rich to aid the development of the community of practice and also to address the differing approaches to learning possessed by students. Participants believe that the content should be '*as rich as possible*' otherwise '*after 15 minutes the eyes gloss over*'.

3. **Quizzes / Continuous assessment:** Many of the participants are, not only using quizzes / continuous assessment, to identify struggling students, but also using it as a tool to retain students by showing them their abilities before the final examination and the stress which it creates. As one participant stated, *'once students see they are capable of doing things... they tend to be more engaged'*. The continuous assessment can also reduce the pressure that students might feel as *'many courses have a 100% exam at the end, the student comes to it feeling more and more pressure'*. Hence, the confidence boosting of the continuous assessment cannot be underestimated.

The reactive strategies that are being employed are:

1. **Communication:** All participants referred to communication with the student as being the first step they would take if they identified a struggling student. Additionally communication amongst the teaching team and the larger institution were also identified as being key to retaining students. It should be noted that the two participants who mentioned communication with the wider teaching team (demonstrators and personal academic tutors) worked for large universities with numerous post-graduate students who fulfil these roles. This may not be feasible in smaller institutions.
2. **Institutional aid:** Participants appeared to be unaware of the institutional aids that were available to struggling online students and as such would be unable to recommend students facilities such as counsellors, financial hardship funds, etc.

5.2.4 Research Question 3: Discussion

All participants in this study looked at prevention of student dropout as being a long-term strategy that begins with course design and development and continues through to delivering the course. Proactive retentive strategies are implemented at the design, development, and management stages, while reactive strategies are implemented after the identification of a student in difficulty. In terms of proactive strategies at the design and development stages it is necessary for material to be developed with two factors in mind, quality and the sense of community.

With regard to material quality it is essential that the material created fulfils three criteria. It must be relevant, interactive and rich. Relevance is of particular importance to the participants as their students are, in the main, mature practitioners of digital forensics and cybersecurity. Kahu et al. (2013) found that mature students were able to better integrate work experience and learning, and when done, these mature students became more motivated. Hence, ensuring that material is relevant and realistic will give learners an opportunity to more thoroughly engage with learning, and hence reduce the retention risk.

At the course design stage, not only is the content decided upon but the method of presenting that content should be decided upon also. VLEs allow for any web content to be included, hence allowing for rich media. Rich media refers to '*specific types of installations on the web that are more functional and interactive and contain video and or interactive features*' (Techopedia, 2020, p. 1). Studies have shown that this rich media can boost presence in an online course (Symonds et al., 2010) and that multiple media channels suit students of differing learning styles (Oregon et al., 2018) and allow for more information assimilation through use of different cognitive channels (Mayer, 2005).

The development of a sense of community in an online learning experience, begins long before the first student logs in. Numerous design decisions are made during course development which will help or hinder community spirit in the final course delivery. For instance will the course utilise asynchronous or synchronous teaching techniques? All participants in this study believed that synchronous teaching was much better at creating the sense of community than asynchronous teaching was. Technologies such as Zoom (Zoom, 2020) or Adobe Connect (Adobe, 2020) have made synchronous lecturing not only possible, but simple. However, the synchronous nature of these lessons limits participation from students as it limits the flexibility of learning, one of the main benefits of online education. Most participants in this study utilise a small number of synchronous review style sessions during a course, but the bulk of teaching is conducted through asynchronous methods.

Another design decision that is made for every online course is how communication will be handled. Generally the VLE provides every course with a discussion forum by default,

however, the educator must still decide how to use this forum. At the course design stage decisions can be made as to how the forum will be used. Should it be used solely as a place for course announcements and nothing more? If used as a more informal place it will encourage the sense of community. For instance, students can be asked to participate in a discussion about some topic on the course. Students should be asked to introduce themselves, and also they should be encouraged to speak about non-course material. In physical learning spaces students speak about much more than just course material, why should this be different in the virtual learning space?

The final proactive strategy that educators can implement at the course design phase is that of continual knowledge checks. As stated in the previous section, these knowledge checks can be used to identify students who are a retention risk (Figure 7), however, according to a number of participants, they also have the effect of boosting student confidence and reducing the fear of the final examination that might be felt by some students. The use of regular knowledge checks, either as part of continuous assessment or just as exercises will, in general, aid student retention.

Once a course is running, the first step to prevent student dropout is to identify those students who are at-risk. The strategies for doing that were discussed as part of research question 2. Once a student has been identified as being at-risk, all participants stated that they would contact that student. If this contact happens at a sufficiently early time (i.e. if the student has been identified in time) it may be successful. This means that the reactive nature of prevention is reliant on the identification stage in order to be successful.

5.3 Summary

This chapter provided the analysis of the interviews conducted with the participants during the course of this study. The chapter began by describing the participants that agreed to be interviewed for this study and the courses that they were currently teaching. Thematic analysis was performed on the recorded interviews. This chapter describes the process for this, and also provides the actual analysis of the interviews in terms of research questions two

and three, relating to the identification of struggling students and the prevention of their drop out respectively. The results of the analysis for each question are then discussed.

