# The Implementation of a Rich Formative Assessment Environment in Mathematics and Related Subjects

A thesis submitted to

The University of Manchester for the degree of Doctor of Philosophy

in the Faculty of Engineering and Physical Sciences

2015

ELIZABETH ACOSTA GONZAGA

SCHOOL OF PHYSICS AND ASTRONOMY

# Table of Contents

# Appendices

# List of Figures

# List of Tables

ABSTRACT

The University of Manchester

Abstract of thesis submitted by **Elizabeth Acosta Gonzaga** for the degree of Doctor of Philosophy and entitled **The Implementation of a Rich Formative Assessment Environment in Mathematics and Related Subjects**.

Month and Year of Submission: March, 2015.

Currently, Information and Communication Technology (ICT) is available to almost everyone. In the last few decades its impact has widely expanded in all fields including the educational one. There are important attitudinal factors affecting the usage of learning technology successfully. Information system researchers have identified the importance of personal factors, such as attitudes, beliefs, culture and behaviours in technology acceptance (Davis, Bagozzi, & Warshaw, 1992). Previous research has analysed these factors in the acceptance of educational technology in higher education (Liu, Liao, & Pratt, 2009), (Teo, 2009), (Terzis & Economides, 2011), (Cheung & Vogel, 2013). However, the factors playing a role in a mathematical context have not been fully analysed.

This research explores the role of several attitudinal factors in the acceptance of educational technology for the assessment process in mathematics in a higher education setting. We examine the effects of formative on-line feedback on the adoption of educational technology by analysing both teachers' and students' opinions regarding enriched formative on-line assessment for mathematics. Our results suggest that UK students find on-line feedback is more enjoyable and useful than traditional feedback. Attitude and enjoyment are two important factors influencing usage intentions. Results also show that two influential factors to facilitate using the on-line testing environment are that feeling confident about computers and the availability of information technology services. In Mexico, findings reveal that students' attitude has the strongest influence on usage intentions which means that their feelings and opinions are important. They also think that on-line feedback is more enjoyable than face-to-face feedback. Students who are proficient in on-line testing are more likely to find it easy to use. Students also agree that when an on-line environment is easy to use, it is also more likely that they consider it as useful and enjoyable. Taking into account these results can be the best way to design a mathematics e-assessment activity for UK and Mexican students. UK instructors agree that on-line assessments are useful tools to enrich instructional strategies. Teachers in Mexico have a similar opinion. Both results show that they really value providing on-line feedback.

# DECLARATION

The work in this thesis is based on research carried out at the School of Physics and Astronomy, University of Manchester, England. No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

# COPYRIGHT STATEMENT

# ACKNOWLEDGEMENTS

I would like to convey my most sincere gratitude to my supervisor, Professor Niels R. Walet for all his enthusiastic guidance, enormous support, great patience and kind advice he has provided throughout my PhD.

I wish to gratefully acknowledge the PhD scholarship and associated financial support provided by The National Polytechnic Institute (IPN) and The National Council on Science and Technology of Mexico (CONACYT).

And also to my family and friends.

# Chapter 1. Preliminaries

## 1.1 Introduction

Currently, Information and Communication Technology (ICT) is available to almost everyone. In the last few decades its impact has widely expanded in all fields including the educational one. There are important attitudinal factors affecting the usage of learning technology successfully. Information system researchers have identified the importance of personal factors, such as attitudes, beliefs, culture and behaviours in technology acceptance (Davis et al., 1992). Previous research has analysed these factors in the acceptance of educational technology in higher education (Teo, 2009), (Liu et al., 2009), (Terzis & Economides, 2011), (Cheung & Vogel, 2013). However, the factors playing a role in a mathematical context have not been fully analysed.

This research explores the role of several attitudinal factors in the acceptance of educational technology for the assessment process in mathematics in a higher education setting. We examine the effects of formative on-line feedback on the adoption of educational technology by analysing both teachers' and students' opinions regarding enriched formative on-line assessment for mathematics. Our results suggest that UK students find on-line feedback is more enjoyable and useful than traditional feedback. Attitude and enjoyment are two important factors influencing usage intentions. Results also show that two influential factors to facilitate using the on-line testing environment are that feeling confident about computers and the availability of information technology services. In Mexico, findings reveal that students' attitude has the strongest influence on usage intentions which means that their feelings and opinions are important. They also think that on-line feedback is more enjoyable than face-to-face feedback. Students who are proficient in on-line testing are more likely to find it easy to use. Students also agree that when an on-line environment is easy to use, it is also more likely that they consider it as useful and enjoyable. Taking into account these results can be the best way to design a mathematics e-assessment activity for UK and Mexican students. UK instructors agree that on-line assessments are useful tools to enrich instructional strategies. Teachers in Mexico have a similar opinion. Both results show that they really value providing on-line feedback.

This thesis is organized into ten chapters.

Chapter 2 gives some background on the foundations of learning theories, assessment and feedback, as a way to introduce the reader to the basics regarding the educational assessment.

In chapter 3 we briefly explain how the usage of ICT has spread widely into the teaching and learning processes. Also, how the technology is actually used to support the educational processes is discussed.

Chapter 4 gives a brief introduction to structural equation modelling (SEM), which we use as a statistical technique to find significant relationships for the models employed in the thesis. This chapter also describes the data analysis procedures.

Chapter 5 explores the lecturer's opinions in the Faculty of Engineering & Physical Sciences at the University of Manchester regarding the adoption of on-line assessments in mathematical subjects. This chapter analyses the teacher's attitude and intention to use on-line assessment technologies in the future.

Chapter 6 examines students' perceptions and feelings regarding mathematical on-line assessment at the University of Manchester. We discover significant relationships which help us to find the key factors that allow a positive adoption of electronic assessments.

Chapter 7 explains teacher's thoughts about doing on-line assessments in the UPIICSA at the National Polytechnic Institute (IPN) in Mexico. We analyse several important attitudinal factors such as usage intention and perceived usefulness which reflects on the Manchester work on-line feedback that can be involved in a successful uptake for STEM subjects.

Chapter 8 discusses student's opinions regarding on-line exams for STEM subjects in the UPIICSA at National Polytechnic Institute (IPN) in Mexico. We have also asked for significant attitudinal factors involved in the uptake of on-line assessments.

Chapter 9 investigates significant attitudinal factors on the adoption of educational technology for mathematics students at the UPIICSA at the IPN in Mexico.

Finally, chapter 10 makes a comparison between the issues that are significant for students in the UK and for students in Mexico. This chapter makes conclusive notes regarding what is important in both cultures.

# Chapter 2. Elements of the Assessment Process

## 2.1 Introduction

The global education context has changed in the last few decades, especially in higher education. The process of teaching and learning is very different from the way it was taught before. This is the result of several factors: On the one hand, there are enormous challenges facing universities and colleges, including an increased demand for quality educational services. A growing number of students demanding education has forced universities to create a new era of education quality assurance and accountability. On the other hand, the continuous breakthroughs in ICT has imposed a revolution in all fields of human activity, and higher education has not been an exception. ICT has changed the way in which we live and communicate. While many may consider that ICT has not brought enough or no benefits, the fact is that ICT has changed teaching, learning and assessment methods at all levels of education, especially in the universities.

These days university teachers have to deal with various new challenges including the fact that large classes are common, that workloads are increasing, and the importance of quality assurance (i.e., quality management) of education and assessment (Reiners, Dreher, & Dreher, 2011).

From an economic perspective universities are facing an increasing pressure to be more responsible for their own finances. From a realistic perspective, any improvement in quality of teaching is often mitigated by economic pressures that result in teachers being required to do more with less. The ‘do more’ often involves meeting strategic goals to improve teaching and learning outcomes and research targets (ranked nationally and internationally), in addition to multifaceted administrative tasks. The ‘with less’ often involves restricted budgets and, due to increasing task demands less time available in an educator’s working week, often requiring staff to work unpaid overtime (Dreher, Reiners, & Dreher, 2011).

On the other hand, companies and organizations increasingly demand new capacities and skills from their workforce; higher-order thinking and social skills are required in the managing of technology, employees should also be able to obtain, transmit and interpret information quickly and effectively.

Educational institutions can play an important role through teaching, research, and innovation and through influence on staff and students. Nevertheless, there are crucial issues that they have to confront: the growing demand for higher education is not only a challenge that developed countries have, it is a global phenomenon that all countries are facing. In the case of emerging countries, it has been a problem created to a large extent due to the lack of public policies as well as the limited economic support that political authorities provide for learning, research, and innovation. These have substantially limited economic and social growth. The technological gap facing universities in developing countries is also a major challenge.

The aim of education should be to teach people to be responsible individuals, critical thinkers and productive workers who can contribute to the well-being of society. Education is a process that involves acquiring learning, developing skills and attitudes and that attempts to create a stable and persistent change in what a person or group of people know and can do. To determine if this goal has been reached assessment of learning is the key.

There has been a growing emphasis on finding ways to improve teaching practices, especially considering the assessment impact of student learning. It was only a few years ago that the educational assessment context began to change; it recently has been recognized as a key process in teaching and learning that enhances student learning. The important role of assessment and feedback in learning and teaching in higher education has been well recognised in literature (Brown, Bull, & Pendlebury, 1997), (Gibbs & Simpson, 2004), (Nicol & Macfarlane-Dick, 2006), (Bloxham & Boyd, 2007).

Even though its role has been more recognized, some educational systems have not put in enough effort to develop strategies to enhance the educational assessment process in higher education. In the words of a student's report, “we would like to see all universities and colleges implement a systematic policy to leverage technology to provide innovative methods of assessment and feedback” (National Student Forum-Annual Report, 2009). In some universities, the assessment practices have not changed much beyond the traditional pen-and-paper tests; others, are more focused on reporting a final grade of the student' learning at the end of course. In short, universities have used the same assessment methods for a long time.

However, the current economic needs of society require that universities develop new skills in students such as Williams (2008) mention information analysis, collaborative working and ‘just-in-time learning'. Traditional methods of assessment seem obsolete for assessing new skills; they are the greatest obstacle to innovation and there is a prevailing necessity to implement and validate new assessment methods (Williams, 2008). If we also consider students' opinions, the study of Iannone & Simpson (2013) explore the perception of mathematics students about assessment methods in higher education. Their findings reveal that students perceive traditional assessment (closed-book examinations) as the main discriminator of mathematical ability. In addition, traditional tests are not very useful to measure student's intellectual performance (Clarke-Midura & Dede, 2010).

The design and implementation of new teaching and learning methods require greater effort, time and resources than traditional methods. Since heterogeneity and wide scope of technology, innovative assessment methods are more complex and difficult to implement (Williams 2008). Given that innovative assessment can take a wide variety of forms, Maclellan (2004) argues that its use to validate learning (in high-stake assessments) can be problematic. She emphasises this concern by explaining two issues, *task specification* allows the use of irrelevant variables that can make it difficult to make a clear distinction between the purpose of the assessment and the skill required to answer the assessment task in the right way. And the context of *consistency of marking* it is difficult to assign assessment criteria correctly. How people can interpret these may be different. She advises that the validity of innovative methods has to be considered.

Assessment is defined as a measurement of the learner's achievement and progress in a learning process (Gikandi et al., 2011). It involves activities focused on measuring characteristics of human learners such as learning, motivation and attitudes. The principal goal of the assessment process is to determine students' skills, knowledge, understanding and abilities. It can be used to promote learning as well as to ensure that students meet the intended learning goals (Quality Assurance Agency for Higher Education 2006), (Stödberg, 2012).

This process must be recognized as the central point of an interactive teaching and learning process. It is the key factor in encouraging the assessment of higher order thinking, social skills and group work (Buzetto-More & Alade, 2006). It is a core component of effective learning (Gikandi, Morrow, & Davis, 2011) that can enhance students' learning by making it more efficient, transparent and fair. Brown et al. (1997) state "if you want to change student learning then change the methods of assessment".

There is commonly a confusion between the terms "evaluation" and "assessment". The first one refers to activities focused on estimating the outcomes and worth of products, programs, and projects, whereas the second is used to refer to operations associated with measuring achievements of persons in relation to desirable outcomes. In short, we assess people and evaluate things (Black & Wiliam, 1998b).

Assessment is important for all participants in the educational process, although from different perspectives. Students receive useful feedback from lecturers and can also gain credits at the end of the course. Teachers need to measure students' learning aims which also provides the educational administrator/manager with operational and performance data (Reiners et al., 2011). Universities also can enhance their institutional effectiveness (Buzetto-More & Alade, 2006).

As a result, Kellough & Kellough (1998) identify seven purposes of assessment: *1. To improve student learning; 2. To identify students' strengths and weaknesses; 3. To review, assess, and improve the effectiveness of different teaching strategies; 4. To* *review, assess, and improve the effectiveness of curricular programs; 5. To improve teaching effectiveness; 6. To provide useful administrative data that will expedite decision making; and 7. To communicate with stakeholders.*

According to Martell & Calderon (2005) assessing student learning is an on-going cyclical process or loop that involves the identification of outcomes, the gathering and analysing of data, discussion, suggesting improvements, implementing changes, and reflection. In this context the term "closing the loop" is widely used, which defines a continuous process that uses assessment data to improve student knowledge.

Moreover, the assessment should be intended to go beyond the teaching-learning process. It may be viewed as an integral part of the process of being a student. As Gibbs & Lucas (1996) state, assessment and feedback are not only central to learning but also to the student experience. Students have realised this for some time, as they recognised that assessment and feedback are areas which need attention across the HE sector, as explained in the National Student Forum - Annual Report (2009).

The type, frequency and format of assessment practices depends on the underlying educational learning theory and the educational objectives (Dreher et al., 2011). Therefore, we will briefly summarize the main educational learning theories in the next section.

## 2.2 Educational Learning Theories

Psychologists, pedagogues and educators have developed a set of theoretical approaches that explain the mechanisms involved in learning. These approaches have evolved over time, being influenced by their context, such as society and technology. These theories have contributed insights that have allowed us to enhance people's learning. They have been applied in different areas of human knowledge, but mainly in schools, suggesting that students can learn using several

*Behaviourism* was widely applied in the first half of the 20th century. It contains conditioning theories of learning, which consider three main learning processes: classical conditioning, operant conditioning, and contiguous conditioning (Schunk, 2012).

The best-known conditioning theory is the operant conditioning developed by Skinner (1976). It relies on the mechanisms of stimulus and response. Here the environment, situations and events serve as stimuli for a response. This theory uses a reinforcement mechanism to strengthen the learned response when a given stimulus is present. The operant conditioning theory is basically a three-term model involving a discriminative stimulus (antecedent), response (behaviour), and reinforcing stimulus (consequence) (Schunk, 2012). In a school context, it means the presentation of a problem (stimulus) is followed by the contribution on the part of the student to the solution (response) being the feedback which provides the reinforcement. It is assumed that the consequences of behaviours determine the likelihood that students will respond to antecedents. Consequences that are reinforcing increase behaviour; consequences that are punishing decrease behaviour (Schunk, 2012).

In order for learning to take place, students must adapt to the instructional environment, giving the answers expected. Hence, it is important to encourage students to participate in the learning process. The teacher can engage them through stimuli that strengthen students' involvement (reinforcers). To give students an enjoyable learning and development environment it is preferable to provide positive reinforcement (Gúzman & Hernández, 1993).

Behaviourism considers students as a “black box”, which receive the knowledge whereas the teacher plays the role of dispenser of knowledge and feedback (Hung, 2001). Therefore, the fundamental idea of behaviourism for achieving learning is that students can learn by acting (obtaining a desired skill/knowledge by performing a learning activity). In order to teach students, it is necessary to trigger students’ participation by giving them appropriate feedback and reinforcement properly.

This theory assumes that techniques and procedures that impact learning include modelling (acquiring new skills and behaviours by observing another person who perform the behaviour/skill to be acquired), reinforcing (strengthen those behaviours to allow getting a desired behaviour) and imitation (reproducing a behaviour shown in a model). The educational strategy includes a detailed and clear presentation of the instructional objectives, which must specify the desired behaviour of the student. This can be done through an analysis of the activities and tasks that form the steps to achieve the skill required (Gúzman & Hernández, 1993).

Since, from a behaviourist perspective, the function of assessment is to identify students' educational needs, this suggests developing an adequate methodology to assess the results of instruction. This approach requires continuous assessment using objective instruments, that serve as the basis of teaching. The design and use of a correct instrument (exam, etc.) is essential in order to prove one has achieved behavioural aims. The assessment practices are used as judgements rather than as criterions (standards). The skills and abilities of the students are assessed in terms of an absolute skill level and acquired knowledge (grade of dominance).

In this approach Skinner (1976) also argues that a teacher's skills for teaching are not innate. He suggests that they are rather a set of knowledge and skills that can be acquired through training.

The main objection to this theory is that students have to demonstrate an observable behaviour as an evaluation mechanism of learning, without considering students' thoughts.

*Cognitivism*. The cognitive theory of learning was developed on the basis of Gestalt psychology in the work of Ausubel (1963) and Bruner, Goodnow, & Austin (1956). It is built on the assumption that the mind is perceived as an information processor with short-term and long-term memories, including a working memory (Hung, 2001).

Unlike behaviourism that asks “what did the student learn to do?” cognitivism asks “how did the student learn to perceive a situation?”. The foundation of this theory assumes that learning is not just adding new knowledge and removing others. It is a deep understanding of a situation in a new perspective, changing a pattern, shape (Gestalt) to another with the possibility that this change can occur through a new experience or reflection (Gúzman & Hernández, 1993).

A basic principle of this approach is the theory of human information processing developed in the early 1950s. This assumes that the mechanisms used by the mind to extract and process information collected from the environment can be represented as machines. As a result of this reasoning, the idea of artificial intelligence was first introduced. Another underlying principle is the meaningful learning theory of Ausubel (1963) that explains the functioning of the learners' cognitive structures and mechanisms to build meaningful learning.

The cognitive theory proposes the development of learning strategies focusing on promoting self-learning in students (learning to learn), acquiring skills searching and managing information in order to achieve autonomy in learning. It supports strategies for teaching creativity as well as programs "to teach to think" (Nickerson, Perkins, & Smith, 1985) whose purpose is to promote skills such as analysis, inductive and deductive reasoning, synthesis, problem solving, classification and critical

Cognitivists argue that since new knowledge is constantly produced, it is essential that the curricula of schools include strategies to teach students how to think, not just learn by receiving knowledge. For decades, schools have played the role of transmitting knowledge. However, this theory proposes including learning strategies that develop students' skills to find and produce knowledge, to promote curiosity, creativity, reasoning and imagination.

In this framework it is essential to investigate students' "cognitive styles", to find out what are his/her knowledge and mental models, and use these as a support to build new learning. The student is an active information processor, responsible for his/her own learning. It establishes that to achieve knowledge and skills it is essential to practise them.

Cognitive theory focuses on a student motivation to learn, helping them to seek ways and means to satisfy his/her own intellectual curiosity and not just under teacher's pressure or gaining a grade. It focuses on evaluating students' thinking and reasoning skills and not just the ability to manage information or master the content.

In the educational context this theory points out that the assessment process plays an important role to improve teaching, given that it allows a continuous checking to detect right answers and mistakes. However, it is essential to use a correct instrument, proving achievement of behavioural aims.

The role of the assessment is to identify student's psycho-educational issues in order to design an adequate instructional sequence that enables the assessing of students needs and performance (the results of the instruction). This perspective sees the assessment practice as judgements more than criterions (standards).

*Social-cognitivism*. This theory was mainly developed by Bandura (1986). The basis of his work is the idea that human learning occurs in a social environment. By observing others, people acquire knowledge, rules, skills, strategies, beliefs, and attitudes. Individuals also learn from the usefulness and appropriateness of behaviours and the consequences of modelled behaviours, and they act in accordance with beliefs about their capabilities and the expected outcomes of their actions (Schunk, 2012). The theory makes a distinction between en-active and vicarious learning. Here en-active means learning by “doing” and vicarious means learning by “observing” or “listening”. The learning of complex skills occurs through a combination of en-active and vicarious learning. Learners begin acquiring the skill by observing models and practising that skill.

This theory points out that observing does not guarantee learning or achieving the skill that demonstrates the learning. Instead, it states that one should provide information about probable consequences of actions and motivate observers to act accordingly (Schunk, 2012).

Bandura establishes that one of the key assumptions of this theory is the concept of self-efficacy. He defines it as “personal beliefs about one’s capabilities to learn or perform actions at designated levels". In getting self-efficacy, learners assess their skills and their capabilities to translate them into actions. This concept will be used in the research models of this thesis.

*Constructivism and Social Constructivism*. The main principles underlying constructivism were determined by the interaction of the cognitive processes -thinking and learning- and physical and social contexts. The theory has mainly been developed in the field of human development, especially under the influence of the theories of Piaget, Bruner and Vygotsky.

The theory of Piaget (1977) postulates that learning is a process of accommodation, assimilation and equilibrium of knowledge. Learners create their knowledge in an interconnected and complex process; many different representation of content are necessary to create a deep learning. Bruner et al. (1956) emphasize how the brain obtains and processes information through encoding and retrieval from memory. Both Piaget and Bruner stress ideas about how the mind constructs knowledge. Although there may be many different versions of what constructivism entails, the general view held is that learning is an active process of constructing rather than acquiring knowledge (Hung, 2001).