Chapter 6: Conclusions & Recommendations

This chapter provides the conclusions from the research that was undertaken during this project. It also provides a set of recommendations, both for educators and institutions which can be used to aid the identification of students who are a retention risk, and to increase the retention rate in online education. It then discusses some of the limitations of this study and proceeds to identify some areas that warrant further research.

6.1 Conclusions

The aim of this research was to determine best practice for identifying students who were at-risk of dropping out of online education and assisting said students to complete their studies. In order to do this, three research questions were considered. This section provides the conclusions of this research in terms of each of these research questions.

6.1.1 *Research Question 1*

The first research question asked what data is provided by VLEs that might help educators to measure student engagement? To place this in context of the overall research aim, engaged students are less likely to drop out than those students who are not engaging with the material (Crossling et al., 2009). To answer this question a document analysis was conducted for the three most popular VLEs in use today, this showed that in all cases there is a large volume of data which can be used to measure engagement. These ranged from simple metrics such as the number of logins or page views to modules employing artificial intelligence techniques to determine those students who are an attrition risk. From this we see that the VLE providers are taking the issue of student tracking seriously and are providing multiple tools that can be used to predict student retention.

6.1.2 *Research Question 2*

The second research question asked to what extent educators were able to identify students who were at-risk of exiting a course prematurely in online education. To answer this question (and question 3) interviews were conducted with five online educators in higher education to gather their knowledge of the domain. Participants were asked about the techniques that they used to attempt to identify those students who were at-risk.

The first finding from this was that these experienced educators were not aware of the volume of data that is available to them in the VLEs. For instance, it was discovered that candidates were aware of only 30% - 50% of the features that were present in the VLE(s) that they were using (Figure 6). This in itself is limiting the effectiveness of the participants in terms of identifying those students who are at-risk. This may be related to the fact that only one of the five participants had received any training in how to use the VLE, while these participants have above average technical knowledge, it is still upsetting to think that training has not been provided to the majority of them in how to use the technological solution. Additionally all participants mentioned the time constraints that they are under in the workplace, and how the identification of students who are struggling has become more difficult due to added time pressures.

On a more positive note, the participants are actively attempting to identify students who are struggling and on doing so are actively communicating with said students to try to help them. All of the participants have created their own identification strategies based on continual knowledge checks embedded throughout the course, creating a quick method of identifying at-risk students. In summary, the will to identify struggling students is present, but the ability and the time might not be.

6.1.3 Research Question 3

The third research question asked what preventative measures are being used by educators to attempt to improve student retention in the online learning environment? Again, this was answered through interviews conducted with the participants. From the participant responses it appears that there are two broad categories of intervention: proactive and reactive strategies. Proactive strategies can be employed during course design and development as well as during course delivery, whereas reactive strategies are employed when an at-risk student has been identified.

Participants identified three main proactive strategies to aid retention which were: building the online community; regular knowledge checks; and the use of rich media. The online community can be built using discussions, synchronous lectures and social media for example. The cognitive theory of multimedia learning (Mayer, 2005) states that students absorb more

information through the use of visual and auditory channels. Hence, the use of rich media will allow students to absorb more information and perform better in the course. Regular knowledge checks provide a confidence boost to the student and allow them to feel better prepared for the examination and therefore less likely to drop out. It's interesting to note that participants felt that these regular knowledge checks were important, not only as a preventative strategy, but in terms of identification also. Indeed it can argued that one of the most fundamental proactive preventative steps that can be taken is that of identification (research question 2) of struggling students. Without identification we will not be able to prevent students from dropping out.

Once the student has been identified we then begin to deal with the reactive strategies. In all cases, participants believed that communication with the student was the key point here. All participants suggested that they would immediately contact the identified student to see what the issues were. Additionally contacting teaching colleagues (such as the larger instructional team e.g. tutors / demonstrators) and senior academic figures (such as programme directors) were also mentioned. Overall, participants considered proactive prevention strategies to be more worthwhile than reactive strategies.

6.2 Recommendations

The ability to transfer theory to practice is one of the cornerstones of successful research. To this end the following set of recommendations are provided. The recommendations are divided amongst those for the institution and the educator. Institutions that are providing, or wish to provide, online education should, as a minimum, provide the following:

1. **VLE Training:** Educators need to be aware of the facilities for monitoring student engagement that are present in the VLE. This can only be achieved with training for those educators who are involved in online education. However, this training should not be a one-off occurrence, additional training should be provided when new functionality is introduced to the VLE to ensure that all educators are current in VLE usage.

2. **Key Performance Indicators (KPI):** Most institutions have KPI related to student attendance in the traditional classroom. Educators are aware that if a student misses a particular number of classes there are a set of procedures that need to be followed. Generally these KPI do not exist for online education. One participant thought educators would be more comfortable if these KPI existed instead of relying on their *'own gut feeling'*.

3. **Support services for online students:** Traditional brick-and-mortar universities provide numerous support mechanisms for students (such as financial, medical, mental health, etc.). Participants in this study seem to be generally unaware of what, if any, support services are in place in the institution for online students. Institutions should address this as a matter of urgency, potentially including it in the VLE training course.

At the course level it is recommended that educators should:

1. **Build a greater sense of community:** This can be achieved in numerous ways such as:
 - a. Increasing the use of forums, specifically for ice-breakers and non-course related conversations which might help foster a greater sense of community;
 - b. Increasing direct communication between instructor and students;
 - c. Increasing the use of synchronous teaching methods.

2. **Build knowledge checks into courses:** Participants consider regular quizzes (or continuous assessment) to be the best means of identifying students who are struggling and also encouraging students to continue on a course of study.

3. **Ensure relevance and richness of training material:** Media richness and interaction will suit more learning styles than voice over PowerPoint, or text alone. It will also aid students in retaining the information. Material should also be relevant and applicable in the real-world to increase student motivation.

6.3 Limitations

All research has limitations which must be acknowledged. In the case of this study these limitations are:

1. **Participants' technical proficiency:** All of the participants in this study are lecturing in digital forensics and cybersecurity, as such, their technical abilities would be above those of an 'average' lecturer. Hence their ability to use the VLE and to access all of the information would be better than most educators. On saying that, the participants still failed to identify many of the tracking features provided in the VLEs that they have been using for a number of years.
2. **Student profile:** The participants' students were, in the main, mature students in full-time employment. Generally these students were employed by law enforcement agencies. Mature students are often highly motivated (Kahu et al., 2013) and as such, may be less likely to drop out of a course than other students would.

6.4 Future Work

There are two possible directions for future work from this research. The first is to perform this study again with a different set of participants. The identified limitations of this study were related to the technical proficiency of the interview participants and the profile of the participants' students. This study should be conducted again with a different group of participants, lecturers in online environments who do not have a technical background and who are teaching undergraduate students. This may provide different results to those obtained in this study.

This study resulted in three institutional and three course level recommendations. Another direction for future research is to utilise these recommendations at course level and evaluate retention to determine if the recommendations are effective.

6.5 Summary

This chapter offered the overall conclusions from this research and suggested recommendations, both for the educator and the institution that would help to reduce the retention problem in online higher education. This chapter also discussed some of the limitations of the current study and finally recommended some possible directions for further research in this area.

References

Adobe (2020) *Adobe Web Conferencing Software - Adobe Connect*. Available at: <https://www.adobe.com/products/adobeconnect.html> (Accessed: 20 February 2020).

Allen, I. E. and Seaman, J. (2007) *Online Nation: Five Years of Growth in Online Learning*. Available at: <https://files.eric.ed.gov/fulltext/ED529699.pdf> (Accessed: 15 January 2020).

Allen, I. E. and Seaman, J. (2010) *Learning on Demand: Online Education in the United States, 2009*. United States: Sloan Consortium.

Analytics Edge (2020) *Misunderstood Metrics: Time on Page / Average Session Duration*. Available at: <https://help.analyticsedge.com/article/misunderstood-metrics-time-on-page-session-duration/> (Accessed: 28 January 2020).

Baker, S. and Inventado, P. S. (2016) 'Educational data mining and learning analytics: Potentials and possibilities for online education'. In Veletsianos, G. (Ed.) *Emergence and Innovation in Digital Learning*, pp. 83 – 98, doi:10.15215/aupress/9781771991490.01.

Blackboard (2020a) *Blackboard Learn: Learning Management System*. Available at: <https://www.blackboard.com/blackboard-learn/index.html> (Accessed: 23 January 2020).

Blackboard (2020b) *Track Performance*. Available at: <https://help.blackboard.com/Learn/Instructor/Performance> (Accessed: 27 January 2020).

Bowen, G. A. (2009) 'Document Analysis as a Qualitative Research Method'. *Qualitative Research Journal*, 9(2), pp. 27 – 40.

Broer, J. and Breiter, A. (2015) 'Potentials of Gamification in Learning Management Systems: A Qualitative Evaluation'. In *Design for Teaching and Learning in a Networked World*. Springer, pp. 389 – 394.

Busato, V. V., Prins, F. J., Elshout, J. J. and Hamakar, C. (2000) 'Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education'. *Personality and Individual Differences*, 29(6), pp. 1057 – 1068.

Crede, M., Roch, S. G. and Kieszczynka, U. M. (2010) 'Class Attendance in College: A Meta-Analytic review of the relationship of class attendance with grades and student characteristics'. *Review of Educational Research*, 80(2), pp. 272 – 295.

Crossling, G., Heagney, M. and Thomas, L. (2009) 'Improving student retention in higher education: Improving teaching and learning'. *Australian Universities' Review*, 51(2), pp. 9 – 18.

D2L (2020a) *D2L - Creators of the Brightspace Learning Management System Software (LMS Software)*. Available at: <https://www.d2l.com/> (Accessed: 27 January 2020).

Day, L., Hanson, K., Maltby, J., Proctor, C. and Wood, A. (2010) 'Hope uniquely predicts objective academic achievement above intelligence, personality, and previous academic achievement'. *Journal of Research in Personality*, 44(4), pp. 550 – 553.