The theory of Piaget makes the assumptions that the learner is a constructor of his/her own knowledge. He/she learns from their own experience. The student builds his/her own learning based on previous and current knowledge. Learners play an active role in the acquisition of knowledge. This approach transforms students from passive into active constructors of knowledge and it requires that the teacher gives to students an appropriate guide, environment and tools in order to obtain deeper learning.

Piaget's theory proposes that students build confidence in their own ideas by making their own decisions and accepting their own mistakes. Teachers help to build student's self-confidence. For Piaget the emphasis of teaching focuses on the activity, initiative and curiosity of the students in the light of new knowledge. The teacher should promote an atmosphere of reciprocal respect and confidence in which students can learn through problem-solving approach and cognitive challenge. The interactions between students are important as they can exchange and confront their ideas. This promotes the socio-cognitive conflict that causes restructuring of their short or medium schemes. Gúzman & Hernández (1993) state the benefits of this approach as:

-It is possible to achieve a significant learning as it is directly built by the student.
-The learning achieved by the student can be transferred to new situations, unlike what happens when new knowledge is superficially gained.
-If a student feels he/she is able to produce valuable knowledge, this leads to improving his/her self-esteem and self-concept.

More recently, the social aspects of constructivism commonly linked to the work of Vygotsky (1962) have been emphasized. This is known as social constructivism, where one considers learning as a social process, where both the cultural background and social context play an important role. Vygostsky emphasizes the critical importance of interaction with people – parents, teachers, classmates – in cognitive development (Hung, 2001). This occurs at an interpersonal level through social interactions, where the knowledge is internalised. After learners have obtained a clear understanding of new knowledge and insights, they build their own thoughts and ideas.

Both approaches, constructivism and social constructivism, emphasise the social and individual dimensions of cognition. In summary, Hung (2001) uses the follow three points as a summary of both theories: "(1) Learning is an active process of constructing rather than acquiring knowledge; (2) Knowledge can be socially constructed where the social interaction may include just oneself; (3) The interpretation of knowledge is dependent on (a) the prior knowledge and beliefs held in one's own mind and (b) the cultural and social context through which the knowledge was constructed".

In the educational context, social constructivism introduces the concept of Zone of Proximal Development (ZPD) that we have depicted in Figure 2.1. This is a "zone" in which optimal learning occurs through a process that considers a student's previous knowledge before adding new pieces of knowledge (a process known as scaffolding) in the ZPD. It is a process of continual and constant assistance where a teacher provides support to students in order to engage and increase student's interest in their acquisition of new knowledge, or correcting their misunderstandings. This concept is a social-collaborative process in which the student is a communicator who shares information and knowledge with others in a social process creating a new scaffolding mechanism.

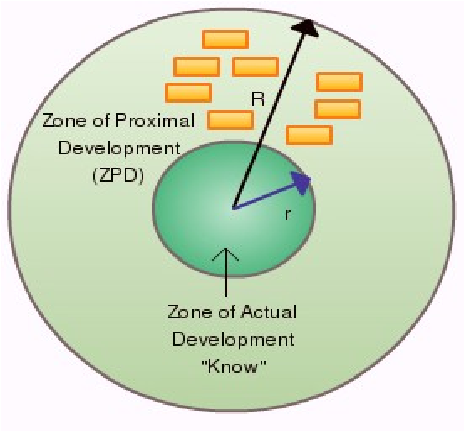


*Figure 2.1. Diagrammatic illustration of social constructivist theory (r: radius of core actual knowledge development; R: radius of potential ZPD boundary; shaded boxes: scaffolding mechanisms) taken from (Bryceson, 2007)*

Figure 2.1 illustrates that teachers can play an essential role in guiding and transmitting socio-cultural knowledge to students helping them to learn and internalize knowledge, hence the social interactions that students can make with others are essential for their cognitive and social-cultural development. For the development of the ZPD, it is required that the teacher is an expert in the domain of knowledge and is sensitive to the progress that students achieve. In this way the teacher plays the role of "director" creating a support system (scaffolding). When students have learned and internalized new knowledge, the teacher becomes a "viewer".

Vygotsky proposes (see Figure 2.1) that the ZDP occurs when the teacher focuses on guiding students from lower-levels to higher-levels of the zone, paying particular attention to students' cognitive issues to guide them to a desired performance; at the same time, responsibility and control are transferred to the student. At the beginning students are not able to understand or perform a task by themselves, but after guidance by the teacher, they are able to understand and develop the activity without external help.

In this theory educational assessment practices focus on determining the potential development level of a student. The teacher provides support through guidelines previously analysed and categorized according to the performance level shown by students in performing a task or activity. Students who require more guidance will have a smaller learning potential compared with those who require less advice. This process diagnoses an individual's learning potential and determines the educational practices that align their learning with their cognitive development.

Table 2.1 summarises key concepts in educational learning theories.

*Table 2.1. Key concepts in learning theories taken from (Hung, 2001)*

| | Behaviourist | Cognitivist | Constructivist | Social constructivist |
|---|---|---|---|---|
| Learning | Stimulus and response | Transmitting and processing of knowledge and strategies | Personal discovery and experimentations | Mediation of different perspectives through language |
| Type of learning | Memorizing and responding | Memorizing and application of rules | Problem solving in realistic and investigative situations | Collaborative learning and problem solving |
| Instructional strategies | Present for practice and feedback | Plan for cognitive learning strategies | Provided for active and self-regulated learner | Provide for scaffolds in the learning process |
| Key concepts | Reinforcement | Reproduction and elaboration | Personal discovery generally from first principles | Discovering different perspectives and shared meanings |

In designing effective educational assessment strategics, the assessment depends on the educational learning theory employed (behaviourism, cognitivism, social-cognitive, constructivism) and the educational objectives. These objectives should be designed in a way that reflect the knowledge/skill to be assessed, which it is also heavily dependent on the purpose, summative assessment or formative assessment that we will explain in the next section. Bloom’s taxonomy of educational objectives can be used to guide the setting of objectives and consequently their assessment in both summative and formative modes (Dreher et al., 2011). We will now explain the taxonomy briefly.

## 2.3 Bloom's Taxonomy

Bloom’s taxonomy of educational objectives (Bloom, 1956) is a framework for classifying statements of what we expect or intend students to learn as a result of instruction. It was designed as a framework to facilitate the exchange of test items among faculty at various universities in order to create banks of items. It describes a scheme for classifying educational goals, objectives and standards (Krathwohl, 2002). For Bloom, learning is a process that includes three main domains: cognitive (thinking), affective (feeling emotion) and psychomotor (doing), where the assessment process takes place in the cognitive domain that we have schematically represented in Figure 2.2. The Taxonomy gives definitions for each of the six main levels in the cognitive domain: knowledge, comprehension, application, analysis, synthesis, and evaluation. The cognitive domain is also subdivided into three categories: recall, interpretation and problem-solving.

The recall category is the lowest in the learning process, and involves activities such as memorizing and remembering without necessarily understanding (the knowledge level). At this category, the student restates and summarizes exercises (the comprehension level). The second category, interpretation, requires that student comprehends the topic and transfers it to practical situations (the application level). This demands recognition of relevant components and good design in the learning material (the analysis level). The last category, problem-solving, requires that student combines information to produce new knowledge (the synthesis level) and make appropriate decisions in order to choose the best one (the evaluation level).

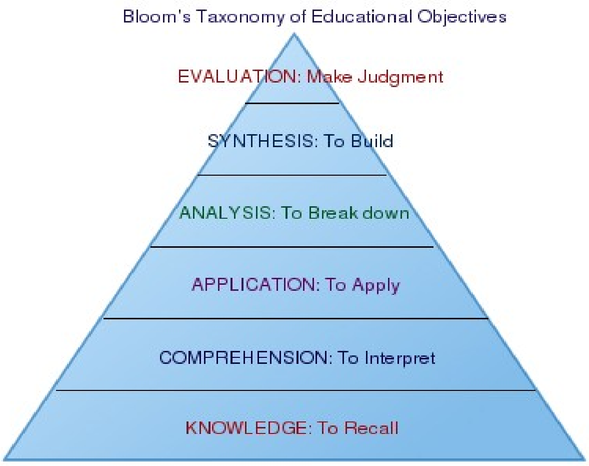


*Figure 2.2. Bloom's taxonomy of educational objectives taken from (Buzzetto-More & Alade, 2006)*

The recall category can be evaluated by multiple choice or direct questions. Furthermore, assessing higher levels of knowledge requires that the student demonstrates skills of synthesis and analysis; these can be best assessed using essays and practical projects. This type of assessments are more demanding for both educators and students; students have to demonstrate synthesis and evaluation, and since this assessment requires structured formats that are more difficult to develop and assess and to provide feedback on, teachers find this more challenging (Dreher et al., 2011).

Technology has transformed educational approaches bringing important improvements in teaching, learning and assessment methods. We will mention briefly how ICT has contributed to enhance learning theories. For instance, Bryceson (2007) explains how technology has enriched the development of the ZPD concept (the scaffolding process). Web-based tools such as threaded computer conferencing, frequently-asked question lists, collaborative workspaces and on-line chat have been demonstrated to be a useful mechanism for the development the ZPD concept.

Within the social constructivist approach, the design of teaching activities using technology such as on-line discussion boards provided by web-based learning management system/virtual learning environment (LMS/VLE) are a successful socialisation mechanism for students that allows the development of communities of learning. Additionally, class materials that contain challenging tasks are viewed positively and accepted by students, because they consider that these materials can encourage them to become more involved in an on-line course. Using well-designed materials is considered an important scaffolding mechanism that can engage the student, especially at the beginning of a course, when students are somewhat isolated. Likewise, the actual tendency is towards the use of socially-oriented environments to communicate with others fostering knowledge construction.

## 2.4 Feedback in the Assessment Process: Formative and Summative

The concept of feedback is important for the assessment process since it is a way in which students and teachers communicate. Feedback is a comparison between student's present state and goals and standards. This is then used to determine if they should continue as is or whether a type of change is required. Feedback can occur between teacher and student, student and student (peer feedback) or learners that generate their own feedback (self-assessment) in order to check their engagement with learning activities and task (Nicol & Macfarlane-Dick, 2006). The literature divides feedback and assessment into two main categories, formative and summative.

Feedback is especially important to students as it is the key place where students can obtain comments and support from teachers. Cross (1996) remarks “one of the basic principles of learning is that learners need feedback. They need to know what they are trying to accomplish, and then they need to know how close they are coming to the goal". In order to develop better strategies to improve student learning, these researchers suggest three conditions for developing adequate feedback which are: (1) a knowledge of standards, (2) the necessity to compare these standards to one’s own work and, (3) taking action to close the gap.

The concept of “closing the gap” is commonly used in formative practices to improve student's learning. It means that using feedback teachers can close the distance between current and desired student's performance (Heinrich et al., 2009). It involves teacher's timely feedback so that students can take it into account while working on their following assignment.

*Formative Assessment* has been recognized in recent years as a key strategy in enhancing students' learning. Black & Wiliam (1998a) gives evidence that improving formative assessment in the classroom leads to higher achievement in students. Universities and colleges have developed strategies for implementing formative assessments, particularly as an approach to teach students how to learn to reflect.

Formative assessment is a process where teachers obtain information about the state of a student's learning, and they use this information to determine specific strategies in order to adjust or improve student's learning. Formative assessment is used to indicate the continuous process of communication between the teacher and the student. In this process teachers evaluate students' performance by reviewing student assignments, making corrections and giving suggestions to students in order to improve their assignments. Formative assessment gives evidence that allows one to adapt the teaching activities to meet students' needs.

The main objective of formative assessment is to make students aware how effective they are as learners through reflection and regular feedback. The process involves the concept of feedback where students can learn from their own reflection (self-reflection) in a self-assessment process and/or from peer interactions (peer-assessment) (Hodgson & Pang, 2012). The concept of self-regulation refers to the degree to which students can regulate aspects of their thinking, motivation and behaviour during their learning (Pintrich & Zusho, 2002). Students that are more self-regulated are able to produce better feedback to achieve their goals (Nicol & Macfarlane-Dick, 2006).

Timeliness is a key attribute for formative assessment. This means that teachers should give their recommendations whilst the student is still interested and engaged in doing assessment tasks. Formative feedback alerts students of their mistakes so that they can improve in areas of weaknesses and avoid repeating the same mistakes (self-evident) (Gill & Greenhow, 2008). This is also pointed out by Taras (2002) who stresses that marks should be given to students at the right moment; marks given to students before understanding the feedback can interfere with student's judgements making self-assessment difficult. Marks can be suitable for formative proposals, but not in isolation and not before feedback and judgements have been understood and internalized, especially since bad marks without feedback can affect a students' learning and confidence.

These attributes give insights in how the formative assessment process can improve the communication, thus determining whether it is effective. These attributes should be considered during the design of formative assessment activities.

Formative assessment has shown to be very useful to students when they can learn from peer comments and make further revisions before their final submission. Formative feedback implies a partnership and reciprocal relation; it could be said that the relationship finishes when students produce a correct assignment, which shows that learning goals have been addressed correctly.

*Summative assessment* is the conventional form of the assessment practice. This plays a valuable role during students' learning process. Summative assessment means the final evaluation of student's learning during one course or assignment, frequently with a mark attached. In contrast to formative assessment, it is commonly used as a final mark or grade where no feedback is provided, containing tasks such as examinations, tests or final essays/assignments. Both summative and formative assessment practices can be summarized in the words of Scriven (1991) “when the cook tastes the soup, that's formative; when the guests taste the soup, that's summative”.

However, summative assessment is commonly seen as a learning approach that may encourage surface learning and low order thinking because it is frequently used to assess declarative knowledge and basic application with no evidence of personal reflection and deep understanding (Smith & Wood, 2000), (Gikandi et al., 2011).

A common idea is the belief that if a student gets a good mark, he/she has obtained good knowledge and skills. As a result, students commonly perceive assessment tasks with awarded grades as more important. Taras (2002) suggests that in a higher education context the message that teachers are giving is wrong, because it stresses obtaining grades over learning. As a useful suggestion, she advises to be more supportive of students' participation in the process through peer and self-assessment.

An adequate assessment process is one that includes both formative components into summative assessment. For example, if the design of a task involves explaining some references to a topic, students can write a draft of their work and discuss it among themselves to get formative feedback before sending it to the teacher (summative).

Nevertheless, Reiners et al. (2011) stress that when designing high quality formative assessments, there are some issues to take into account. The educational model is often restricted by an increased workload for teachers, economic pressure (for administrators) in a competitive market and dissatisfaction (for students) with poor quantity-quality ratios where assessment activities are evaluated on simplistic levels. They point out that a successful strategy must find the right balance between quality control regarding both formative and summative assessment, educators' skills and effort, time and costs, including also the universities' resources.

## 2.5 Conclusions

We have reviewed learning theories as the foundations of the design of the process of teaching, learning and assessment and how they have influenced the way in which educational institutions have incorporated these processes in their curricula. As the design of a curricula includes defining the learning objectives, Bloom's taxonomy plays an important role in determining how to incorporate and evaluate them. Depending on the learning objectives, the instructional strategy will determine the type of assessment, formative or summative.

Formative assessment practice is important for improving students learning and performance, through supportive activities between teachers and students or between students (peers to peers). Formative assessment can have a remarkable impact on how students are acquiring knowledge and building abilities and skills.

Through reflection on their own learning, students can assess their activities, comparing them with the standard required in order to determine if an adjustment is necessary. Summative assessment reports a grade to teachers and academic managers, showing how well students are doing.

Although ICT has brought remarkable improvements to the process of teaching and learning in educational institutions, assessment technologies have not taken sufficient advantage. The assessment process has been an area slower to incorporate new technologies than other aspects of teaching (Byrnes & Ellis, 2006); (Ellis & Goodyear, 2010). The assessment technologies are only in their infancy and will take time to grow.

There is an urgent need to reconsider the role of assessment and feedback practices in teaching and learning methods. The redesign of the curriculum activities is crucial, as Fluck (2010) stresses "student learning is unlikely to change in response to ICT without a transformation of assessment practices". Nevertheless, incorporating educational technologies following the appropriate learning objectives, expressed for the right pedagogical reasons, is the correct way to widely transform and enrich these practices.

Once we recognize that technology is the essential enabler for education, it will be in turn the base for building a more equal society. Emerging technology is creating significant growth opportunities in teaching, learning and assessment in universities. Therefore, the role that government and policy makers play is to facilitate and stimulate the development of innovative technologies and this is crucial.

Emergent economies need to set punctual initiatives to allow new opportunities for education, research and innovation. To realise what has to be done in developed economies for encouraging major investments in these areas is a way to get the most opportunities. Establishing agreements and collaborations between both economies for supporting education and research in countries that need this is the most important way to support high quality education needs.

# Chapter 3.Technology-Enhanced Assessment Process

## 3.1 Introduction

ICT has touched and transformed all fields of human activities. If we add the fact that technological capability is increasing exponentially, we can expect an even bigger impact on all human activities in the future. In the educational context, ICT has played an important role, transforming the way instructors teach and students learn. ICT has helped to enhance educational methods and approaches, making them more efficient. Recently, we have seen an increasing emphasis embedding them on all educational levels.

However, some might argue that educational technologies have not changed the way the teachers instruct or the extent to which the students learn. Nevertheless, the educational tools that students use to learn are changing. These have modified the process of teaching and learning per se. The influence of ICT on the processes of teaching and learning has been significant. Researchers have turned to ICT as a way to fulfilling the requirements for learning in a modern society, and this has created great demand from a diverse range of actors ranging from businesses to institutes of higher education (Sun, Tsai, Finger, Chen, & Yeh, 2008). There are many studies regarding the benefits of ICT on the teaching and learning process. For instance, Wang & Wang (2009) point out that ICT enables communication between instructors and students by serving as a platform to facilitate teaching and learning. Also, Gunasekaran, McNeil & Shaul (2002) state that ICT encourages interaction and communication between students and instructors. The fact is that technological tools are also becoming part of the equipment required for 21st century education.

E-learning becomes an attractive learning strategy particularly when individuals cannot obtain the education they want from local sources such as when students require a specialized course that is not part of the curricula of the university, or when people have to deal with daily responsibilities either at home or at work that do not allow them to attend a course physically.

In the last few decades ICT has played a key role in the real world. Currently, organizations require a workforce consisting of people who have the abilities and skills to be able to make their own decisions, high performance in teamwork, and the ability to manage effectively complex situations, which suggests “that the ability to use technology will become a standard job-entry requirement” (Bennett, 2002). Moreover, ICT is supporting global businesses in the process of employee learning and development.

In recent years, there has been an important number of innovations in education such as the enormous proliferation of MOOCs (Massive Open Online Courses). These have the potential to provide education on a global scale. This enables the instructor to spend less time lecturing and more time interacting with the students. They serve as an effective means to scale education to students when other (in-person) forms of instruction are unavailable (Cooper & Sahami, 2013).

On the other hand, the emergence of educational analytics, which means using analytics in the educational context, is giving measurable advances for teaching and learning. These techniques allow us to make evidence-based decisions, actions and personalisation in diverse areas of education. These innovative tools are bringing important benefits to all stakeholders such as students, teachers, lecturers, departments until educational institutions as well as regional- and national-level stakeholders (Van Harmelen, 2012).

The impact of ICT has widely expanded in all fields; we are especially interested in applications for teaching mathematics and allied subjects. The work of Gunasekaran et al. (2002) gives evidence of the effectiveness of using ICT to do this. Their study explains how researchers (Larson & Bruning, 1996) examine perceptions in an interactive collaborative mathematics course. Their conclusions show that “the distance learning format gives teachers access to more resources, is useful for under-achieving students, and is an effective way to implement national curriculum and instruction standards”. Likewise, the study of McCollum (1997) describes a professor at California State University who divided a statistics course, teaching one group in a traditional way and another in an on-line version of the course using web-based tools (website, e-mail, and an electronic chat room). The students who took the on-line course did better than the others.

Even though some technologies, environments and tools have been developed to support the learning and teaching processes, the assessment process is still in its early stages. Some policy-makers and senior management have redesigned the assessment practices in the universities using ICT and have achieved favourable results. For instance, the work of Heinrich, Milne & Moore (2009) shows some benefits in employing technology such as improved marking quality and feedback, support for human markers, insight into student understanding through quizzes and tests, ease of electronic submission and handling of assignments. Dreher et al. (2011) argue that automated assessments are technological tools that carry the potential to improve the assessment process for all stakeholders. Students can receive immediate and objective feedback, educators can focus on teaching and giving formative feedback, and administration/management can be performed at lower costs.

Whitelock & Watt (2008) point out that ICT has also contributed significantly to the educational assessment process. They mention that “the benefits gained include student retention, enhanced quality of feedback, flexibility for distance learning, strategies to cope with large student numbers, objectivity in marking and more effective use of virtual learning environments”. In fact, ICT can make a huge difference in the educational process by introducing new ways of learning, teaching and assessment by using novel technological tools.

Taking into account that students as a “digital natives” engage in an educational system that was designed in a pre-digital era, they nevertheless need to teach themselves modern life-skills through participation in the networked society and must learn industry-relevant skills and knowledge on the job. Social-technological innovations are the gateway to the future for universities. It is therefore important to examine the adoption of and resistance to educational innovations in universities (Dreher et al., 2011).