Department of Business, Enterprise and Innovation (2020) *Data Protection and the General Data Protection Regulation (GDPR)*. Available at: <https://dbei.gov.ie/en/Data-Protection/> (Accessed: 24 January 2020).

Department of Education (2011) *National Strategy for Higher Education to 2030*. Ireland: Government Publications Office.

Dictionary.com (2020) *Retain*. Available at: <https://www.dictionary.com/browse/retain> (Accessed: 22 January 2020).

Dutt, A., Ismail, M.A. and Herawan, T. (2017) 'A systematic review on educational data mining'. *IEEE Access*, 5, pp. 15991 – 16005.

EDM Society (2020) *Education data mining*. Available at: <http://educationaldatamining.org/> (Accessed: 23 January 2020).

Edutechnica (2020) *7th Annual LMS Data Update*. Available at: <https://edutechnica.com/2019/10/07/7th-annual-lms-data-update/> (Accessed: 27 January 2020).

Fike, D.S. and Fike, R. (2008) 'Predictors of first-year student retention in the community college'. *Community college review*, 36(2), pp. 68 – 88.

Friedman, B. A. and Mandel, R. G. (2011) 'Motivation predictors of college student academic performance and retention'. *Journal of College Student Retention: Research, Theory & Practice*, 13(1), pp. 1 – 15.

Gabriel, D. (2011) 'Methods and methodology'. *Dr. Deborah Gabriel - Research Guides*, 13 May. Available at: <https://deborahgabriel.com/2011/05/13/methods-and-methodology/> (Accessed: 23 January 2020).

Garrison, D. R., Andersen, T. and Archer, W. (2010) 'The First Decade of the Community of Inquiry Framework: A Retrospective'. *The Internet and Higher Education*, 13(1-2), pp. 5 – 9.

Gaytan, J. (2015) 'Comparing Faculty and Student Perceptions Regarding Factors that Affect Student Retention in Online Education'. *American Journal of Distance Education*, 29(1), pp. 56 – 66.

Ginder, S. A., Kelly-Reid, J. E. and Mann, F. B. (2019) *Enrolment and Employees in Postsecondary Institutions, Fall 2017; and Financial Statistics and Academic Libraries, Fiscal Year 2017: First Look (Provisional Data)*. Available at: <https://nces.ed.gov/pubs2019/2019021REV.pdf> (Accessed: 16 January 2020).

Glynn, S.M., Brickman, P., Armstrong, N. and Taasobshirazi, G. (2011) 'Science motivation questionnaire II: Validation with science majors and nonscience majors'. *Journal of research in science teaching*, 48(10), pp. 1159 – 1176.

Hämäläinen, W. and Vinni, M. (2006) 'Comparison of machine learning methods for intelligent tutoring systems'. In *Proceedings of the International Conference on Intelligent Tutoring Systems*, Berlin: Springer, pp. 525 – 534.

Hammersley, M. (2012) *Methodological Paradigms in Educational Research*. Available at: <https://bera.ac.uk/publication/methodological-paradigms-in-educational-research> (Accessed: 25 January 2020).

Hart, C. (2018) *Doing a Literature Review: Releasing the Social Science Research Imagination*. London: Sage.

Harris, H. S. and Martin, E. W. (2012) 'Student motivations for choosing online classes'. *International Journal for the Scholarship of Teaching and Learning*, 6(2).

Higher Education Authority (2018) *Higher Education Authority 2018 - 2022 Strategic Plan*. Available at: <https://hea.ie/assets/uploads/2017/04/HEA-2018-2022-Strategic-Plan-FINAL.pdf> (Accessed: 16 January 2020).

Houston, D., Meyer, L. H. and Paewai, S. (2006) 'Academic staff workloads and job satisfaction: Expectations and values in academe'. *Journal of Higher Education Policy and Management*, 28(1), pp. 17 – 30.

Hrastinski, S. (2008) 'Asynchronous and synchronous e-learning'. *Educause Quarterly*, 31(4), pp. 51 – 55.

Instructure (2020a) *Canvas the Learning Management Platform*. Available at: <https://www.instructure.com/canvas/en-gb> (Accessed: 23 January 2020).

Instructure (2020b) *How do I view Course Analytics*. Available at: <https://community.canvaslms.com/docs/DOC-10299-415266790> (Accessed: 27 January 2020).

Jaggers, S. S. (2014) 'Choosing between online and face-to-face courses: Community college student voices'. *American Journal of Distance Education*, 28(1), pp. 27 – 38.

Jiang, S., Williams, A., Schenke, K., Warschauer, M. and O'Dowd, D. (2014) 'Predicting MOOC performance with week 1 behavior'. In Stamper, J., Pardos, Z., Mavrikis, M. and McLaren, B. M. (eds) *Proceedings of the 7th International Conference on Educational data mining*, pp. 273 – 275.

JISC (2019) *Case Study I: Predictive Analytics at Nottingham Trent University*. Available at: <https://analytics.jiscinvolve.org/wp/files/2016/04/CASE-STUDY-I-Nottingham-Trent-University.pdf> (Accessed: 22 February 2020).