However, the gap between understanding the benefits of on-line assessment, and having staff engaging with it in day-to-day assessment activities is significant. Thus, it is crucial to understand which are the most common factors that affect the uptake of electronic assessment technologies. The key studies were chosen we then selected the most common factors included in Table A.1 in the Appendix A. Table A.1 shows only a small section of the original table. We classified the crucial factors which might help to build a successful implementation of electronic assessment technologies. It is also important to consider that a teacher's engagement is determined by individual educational beliefs or disciplinary differences, attitudes working with technologies and self-efficacy (Chew, Jones, & Blackey, 2010) that we will analyse in the case of study.

## 3.2 e-Assessment

The specific process of assessment using ICT has come to be known as electronic assessment or e-assessment. It includes the entire assessment process from designing assignments to storing the results. It involves the assessment process such as coursework submission, peer-assessment, grading and feedback, traditional examination and quizzes from the perspective of students, tutors, learning establishments, awarding bodies and regulators, and the general public (JISC, 2007).

The crucial role that technology plays is building a useful link between the processes of teaching and learning and assessment creating new approaches and opportunities for enhancing learning goals. As Bennett points out in his paper written as early as in 1998, computer-based assessment opens up new opportunities for innovation in testing and assessment (Bennett, 1998).

However, e-assessment practices must not only be seen as an electronic tool embedded in the same traditional teaching methods; it has to reach further objectives and has to be a carefully planned process. It has to mainly be designed following pedagogical principles rather than just embedded innovative technology (Whitelock & Brasher, 2006) and/or to deliver only an automated version of item-based paper-and-pencil tests.

ICT opens new possibilities for innovative assessment practices. Universities might capitalize on the full power of ICT to innovate by providing a richer experience of student learning. Universities are becoming more aware of this and are transforming and enriching their practices by using digital assessment technologies. E-assessment also represents an attractive option for institutions looking to address the logistical problems associated with the increase in student numbers entering higher education (Walker, Topping, & Rodrigues, 2008). Furthermore, e-assessment also helps to speed up educational processes by eliminating paper-based processes such as printing and shipping, which represent a cost to both universities and students. It thus becomes an attractive strategy for administrative authorities in universities.

Dreher et al. (2011) remark on the pedagogical benefits obtained, particularly in feedback practices, by using e-assessment technologies. The technological tools allow educators to be freed of certain tasks, such as marking hundreds of assessment items, and therefore they have more time and energy to spend on giving more meaningful formative feedback to students. Educators can thus increase the frequency of self-assessment with higher-order learning outcomes to enhance the experience and quality of the learning. Students can in turn be freed to determine their own learning path along defined milestones and assess their learning for successful performance. They mention that the real benefit for students is getting immediate feedback, which enhances their learning performance and also activates their intrinsic motivation within the learning setting. In short, it reduces staff workloads whilst improving the quality of assessment for students.

By taking advantage of the use of e-assessment, (Dreher et al., 2011) also discuss how universities report commercial benefits. In this respect, the reputation that universities want to obtain or maintain is an important issue. Since the quality of education is valued by society in successful graduations and post-graduate job performance, high quality education is a key driver for new student enrolments and a seed of research and business projects, including endowments and sponsorship. Thus the pedagogical benefits of improved assessment methods and outcomes can affect the overall university performance. Moreover, automated assessment can trigger an improvement of the administration and curriculum planning, as researchers (Dreher et al., 2011)  point out precise calculations of financial costs based on the number of students and shorter time spans between exams and results.

Technological assessment is gaining more popularity in enterprises; these are now using it as a way to assess new job candidates. It is also used to evaluate their workers in order to certify job proficiency.

The two main classes of technological tools for educational assessment are broadly classified as *e-testing* and *e-portfolios*. Next, we shall mention some cases of e-testing which show how technology has contributed to provide innovative ways of teaching, learning and assessment. Later, we briefly mention e-portfolios.

Relating to e-testing technologies Hodgson & Pang (2012) discuss how to engage students in formative assessment practices by doing on-line multiple choice questions (MQCs). They report a strategy that help students to reinforce new concepts by encouraging students in activities that allow them to make multiple attempts in the context of a statistics course. The researchers use technology to promote students' participation in on-line tasks on a regular basis. The tasks are useful for students since they can check the correct answers and thus evaluate their own performance. Therefore, they are able to reflect on what is taught in class and think critically, in a process of continuous reflection on their performance. They show how educational technology has supported students in reflection and led them to take greater ownership of their learning.

These researchers state that tests with MCQs (one answer and a few distractors) in on-line learning environments have been widely used as a method of both formative and summative assessment. They stress that these on-line quizzes bring benefits to students by providing timely feedback and that their use motivates students to keep practising during a semester. Also, MCQs can be set to examine a broad spectrum of declarative knowledge of a subject. A special advantage of the use of on-line MCQs for formative assessment practices is that it allows multiple attempts to answer a question, which means that these questions can be used in pre- and post-course tests; commonly a chosen score (highest or average) after a number of attempts can be considered as the final mark for a formative assessment. This means that students learn from feedback following their attempts.

Hodgson & Pang (2012) conclude that on-line formative assessment activities help students to realise of “the gaps in their performance through continuous feedback” from the on-line environment. A web-based environment is a good way of providing learning challenges, particularly for large classes where it is possible to apply a randomised quiz to motivate students to make multiple trials. The researchers conclude that the learning opportunities for students were enhanced. There is “more time for self-regulated learning and reflection on what was learned; students can clarify misconceptions in face-to-face discussion with peers; and peers feel more confident to ask for help in a supportive learning community”. However, the point is that the usage of MCQs supports recall of memorized knowledge without checking deeper understanding as can be included in a taxonomy of educational objectives (Bloom, 1956).

The study by Gill & Greenhow (2008) reports evidence of the benefits of the interaction between students and feedback received on-line while they interact with the computer-aided assessments (CAAs). These researchers focus on providing rich feedback to the students when they answer multiple-choice and responsive numerical input-type questions that compare a student's input, an answer, against that resulting from a coded malrule (an incorrect rule for syntactic transformation of a mathematical expression). The technological tool reports exactly where the error was made and provides a complete solution that allows students to be able to determine their errors. By making students engage with the feedback, they are then able to relate the aspects of the feedback to their written work, such as use of diagrams, presentation of solutions and correct notation of vectors, demonstrating that they have developed organisation and presentation skills. According to the researchers, the study shows that students are able to improve their performance in formative and summative assessments while they are engaged with the CAA assignments, especially by spending time studying the feedback.

These authors remark that students do engage with formative assessment activities, even when no marks are allocated, due to both the quality of the CAA and a structured and supportive environment (lab sessions are scheduled in students' timetables) which shows that when students engage with high-quality feedback, the benefits appear to go further than simply short-term recall.

Other researchers have also explained how a web-based learning tool can help students to improve problem-solving skills and performance. For instance, Crippen & Earl (2007) explain how a web-based testing environment providing worked examples and self-explanation prompts has the potential to improve problem-solving skills and conceptual understanding. They mention the use of worked examples, (detailed problem solutions that contain identifiable qualities and characteristics) are designed to provide students with some structure for understanding what is the solution of a example without giving them a script or algorithm. Researchers worked through these examples with their students; their results suggest the combination of a worked example with a self-explanation promptly produces improvement in performance, problem solving skills, and self-efficacy.

Reiners et al. (2011) have pointed out that automated assessment systems only support memorized knowledge. Nevertheless, the recent technological advances in automated assessment are a convenient option. Emerging technologies on assessment intend to support interpretations of short answer and essay type questions. These educational tools would support interpretation and problem-solving levels (Reiners et al., 2011). For instance, automated essay grading tools (AEG) are computer-based tools to assign grades to essays written in an educational context. These tools are based on natural language processing and normalization techniques which compare students' written words of an essay with a model solution (normalized word vectors and their frequency from the essay are mapped to their corresponding root word in a thesaurus). The research of Nicol & Macfarlane-Dick (2006) shows that essays can assess higher-order learning.

However, Reiners et al. (2011) argue how the success of these innovative tools is being blurred by the idea that these cannot assess higher order tasks as accurately as human beings would do. In the case of assessing student's goals through electronic essays tools, there is a current belief that human markers are superior to computers at the tasks of understanding content and making comparisons between student essays and a model solution. The researchers argue that the use of automated essay grading tools (AEGs) refute the idea that computers cannot do human activities that require higher order thinking. They mention that "while this may be true for many endeavours, it is no longer true for grading essays". As a result they advise that electronic assessment, particularly automated essay grading is an option that works for universities, emphasizing the idea that technology works as accurately as human markers enhancing formative feedback, saving time and money.

Other useful e-testing technologies that have gained broad popularity are plagiarism assessment tools. These are tools that compare a document to a set of 'genuine' reference documents in order to retrieve similar patterns of text. Although these tools "do not assess learning or application of concepts/knowledge" (Reiners et al., 2011), these have been successfully applied in universities as practical and efficient tools to assess the originality of written essays.

An e-portfolio is defined as "the product, created by the learner, a collection of digital artefacts articulating experiences, achievements and learning. Behind any product, or presentation, lie rich and complex processes of planning, synthesising, sharing, discussing, reflecting, giving, receiving and responding to feedback. These processes are the focus of increasing attention, since the process of learning can be as important as the end product" (Gray, 2008). These technologies have proved to be important educational tools, that promote and support learning (Alexiou & Paraskeva, 2010) and teaching leading to more profound forms of learning, adding value to personalised learning that serves as scaffolding approach of understanding and engagement. E-portfolios also facilitate the transition between institutions and stages of education, supporting education and employment, staff appraisal and applications for professional accreditation, and supporting learners based in the workplace (Joyes, Gray, & Hartnell-Young, 2010).

The JISC in the UK is a very useful source for advice on how to implement effective practice in the use of e-portfolio systems and tools, as well as to determine their implications for teaching, learning and assessment. JISC has worked in partnership with other sectors and bodies to develop and provide guidance to institutions on effective e-portfolio practice to support lifelong learning. They have proposed the development of standards and piloted e-portfolio technologies. The main driver for institutional e-portfolio initiatives in the UK is the Personal Developing Planning Policy (QAA, 2001). Although it has not just been the unique driver, according to Joyes et al. (2010) these have also been “the importance of retaining students, widening participation, and increasingly, reflective learning have also contributed to widening interest in e-portfolio tools and technologies”.

E-portfolios systems have proved to be an important tool to enhance quality learning, according to Gray (2008) in the JISC report *effective practice with e-portfolios*, these electronic tools can serve to develop higher-order functions, leading to students becoming independent enquirers; creative thinkers; reflective learners; team workers; self-managers and effective participators, skills that employers and higher education want to see developed in current generations of learners. Moreover, the creation of an e-portfolio involves a critical process: reflection. This process is a key aspect in the development of deeper learning through self-reflection and self-assessment which includes developing activities such as planning, goal-setting and future reflection. The process also helps to build up a range of skills including critical thinking.

Joyes et al. (2010) report on the tangible benefits of the use of e-portfolios. They classify the benefits as efficiency (time-savings for students, academics and, administrators), enhancement (improving quality of evidence and feedback, skill development, satisfaction and increases in recruitment and retention) and transformation (innovation and changes to institutional policy). However, they also point out the drawbacks, since e-portfolio implementation is particularly complex, in part due to the number of stakeholders involved since portfolios can be used in several contexts and purposes. They suggest that there are threshold concepts related to e-portfolio implementation and that developing an understanding of effective practices is not straightforward.

Regardless of the educational technology used in assessment to implement, e-portfolio or e-testing, there are some generic skills and knowledge required according to the *e-assessment: guide to effective practice* (2007). The guide suggests that all staff involved, irrespective of their role, should have (or be trained to have) the following skills and knowledge:

- *A broad understanding of assessment principles.*
- *An understanding of security importance for conduct assessment and a security measures knowledge required for e-assessment (particularly to their centre).*
- *An overall familiarity with the e-assessment environment and delivery platform(s) especially to their centre.*
- *A recognition of possible malpractice in e-assessment and the precautions needed for its prevention.*
- *A legislation awareness relevant to the centre operation. The general regulations of relevant awarding bodies, and regulatory authority guidelines and codes of practice.*

The UK government has proposed several initiatives to recognise the skills and knowledge of people involved in the delivery of e-testing such as the *Level 3 Award for delivering e-testing*. This initiative recognises the importance of key aspects such as security, legislation and regulations of adequate e-assessment practices. The units that make up this Award form part of the Teacher Qualifications Framework developed by (Lifelong Learning UK, 2010).

In order to gain a wider perspective of emerging technologies that will impact the future of educational systems, the Horizons Reports (HRs) are a good starting point. These reports are produced by The New Media Consortium, NMC (2014) and the EDUCAUSE Learning Initiative, ELI (2014). NMC is an international community of experts in educational technology whose role is to help universities, colleges, museums, and organizations drive innovation across their campuses. ELI is a community of higher education institutions and organizations committed to the advancement of learning through the innovative application of technology. Since 2004 both have made yearly predictions of the impact of ICT by using three temporal horizons: the year of the report (short-term predictions), the next two years (mid-term predictions), and the four years following the report (long-term predictions).

Regarding assessment practices, the Horizon Report 2014 (NMC Horizon Report, 2014) discusses the actual trend of learning analytics provides statistical and data mining tools that can improve student services, retention and aims through adaptive learning strategies. On-line learning platforms are generating a large amount of data about student activity and dashboards provide both students and teachers with an overview of this data. This can help students realize how they are doing and help teachers identify students who might need more help and support, making improvements to students' performance and personalizing in the learning experience.

The increasing importance of assessment practices in educational systems is a global phenomenon. Countries such as the UK have seen the e-learning movement and e-assessment as important strategic initiatives whose growth potential is enormous. The UK has set government strategies to include innovative assessment practices. These assessment practices include; assessment feedback, plagiarism/academic integrity, assessing and recording student achievement, and methods of assessment such as essay-type, MCQ, portfolio, etc., and formative versus summative assessments. These practices should have been included in all UK universities by 2010, (Higher Education Academy, 2010). We have included a brief review of the UK awarding bodies and regulators involved in creating strategic initiatives in this area as shown in Appendix B. However, others countries such as Mexico have made insufficient investments in science, technology and innovation. As a result, the potential increase in their economy is inferior to that necessary to reach the level of other countries and achieve a comparable competitiveness to that of other emerging economies. In accordance with all other available indicators the general level of innovation in Mexico is really low, not only in comparison with other Organisation for Economic Co-operation and Development (OECD) countries but also with the most dynamic emerging economies (p*erspectivas OCDE: México reformas para el cambio* OCDE, 2012).

## 3.3 Factors Affecting the Adoption of e-Assessment Technologies

### 3.3.1 Factors Inhibiting e-Assessment Uptake

We will focus on the factors that limit the adoption of ICT at universities in order to possibly avoid repeating the same mistakes. It is believed that not only is it useful to figure out the main barriers, it is also important to know what facilitates adoption; it is also very valuable to focus on what strategies have been successful for universities in order to implement e-assessment.

We will identify facilitators and barriers to the adoption of educational technology for the assessment process. This demonstrates how well a university is prepared for it, whether the university's staff have the skills to carry out the plan and to what the extent all stakeholders are involved in the plan. This will help us to design and implement electronic assessment plans successfully. On the other hand, it also helps in designing future policies for the adoption of e-assessment as well as help to establish a context for commercial agreements related to the assessment by computers.

In order to obtain a clear view, we take into account the model of Ocak (2011), who identifies categories and themes to classify and examine the impediments that face faculty members in the adoption of blended learning environments. The study identifies three categories and eight themes as results of faculty members' problems with blended courses. The categories were classified as instructional processes, community concerns, and technical issues. The themes derived from these categories were identified as *1. complexity of the instruction*, *2. lack of planning and organization*, *3. lack of effective communication*, *4. need for more time*, *5. lack of institutional support*, *6. changing roles*, *7. difficulty for adoption of new technologies and 8. lack of electronic means*. Considering this model as a reference, we identify the categories and themes that affect the adoption of electronic assessment technologies which we have classified as “administrative structures”, “faculty concerns” and “technological infrastructure and systems” categories and their related themes that we have represented in Table 3.1.

*Table 3.1. A model representing the main factors affecting the adoption of e-assessment technologies (own elaboration)*

| | | | |
|---|---|---|---|
| Administrative structures | **Institutional policies**<br>-Lack of incentives (tenure, promotion)<br>-Lack of resources, equipment, infrastructure<br>-Ineffective dissemination of e-assessment | **Administrative structures**<br>-Changing of roles<br>-Lack of support on legal matters, such as plagiarism, data protection, intellectual property rights<br>-Lack of skill and understanding<br>-Health and safety issues<br>-Lack of key roles for administration, support services and departments<br>-Resources withheld by senior management | **Funding**<br>-Not enough investment for technological infrastructure<br>-Lack of new projects<br>-High cost of licences<br><br>**Communication**<br>-Lack of interdepartmental communication<br>-Lack of communication with academic staff |
| Faculty concerns | **Pedagogical research**<br>-Inability to evaluate higher levels skills<br>-Is viewed as a secondary to authoring, marking, reporting<br>-Not appropriate for particular subject<br>-Lack of development of examination procedures<br>-Making ICT the focus instead of pedagogical matters | **Attitudinal issues**<br>-Lack of willingness<br>-Lack of confidence and reliability<br>-Fear of failure<br>-Feelings of isolation<br>-Wrong expectations<br>-Fear of anonymous submission of assignments<br>-Threat of reductions of faculty members<br>-Feelings of discrimination by “non-digital natives” | **Training**<br>-Lack of computer literacy<br>-Lack of expertise in the design<br>-Complexity of instruction<br><br>**Time**<br>-Lack of time |
| Technological infrastructure and systems | **Systems, applications and environments**<br>-Limited functionality<br>-Incompatibility<br>-Limited availability<br>-Lack of reliability<br>-High risk of technological failure<br>-Limited availability of internet-connection | **Pilot projects**<br>-Difficulty of adoption<br>-Insufficient funding<br><br>**Security issues**<br>-Passwords, cheating, impersonation | **Technical support plans**<br>-No backup plans<br>-Lack of follow up<br>-Lack of technical support<br><br>**Upgrade technology**<br>-Lack of plans for renewing systems |

Table 3.1 gives a wide view of the current landscape of electronic assessment. It identifies the main barriers to the adoption that affects all stakeholders involved in the electronic assessment process recognizing the key points which allow us to clearly identify strategies and tactics that may help to minimize the cited barriers.

The table is divided into three categories. We call the top level “administrative structures” which represents policy-makers, senior management and university staff who propose, design and implement the policies and educational plans related to the adoption of technologies, including technologies of e-assessment.

We call the second level “faculty concerns” which represents the faculty members' needs and problems to embrace ICT in their teaching and assessing practices. This includes important matters such as pedagogical and attitudinal issues, training concerns as well as spending time learning educational technologies.

The bottom level called “technological infrastructure and systems” includes matters related to availability and efficiency of the use of ICT resources. These factors are crucial in adopting ICT, particularly for teachers (Bhuasiri, Xaymoungkhoun, Zo, Rho, & Ciganek, 2012). Teachers who perceive that these requirements are attended and also satisfied should be more willing to use assessment technologies.

As seen in Table 3.1, for teachers a common and continuous barrier is the lack of time to learn and use an e-testing environment (Whitelock, Mackenzie, Whitehouse, Ruedel, & Rae, 2006), (Sim, Holifield, & Brown, 2004). However, educational research shows also that the use of ICT can compensate for time spent in grading/scoring practices which are commonly considered as tedious and time consuming practices.

The time invested in creating high-quality materials for e-assessment is another important barrier. This also includes training and experience to develop creative questions (Brasher & Whitelock, 2006). Bull (2000) proposes some strategies to overcome these barriers such as building up banks of questions and to share common questions. Although there is plenty of material available on the web, Bull advises that their quality is often low. In the same way, she points out that security issues, copyright and organization are serious obstacles for the effective use of question banks. Importantly, the activities must be recognised as valid academic products that must be developed by teachers, and should be included in a regular timetable. However, Bull also mentions that until it becomes a mainstream activity the efforts to release time for these activities will be lacking.

Likewise, “difficulty with using systems” (Warburton, 2009) and “lack of adequate computer training policies” (Whitelock et al., 2006) are two common barriers mentioned in literature. Not including plans for renewing systems or inadequate technology upgrade plans is an important barrier. In the day to day of academic activities it is common to hear complaints about "the system has failed" which can cause both teachers and students to feel discouraged in using the systems. The key point here is to implement appropriate policies and plans that keep systems running. This clearly depends on organisational structures, for their implementation and follow-up.

On the other hand, we want to highlight an important finding which must be carefully taken into consideration. It is related to teacher's perception that use of ICT has been imposed by institutions regardless their needs. It makes teachers feel unwilling to use ICT, as they think it can affect their autonomy. To avoid this McCann (2010) advocates guidelines to manage cultural change that includes choosing a leader, defining a vision, starting with pilot projects, motivating teachers with hands-on training and informing them about the system with particular emphasis on results and impact. The above points show that many of the obstacles are related to academic staff (Whitelock et al., 2006).

Other important drawbacks are that the selection of ICT for teaching, learning and assessment is an approach that does not include enough detail of a pedagogical plan, which deters the adequate use of technology (Heinrich et al., 2009). The combination of deep skills and technology and pedagogy knowledge for e-assessment are not common questions. To develop pedagogical and technological strategies and make them accessible to all those involved, is an effective way to cope with it. The provision of training sessions for teachers, resources, advice and guidance is also a way to overcome this barrier. To obtain tangible results in universities will require a clear support by policy-makers at the institutional and national level.

From the students' point of view, although they are very familiar with the use of technology, they are still worried about the security of testing (Cassady & Gridley, 2005), possibilities of cheating (King, Guyette, & Piotrowski, 2009) and the fairness of question banks (Dermo, 2009). If students do not have enough confidence in a test, that can affect their levels of engagement and cooperation (Domino & Domino, 2006).

Other important barriers are highlighted by Bull (2000) in their annual report on Computer-Assisted Assessment (CAA) in the UK. The report points out that a "lack of understanding of the limitations and potential of the method of assessment and the assumption that it is not possible to test higher order skills using CAA" are two important obstacles for the implementation of CAA. The report advocates "to include staff development at a generic and departmental level; to provide good examples of materials in a particular discipline". This is a powerful way to show that CAA can test higher order skills. However, the report advises that because of the high level of skill needed to create such materials, these are difficult to make. This is also shown in the research of Warburton (2009), where it is stated that factors such as "fear of CAA failure, ineffective dissemination of good CAA practice, difficulty using the systems and resources withheld by senior management" are other obstacles.

Concern over security issues is another factor restricting the adoption of electronic assessment technologies. Bull (2000) points out that this results in "*techno panic*, a phenomenon which manifests itself as a demand for much more stringent security measures than would be adopted for paper-based assessments". The causes are: "an inherent unwillingness to participate, resulting in identifying difficulties and reasons for failure; unrealistic expectations of technology; and a misguided belief that students will only consider cheating if they are using technology". The strategies to overcome this issue include encouraging a more sensible approach to security measures and awareness of developments in technology which help the security of examinations. These are also important within the context of assessment strategies and strategies in general.

Aspects of reliability and validity are important in designing electronic exams/test. For educators the main concern is keeping marking reliable, particularly for larger classes (Heinrich et al., 2009). Students are more interested in efficiency, transparency and fairness of their assessment activities, which influences the degree of engagement shown in their studies (Iannone & Simpson, 2013).

### 3.3.2 Factors Driving e-Assessment Uptake

To recognize the factors that facilitate the adoption of technology, we will now consider the strategies and tactics that enable its use. To successfully adopt technology one has to be convinced of its usefulness; to identify which factors are driving each stakeholder to get involved in the assessment process. We have included below the most common factors that encourage their use.

A major factor often mentioned in literature is active institutional and administrative senior management support (Buzzetto-More & Alade, 2006), that strongly supports the proposal of strategies for academic staff development and training (Whitelock et al., 2006), (Warburton 2009), (Heinrich et al., 2009). This is not surprising, as policy-makers and administrators should be the first to be convinced that educational technology can greatly enrich the assessment practices.

The role that policy-makers and senior management play is crucial for educational technology adoption for the e-assessment process. Their support serves as motivation for teachers and students to adopt ICT, as is highlighted in Whitelock et al. (2006), who point out that the main facilitator of effective implementation of e-assessment is the support of the school manager, combined with staff development and pedagogical and technical support. Likewise, Heinrich et al. (2009) notes that teachers need more support from the university management when using automated assessments.

This is also stated in the work of Bhuasiri et al. (2012), whose research tries to identify factors that influence the acceptance of e-learning systems in developing countries. The results of this study are particularly applicable in our research, since e-assessments are an important part of the e-learning technologies. They identify 6 dimensions and 20 critical success factors (CSF) that affect the adoption of e-learning. They define the dimensions; *learner's characteristics*, *instructor's characteristics*, *e-learning environment*, *institution and service quality*, *infrastructure and system quality*, *course and information quality and motivation*; as the principal factors involved in e-learning adoption. Their conclusions show that the “infrastructure and system quality dimension” is the most important from the teachers' perspective. Therefore, it seems that active institutional support is crucial, as we have already pointed out above. Likewise, Heinrich et al. (2009) points out that successful e-assessment adoption depends on the flexibility (willingness) of the academic staff. This is a factor that we will analyse in the case studies in the following chapters.

Another important factor, is the willingness of staff to develop material, which clearly requires specific training to develop teachers' abilities and skills.

The opinions and experiences of educators can influence colleagues' willingness to use a specific educational environment (Heinrich et al., 2009), (Warburton 2008). This suggests that a teacher can agree to use a certain system and disagree to use another one. In this way, his/her opinions can affect colleagues' perceptions and opinions and therefore modify their willingness to use a system. It is another factor that we will analyse in the case studies.

The ideas cited by researchers (Heinrich et al., 2009), (Whitelock et al. 2006) show the importance that pedagogy plays in a technology adoption plan. As Heinrich says “the selection of the technology should be guided by pedagogical design of the assessment” and not as often, the other way around. They also point out that other important factors that facilitate the adoption of electronic assessment are “the removal of geographic limitations, reduction of losing work risk, saving time and resources if printing is not required, the availability of a long-term archive of student work based on the ease of storage of electronic material, and fast return of marked student work” which can also serve as strategies for educators who want to enhance their teaching.

It is noteworthy that the practical benefits for educators are the reduced effort and time spent on assessment practices. When teachers adopt technology in their class, they acquire new skills that improves their performance. Once they have used the technology, they do not want to go back to using traditional practices (Heinrich et al., 2009). Electronic assessment practices facilitate the opportunities for anonymous participation and marking which support group activities, and improve the quality of marking and feedback. Whitelock et al. (2006) mentions that technical support for teachers is an essential facilitator (including technical services and the design of electronic assessment tasks). We will also analyse the impact of this factor on teachers' willingness to adopt e-testing technologies.

More recent research, e.g. Reiners et al. (2011), point out other factors that help the dissemination of strategies of automated assessment technology such as demonstrations, case studies, and hands-on experiences (e.g. 3D Virtual Worlds). Technologies such as advanced plagiarism detection have also been successful promoting advanced automated assessment technologies.

## 3.4 Educational Technology as a Key Element in the Formative Assessment

One of the biggest challenges that higher education faces today is related to the socio-technological change that modern society demands. It requires new ways of teaching, learning and assessing. Here, the role of ICT is crucial; it contributes substantially to transform instructional teaching and assessment practices and builds a useful reciprocal link between them. It is necessary for universities to consider change, since new generations joining universities consist of students considered to be “digital natives” (Generation Y, Generation Z or Generation Net). These students are the product of the current technological culture and commonly use ICT as part of their daily activities, using technologies such as personal computers, and mobile devices (smart-phones, digital tablets). Students use e-mail applications, social networking, instant digital messaging, cloud computing, etc. as part of their daily activities, but are engaged in an educational system designed in a pre-digital era.

We know that educators and students are the main stakeholders in the educational process, and they are who determine the success or failure of the adoption of technology. However, the role of teachers seems to be more important (Selim, 2007), in the sense that, if they decide or are required to carry out part of their teaching activities through the use of technology, students have no choice but to use it. For this reason, teachers need to be willing to adopt these technologies.

It should be mentioned that some teachers have realized that they can improve their teaching and assessment practices by using ICT, but only a few of them are interested in using it. Cuban, Kirkpatrick & Peck (2001) show that although teachers in the United States have access to computers and software, they use them infrequently and not in a sustainable way, which indicates that most teachers are occasional users or non-users. And when this is used in the classroom, most of the time its use is sporadic, without a real redesign of their teaching strategies.

An adequate teaching strategy should include the assessment process. This is considered as the Achilles' heel of all educational processes, due to its role of providing motivation to students through feedback and facilitating learning (Nicol & Macfarlane-Dick, 2006). In fact, doing formative assessment practices using computers encourages immediate feedback, which has positive effects for students in achieving their goals (Brasher & Whitelock, 2006). In most cases, educational technology is functional for both formative and summative assessment practices (Stödberg, 2012).

As a consequence of the socio-technological changes, educational research points out that the current higher education system requires new approaches for assessment and that some of them have already been developed and supported by using technologies. For example, the research of Heinrich et al. (2009) put forward the enhancements achieved particularly doing feedback activities by using electronic assessment technologies which we have represented schematically in Figure 3.1.

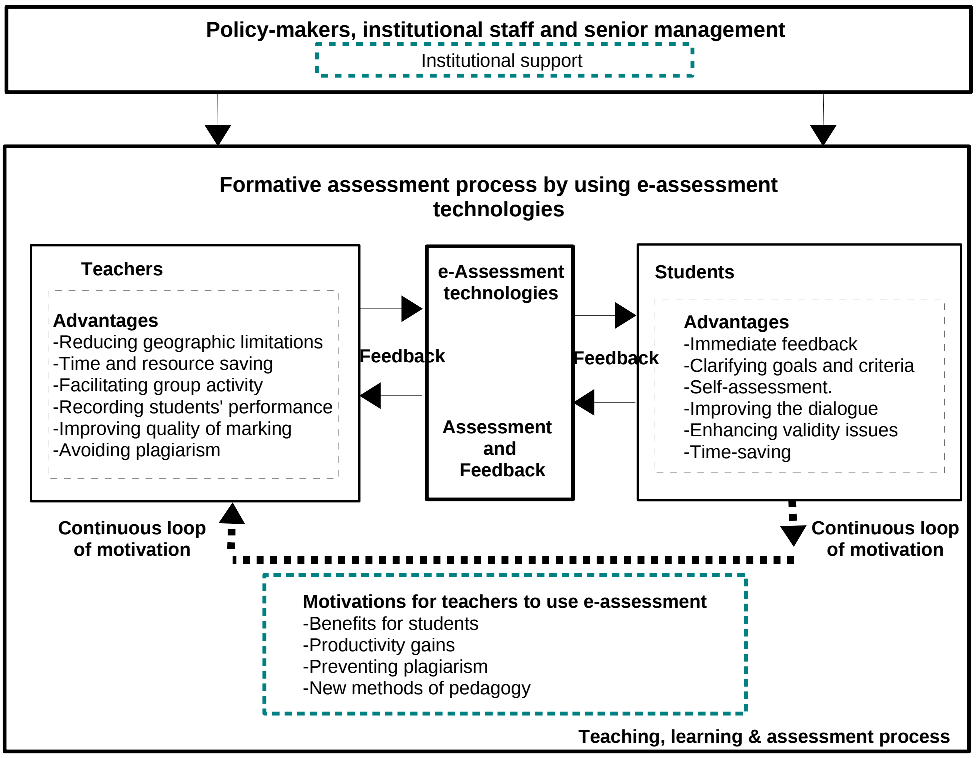


*Figure 3.1. Educational technology as a key facilitator of formative assessment activities to enhance the teaching and learning process*

Here, we want to highlight the benefits that both educators and students can get by using e-assessment technologies. It shows two main blocks; the first represents the institutional support standing for policy-makers, senior management and administrators. The second depicts the process of teaching, learning and assessment represented by teachers and students and their related activities.

The block called *teachers* represents the advantages obtained by teachers using formative assessment activities and feedback. The block called *students* shows how they also gain benefits by getting guidance and immediate support from teachers and/or by doing self-assessment activities.

Teachers also benefit by using innovative teaching and assessment methods, time-savings, and productivity gains. These advantages not only enhance their teaching and assessment practices, but also improve their academic performance. All of these tangible advantages serve as a trigger to encourage teachers to adopt and/or keep using ICT. We have represented this as a continuous loop that links both blocks. We will analyse this aspect in detail in the following chapters.

The role an institutional staff plays is crucial. Figure 3.1 shows how the institutional support can influence directly the whole teaching, learning and assessment process. Policy-makers and senior management have the authority to make technology a common tool for teachers and students. The institutional support can encourage teachers to adopt technologies in their assessment practices. As Whitelock et al. (2006) state, “the facilitators of effective implementation of e-assessment is the support of school managers with steady staff development, pedagogical and technical support for teachers”.

Chew et al. (2010) suggest how some evidence based on good practices was successful to enhance staff practices. They embedded them by following top-down, horizontal and bottom-up strategies. The following two are top-down strategies: (1) Strong support from the senior management university team, means a clear institutional strategy that highlights technology enhanced learning, teaching and assessment and disseminating top-down messages through the institutional websites, blogs and committee meetings. (2) Using a strategy to engage, first, senior members and heads of division/departments. In turn, they can encourage the whole division/department because they are convinced of the benefits of on-line assessment for both staff and students. This strategy includes peer encouragement and discussions.

The “horizontal” strategies, are particularly effective to stimulate self-initiative and willingness. (1) Short and simple user guides and frequently asked questions (FAQ). (2) Disciplinary group training using a “disciplinary-specific language”. This strategy stimulates active group discussion and consensus. (3) One-to-one support. This practice is essential as a post training support for those teachers who lack technological competence. (4) Funded research projects. This is a key method for engaging staff to “use research to support the change”. They are specially motivated as a result of writing a collaborative publication. Also a funding for research is helpful to both young and senior academics.

The final one is a bottom-up strategy. Student-driven engagement has been the most efficient and effective practice and has to be a strategic priority. This strategy can improve the nature and quality of the students' experience.

Taking into account teacher's needs, the research of Reiners et al. (2011) asked teaching staff who had already used automated assessment about the features that they would look for when choosing or using an automated assessment or marking tool. Based on a qualitative analysis of their open-ended responses 7 desired elements emerged: 1. efficiency (shorter time or higher quality in the same time); 2. ease of use; 3. accuracy and reliability without manual verification of each assessment; 4. enhanced feedback for the students and reports for staff and administration; 5. advanced pedagogical opportunities such as assessing higher order thinking skills; 6. greater flexibility and individualization while setting up assessments; and 7. commitment from the institution to apply automated assessment.

Teachers also state that they are looking for integration with existing systems, and administrative features to help organize and archive assessments. Additionally, those staff members who already had used automated essay grading were asked what they found useful in this technology. Their answer was freeing time/energy for other educational tasks; marking the assessment in a shorter time, increasing the accuracy of assessment; reducing the cost and improving the feedback to students.

We want to highlight some helpful insights. To build a successful implementation of e-assessment we must identify which are the principal groups of stakeholders concerned with the assessment process. Heinrich et al. (2009) recognize them as students, teachers and institutional support and management staff, all of which play a different role inside the process; this knowledge provides some features to make a useful electronic assessment project.

In this sense, Buzzetto-More & Alade (2006), point out that a project starts by identifying learning goals and measurable objectives, with carefully planned approaches and activities. The use of specific traits that help define measurable objectives should be related with concepts of a taxonomy of educational objectives such as recommended by Bloom (1956). The role of institutional and administrative support staff is thus crucial.

Likewise, the design of an electronic assessment project should consider curricular alignment, which means that the program should evaluate teaching and learning aims to the required skills and abilities. Educators and institutional staff should determine which are the principal objectives obtainable from the design of the curriculum and how this design can affect the way that the student learns and is evaluated. This means that building a successful link between the educational processes and technology must be a well-detailed process which includes the use of technology for the correct pedagogical reasons (e.g. Heinrich et al., 2009). They state that “the use of technology for its own sake will not improve educational assessment” (Committee on the Foundations of Assessment 2001).

## 3.5 Conclusions

In this chapter we have reviewed the literature on the current state of knowledge and practice regarding electronic assessment, and wish to conclude with some useful insights that summarize the adoption of ICT in the assessment process. Knowing these factors -positive or negative- enables us to determine when the staff are well prepared to implement a plan to adopt technology, as well as the grade of willingness of the main stakeholders.

For teachers the lack of time to develop questions or even to learn the software (Dermo, 2007); (McCann, 2010); (Whitelock & Brasher, 2006); (Warburton, 2009) are important barriers for the adoption of e-assessment. Nevertheless, the literature shows that adopting electronic assessment practices can help teachers save time (Whitelock et al., 2006), which compensates for the time spent in learning and developing e-assessment strategies.

Teachers' training in computer literacy and test construction is another important lesson. (Sim et al., 2004), (Warburton, 2009), (Purvis, Aspden, Bannister, & Helm, 2011), (Dermo, 2007), (Ashton, Beevers, & Thomas, 2008). Another important factor is the design and development of a technology plan that considers sufficient details of pedagogical strategies.

It is fundamental to include aspects of validity and reliability in designing a useful plan for adopting technology to educational assessment, because it depends to a large extent on the level of trust and confidence that students embrace in the assessment practices. This will also be reflected in students' efforts in their learning (Iannone & Simpson, 2013).

In the research of Bhuasiri et al. (2012), the infrastructure and system quality are the most significant categories from a faculty perspective at the universities and found to be also at the educational organizations. Hence, it might be interesting for policy-makers and senior management to initiate strategies regarding funding the development of new educational projects that enhance the assessment practices by the use of ICT. As Whitelock et al. (2006) remark, a successful implementation of electronic assessment depends on active institutional and administrative support. It must not be forgotten that the adequate design of electronic assessment methods must include technology for the right pedagogical reasons as educational research advocates, the use of technology for its own sake does not improve educational assessment (Heinrich et al., 2009). As is shown in the experience of other countries such as UK, where e-learning has been recognized as a movement with a huge growth. The UK government has focused on developing new initiatives to recognize the electronic assessment process as an important strategic initiative. Policy-makers and senior management have the power to foment and create these changes. Also, teachers must be involved in a steady and continuous change. Their strategies must be extended to include all stakeholders involved in the educational processes.

Not surprisingly, teachers also need to take into account that there are wide political and business issues in the background that affect the appropriate development of electronic assessment at universities. To deal with this situation, policy-makers and senior management must learn the best strategies to obtain real progress to all related stakeholders.

ICT has revolutionized the education system by making it more accessible to modern society. This should be an advantage to students, teachers and universities. Technology enables education to be available to more students, including those from the social stratum of the needy. This not only can meet actual demands for higher education, but also offers innovative teaching, learning and assessment methods that undoubtedly will be attractive to new generations of students.

Educational technology also plays a crucial role in automating each task and stage of the assessment process and although its use and effectiveness has already been proven, educational literature shows that there is still a huge need to develop these innovative approaches. Research in e-assessment includes a huge variety of perspectives that leads to continuous change. There is no single solution to the challenge of effective education. However, the willingness to adopt innovative educational assessment methods will indeed make a positive difference to students' learning. We conclude by citing the idea of Stödberg (2012), who points out that knowledge in this area is quickly expanding and there is a need for more studies related to e-assessment. There are still many opportunities that are not being taking advantage of, and the emerging research should be constructed with the aim of proposing specific strategies for developing new approaches to e-assessment.

# Chapter 4. Modelling Techniques for Attitudinal Surveys

## 4.1 Structural Equation Modelling and PLS-SEM

### 4.1.1 Introduction

In this chapter we shall discuss the statistical techniques we apply in the later chapters. We need to select a statistical technique to analyse relationships between variables of the models in the case studies. This technique should be a reliable and versatile approach that also has the capacity to manage small sample size. We decide to use a structural equation modelling (SEM) approach for this purpose.

SEM is a widely used multivariate analytical approach. It is a versatile method to simultaneously test and estimate complex causal relationships among variables, even when the relationships are hypothetical or not directly observable (Williams, Vandenberg, & Edwards, 2009). Combining factor analysis and linear regression models, SEM statistically examines the relationships between theory-based latent variables and their indicator variables by measuring directly observable indicator variables (Hair, Black, Babin, & Anderson, 2010). SEM is similar to multiple regression in the sense that both techniques test linear relationships between variables.

SEM is also able to simultaneously examine multi-level dependence relationships. This means “a dependent variable becomes an independent variable in subsequent relationships within the same analysis” (Shook, Ketchen, Cycyota, & Crockett, 2003) as well as having relationships between multiple dependent variables (Jöreskog, Sörbom, du Toit, & du Toit, 1999).

In recent years SEM has been increasingly applied in social sciences research to investigate complex and intricate relationships that previously could not be easily untangled and examined. Nowadays, SEM seems to be a more attractive approach than the classical approaches such as ANOVA because it more effectively evaluates measurement models and structural paths (relationship between the latent variables in the structural model), particularly when the structural model involves multiple dependent variables, latent constructs based on multi-item indicator variables, and multiple stages/levels of constructs in a structural model (Astrachan, Patel, & Wanzenried, 2014).

Let us look at this in more detail. First of all, we are dealing with latent variables (constructs) and complex models. Latent variables may be measurable directly by an observable indicator variable (available data, e.g. responses to survey questions that are used in a measurement model to determine the latent variables) (Hair, Hult, Ringle, & Sarstedt, 2014). However, the indicator variables may not reflect the latent variable completely accurately which means the measurement will contain errors. By explicitly assessing the error in a structural model, SEM “provides a powerful means of simultaneously assessing the quality of measurement and examining causal relationships among constructs” (Wang & Wang, 2012). So, while multiple regression analysis assumes there is no error in the data, SEM recognizes and accounts for the error in each measured item in an effort to improve the accuracy. Furthermore, SEM considers complex models to find an optimal model that reduces cross-loadings. A cross-loading is an indicator's outer loading on the associated construct that should be greater than the loadings of other constructs, (i.e. indicator variables should significantly load only on one construct). If cross-loadings do exist (i.e. indicator variable loads on multiple constructs) this provides further evidence for not accomplishing discriminant validity (Hair et al., 2014) (Astrachan et al., 2014). SEM also identifies the higher loadings (single regression of each indicator variable) for relevant measures.