Johnson, R.B. and Onwuegbuzie, A. J. (2004) 'Mixed methods research: A research paradigm whose time has come'. *Educational Researcher*, 33(7), pp. 14 – 26.

Kahu, E. R., Stephens, C. V., Leach, L. and Zepke, N. (2013) 'The engagement of mature distance students'. *Higher Education Research and Development*, 32(5), pp. 791 – 804.

Kothari, C.R. (2004) *Research Methodology: Methods and Techniques*, 2nd Edition. India: New Age International Publishers.

Lao, L. and Wildemuth, B. M. (2009) 'Semi-structured Interviews'. In Wildemuth, B. M. (ed.) *Applications of social research methods to questions in information and library sciences*. California: Libraries Unlimited, pp. 232 – 241.

Lau, L.K. (2003) 'Institutional factors affecting student retention'. *Education - Indianapolis then Chula Vista*, 124(1), pp. 126 – 136.

Lisciandro, J.G. and Gibbs, G. (2016) "'OnTrack" to University: Understanding Mechanisms of Student Retention in an Australian Pre-University Enabling Program'. *Australian Journal of Adult Learning*, 56(2), pp. 198 –224.

Liston, M., Pigott, V., Frawley, D. and Carroll, D. (2018) *A Study of Progression in Irish Higher Education 2014/15 to 2015/2016*. Available at: <https://hea.ie/assets/uploads/2018/05/HEA-Progression-Report-2018-FINAL.pdf> (Accessed: 22 January 2020).

Mayer, R. E. (2005) 'Cognitive Theory of Multimedia Learning'. In Mayer, R. E. (ed.) *The Cambridge handbook of multimedia learning*. United Kingdom: Cambridge University Press. pp. 43 – 71.

McKenzie, K. and Schweitzer, R. D. (2001) 'Who succeeds at university? Factors predicting academic performance in first year Australian university students'. *Higher Education Research and Development*, 20, pp. 21 –33.

McKenzie, K., Gow, K. and Schweitzer, R. (2004) 'Exploring first-year academic achievement through structural equation modelling'. *Higher Education Research and Development*, 23(1), pp. 95 – 112.

Miles, M., Huberman, M. and Saldana, J. (2019) *Qualitative Data Analysis: A Methods Sourcebook*. 4th edition. London: Sage.

Minaei-Bidgoli, B., Kashy, D.A., Kortemeyer, G. and Punch, W.F. (2003) 'Predicting student performance: an application of data mining methods with an educational web-based system'. In *33rd Annual Frontiers in Education, 2003. FIE 2003*. pp. T2A-13.

Moodle (2020a) *Moodle: Open Source Learning Platform*. Available at: <https://moodle.org/> (Accessed: 23 January 2020).

Moodle (2020b) *Tracking Progress*. Available at: https://docs.moodle.org/38/en/Tracking_progress (Accessed: 27 January 2020).

Moodle (2020c) *Students at risk of dropping out*. Available at: https://docs.moodle.org/38/en/Students_at_risk_of_dropping_out (Accessed: 28 January 2020).

Moore, J. L., Dickson-Deane, C. and Galyen, K. (2011) 'e-Learning, online learning and distance learning environments: Are they the same?'. *The Internet and Higher Education*, 14(2), pp. 129 – 135.

Muilenburg, L. Y. and Berge, Z. L. (2005) 'Student barriers to online learning: A factor analytic study'. *Distance education*, 26(1), pp. 29 – 48.

O'Keefe, P. (2013) 'A sense of belonging: Improving student retention'. *College Student Journal*, 47(4), pp. 605 – 613.

Oregon, E., McCoy, L. and Carmon-Johnson, L. (2018) 'Case Analysis: Exploring the Application of Using Rich Media Technologies and Social Presence to Decrease Attrition in an Online Graduate Program'. *Journal of Educators Online*, 15(2), pp. 2 – 11.

Park, H. S. and Bauer, S. (2002) 'Parenting practices, ethnicity, socioeconomic status and academic achievement in adolescents'. *School Psychology International*, 23(4), pp. 386 – 396.

Patterson, B. and McFadden, C. (2009) 'Attrition in online and campus degree programs'. *Online Journal of Distance Learning Administration*, 12(2).

Pintrich, P. R., Smith, D. A., Garcia, T. and McKeachie, W. J. (1993) 'Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ)'. *Educational and psychological measurement*, 53(3), pp. 801 – 813.

Ramrathan, L., Le Grange, L. and Shawa, L. B. (2017) 'Ethics in Educational Research'. *Education studies for initial teacher education*, pp. 432 – 443.

Robinson, C., Yeomans, M., Reich, J., Hulleman, C. and Gehlbach, H. (2016) 'Forecasting student achievement in MOOCs with natural language processing'. In *Proceedings of the sixth international conference on learning analytics & knowledge*, pp. 383 – 387.

Romero, C., Ventura, S., Espejo, P.G. and Hervás, C. (2008) Data mining algorithms to classify students. In Joazeiro de Baker, R. S., Barnes, T. and Beck, J. E. (eds) *Educational data mining 2008*, pp. 8 – 17.