Secondly, we can analyse direct, indirect, and total effects. Direct effects include relationships between independent and dependent variables, indirect effects involve relationships between independent and dependent variables that are mediated or moderated by some other variable. Total effects relate to the sum of two or more direct or indirect effects. In contrast to multiple regression analysis which cannot directly deal with the measurement issues of multi-item constructs, SEM is specifically designed to improve multi-item measurement models by directly accounting for error. Thirdly, when assessing structural models. SEM allows simultaneous analysis of all structural relationships, relationships or paths among numerous variables. This is different from regression where one can evaluate structural relationships using path analysis (examining each trajectory separately). SEM is an inherently simpler approach that leads to more accurate results. After reviewing these reasons, it was decided to use SEM in this study.

Applying SEM uses a two-tier process. The first tier involves estimating the measurement model for all latent variables. It measures how well the observed indicators fit the unobserved (latent) variables. In the second tier, the structural model covers the relationships among hypothetical latent variables also called constructs. These variables represent feelings, attitudes, and opinions of a person. Latent variables that only predict other latent variables are called exogenous variables. A latent variable that is a dependent variable in at least one causal relationship is called an endogenous variable. The structural model is estimated by determining the relationships among the exogenous and endogenous latent variables. The relationships between constructs are hypothesized in accordance with theoretical and logical reasoning (Götz, Liehr-Gobbers, & Krafft, 2010).

For example, the outer model relationships are considered to be linear. In this model each indicator variable (X) reflects its latent variable (Y). Each indicator variable is related to its latent variable by a simple regression. In mathematical notation, we have that:

$$X_i = (\lambda_{0i} + \lambda_i \cdot Y_j + error_i) \tag{1}$$

The coefficients $\lambda_i$ are called loadings, $\lambda_{0i}$ is the intercept term, and the error terms account for the residuals. The inner model can be expressed by the following equation:

$$Y_j = (\beta_{0j} + \sum_i \beta_{ji} \cdot Y_i + error_j) \tag{2}$$

Y is a latent variable that is calculated as a linear combination of its block of indicators. *i* of $Y_i$ refers to all the latent variables that are supposed to predict $Y_j$. The $\beta_{ji}$ represent the regression parameters (path coefficients), where $\beta_{0j}$ is the intercept term. The unexplained variance is captured by the error term.

The measurement and the structural models are evaluated using PLS-SEM analysis (SmartPLS 2.0) (Ringle, Wende, & Will, 2005); a set of techniques based on the partial least-squares (PLS) method to measure hypothetical variables. A PLS-SEM fit is based upon accounting for explained variance in the endogenous constructs (Hair et al., 2014). Also, PLS-SEM analyses can easily incorporate single-item measures, and can obtain solutions to highly complex models, i.e., models with a large number of constructs, indicators and structural relationships (Hair et al., 2014); (Hair, Ringle, & Sarstedt, 2013) (Astrachan et al., 2014).

Terzis & Economides (2011) point it out that this approach is particularly suitable for: a) small sample size, b) testing theories in early stages of development (Fornell & Bookstein, 1982) and c) better for prediction (as compared with the covariance-based technique, CB-SEM). In the field of educational technology acceptance there are many studies applying PLS analysis, e.g., (Gong, Xu, & Yu, 2004), (Terzis, Moridis, & Economides, 2012), (Agudo-Peregrina, Hernández-García, & Pascual-Miguel, 2014), (Sánchez-Franco, Peral-Peral, & Villarejo-Ramos, 2014).

Regarding the sample size, “PLS-SEM works efficiently with small sample sizes and complex models and makes practically no assumptions about the underlying data [distributions]” (Hair et al., 2013), (Hair et al., 2014). In PLS-SEM, the guideline is that sample size should be ten times the number of arrows pointing at a construct (Hair et al., 2014).

## 4.2 Measures of the Model Quality

### 4.2.1 Measurement Model

The measurement or outer model specifies the relationship between observable variables and the underlying constructs to evaluate their overall quality. According to literature there are several criteria for validating a measurement model.

*Construct validity*. Means that the observed pattern (how things are) corresponds with our theoretical pattern (how we think things work). It is important that all the constructs' indicators jointly measure the construct adequately. This is assessed by obtaining the construct reliability; this requires indicators assigned to the same construct to reveal a strong mutual association. A composite reliability index (factor reliability) ($\rho_c$) is used to check how well a construct is measured by its assigned indicators (Götz et al., 2010).

According to Fornell & Larcker (1981), we can to measure composite reliability $\rho_c$ as

$$\rho_c = \frac{(\sum_i \lambda_i)^2}{(\sum_i \lambda_i)^2 + \sum_i var(\varepsilon_i)} \qquad (3)$$

$\lambda_i$ indicates the loading of indicator variable *i* of a latent variable, $\varepsilon_i$ indicates the measurement error of indicator variable *i*. The composite reliability can vary between 0 and 1 and values larger than 0.6 are judged as acceptable (Bagozzi & Yi, 1988). Other researchers suggest that a composite reliability value greater than 0.7 is considered adequate (Agarwal & Karahanna, 2000; Compeau, Higgins, & Huff, 1999). The composite reliability is similar to Cronbach's alpha.

*Convergent validity*. This indicates when each indicator correlates strongly with alternative indicators of the same theoretical construct. An accepted measure to analyse convergent validity is the average variance extracted (AVE), defined in Fornell & Larcker (1981). AVE explains the variance of its indicators that is captured by the underlying construct. An AVE of more than 0.5 is considered sufficient (Götz et al., 2010).

*Indicator reliability*: This shows how much of the variation of an indicator can be explained by the theoretical construct. A usual criterion is that more than 50% of an indicator's variance should be explained by the underlying construct. This mean that indicator's loadings larger than .7 are acceptable. This criterion also implies that the shared variance between a theoretical construct and its indicator is larger than the variance of the measurement error (Götz et al., 2010). Empirical research can include weak loadings, particularly when new scales are applied (Hulland, 1999). However, when a measurement model includes proven indicators loadings smaller than .4 (Hulland, 1999) suggests eliminating them from the measurement model within the PLS model.

*Discriminant validity.* Is shown when each measurement indicator correlates weakly with all others constructs except for the one to which it is theoretically associated. In other words, the indicators should load more strongly on their corresponding construct than another construct of the model. It is confirmed when the square root of AVE of a construct is greater than any correlation with another variable, also known as Fornell–Larcker criterion (Fornell & Larcker, 1981).

### 4.2.2 Structural Equation Model

After the constructs have been confirmed as reliable and valid, the next step is to use the structural model to identify patterns in the relationships. The PLS method assesses the structural model's quality by the endogenous variables' determination coefficient ($R^2$). The assessment of the model's quality should also be based on the path coefficients' directions and significance levels (Chin, 1998). $R^2$ reflects the level or share of the latent construct's explained variance. $R^2$ can take values between 0 and 1. The acceptable threshold value depends on the individual study. However, the larger $R^2$ is, the larger the percentage of variance explained (Götz et al., 2010). For social science research the following categories are sometimes applied: weak (.25), moderate (.50), or substantial (.75) (Hair et al., 2010) (Astrachan et al., 2014). Or as Hair, Sarstedt, Ringle & Mena (2012) say the “acceptable level depends on research context”.

The goodness of the path coefficients estimated in PLS can be tested by means of t-statistics, which can be obtained by the bootstrapping procedure used in SmartPLS 2.0 (Ringle, Wende, & Will, 2005), and provides confidence intervals for the parameter estimate. Bootstrapping manages the observed sample as if it represents the population. The procedure creates a large, pre-specified number of bootstrap samples. Each bootstrap sample has the same number of cases as the original sample. Bootstrap samples are created by randomly drawing cases with replacement from the original sample. By provided the mean and standard error of each path, PLS performs a Student' t-test in order to determine the significance of path model relationships (Henseler, Ringle, & Sinkovics, 2009). Paths that are insignificant, or show signs contrary to the hypothesized direction, do not support a prior hypothesis, while significant paths showing the hypothesized direction empirically support the proposed causal relationship. The hypotheses are tested by quantifying the structural equation paths' significance with an appropriate re-sampling method and by examining all the hypothesized relationships' absolute values.

Therefore the structural model and hypotheses are assessed mainly by two criteria: (1) by examining the variance measured for $R^2$ by the antecedent constructs. (2) the significance of the path coefficients and total effects by using the bootstrapping procedure and calculating the t-values (Terzis & Economides, 2011).

## 4.3 Conclusions

We consider PLS-SEM an adequate approach to analyse causal relationships among latent and indicator variables for this research. Since the aim of out work is to analyse new relationships among latent variables, this technique is totally adequate. It is also useful since it has the ability to analyse small sample sizes.

# Chapter 5. Case Study: Lecturers in the University of Manchester

## 5.1 Introduction

In this chapter we will investigate in detail how the factors hindering adoption of ICT in STEM assessments can be quantified in a group of instructors. Although there are plenty of studies that have demonstrated the improvements that technology brings to the teaching and learning processes in STEM, the assessment process has been somewhat neglected. As previously mentioned, it is recognised that the assessment is an important stage in the processes of teaching and learning. Using technology for effective formative assessment can provide significant advantages in improving understanding and performance of the students. On-line summative feedback supports students to achieve their goals by immediately sending them the result of the involved effort during the course. Such advantages have been studied, by, e.g. Terzis et al. (2012) who emphasize that web-based assessment technologies, either formative or summative, offer many advantages, such as: "(A) high interaction and adaptation with test-takers, (B) real-time feedback, (C) real-time score reports, (D) more efficient managing, setting, and delivering of exams, (E) easier data management, (F) cost reduction, (G) self-evaluation and recognition of students' strengths and weaknesses".

Formative electronic assessment has become more and more important in higher education and given that undergraduate classes can consist of several hundred students it is not always possible for the faculty to meet with individual students and guide their learning. This issue combined with the diversity of students' academic backgrounds supports the need for a formative electronic assessment (Miller, 2009).

In teaching mathematical subjects, technology has facilitated formative and summative practices providing the students with timely, informative feedback that helps them to build/reinforce mathematical skills and abilities. In short, technology can sustainably enrich the students learning in mathematics. However, in order to reach these benefits, students and teachers should have a common set of goals and targets for its use. As mentioned in the previous chapter, there are some concerns in this area. This is very important since the success of any initiative to implement technology in an educational programme depends strongly upon the support and attitude of the teachers involved. It has been suggested that if teachers believe or perceive that computers do not satisfy their own or their students' needs, they are likely to resist any attempts to introduce technology into their teaching and learning (Askar & Umay 2001). Lecturers' attitudes towards computers, whether positive or negative, (Teo, Lee, & Chai, 2008) affect how they respond to technologies. This in turn affects the way students view the importance of computers in schools affecting current and future computer usage. Therefore, examining the factors determining lecturers' attitude and intentions should answer some questions relating to acceptance and usage of on-line learning environments in teaching and learning mathematics. In other words, we would like to enhance students' appreciation through stimulating the instructor´s use of web-based assessment. Opinions of those lecturers already using the on-line platform and those that are not yet using it but might be interested in using it are taken into account.

In order to understand these factors, we base our models on previous research which has considered the use of several “constructs” for the study of a person's acceptance, usefulness, perception, feelings and attitude towards ICT. We have included the concept as a group of variables (indicators) measuring a person's attitudinal features. Our model for lecturers is mainly based on the Technology Acceptance Model (TAM) which has proved to be a robust model to determine the attitude and intention to use technology in educational settings to take account of others factors, some additional categories have also been included.

According to the TAM, the acceptance and usage of technology is mainly determined by a person's attitude and intention to use it. That is the reason why researchers have increasingly addressed the role of motivational factors such as self-efficacy, attitude, faculty’s roles, training, and tenure, as all contribute to the success or failure of distance education efforts and technology integration (Howell, Saba, Lindsay, & Williams, 2004).

## 5.2 Technological Acceptance Model

Any organization has to deal with the problem of how to make good investments. This issue becomes especially important when an organization needs to invest in technology due to the large failure rate of ICT investments. There is a lot involved in this, because investments in emerging information technologies may be of the order of hundreds of thousands or even millions of pounds sterling. That is one of the reasons why research and practise in the Information Systems (IS) field had a peak in previous decades.

In which form of ICT should an institution invest? An organization should invest its money in technology in order to efficiently deal with important issues such as pressure to cut costs, pressure to produce more without increasing costs, or to improve services or products quality in order to stay in business (Legris, Ingham, & Collerette, 2003). A successful investment in technology can also lead to enhanced productivity, whereas a failed decision can lead to undesirable consequences such as financial losses and dissatisfaction among employees (Venkatesh, 2000). A good investment in the appropriate information technology is a key factor in a competitive world. Even when organizations have already reached an advantageous position, they are often worried about how to invest in ICT without compromising themselves.

The second important reason why IS research has grown is due the underutilization of ICT. Venkatesh (2000) gives an example of how the American “Internal Revenue Service (IRS)” invested about 4 billion of dollars in a system aimed at simplifying the processing of tax returns for 1996 by computerizing the process. The following year the IRS was forced to revert to a manual method of processing returns. The main reason was that the users found the system to be too difficult to use. Low usage of installed systems has in general been identified as a major factor in lacklustre returns from organizational investments in information technology. Emerging technology cannot improve the effectiveness of people if it is not accepted and used. Therefore, understanding the conditions under which ICT will be embraced by organizations and their staff remains a high-priority research issue.

For these reasons organizations do like to find a reliable model in so they can invest wisely in IS and assuring their investments. From the mid-eighties several models were developed that can help in predicting system use (Legris et al., 2003). One of those is the Technology Acceptance Model (TAM), proposed by Davis (1989) as a model to predict IS adoption. They developed TAM based on the Theory of Reasoned Action (TRA). TRA was first proposed by Fishbein & Ajzen (1975) as a model to explain and predict the behaviour of people in a specific situation (See Figure 5.1).

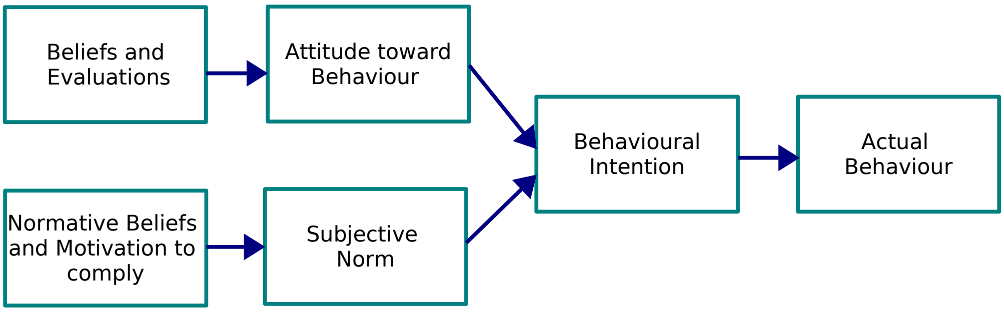


*Figure 5.1. A graphical representation of the "reasoned action" model taken from (Davis et al., 1989)*

TRA states that a person's actual behaviour can be explained by his/her intention and beliefs, and that, intentions can be explained by both his/her attitude and subjective norms. They define attitude as "the degree of a person's favourable or unfavourable evaluation or appraisal of the behaviour in question", suggesting that the attitude of an individual toward behaviour (A) can be measured by considering the sum of the product of the beliefs ($b_i$) about consequences of performing that behaviour, and an evaluation ($e_i$) of those consequences;

$$A = \sum_i b_i e_i \qquad (4)$$

Fishbein & Ajzen (1975) define the subjective norm (SN) as the "person's perception that most people who are important to him or her think he or she should or not should perform the behaviour in question". This idea suggests that SN is the sum of the product of an individual's normative beliefs ($nb_i$), or value judgements that is perceived expectations of other individuals or groups, and his or her motivation to comply ($mc_i$) (Chuttur, 2009),

$$SN = \sum_i nb_i mc_i \qquad (5)$$

Then, the behavioural intention of an individual to perform a behaviour could be given by the equation 6 where A is a measure of attitude towards the behaviour, and SN is a measure of the subjective norm associated with the behaviour in question.

They assume that an individual's behavioural intention can be determined by adding a measure of that person's attitude toward the behaviour and a measure of the subjective norm associated with the behaviour,

$$BI = A + SN \tag{6}$$

Davis (1989) realised that the actual use of a system can be considered as a behaviour. Also, instead of considering an individual's beliefs, he identified, using previous research, the particular beliefs that are operative in the context of computer user behaviour. His proposed model provides diagnostic insight into how system characteristics influence user attitudes and usage.

Davis developed TAM in the mid-1980s under contract with IBM Canada, Ltd. where it was used to evaluate the market potential for a variety of emerging applications in the area of multimedia, image processing, and pen-based computing in order to guide investments in new product development (Davis & Venkatesh, 1996).

The TAM assumes (see Figure 5.2) that “perceived usefulness” and “perceived ease of use” are primary drivers for technology acceptance. Perceived usefulness is defined as “the degree to which an individual believes that using a particular system would enhance his/her job performance”. Perceived ease of use is defined as “the degree to which an individual believes that using a particular system would be free of physical and mental effort” (Davis, 1989). The TAM also assumes that perceived usefulness will be influenced by perceived ease of use because, as Davis concludes, if “two systems perform the identical set of functions, a user should find the one that is easier to use more useful. A designer should therefore be able to enhance perceived usefulness either by adding new functional capabilities to a system, or by making it easier to invoke the functions which already exists” (Davis, 1993). TAM states that “perceived usefulness” has a direct effect on an individual's behavioural intention toward using a system, and “perceived ease of use” affects indirectly through “perceived usefulness” (Davis, 1989). That means “perceived usefulness” mediates the effect of “perceived ease of use” on behavioural intention. Several empirical studies have supported this idea (Venkatesh & Davis, 2000), (Wu & Chen, 2005). Also, studies such as Lee (2009) and Liu, Liao & Pratt (2009) provide evidence of the effect of "perceived usefulness" on the behavioural intention towards technology. Therefore, both PU and PEU are considered important predictors of an intention to use the system, and/or actual system usage.

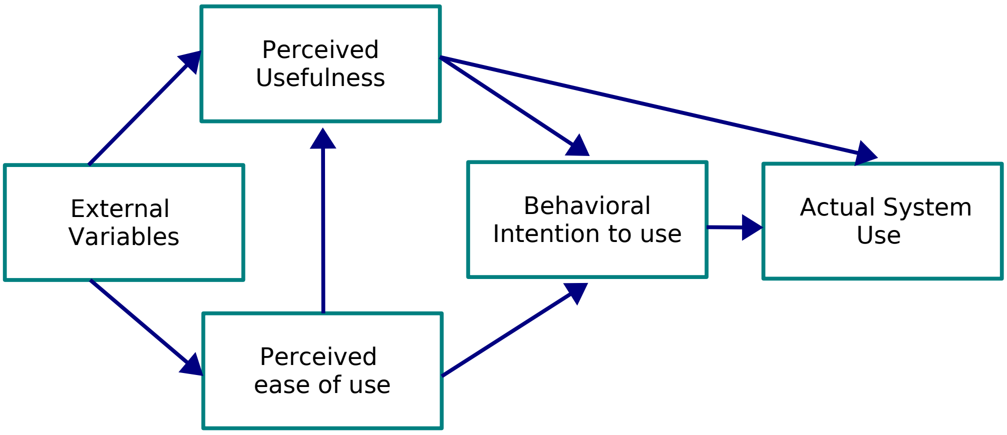


*Figure 5.2. The Technological Acceptance Model (Davis, 1989) addresses why users accept or reject ICT and how user acceptance is influenced by external variables*

The objective of the original TAM was to provide a basis for tracing the effects of external variables on internal beliefs, attitudes and intentions. The "perceived usefulness" and "perceived ease of use" are relevant when trying to anticipate who will adopt and begin using certain technology.

Significant progress has been made over the last decade in explaining and predicting user acceptance of information technology at work. In particular, substantial theoretical and empirical support has accumulated in favour of the TAM. The TAM has demonstrated its robustness in diverse technologies (e.g. word processing, e-mail, web-based applications) in diverse situations (e.g. time and culture) with diverse control factors (e.g. gender, organizational type and size) and diverse subjects (e.g. undergraduate students, MBAs, and knowledge workers). This may explain why TAM, which was proposed in 1989, by 2003 had 698 journal citations (Lee, Kozar, & Larsen, 2003).

In the original form of the model, Davis argued that TAM was designed to understand the causal chain linking external variables to its user acceptance and actual use in a workplace. “External variables such as objective system design characteristics, training, computer self-efficacy, user involvement in design, and the nature of the implementation process are theorized to influence behavioural intention to use, and ultimately usage, indirectly via their influence on perceived usefulness and perceived ease of use.” (Davis & Venkatesh, 1996).

Thus, it is thought that perceived usefulness can be strongly affected by attitudinal beliefs (external variables). Previous studies found that there are important human factors (attitudinal factors) that successfully affect the usage of technology, for instance, in an educational context Benson Soong, Chuan Chan, Chai Chua & Fong Loh (2001) find that instructors' and students' technical competence and their mindset about on-line learning, level of collaboration in the course, and level of perceived IT infrastructure are crucial factors for successful applications of on-line course resources. Likewise, in the work by Bhuasiri et al. (2012), which also includes results first reported by Volery & Lord (2000), it is concluded that technological factors (ease of access, support interaction, design, etc.), instructors' characteristics (attitude toward students, teaching style, technical competence, encouragement of students interaction, etc.) and students' characteristics affect effectiveness of on-line delivery. In addition, other factors linked to instructors are mentioned in the study by Webster & Hackley (1997) such as teaching learning style, attitude, and control of the technology. All of these affect a person's attitude towards using technology. These attitudes determine behavioural intentions, which in turn lead to actual system use.

A wide range of other studies in different fields have extended TAM by adding several variables. For instance, Warkentin, Gefen, Pavlou & Rose (2002) examine citizens' adoption of e-government in different countries by integrating TAM with trust, perceived risk, perceived behaviour control, and culture. Ilias, Razak & Yasoa (2009) extend TAM with perceived credibility, information system quality, as well as information quality to investigate taxpayers' attitude in using e-filing system considering demographic factors.

In an educational context TAM has been widely applied to predict the adoption and usage of educational technologies. Researchers have also extended TAM by including additional constructs. For example, (Teo, 2009) examines the level of technology acceptance pre-service teachers by including technological complexity and facilitating conditions. A model proposed by (Cheung & Vogel, 2013) incorporates perceived resources, compatibility, sharing knowledge and subjective norms (represented by peers, instructors and mass media) as additional constructs to explain the factors that influence students' intentions to use Google Applications for collaborative learning. In the field of the assessment (Terzis & Economides, 2011) analyse the effect of students' behavioural intention to use a computer-based assessment (CBA) by adding content and goal expectancy as two new constructs.

Researchers who have applied TAM in e-learning studies have indeed found that “perceived usefulness” and “perceived ease of use” have significant effects on an individual's behavioural intention to use e-learning systems (Liu et al., 2009); (Ong, Lai, & Wang, 2004). Therefore, TAM has demonstrated that it is a useful model to predict the acceptance and usage of technology in the educational field. This is the basis of our research. Thus, we include “perceived usefulness” (PU) as a construct that impacts the lecturers' attitude. Likewise “perceived ease of use” (PEU) directly affects perceived usefulness and indirectly lecturers' attitudes and usage intentions.

We also make use of the theory of planned behaviour (TPB) (Ajzen, 1991). TPB is also based on the theory reasoned action (TRA) which holds that “attitude”, “subjective norms”, and “perceived behavioural control” are direct determinants of intentions that influence behaviour. TPB includes subjective norms (SN), such as “the perceived social pressure to perform or not to perform the behaviour” (Ajzen, 1991). The effects of social influence on behavioural intention are direct (Teo et al., 2008). TPB states that attitude impacts a user's behavioural intention. (Piccoli, Ahmad, & Ives, 2001) find that instructors' attitudes toward e-learning positively influence results of e-learning since instructors are major actors in learning activities.

Likewise, we also include in our model the construct of “self-efficacy”, based on social cognitive theory (SCT) (Bandura, 1986). This is a widely accepted and empirically validated model for understanding and predicting human behaviour and identifying methods in which behaviour can be changed. Many researchers have applied it as a theoretical framework to predict and explain an individual's behaviour in a variety of contexts involving cognitive, social, motor, health, instructional, and self-regulatory skills. SCT proposes two cognitive factors: self-efficacy and outcomes expectations that influence individual behaviour. In this sense, self-efficacy is defined "as one's judgements and beliefs of his/her confidence and capability to perform a specific behaviour" (Bandura, 1986). Self-efficacy has been shown to predict behavioural change with different types of participants in various settings. Therefore, in order to understand the impact of self-efficacy has on lecturers' behavioural attitudes, it was decided to include it as a factor in the model. It is included as "computer self-efficacy" (CSE).

Usefulness plays a crucial role in determining the attitude of lecturers towards on-line testing tools. Nevertheless, there is another important concept, incentives, or motives for integrating computers in their teaching. The research of Crumpacker (2001) explains that the instructor's effort level can be affected by incentives; as he explains, they generally increase motivation, and obstacles usually decrease motivation.

Likewise, the study of Baylor & Ritchie (2002), identifies the concept of *teachers morale* as relevant when integrating computers in their teaching. They define this concept based on the study of Hadley & Sheingold (1993) who identify two trends: *student accomplishment, rather than their own external rewards, is most motivating for the teachers, followed by students' being able to use computers as a tool for their own purposes*. As they state, ''in the daily professional life of these teachers, it is the psychic pay-off of student's learning and engagement that appears to matter most'' (p. 281). Thus, the natural desire for making the students' learning experience more valuable as well as the desire for enriching their teaching activities are both important.

Crumpacker (2001) argues that it is equally important for instructors to possess the necessary skills. He explains that a context where incentives are attainable and obstacles are negligible elicits a higher effort that could result in students receiving a higher quality of education. To this end, identifying instructor-perceived incentives and obstacles is important.

Therefore, it is necessary to determine what really motivates lecturers to use technology in the context of assessment. One of the factors could be the natural desire of teachers to enhance student learning. In terms of assessment activities, the appropriate use of well-designed feedback can enhance students' learning by helping them to learn more efficiently; this means improving knowledge, skills, and attitudes and/or correcting misunderstandings or wrong-headed approaches to solving problems. Therefore, to what extent the lecturers' belief that on-line tools such as on-line feedback can serve for enhancing students learning, which in turn, can serve as an incentive. Thus, a “feedback” factor (FE) has been included as “the lecturers' belief that they can enhance their teaching by using on-line feedback”, consequently, enriching students learning. It has been included as a new construct. The hypothesis was that feedback would have a direct positive effect on lecturers' attitude and behavioural intentions to use web-based assessment for teaching mathematics.

Perceived system satisfaction (PSS) was also included as a construct in the model. This factor was taken from the study by Liaw, Huang & Chen (2007) who show that perceived system satisfaction can be a crucial factor that influences instructors' perceived usefulness toward e-learning. The perceived satisfaction is defined as user acceptance of information systems and the degree of comfort involved in using them. Thus, a greater degree of satisfaction toward an information system implies a higher degree of willingness to use it (Liaw & Huang, 2013).

## 5.3 Research Methods

### 5.3.1 Designing Research Models

In order to discover lecturers' usage intentions we propose the following empirical research model. The model proposed is a model that groups indicators (items) into “constructs” also called “latent variables” that are research abstractions that cannot be measured directly such as beliefs and perceptions. Based on the theoretical frameworks described above we propose the constructs and indicators listed in Table

5.1. They are grouped according to the construct evaluated.

*Table 5.1. Distribution of constructs and indicators proposed to lecturer's model*

| **Indicators** | | **Construct** |
|---|---|---|
| AT1, AT2, AT3, AT4 | belongs to | AT |
| PU1, PU2, PU3, PU4 | belongs to | PU |
| FE1, FE2, FE2, FE4 | belongs to | FE |
| PEU1, PUE2, PUE3, PU4 | belongs to | PEU |
| CSE1, CSE2, CSE3, CSE4 | belongs to | CSE |
| BI1, BI2, BI3, BI4 | belongs to | BI |
| SN1, SN2, SN3, SN4 | belongs to | SN |
| PSS1, PSS2, PSS3, PSS4 | belongs to | PSS |

*Proposing hypotheses.*

Based on the theoretical frameworks mentioned before, the following hypotheses were proposed:

- *H1: states there is a causal relationship between perceived usefulness and attitude to use web-based assessment.*
- *H2: states there is a causal relationship between feedback and attitude to use on-line assessment.*
- *H3: states there is a causal relationship between feedback and behaviour intention to use web-based assessment.*
- *H4: states there is a causal relationship between perceived ease of use and perceived usefulness to use on-line assessment.*
- *H5: states there is a causal relationship between computer self-efficacy and perceived ease of use web-based assessment.*
- *H6: states there is a causal relationship between computer self-efficacy and perceived system satisfaction to use web-based assessment.*
- *H7: states there is a causal relationship between perceived system satisfaction and perceived usefulness to use on-line assessment.*
- *H8: states there is a causal relationship between attitude and behavioural*

*intention to use web-based assessment.*

- *H9: states there is a causal relationship between subjective norm and behavioural intention to use on-line assessment.*

Figure 5.3 represents schematically the causal relationships proposed by the hypothesis within the model.

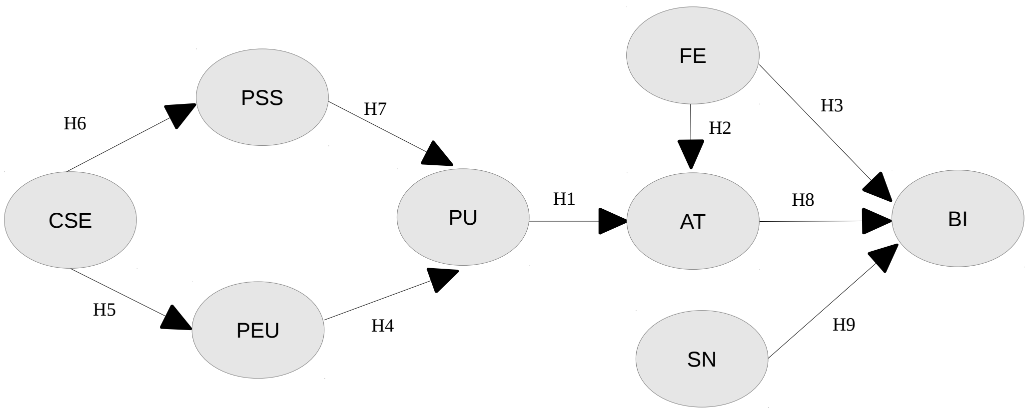


*Figure 5.3. A schematic representation of the path model integrated by eight theoretical constructs and their hypotheses to discover lecturers' usage intentions*

## 5.4 Instrumentation

### 5.4.1 Developing the Scale

Some of items used in this research are based on previous studies. Others are an empirical proposal and designed to take into account the belief/perception we wish to measure. The original wording of some scale items was adapted to fit the context of educational electronic assessment. All items were assessed by a 5-point-Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The constructs and their indicators are listed in Table 5.2.

*Table 5.2. Scale used and representative study*

| **Construct/indicator** | **Representative study** |
|---|---|
| **Attitude (AT)** | adapted from (Fishbein, & Ajzen, 1975) |
| AT1. I think that using electronic assessment is useful | |
| AT2. I am convinced that university policies on using on-line tools in teaching are a good way forward | |
| AT3. As a teacher, I am here for teaching not to write on-line assessments | |
| AT4. I do not believe in using computer-based assessments | |
| **Perceived Usefulness (PU)** | |
| PU1. I think that on-line assessment is helpful as part of my teaching | |
| PU2. I find it useful to use blended learning in my teaching (i.e., face-to-face combined with on-line) | |
| PU3. I think that on-line tests/exams are useful to assess students' ability in applying mathematical techniques | |
| PU4. I can save time by having my students take on-line tests | |
| **Feedback (FE)** | self-developed |
| FE1. I believe that on-line feedback encourages students to do better in their studies | |
| FE2. I think that electronic assessments are useful for students, because they get their grades faster | |
| FE3. I think that students enjoy when I am using cutting edge technology in on-line tests | |
| FE4. I find blended learning environments useful because students can get on-line support from others students more easily | |
| **Perceived Ease of Use (PEU)** | adapted from (Davis, 1989) |
| PEU1. I find electronic assessment easy to use | |
| PEU2. It is straightforward to learn how to use a virtual learning environment (e.g. blackboard) for teaching | |
| PEU3. I find it straightforward to use computers in my teaching | |
| PEU4. I find it straightforward to use computers for assessing the mathematical skills of my students | |
| **Computer Self-Efficacy (CSE)** | adapted from (Bandura, 1986) |
| CSE1. I feel confident using electronic assessment in my teaching | adapted from (Compeau & Higgins, 1995) |
| CSE2. I am comfortable developing on-line assessments | |

| | |
|---|---|
| CSE3. I feel confident using blended technologies for teaching | |
| CSE4. I have successfully used electronic assessments in the past | |
| **Behaviour Intention (BI)** | adapted from (Davis, 1989) |
| BI1. I intend to use a mixture of face-to-face and on-line learning to assist my teaching | |
| BI2. I plan to use electronic tools for assessing problems with mathematical content | |
| BI3. I am committed to use on-line assignments to assess the mathematical skills of my students | |
| BI4. I intend to understand better the university policies on using on-line tools in my teaching | |
| **Subjective Norms (SN)** | adapted from (Fishbein, & Ajzen, 1975) |
| SI1. I am following the university policies on using on-line tools in my teaching | |
| SI2. People around me are positive about using electronic assessment | |
| SI3. If I need support using (blended) e-Learning environments, a colleague is available to help me | |
| SI4. University staff have supported me in the use of computers in my teaching | |
| **Perceived System Satisfaction (PSS)** | self-developed |
| PSS1. I am content using a mix of on-line and classroom instruction for teaching | |
| PSS2. It is good use of my time trying to learn building electronic assessments | |
| PSS3. It is harder to develop an electronic assessments than a paper-based one | |
| PSS4. On-line tests can be solved by using flaws in the software | |

## 5.5 Data Collection Procedures

Lecturers in this study used the on-line platform during the February to June 2012 semester. After that, an on-line questionnaire was used to gather lecturers' responses to our survey. The total of 35 responses to the questionnaire were obtained, 2 were removed as incomplete resulting in a dataset of 33. The gender balance was 26.47% female, 67.65% male and no answer 5.88%. As can be seen the result shows most of them are male, which is not surprising because our study was applied to STEM subjects where males lecturing staff is predominant. Regarding the distribution of lecturers' responses by school, results show that most of the surveyed lecturers belong to school of physics and astronomy (33%), followed by the school of chemistry (25%), the school of earth, atmospheric and environmental sciences (19%), the school of chemical engineering and analytical science (14%), and the school of electrical & electronic engineering (8%).

The information collected indicates that the majority of the surveyed lecturers have been lecturing for many years, indicating that these teachers are experienced. Most of them have over 15 years teaching experience (64%) or have from 6 to 10 years teaching experience (15%). These were (9%) for 1-5 years, the (6%) for 11-15 years, and the (6%) for less than 1 year. These results are interesting for our study, since most of lecturers surveyed are experienced as a lecturer. This could affect our results since experienced lecturers are used to teaching in the same way for many years.

## 5.6 Data Analysis Techniques

Table 5.3 below shows the descriptive statistics for each dimension including the values for each item evaluated. Each variable is grouped according to the subject evaluated. These results can help us to gain preliminary insights from the data. The values of the each item are obtained by assigning a numerical value to each entry of the Likert scale from (5) to (1), where a higher number means a more positive response.

*Table 5.3. Descriptive statistics (mean and standard deviation) of the items in the lecturer's scale*

| **Attitude (AT)** | **Mean ()** | **Standard deviation ()** |
|---|---|---|
| AT1. I think that using electronic assessment is useful | 3.36 | 1.14 |
| AT2. I am convinced that university policies on using on-line tools in teaching are a good way forward | 2.67 | 0.89 |
| AT3. As a teacher, I am here for teaching not to write on-line assessments (R) | 2.91 | 1.07 |
| AT4. I do not believe in using computer-based assessments (R) | 3.64 | 1.08 |
| **Perceived Usefulness (PU)** | | |
| PU1. I think that on-line assessment is helpful as part of my teaching | 3.21 | 1.14 |
| PU2. I find it useful to use blended learning in my teaching (i.e., face-to-face combined with on-line) | 3.39 | 1.14 |
| PU3. I think that on-line tests/exams are useful to assess students' ability in applying mathematical techniques | 3.09 | 1.23 |
| PU4. I can save time by having my students take on-line tests | 2.79 | 1.19 |
| **Feedback (FE)** | | |
| FE1. I believe that on-line feedback encourages students to do better in their studies | 3.30 | 0.88 |
| FE2. I think that electronic assessments are useful for students, because they get their grades faster | 3.30 | 0.92 |
| FE3. I think that students enjoy when I am using cutting edge technology in on-line tests | 2.88 | 1.08 |
| FE4. I find blended learning environments useful because students can get on-line support from others students more easily | 2.91 | 0.95 |
| **Perceived Ease of Use (PEU)** | | |
| PEU1. I find electronic assessment easy to use | 2.55 | 0.97 |
| PEU2. It is straightforward to learn how to use a virtual learning environment (e.g. blackboard) for teaching | 3.24 | 1.30 |
| PEU3. I find it straightforward to use computers in my teaching | 3.48 | 1.20 |
| PEU4. I find it straightforward to use computers for assessing the mathematical skills of my students | 2.55 | 1.15 |
| **Computer Self-Efficacy (CSE)** | | |
| CSE1. I feel confident using electronic assessment in my teaching | 2.91 | 1.21 |
| CSE2. I am comfortable developing on-line assessments | 2.61 | 1.20 |
| CSE3. I feel confident using blended technologies for teaching | 3.24 | 1.20 |
| CSE4. I have successfully used electronic assessments in the past | 2.88 | 1.39 |
| **Behaviour Intention (BI)** | | |
| BI1. I intend to use a mixture of face-to-face and on-line learning to assist my teaching | 3.36 | 1.14 |
| BI2. I plan to use electronic tools for assessing problems with mathematical | 2.58 | 1.12 |

| | | |
|---|---|---|
| content | | |
| BI3. I am committed to use on-line assignments to assess the mathematical skills of my students | 2.36 | 1.03 |
| BI4. I intend to understand better the university policies on using on-line tools in my teaching | 2.67 | 1.05 |
| **Subjective Norms (SN)** | | |
| SN1. I am following the university policies on using on-line tools in my teaching | 3.24 | 0.61 |
| SN2. People around me are positive about using electronic assessment | 2.39 | 0.79 |
| SN3. If I need support using (blended) e-Learning environments, a colleague is available to help me | 3.06 | 1.22 |
| SN4. University staff have supported me in the use of computers in my teaching | 3.24 | 1.25 |
| **Perceived System Satisfaction (PSS)** | | |
| PSS1. I am content using a mix of on-line and classroom instruction for teaching | 3.36 | 1.17 |
| PSS2. It is good use of my time trying to learn building electronic assessments | 2.55 | 1.15 |
| PSS3. It is harder to develop an electronic assessments than a paper-based one (R) | 1.91 | 0.80 |
| PSS4. On-line tests can be solved by using flaws in the software (R) | 2.91 | 0.98 |

(R) Reverse code

For individual scores, a mean below 2.00 indicates a very negative attitude; between 2.00 and 2.75 indicates a slightly negative attitude; a mean above 4.00 indicates a very positive attitude; from 2.75 through 3.25 reflects an attitude of ambivalence; and a mean from 3.25 through 4.00 indicate a slightly positive attitude.

Considering all mean values above 3.25 as positive, we can conclude that lecturers are content to use blended learning in their teaching, consider it as a useful method of teaching, and are willing to use it in the future. They consider themselves capable to incorporate computers easily in their teaching activities.

They also perceive that using on-line assessments are useful. They think that it is useful for students, since it allows them to obtain grades faster. They also consider that on-line feedback can be a helpful way of encouraging students to improve performance in their studies.

In order to better understand the lecturers' opinion about attitude and usage intention, these constructs were closely analysed. Figure 5.4 shows the responses of each item of the attitude construct.

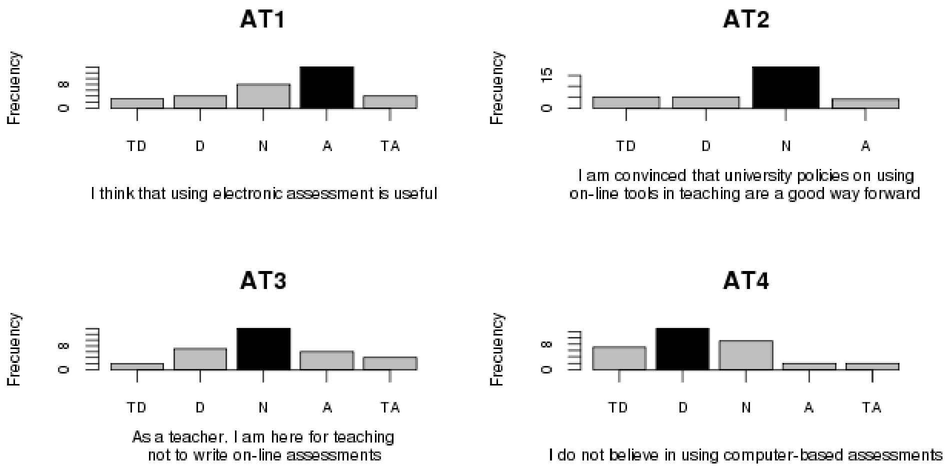


*Figure 5.4. Summary of lecturers' responses for the attitude construct*

The results of item AT1 indicate that 55% of teachers show a favourable attitude towards the use of on-line exams. This represents just half of the respondents; of those that are in favour, 12% strongly agreed and 42% agreed. Lecturers who maintain a neutral response (neither agreed nor disagreed) represent 24% and lecturers who think that electronic evaluations are not useful, either with opinions from disagree to strongly disagree accounts for 21%. These results show lecturers are mostly in favour of on-line assessments.

The response for the AT2 item shows that 58% of lecturers have a sceptical attitude towards university policies. This means that they are neither for nor against the way that the university sets its priorities to use technological tools for electronic assessment. 30% of them disagree or strongly disagree; they are not in favour of these policies and only 12% of them are in favour. Notably, none of them totally agree.

The results of item AT3 shows that 43% are not convinced that writing on-line assessments as part of their academic activities does pose extra work. There is a slight margin of difference between those who are in favour (30%) against those who are not in favour (27%).

Interesting results emerge from item AT4 which in general shows that lecturers have a positive attitude. 61% of them believe in the use of on-line exams either with answers that agree or strongly agree versus 12% that show an attitude not in favour; 27% of them responded neutrally.