Romero, C., López, M.I., Luna, J.M. and Ventura, S. (2013) 'Predicting students' final performance from participation in on-line discussion forums'. *Computers & Education*, 68, pp. 458 – 472.

Sanders, J. (2006) 'What is E-Learning?'. In Sanders, J. (ed.) *E-Learning for GP Educators*, Oxford: Radcliffe Publishing, pp. 1 – 6.

Schneider, M. and Preckel, F. (2017) 'Variables associated with achievement in higher education: A systematic review meta-analysis'. *Psychological Bulletin*, 143(6), pp. 565 – 661.

Sclater, N., Peasgood, A. and Mullan, J. (2016) *Learning analytics in higher education*. London: Jisc. Available at: <https://www.jisc.ac.uk/sites/default/files/learning-analytics-in-he-v3.pdf> (Accessed: 23 January 2020).

Shahiri, A.M. and Husain, W. (2015) 'A review on predicting student's performance using data mining techniques'. *Procedia Computer Science*, 72, pp. 414 – 422.

Sife, A., Lwoga, E. and Sanga, C. (2007) 'New technologies for teaching and learning: Challenges for higher learning institutions in developing countries'. *International Journal of Education and Development using ICT*, 3(2), pp. 57 – 67.

Sirin, S. R., (2005) 'Socioeconomic status and academic achievement: A meta-analytic review of research'. *Review of educational research*, 75(3), pp. 417 – 453.

Spurgin, K. and Wildemuth, B. M. (2009) 'Content Analysis'. In Wildemuth, B. M. (ed.) *Applications of social research methods to questions in information and library sciences*. California: Libraries Unlimited, pp. 297 – 307.

Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y. and Yeh, D. (2008) 'What drives a successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction'. *Computers & Education*, 50(4), pp. 1183 – 1202.

Swecker, H.K., Fifolt, M. and Searby, L. (2013) 'Academic advising and first-generation college students: A quantitative study on student retention'. *NACADA Journal*, 33(1), pp. 46 – 53.

Symonds, S., Jamieson, A., Bell, A., Wood, B., Ryan, A. and Patterson, L. (2010) 'Taking Sociology online: Boosting teacher presence and student engagement through rich media'. *Curriculum, Technology & Transformation for an Unknown Future: Proceedings ASCILITE Sydney*, pp. 948 – 950.

Techopedia (2020) *Rich Media*. Available at: <https://www.techopedia.com/definition/9202/rich-media> (Accessed: 20 February 2020).

Tinto, V. (2006) 'Research and practice of student retention: What next?'. *Journal of College Student Retention: Research, Theory & Practice*, 8(1), pp. 1 – 19.

Trinity College Dublin (2014) *Strategic Plan 2014 - 2019*. Available at: <https://www.tcd.ie/strategy/strategic-plan-201419.pdf> (Accessed: 16 January 2020).

University College Cork (2017) *Strategic Plan 2017 - 2022*. Available at: <https://www.ucc.ie/en/media/support/strategicplanning/UCCStrategicPlan2017-2022.pdf> (Accessed: 16 January 2020).

US Bureau of Labor Statistics (2019) *Unemployment Rates and Earnings by Educational Attainment*. Available at: <https://www.bls.gov/emp/chart-unemployment-earnings-education.htm> (Accessed: 25 January 2020).

Walliman, N. (2017) *Research Methods: The Basics*. London: Routledge.

Welsh, E. T., Wanberg, C. R., Brown, K. G. and Simmering, M. J. (2003) 'E-Learning: emerging uses, empirical results and future directions'. *International Journal of Training and Development*, 7(4), pp. 245 – 258.

Williams, D. R., Priest, N. and Anderson, N. B. (2016) 'Understanding associations among race, socioeconomic status, and health: Patterns and prospects'. *Health Psychology*, 35(4), pp. 407 – 415.

Zhang, Y. and Wildemuth, B. M. (2009a) 'Unstructured Interviews'. In Wildemuth, B. M. (ed.) *Applications of social research methods to questions in information and library sciences*. California: Libraries Unlimited, pp. 222 – 231.

Zhang, Y. and Wildemuth, B. M. (2009b) 'Qualitative Analysis of Content'. In Wildemuth, B. M. (ed.) *Applications of social research methods to questions in information and library sciences*. California: Libraries Unlimited, pp. 308 – 319.

Zoom (2020) *Video Conferencing, Web Conferencing, Webinars, Screen Sharing - Zoom*. Available at: <https://zoom.us/> (Accessed: 20 February 2020).

Appendix A: Ethical Approval



ETHICS COMMITTEE FEEDBACK - 18th January 2020

Feedback for Student: Fergus Toolan

Student Number: 2932232

Supervisor: Angela O'Keeffe

On review of your amended ethics submission The Master of Arts in Training and Education (MATE)

*Faculty Ethics Committee (FEC) has approved this proposed study you may proceed.