It seems that lecturers in the main have a positive attitude towards using on-line testing. However, these results are interesting compared with the results of the variable BI which is deployed in Figure 5.5.

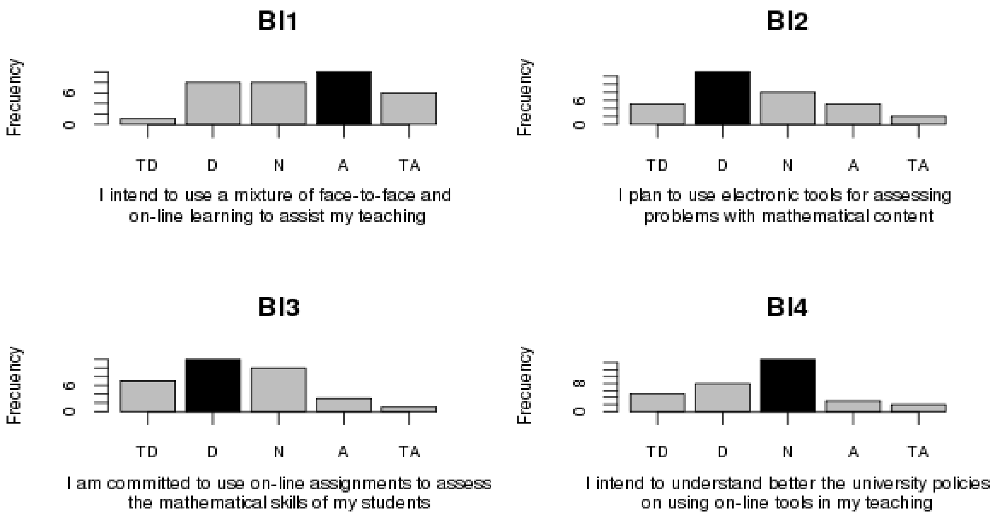


*Figure 5.5. Summary of lecturers' responses for the behaviour intention construct*

The results of the item BI1 show that lecturers are willing to try to combine the traditional teaching and electronic tools (30%). It is noteworthy that only 18% of them are fully committed. Both responses represent 48%, indicating a willingness of lecturers to this question. 24% of respondents neither agree nor disagree, 24% disagreed and 3% definitely disagreed.

In contrast, the results of item BI2 show that 55% of teachers disagree (39% disagreed and 15% strongly disagreed) indicating that more than half of them are not convinced about using on-line tests as a tool to assess mathematical ability, which is corroborated by the percentage of teachers who are in favour (15%) and those who are fully convinced (6%). The results of this item show a contrast to the previous item.

These results were favourable. However, when asked about applying these in a mathematical context their attitude changed to not favourable.

The responses of BI3 show that only 12% agreed with the idea of evaluating students' math skills by assigning on-line tasks, either with answers that agree or strongly agree. It is noteworthy that 36% disagreement is a slightly higher amount than the 21% of those who strongly disagree. In combining both results shows that 58% of teachers do not show a favourable attitude to this item.

As regards the results for item BI2, it is our belief these item also reflect disapproval by the lecturers as it asks about their intention to use electronic tools to evaluate mathematical problems. It seems that including aspects related to evaluating mathematical content or development of mathematical skills significantly affects the response of lecturers as shown by these items; this shows some distrust by the lecturer to use on-line math tests/exams.

The answers for BI4 shows that 45% of lecturers have an attitude of indifference. 40% of them are definitely not in favour, which could mean that lecturers believe the strategies that educational managers set regarding the use of technological tools for teaching are not important for them and only 15 % are in favour.

## 5.7 Findings and Results

### 5.7.1 Evaluation of the Measurement Model (Model Quality)

Regarding the sample size, “PLS-SEM works efficiently with small sample sizes and complex models and makes practically no assumptions about the underlying data [distributions]” (Hair et al., 2013), (Hair et al., 2014). In PLS-SEM, the guideline is that sample size should be ten times the number of arrows pointing at a construct (Hair et al., 2014). The proposed model has three independent variables impacting a dependent variable (behavioural intention); therefore this requirement is achieved.

*Table 5.4. Construct, convergent and discriminant validity coefficients (bivariate correlations. In the main diagonal, the squared-root AVE of each construct) for the measurement model*

| | AVE | Composite Reliability | Cronbach's Alpha | AT | BI | CSE | FE | PEU | PSS | PU | SN |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **AT** | 0.69 | 0.9 | 0.85 | **0.83** | | | | | | | |
| **BI** | 0.61 | 0.86 | 0.77 | 0.64 | **0.78** | | | | | | |
| **CSE** | 0.76 | 0.93 | 0.9 | 0.67 | 0.49 | **0.87** | | | | | |
| **FE** | 0.6 | 0.86 | 0.78 | 0.85 | 0.76 | 0.48 | **0.77** | | | | |
| **PEU** | 0.53 | 0.81 | 0.72 | 0.59 | 0.72 | 0.56 | 0.53 | **0.73** | | | |
| **PSS** | 0.44 | 0.73 | 0.52 | 0.69 | 0.63 | 0.69 | 0.6 | 0.66 | **0.66** | | |
| **PU** | 0.68 | 0.89 | 0.84 | 0.84 | 0.69 | 0.75 | 0.77 | 0.61 | 0.61 | **0.82** | |
| **SN** | 0.53 | 0.81 | 0.7 | 0.67 | 0.42 | 0.44 | 0.55 | 0.36 | 0.57 | 0.57 | **0.73** |

Table 5.4 shows the coefficients for the composite reliability of each construct which is over 0.7. In all cases it also shows the coefficients of the average variance extracted (AVE) that all cases exceed the adequate value except to PSS construct that also displayed a low value of the Cronbach's alpha reliability.

Table 5.4 also displays the criterion to assess discriminant validity. It shows the correlations of the variables and the square roots of the AVEs which are the diagonal elements of the table. All the square roots of the AVEs are higher than any other correlation except the relationship of AT with FE, PU, and PSS with AT, CSE, PEU.

The AVE ranged from .53 to .76, confirming convergent validity and implicitly, content validity in all constructs except PSS. However, composite reliabilities ranged from .73 to .93 demonstrating reliability for all constructs.

Results of indicator reliability items BI1(0.51), PEU2(0.53), PEU3(0.61), PSS3R(0.58), PSS4R(0.22), FE2(0.68), SN1(0.46) show low factor loadings.

After the evaluation of the measurement model, it is seen that items (BI1, PEU2, PEU3, PSS3R, PSS4R, PUS2, SI1) do not accomplish the recommended threshold indicator reliability (of .4), we also see that the correlation values of the constructs FE, PSS, PU are higher than the AVE values, thus not reaching discriminant reliability. In order to accomplish discriminant reliability, the study of Nunnally (1978) advises that indicators with loadings lower than a threshold of 0.7 should be removed from the subsequent analysis. This matches with the above-mentioned criteria (Hulland, 1999), where it is suggested that indicator loadings smaller than 0.4 should be eliminated from the measurement model.

*Table 5.5. Adjusted model. Measures of quality (construct, convergent and discriminant validities). Discriminant validity (bivariate correlations. In the main diagonal, the squared-root AVE of each construct)*

| | AVE | Composite Reliability | Cronbach's Alpha | AT | BI | CSE | FE | PEU | PSS | PU | SN |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **AT** | 0.82 | 0.93 | 0.89 | **0.91** | | | | | | | |
| **BI** | 0.76 | 0.9 | 0.83 | 0.44 | **0.87** | | | | | | |
| **CSE** | 0.76 | 0.93 | 0.9 | 0.66 | 0.37 | **0.87** | | | | | |
| **FE** | 0.67 | 0.86 | 0.76 | 0.77 | 0.77 | 0.46 | **0.82** | | | | |
| **PEU** | 0.81 | 0.89 | 0.77 | 0.44 | 0.75 | 0.53 | 0.58 | **0.9** | | | |
| **PSS** | 0.76 | 0.86 | 0.69 | 0.59 | 0.49 | 0.66 | 0.59 | 0.52 | **0.87** | | |
| **PU** | 0.68 | 0.89 | 0.84 | 0.77 | 0.62 | 0.75 | 0.75 | 0.66 | 0.6 | **0.82** | |
| **SN** | 0.64 | 0.84 | 0.72 | 0.64 | 0.39 | 0.43 | 0.57 | 0.36 | 0.56 | 0.57 | **0.80** |

Therefore, a systematic process of examining the loadings and removing indicators with loadings below .70 was followed (Hair et al., 2010). We delete those indicators with the lowest indicator reliability to analyse the impact of indicator deletion on AVE and composite reliability. Deleting the indicators PSS3R (0.34), PSS4R (0.05) of the PSS construct shows a lower value, AVE=0.44, than the threshold suggested. Also, we remove the indicators BI1(0.26), PEU2(0.28), PEU3(0.37), FE2(0.46), SN1(.021) that showed the lowest indicator loadings. It is common to deal with this issue since some of these indicators are construct empirical measures or these are standard scales initially adopted in causal modelling (Barclay, Higgins, & Thompson, 1995).

As an alternative method to assess discriminant validity the cross-loadings were examined. All indicator loadings were higher in their construct than the respective cross-loadings, except indicators BI1 and PSS4R, providing further evidence for not accomplishing discriminant validity (Hair et al., 2014), (Astrachan et al., 2014).

*Adjusted model*. After several iterations and the removal of weaker items (BI1, PSS4R, PSS3R, SI1, PEU2, PEU3, AT2, FE2) new outcomes were obtained and the fitted the model is as shown in Table 5.5.

This model meets the criterion for composite reliability. The model also demonstrates convergent validity which is also confirmed by the Cronbach's alpha coefficients. Values that meet the cut-off recommended to achieve discriminant validity were obtained by following the Fornell–Larcker criterion (Fornell & Larcker, 1981). Table 5.5 shows the correlations of the variables and the square roots of the AVEs which are the diagonal elements of the matrix on the right. After fitting the measurement model, all the square roots of the AVEs are higher than any other correlation. Thus, overall discriminant validity was achieved with in this PLS-SEM analysis.

Reliability and validity of the measurement model was proved by its internal consistency, convergent validity and discriminant validity. The overall model fit for the measurement model was within recommended ranges. Thus, the model achieves a good fit.

### 5.7.2 Evaluation of the Structural Equation Model

After the constructs have been confirmed as reliable and valid, the next step is to assess the structural model results to identify patterns in the data relationships.

Figure 5.6 shows the path coefficients and coefficients of determination ($R^2$ values) the lecturer's model. The $R^2$ values ranges from 0 to 1 with higher levels indicating higher levels of predictive accuracy. It shows the final structural model fit.

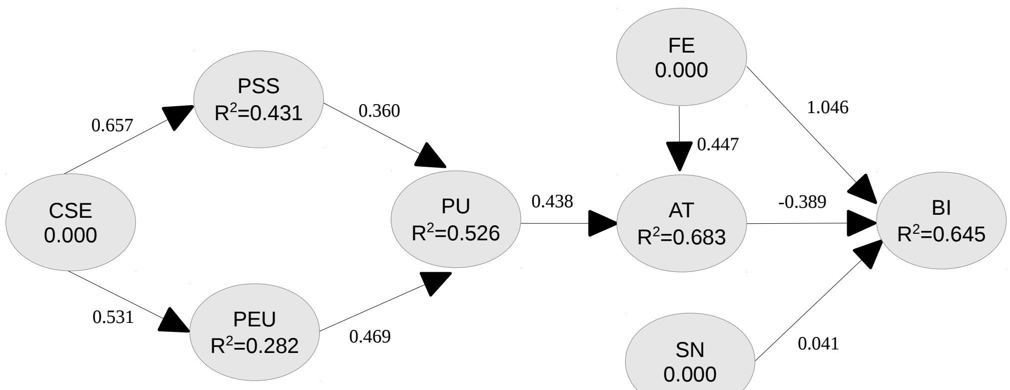


*Figure 5.6. The structural model used for hypotheses testing (path coefficients)*

### 5.7.2.1 Testing Hypotheses

A bootstrap procedure with 200 re-samples was used to test the statistical significance of the relations in the model. The results for the hypotheses are summarized in Table 5.6 and Figure 5.7.

*Table 5.6. Results of hypotheses test (summary of path coefficients and significance levels)*

| Hypotheses | Hypotheses path | Path coefficients | T-values significance | Accept/reject |
|---|---|---|---|---|
| H1 | PU → AT | 0.44 | 2.99*** | Accept |
| H2 | FE → AT | 0.45 | 2.89*** | Accept |
| H3 | FE → BI | 1.05 | 8.48*** | Accept |
| H4 | PEU → PU | 0.47 | 3.47*** | Accept |
| H5 | CSE → PEU | 0.53 | 4.43*** | Accept |
| H6 | CSE → PSS | 0.43 | 5.83*** | Accept |
| H7 | PSS → PU | 0.36 | 2.54** | Accept |
| H8 | AT → BI | -0.39 | 2.51* | Reject |
| H9 | SN → BI | 0.04 | 0.3 | Reject |

*Critical t-values for a two-tailed test are: <1.96 (p > .05*), 1.96 (p = .05**), and 2.58 (p = .001***).*

Figure 5.7 represented the hypothesized relationships among constructs, t-values and $R^2$ values schematically.

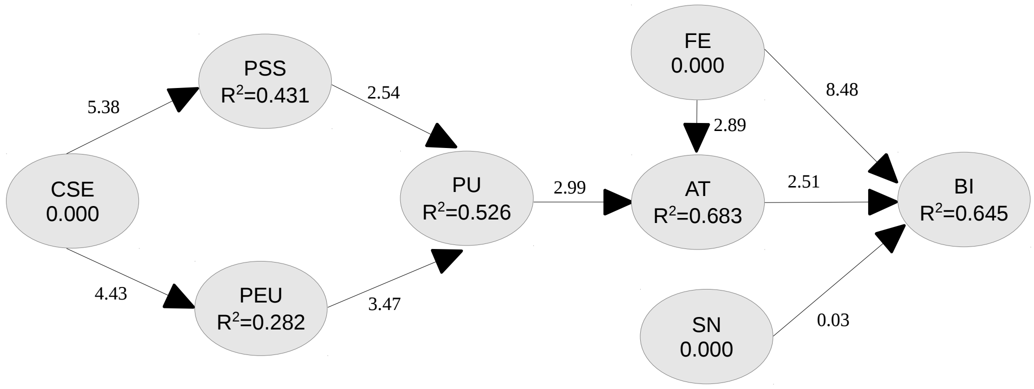


*Figure 5.7. The structural model used for hypotheses testing (t-values). Using a two-tailed t-test with a significance level of 5%, the path coefficients will be significant if the t-value is larger than 1.96.*

### 5.7.2.2 Hypotheses Examination

The next step was to examine the $R^2$ values for the endogenous constructs. Behavioural intention, the primary outcome measure of the model, was close to substantial, with an $R^2$ = .65. Attitude was also close to substantial with an $R^2$ = .68; perceived usefulness was moderate with an $R^2$ = .53; perceived system satisfaction was slightly below moderate $R^2$ = .43, while perceived ease of use was weak with an $R^2$ = .28.

The analysis of path coefficients and levels of significance shows that seven hypotheses were accepted, except the hypothesized path relationship between AT and BI (β: -0.39) and SN and BI (β: 0.04).

Hypothesis 1a predicts a causal relationship between perceived usefulness and attitude. The results shows that perceived usefulness had a significance impact on attitude with a path coefficient of (β: 0.44, p = .001); therefore the hypothesis is supported. The causal relationship between feedback and attitude (H2) is accepted (β: 0.45, p = .001). H3 shows that feedback is significant determinant of behaviour intention (β: 1.05, p = .001); therefore the hypotheses is strongly accepted. H4 predicts a causal relationship between perceived ease of use and perceived usefulness (β: 0.47, p = .001); it is accepted. The causal relationship between computer self-efficacy and perceived ease of use (H5) is accepted at (β: 0.53, p = .001). H6 states a causal relationship between computer self-efficacy and perceived system satisfaction (β: 0.43, $p = .001$). The causal relationship between perceived system satisfaction and perceived usefulness (H7) is also accepted at (β: 0.36, $p = .05$). H8 states a causal relationship between attitude and behavioural intention, the result (β: -0.39, $p = .05$) shows that this hypothesis is not accepted. H9 predicts a relationship between social influence and behavioral intention (β: 0.04, $p > .05$); this hypothesis was not supported by the data.

## 5.8 Discussion

The results show that a large amount of the variance in the endogenous construct behavioural intention is explained by only three constructs: attitude, feedback and social influence. These constructs together explain 65% of the variance $R^2$. This clearly illustrates that feedback (the lecturers' belief that electronic feedback can enhance students' learning) influences their intention to use on-line assessments more than their attitude towards the tool (the degree of a lecturers' favourable or unfavourable evaluation of the behaviour in question) and social influence. In other words, beliefs such as *on-line feedback serves as an encouragement to students in their learning*, *allowing them to gain support from classmates and students enjoy doing innovative on-line tests*, all contribute significantly more to usage intention than the attitude and opinions from colleagues and authorities. This also confirms what Venkatesh & Davis (2000), Venkatesh, Morris, Davis & Davis (2003) state; social influence can have a significant effect on intention in a mandatory setting, but not in a voluntary setting. It seems that lecturers in this study view the use of on-line testing tools as optional. This is also found in the research of Pynoo et al. (2012) who investigate teachers' acceptance and use of a voluntary educational portal. Their findings indicate that in such a setting, where teachers have a large degree of autonomy during teaching and preparation, including the choice of the technology, they experience no influence from their colleagues or institution.

Therefore, the results demonstrate that lecturers agree that on-line feedback is a versatile tool that allows them to adapt their instructional strategies for enhancing students' learning and performance, which in turn, is a good incentive for lecturers to be willing to use technology reflecting a nice sense of altruism.

Their responses show that their attitude towards using computers for assessment (*I do not believe in using computer-based assessments*) such as to author on-line tests (*I am here for teaching not to write on-line assessments*) as well social factors such as the opinions, suggestions from people to use the system (*people around me are positive about using electronic assessment*), (*if I need support using e-learning environments*, *a colleague is available to help me*) and the influence from educational managers (*university staff have supported me in the use of computers in my teaching*) do not substantially influence their usage intention. Therefore, this study shows that the attitude towards the use of testing was not a significant predictor of intention to use technology. This confirms what Davis (1989) proposes: attitude is only a modest factor in predicting technology acceptance and that individuals may use a technology even if they do not have a positive attitude toward technology per se, as long as it is perceived to be useful and/or easy to use in ways that enhance their productivity. This is also consistent with the research by Venkatesh et al. (2003) who compare eight user acceptance models and suggest that the attitudinal constructs are mainly significant when specific constructs related to performance and effort (similar to perceived usefulness and perceived ease of use from TAM) are not included in the model. This proves that lecturers find on-line testing very useful and they are willing to use it even if they do not have a positive attitude, explaining the lack of a significant relationship between attitude and usage intention.

In turn, attitude was predicted by perceived usefulness and feedback; together, these constructs explained 68% of the variance in attitude. The path coefficients of these constructs (perceived usefulness β: 0.44, feedback β: 0.45) shows that both influences are almost equal determinants for predicting a lecturers' attitude.

The effect of perceived usefulness on attitude means that factors such as practicality, helpfulness and time saving are important determinants for predicting a positive attitude. Lectures think that it is convenient to use technology to assess students' mathematical ability. Even though the questions included in this construct are empirical items, the outcomes demonstrate that these match with previous studies that show positive effects of perceived usefulness on attitude from the TAM (Ming Chi, 2010), (Gong et al., 2004); lecturers recognise that using technology for teaching represents advantages for students' learning. This matches with the study of Teo (2010), Teo et al. (2008), Legris et al. (2003). The positive effects of feedback on attitude prove feedback is an important determinant for predicting positive lecturers' attitude to use technology.

Perceived usefulness was predicted by perceived ease of use and perceived satisfaction with the system. These constructs explain 53% of the variance in perceived usefulness. Both constructs are influential for predicting usefulness. Nevertheless, the ease of use has more effect than the system satisfaction. It might mean that when lecturers think it is easy to use an on-line environment in testing, it is also more likely that they consider it as useful. In other words, the more practical and easier the system is (lecturers need to spend less time using it) the more useful it is, which also has an impact on attitude.

Computer self-efficacy explains 28% of the variance in perceived ease of use and 43% of the variance in perceived system satisfaction. This shows that if lecturers feel competent to use an on-line assessment system, it is more likely that they think they are able to use it easily. Nevertheless, as we have shown that more variance is explained by the perceived system satisfaction construct, it suggests that lecturers find the on-line environment more enjoyable than easy of use.

These results should be helpful for educational managers to elaborate implementation strategies and put in place effective support structures for lecturers to successfully experience the use of technology to cultivate positive computer attitudes, that in turn will ensure further and continual usage over time (Teo et al., 2008).

## 5.9 Conclusions

We have investigated instructors' attitudes and behavioural intentions to on-line assessment in mathematical subjects. We started this analysis based on the fact that people attempt to adopt a behaviour or a technology based on their beliefs about the consequences of adoption. Therefore, considering what matters to lecturers in order to get a willingness to use on-line assessments is a good starting point. Given that

TAM has been successfully accepted as a robust and efficient model to determine the acceptance and usage of technology, it is appropriate for examining what is involved in lecturers' usage intention.