Peter Gillis

Dissertation Module Lead

Appendix B: Participant Information Sheet

Dear _____

My name is Fergus Toolan and I am a student in Griffith College Dublin's Masters in Training and Education Programme. As part of my dissertation I would like to interview you in relation to your experience with using online learning and its abilities to help you to predict student performance / engagement.

If you agree to partake in this study the following will happen:

1. We will agree a suitable time to meet (via Skype) for an interview.
2. All interviews will be recorded and should last no more than 1 hour.
3. I will ask questions relating to your experiences with online teaching, and in particular how you identify students who might be in danger of dropping out or failing.

After the interview I will transcribe the recordings (removing any personally identifying information that may be present). The recordings and transcriptions will be stored in encrypted containers on my computer. I give my assurance that all information gathered as part of this research will be destroyed after graduation or in 24 months, whichever is sooner.

You are free to withdraw from the study at any time up to the analysis of findings, at which stage the data will be anonymous. Up to this point you are free to withdraw your data without giving a reason for withdrawing, and without your withdrawal having any adverse effect for you.

If, at any time, you have concerns about the study or what you may have said during the interview you can contact my supervisor Dr. Angela O'Keefe at angela.okeefe@griffith.ie.

Thank you for taking the time to read this letter.



Dr. Fergus Toolan

Appendix C: Consent Form

Strategies for Improving Retention in Online Learning

Consent to take part in research

- I _____ voluntarily agree to participate in this study.
- I understand that I am free to withdraw from the study at any time up to the analysis of findings, at which stage the data will be anonymous. Up to this point I am free to withdraw my data without giving a reason for withdrawing, and without my withdrawal having any adverse effect for me.
- I have had the purpose and nature of the study explained to me in writing and I have had the opportunity to ask questions about the study.
- I understand that participation involves being interviewed by the researcher about my experiences in online teaching.
- I understand that I will not benefit from participating in this research.
- I agree to my interview being audio-recorded.
- I understand that all information I provide for this study will be treated confidentially.
- I understand that in any report on the results of this research my identity will remain anonymous. This will be done by changing my name and disguising any details of my interview which may reveal my identity or the identity of people I speak about or the institution at which I work.
- I understand that disguised extracts from my interview may be quoted in dissertations, conference presentations, or published papers.
- I understand that signed consent forms and original audio recordings will be retained in the researchers office (for signed forms) and in encrypted containers on the researchers computer (for recordings and interview transcriptions).
- I understand that all information gathered as part of this research will be destroyed after graduation or in 24 months, whichever is sooner.

- I understand that under freedom of information legislation I am entitled to access the information I have provided at any time while it is in storage as specified above.
- I understand that I am free to contact any of the people involved in the research to seek further clarification and information. These are:
 - Dr. Fergus Toolan – Norwegian Police University College & Student in the Masters in Training and Education in Griffith College Dublin: fergus.toolan@gmail.com.
 - Dr. Angela O’Keefe – Dissertation Supervisor, Griffith College Dublin: angela.okeefe@griffith.ie.

Signature of research participant

Signature of participant

Date

Signature of researcher

I believe the participant is giving informed consent to participate in this study

Signature of researcher

Date

Appendix D: Interview Question Guide

Initial Starting questions (for all participants)

- How many years have you been teaching in higher education?
- How many years have you been teaching in the online environment in higher education?
- Approximately what percentage of your teaching is conducted in the online environment?
- Are you a full or part time teacher?
- Do you have a qualification in T&L, or have you taken any courses?
 - If so please provide the level

Suggested guiding questions for the interview are:

- Tell me about the courses that you teach online

Follow-up questions might include:

- What is the profile of the 'average' student – e.g. mature part-time; etc?
- What is the current retention rate?
- Does this compare well / poorly to that of traditional classroom education?
- Do you use online learning in a synchronous / asynchronous manner (or maybe both). Can you compare these as a means of identifying at-risk students?
- What type of content do you create, just text or is video included? What level of interaction is there in this content? Do these affect the identification / prevention of attrition?

- Tell me about the VLE that you use

Follow-up questions might include:

- What VLE do you use?
- Are you aware of the logging functionality that is present?
- Can you access all information that you wish about a particular student?
- Have you received any training in how to use the VLE in question?

- What data are you aware of that is available to you in the online environment that might aid you in predicting student achievement?

Follow-up questions / pointers might include:

- Prompting participants with information gathered in answering RQ1. This would include: logins; discussion posts; time-in-course; at-risk modules; etc. A list of prompts can be found in the comparison of VLEs in Chapter 4 (Analysis).

- How often do you consult this data?

Follow-up questions might include:

- Do you think this is frequent enough?
- Do you think that you have sufficient knowledge of what is stored in the VLE to be able to access all of the data that is present?

- What do you consider to be the most successful technique to identify at-risk students?

- If you identify a student who is not participating in a course what do you do?

Follow-up questions might include:

- Do you contact them?
- Are there facilities in place in the institution to help these students?
- Are these facilities identical to those for classroom students... is this appropriate?

- Do you consider that there is a sense of community present in the online learning environment?

Follow-up questions might include:

- How would you attempt to build a sense of community?
- Do you use group exercises?
- Do you encourage discussion board participation?
- What do you think of synchronous vs. asynchronous in terms of building a sense of community?

Appendix E: Thematic Analysis Research Question 2

Lackof Training	Time Pressure	Data Knowledge	Quizzes / CA	Institutional Knowledge
<p>When asked about training they received only one participant had any training in VLE usage. Others responded:</p> <p>'No {Laughs}'</p> <p>'Self-taught'</p> <p>'Nope, never once'</p> <p>'I haven't gotten any training'</p>	<p>Many spoke about desired behaviour but stated that time was not available.</p> <p><i>'When I first started I had loads of spreadsheets ... and everything else kind of took over and it got pushed aside'</i></p> <p><i>'that would be great, but who has the time for that?'</i></p> <p><i>'I'm part-time, if I was full-time I would be doing a bit more'</i></p>	<p>All participants mentioned the number of logins and page views. Only a small number of alternate data sources were identified. When asked if they had all necessary information they said:</p> <p><i>'I'm not sure'</i></p> <p><i>'No, personally'</i></p> <p><i>'Not really'</i></p> <p>Other metrics mentioned included:</p> <ul style="list-style-type: none"> • On-time submission / grades / quizzes (4 participants) • Discussions (3) • Communication (2) • Access to a resource (1) • Time spent online (1) • Meeting attendance (1) • Student looking at feedback (1) 	<p>All participants are using quizzes or CA as an identification strategy. Two participants use quizzes at the end of each module:</p> <p><i>'... have to have done the quiz for this lesson before the next lesson opens up'</i></p> <p><i>'... score high on the quizzes...'</i></p> <p>Three use continuous assessment:</p> <p><i>'If you want engagement, have it as part of the assignment grade'</i></p> <p><i>'... if they've dropped off any assessed exercise'</i></p>	<p>Access to performance data in other modules is missing:</p> <p><i>'... [sol can] hear about stuff in other modules that they're doing'</i></p> <p>A set of standards should be defined for the institution:</p> <p><i>'Because otherwise I think ... just doing it on my own gut feeling.'</i></p>

Appendix F: Thematic Analysis Research Question 3

Proactive		
Community	Quality of Material	Quizzes / CA
<p><i>'I think there should be a greater community between the students'</i></p> <p>Three strategies were mentioned for building the sense of community:</p> <p>Discussion Board All participants make use of the the discussion board, but very few encourage off topic discussions.</p> <p>Three participants encourage student introductions - <i>'You can use the discussion board as a platform to say, hi my name is...'</i></p> <p>Two participants suggested marks for participation - <i>'... make it part of the portfolio mark and forces [participation]'</i></p> <p>One participant adds external resources (e.g. articles) - <i>'I do try to inject in threads... not directly related .. but adjacent and see if it gets picked up'</i></p> <p>Synchronous Content All thought it best for creating a sense of community. - <i>'[provides for] dialogue rather than monlogue'</i> - <i>'you must have video on ... because you prefer seeing faces'</i></p> <p>Social Media - <i>'... groups on LinkedIn and Facebook complemented what happened on the official course'</i></p>	<p>Most participants considered the quality of material to be important in engaging students. Material must be relevant, interactive, and address different approaches to learning</p> <p>Relevance Material must be relevant and students must see the relevance. - <i>'Is what they're getting true to real-life and useful?'</i></p> <p>Interaction Material should be rich and also interactive - <i>'... after 15 minutes the eyes gloss over ... you are trying to involve the student'</i> - <i>'... there is not anything like an interactive clickable type thing, but there is always an exercise in every lesson'</i></p> <p>Approaches to Learning and Media Richness Participants felt that material should be rich (containing multiple formats) to aid all preferred approaches to learning <i>'I think also it depends on how people learn ...'</i> <i>'... the richer the better, as rich as possible'</i></p>	<p>Quizzes / CA were seen as one of the best methods of identifying risk, but also as one of the best at helping to prevent dropout.</p> <p>Quizzes / CA can be confidence boosting. - <i>'Once students see they are capable of doing things... they tend to be more engaged'</i> - <i>'Many courses have a 100% exam at the end, the student comes up to it feeling more and more pressure'</i></p> <p>Prompt feedback on assignments can also boost confidence and increases the sense of community: - <i>'Prompt feedback ... so the student can get an idea [of where they are] before they do marked assignments'</i></p>

[[Continued Overleaf]]

Reactive	
Communication	Institutional Aid
<p>All participant's first effort to prevent a drop out is to email a student.</p> <ul style="list-style-type: none"> - <i>'Contact them'</i> - <i>'I usually talk to the student myself'</i> - <i>'First of all reach out to the person'</i> <p>Two participants mentioned communication with the rest of the teaching team.</p> <ul style="list-style-type: none"> - <i>'Contact their personal academic tutor'</i> - <i>'Pull demonstrators to one side and say keep an eye on that student'</i> <p>One mentioned informing the programme director</p> <ul style="list-style-type: none"> - <i>'talk to the course head'</i> 	<p>Only one participant (who was a programme head) was aware of the institutional supports were available to students.</p>