The findings indicate that on-line feedback is the most important determinant for predicting lecturers' intention to use on-line assessments. Thus, in order to stimulate usage, it is important to increase their awareness of the advantages that electronic assessment provides to teaching practices. This is an important matter for policy-makers, who can potentially develop useful strategies to convince instructors to adopt technology by demonstrating students' improvements of their learning.

On the other hand, attitude has a negative effect on usage intentions which might mean that even though lecturers perceive that it is useful to perform on-line assessments, they are not totally convinced as the attitude construct shows. Probably, instructors still feel that it is more effective to perform mathematical exercises and exams by hand, rather than by using technology.

They recognize that the positive relationship between usefulness and attitude might mean that on-line assessments are practical and helpful tools that can enhance their productivity. They agreed that these are convenient tools for assessing students' mathematical ability. Lecturers also recognize the positive effects of feedback on usage intentions which shows they think that on-line feedback is a versatile tool that they can use to enrich their instructional strategies.

In conclusion, the results emphasize the significant advantages of including on-line feedback practices to enhance students learning in subjects involving mathematics which also serves as a motivator to the lecturers' favourable intention to use on-line assessment technologies.

# Chapter 6. Case Study: Students in the University of Manchester

## 6.1 Introduction

As mentioned previously the assessment of learning is an important stage in the educational processes. It is the crucial point in where students' learning gets consolidated and produces persistent changes in what students understand. This is the reason why educational assessment requires special attention. In order to provide good quality learning universities can make use of innovative technologies to support the teaching, learning and assessment processes. What should the role of information technologies be? Terzis et al. (2012) point out that the main reasons that learners are pleased using computer-based assessments are the following: "(1) Learners are able to take the assessment anywhere and any-time using a computer. (2) They are able to take the test as many times as they wish. (3) They feel confident regarding the results' accuracy and fairness since the computer does not care who the test taker is. (4) They are able to see their results as soon as they complete the assessment (Cassady & Gridley, 2005). (5) Electronic assessment provides them immediate feedback that helps identify their strengths and weaknesses" (Wilson, Boyd, Chen, & Jamal, 2011), (Crippen & Brooks, 2002).

There are two electronic assessment practices particularly useful for students; the first has the potential to provide them with the opportunity to complete assessments at a place and time that is convenient for them (Bennett, 2001). This also provides students with 'the freedom to explore areas of perceived weakness and to make mistakes without revealing these to those responsible for the final assessment or to peers without this being a deliberate decision on their part' (Challis, 2005). On the other hand, Wilson et al. (2011) explain that (formative) electronic assessment provides students with an important opportunity, the flexibility to evaluate their understanding of a course, the opportunity to set their own learning goals and assess their weaknesses and strengths in order to improve their performance in the course. Realising their strengths enables students to focus on their cognitive development.

To figure out weaknesses will help students to identify key areas of focus for studying. These studies show that students who use the computer-assisted practice quizzes earn significantly higher grades than those students who do not. This demonstrates that computer-assisted formative assessment (used for practice tests) has a positive impact on student performance. When feedback is provided immediately, it not only improves student learning but also encourages independent learning. This in turn leads to self-efficacy (Miller, 2009).

Using the computer to assess mathematical subjects also brings important advantages; for instance, Angus & Watson (2009) illustrate that where students studying mathematics are subject to regular low marks in on-line testing, their learning is significantly improved as measured by an end of semester examination. Indeed, mathematics is particularly suited to an on-line assessment strategy and can provide valuable feedback to students, particularly distance learners (Whitelock & Raw, 2003). These benefits can be extended to students in blended educational contexts. The research of Peltenburg, Van Den Heuvel-Panhuizen & Doig (2009) shows that e-assessment using a dynamic visual tool is able to support students to overcome the difficulties they face in solving subtraction problems. This type of testing allows the teacher to better examine the students' actions and thinking processes than is possible with a paper-and-pencil test (Whitelock, 2009).

Other researchers such as Dettori, Garuti & Lemut (2002) have investigated the possibilities of improving mathematical teaching and learning by using technology in various educational contexts. They report that studying mathematical topics supported by technology highly increases students' motivation towards learning mathematics, which also creates a positive change in students' attitude towards mathematics (Ursini, Sánchez, & Orendain, 2004). There are several studies that study this attitude (Reed, Drijvers, & Kirschner, 2010), (Ursini et al., 2004), (Galbraith & Haines, 1998), which is associated with determining the grade of willingness to learn. This study does not consider this attitude per se. Instead, the focus is on analysing the attitude and intention of using electronic assessment technologies to learn mathematics in the context of higher education.

However, the benefits of on-line assessment can be a factor, provided students have the right attitude towards using computers for learning. In other words, the appropriate computer attitude not only can play an influential role in determining the extent to which students accept it as a learning tool, but also future behaviour towards the computer such as using it for further study and vocational purposes (Rosen & Weil, 1995).

This coincides with what was mentioned in chapter 3 regarding the reasons how and why technologies are adopted by users. In this vein, Venkatesh et al. (2003) suggest including users' attitudes as a key construct in predicting technology acceptance for future use. Therefore, it is important to understand what drives students to adopt electronic assessment.

An instrument to determine it has been developed. It is again an empirical proposal that was only based on prior studies. A model is also developed that predicts the students' feelings towards the use of on-line tests. Even though both questions and model included were new, the results show enough reliability to build useful insights.

Therefore, data was collected from students with the aim of gathering insights regarding the most relevant factors affecting students attitude towards the use of web-based assessment. This study also provides the opportunity to have a better and conclusive understanding of the students' feelings towards the use of on-line testing.

## 6.2 Context

The students studied in this chapter all study in various Schools of the Faculty of Engineering and Physical Sciences at the University of Manchester. As part of the policies set out in the teaching and learning plan of the University, the students are exposed to a variety of on-line materials. Particularly relevant to the subject of this thesis is the use of questions using the Stack engine within the web-based Moodle platform (Sangwin, 2013) that provides feature-rich on-line assessments with malrule based feedback.

Every year, the faculty receives a large amount of students. As the incoming students come from different mathematical backgrounds, it is necessary to determine the previous level of mathematical knowledge, which is done by a diagnostic test. Practice material for the test, and follow-on practice material, are all provided within Stack.

The University of Manchester requires that all students (except the course of Chemistry) have a mathematics A-level. Each school has its own requirements to entry, however all students are required to take the same diagnostic test, then students attend mathematics courses either in the School the Mathematics in the case of an engineering course or in the School of Physics and Astronomy for physics students. In both cases, the University has a variety of on-line learning materials available.

The students practise several mathematical exercises before answering the diagnostic exam: 'first-year student' that consists of a series of mathematical exercises grouped in 12 sections (arithmetic and algebra, geometry and trigonometric, series, functions, polynomials, exponentials and logarithms, differentiation, integration, further algebra, further differentiation, further integration, vectors) that confirm the basic mathematical knowledge that students should know before starting a course. All sections and an example of the section 'differentiation' are presented in Appendix C. The key characteristic of this approach is provided to students with immediate feedback through the web-based platform. In case students answer incorrectly, the platform will give them immediate answers by sending a series of hints to help them obtain the correct solution. Students can practise as long as they want/need to reinforce their mathematical skills. This formative feedback represents a big advantage for on-line testing since it opens the possibility of providing students with customised feedback given that the computer can generate this based on the answer given by the student (López, 2009).

In this way, the school obtains evidence of the previous mathematical ability of students with which they can obtain a reliable base to adapt each course to students' needs. Specifically in the School of Physics and Astronomy, from which a large fraction of the respondents originate, students also experience the “mastering physics" product developed by MIT, and now marketed by Pearson (Walet & Birch, 2012).

It is important for the Schools to obtain enough data to build a customised learning strategy to provide students with good quality teaching. While academics may have a wide variety of reasons for selecting particular assessment methods, they need to be aware of their students' perceptions of these methods and how these influence students' learning (Iannone & Simpson, 2013). Lecturers' opinions have already been mentioned. However, in order to obtain a complete view of the context, it is also important to consider what students' think about learning mathematics through technology. What is their opinion? In order to make the most of what technology offers, students should be willing to undertake mathematical exercises and exams using technology to make their learning of mathematics more valuable.

## 6.3 Proposed Constructs

The students' survey was designed in 5 sections. Each section defines a concept (construct) with the aim of evaluating the perception and feelings of students towards web-based assessment, while obtaining a complete context for the issues that influence students attitude towards on-line exercises and take exams. We shall work at the concepts "affective factors", "utility", "suitability" "reliability" and "feedback". A brief description of them is included below. As mentioned in the previous chapter the term "construct"  denotes a group of indicators or items.

*Affective factors*. (Fishbein, & Ajzen, 1975) define attitude as ‘‘the degree of a person's favourable or unfavourable evaluation or appraisal of the behaviour in question’’. Attitude involves direction (positive or negative), intensity (high or low); it consists of several elements such as cognitions or beliefs, feelings or emotions associated with evaluations and behavioural tendencies (Ursini et al., 2004). This can be clearly seen in what Segers, Dochy & Cascallar (2003) called the 'pre-assessment effect' which impacts how students' learning is affected by an assessment task, and is the result of how students perceive that task, its fairness, validity, and the values it embraces. In other words, the way students perceive how assessment tasks are designed (aspects such as utility, suitability, reliability) can determine their attitude towards assessment activities and thus it may also influence the grade obtained. In terms of this study, affective factors such as “the feelings of students doing electronic assessment” are defined.

*Utility.* Aspects of learning such as freedom, ubiquity or time-saving are considered as important to students. The fact that students can do maths exercises any time, any place makes them willing to perform a task. In fact, students find electronic assessment useful since they can practise and minimize their weaknesses in any subject or specific task and to evaluate their performance (Kaklauskas et al., 2010). Therefore, according to the aims of this study, to know students' perception about the electronic assessment utility is required. Utility is defined as “the degree in which e-assessment can enhance students learning“.

*Suitability.* Another important aspect is students' belief whether or not the topic studied is appropriate to be evaluated by using technology. This concept helps to clarify learning goals and criteria of assessment which contributes also to highlight validity issues. It is relevant since the perceptions of the value and validity of assessment affect learning (Scouller, 1998). Suitability is defined as “the students' belief that the technology used is appropriate to evaluate understanding of a specific topic”.

*Feedback*. Chapter 3 has already explained the advantages of integrating feedback in teaching practices. In summary, an essential feature of feedback is that it contributes to self-assessment, making students regulate their learning and adjusting it to their own necessities. Feedback also promotes and improves dialogue with lecturers, classmates, social community, etc. Formative assessment practices through computers foster immediate feedback, which has positive effects for students in achieving their goals (Whitelock & Brasher, 2006). Dreher et al. (2011) mention that the real benefit of assessment methods is getting immediate feedback, which enhances students learning performance and also activates their intrinsic motivation within the learning setting. Feedback is defined as “the students' belief that the support and feedback they receive from the on-line platform will improve their learning”.

*Reliability*. To include aspects of validity and reliability is fundamental for educational assessment practices that depend to a large extent on the level of trust and confidence of students in the assessment practices which will also be reflected in students' efforts in their learning. Students are concerned about the security of testing (Cassady & Gridley, 2005), possibilities of cheating (King et al., 2009) and the fairness of question banks (Dermo, 2009). It is important that students have enough confidence in a test; the level of engagement and cooperation depends on this (Domino & Domino, 2006). Reliability is defined as “the students' belief that the technology assesses their work fairly and accurately”.

## 6.4 Research Methods

*Designing Research Models*. Based on the factors described above a number of constructs for the student's models are proposed. These constructs include several indicators as shown in Table 6.1.

*Table 6.1. Distribution of constructs and indicators proposed to student's model*

| Construct | | Indicators |
|---|---|---|
| AF | is composed of | AF1, AF2, AF3 |
| UT | is composed of | UT1, UT2, UT3 |
| SU | is composed of | SU1, SU2, SU3, SU4, SU5 |
| FE | is composed of | FE1, FE2, FE3, FE4 |
| RE | is composed of | RE1, RE2, RE3 |

*Proposing hypotheses.*

Based on the factors proposed it was hypothesized that:

- *H1: states there is a causal relationship between feedback and reliability.*
- *H2: states there is a causal relationship between feedback and usefulness.*
- *H3: states there is a causal relationship between feedback and affective factors.*
- *H4: states there is a causal relationship between reliability and affective*

*factors.*

- *H5: states there is a causal relationship between usefulness and affective factors.*

The hypotheses can be represented in the following Figure 6.1.

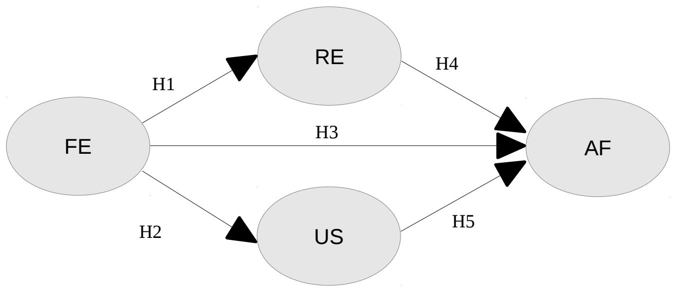


*Figure 6.1. A schematic representation of the theoretical model and hypotheses*

## 6.5 Data Collection Procedures

*Procedure for the questionnaire*. This questionnaire consists of 18 items displayed in the Table 6.2.

Similar to the format of the lecturer's survey, data was gathered from an on-line questionnaire which was run at the same time as that of lecturer's. For this questionnaire a total of 127 responses was obtained, 6 were removed because these were partially filled in resulting in a data set of (71 male and 50 female students).

*Table 6.2. Original scale for student's model*

| | **Affective factors (AF)** |
|---|---|
| **AF1** | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) |
| **AF2** | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) |
| **AF3** | On-line assessments test my computer skills as well as the topic being studied. |
| | **Utility (UT)** |
| **UT1** | I am used to reading lecture notes on-line |
| **UT2** | Electronic assessments help me to get a deeper understanding of the subject. |
| **UT3** | I find it useful that I can do on-line exercises at a time of my choosing. |
| | **Suitability (SU)** |
| **SU1** | I think that on-line tests and exams are appropriate to test my ability in using mathematics. |
| **SU2** | It is easier for me to take an on line test because I can get information using the internet. |
| **SU3** | It is easier to get support from others students in an on-line assessment than in a paper-based one. |
| **SU4** | I can save time by answering a test on-line. |
| **SU5** | I think that I'm using cutting edge technology in my on-line tests. |
| | **Feedback (FE)** |
| **FE1** | On-line feedback encourages me to do better in my studies. |
| **FE2** | I receive sufficient personalised feedback on my on-line tests. |
| **FE3** | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. |
| **FE4** | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. |
| | **Reliability (RE)** |
| **RE1** | It is easier to cheat during an electronic assessment than when doing it on paper. (R) |
| **RE2** | I think that on line tests are fairer because there is less room for human error than a paper-based one. |
| **RE3** | On-line tests can often be passed by using flaws in the software. (R) |

(R) Reverse code

This survey shows that the responders came from 9 academic schools, most of which were studying in the fields of physics and astronomy, mathematics, chemistry, mechanical, aerospace, and civil engineering as can be seen in the Figure 6.2.

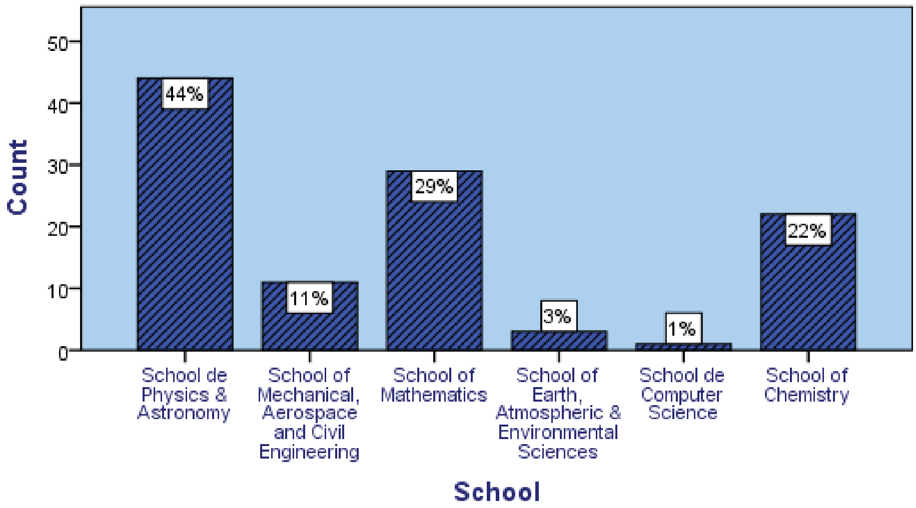


*Figure 6.2. Distribution of students' responses shown by school*

The students study fields are suitable to this research since they are all STEM subjects in a higher education settings.

## 6.6 Data analysis techniques

In order to examine students' emotional attitudes more closely, we analyse the responses related to the affective factors. The items shown in Table 6.3 above belong to this construct. The values of the items are assessed using a Likert scale from (5 = strongly agree) to (1 = strongly disagree) as for lecturers' survey.

Table 6.3 shows the percentage of responses in the survey and then determines students' attitude to these questions, allowing us to obtain some preliminary insights about their thoughts on web-based assessment.

*Table 6.3. Distribution of students' responses for the affective construct*

| I think that is more stressful doing an on-line test/exam than a paper-based one | | |
|---|---|---|
| **Answer** | **Count** | **Percentage** |
| Strongly agree | 17 | 13% |
| Agree | 19 | 15% |
| **Neither agree nor disagree** | 35 | 27% |
| **Disagree** | 34 | 27% |
| Strongly disagree | 9 | 7% |
| Not completed or Not displayed | 14 | 11% |
| It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one | | |
| Strongly agree | 34 | 27% |
| **Agree** | 40 | 31% |
| Neither agree nor disagree | 20 | 16% |
| Disagree | 16 | 13% |
| Strongly disagree | 4 | 3% |
| Not completed or Not displayed | 14 | 11% |
| On-line assessments test my computer skills as well as the topic being studied | | |
| Strongly agree | 5 | 4% |
| Agree | 32 | 25% |
| Neither agree nor disagree | 25 | 20% |
| **Disagree** | 38 | 30% |
| Strongly disagree | 14 | 11% |
| Not completed or Not displayed | 14 | 11% |

The results of the first question showed that 27% of students see no difference between doing electronic or paper-based tests. But 27% think that doing on-line test/exams do not make them feel stressed which probably indicates that they are very familiar with technology. This result contrasts with the second question: although they are used to technology they find it more difficult to resolve an e-test/exam than a paper-based one. The third question indicates they think that if they answer an e-test their knowledge and skills regarding the topic would also be evaluated; this tells us that they may consider the use of technology to evaluate that appropriate topic.

Table 6.4 summarises the descriptive coefficients including mean and standard deviation for each item.

*Table 6.4. Descriptive statistics (mean and standard deviation) of the items in the student's scale*

| | **Affective factors** | Mean () | Std. Devia-tion () |
|---|---|---|---|
| AF1R | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) | 2.99 | 1.18 |
| AF2R | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) | 2.27 | 1.14 |
| AF3 | On-line assessments test my computer skills as well as the topic being studied. | 2.80 | 1.12 |
| | **Utility** | | |
| UT1 | I am used to reading lecture notes on-line | 3.65 | 1.15 |
| UT2 | Electronic assessments help me to get a deeper understanding of the subject. | 3.04 | 1.03 |
| UT3 | I find it useful that I can do on-line exercises at a time of my choosing. | 4.12 | 0.90 |
| | **Suitability** | | |
| SU1 | I think that on-line tests and exams are appropriate to test my ability in using mathematics. | 3.09 | 1.21 |
| SU2 | It is easier for me to take an on line test because I can get information using the internet. | 3.36 | 1.17 |
| SU3 | It is easier to get support from others students in an on-line assessment than in a paper-based one. | 3.06 | 1.20 |
| SU4 | I can save time by answering a test on-line. | 3.30 | 1.06 |
| SU5 | I think that I'm using cutting edge technology in my on-line tests. | 2.51 | 0.93 |
| | **Feedback** | | |
| FE1 | On-line feedback encourages me to do better in my studies. | 3.50 | 0.79 |
| FE2 | I receive sufficient personalised feedback on my on-line tests. | 2.65 | 1.04 |
| FE3 | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. | 3.99 | 0.91 |
| FE4 | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. | 4.01 | 1.00 |
| | **Reliability** | | |
| RE1R | It is easier to cheat during an electronic assessment than when doing it on paper. (R) | 2.48 | 1.12 |
| RE2 | I think that on line tests are fairer because there is less room for human error than a paper-based one. | 2.39 | 1.04 |
| RE3R | On-line tests can often be passed by using flaws in the software. (R) | 3.12 | 0.96 |

The values in Table 6.4 represent the average responses. Thus higher values mean there are a greater number of positive answers (mean>3.0) while the lower values represent negative responses (mean<3.0). The results show that question UT3 'I find it useful that I can do on-line exercises at a time of my choosing' (mean=4.01) has the highest mean value indicating that students consider it is very important to feel free to practise on-line exercises as they can/want. This feature is followed by question FE4 'Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic' (mean=4.01), and question FE3 'Immediate on-line feedback can help me resolve doubts about the material I am studying faster' (mean=3.99), these results show that it is very important to students to receive appropriate feedback.

On the other hand, the questions with lower values are 'It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one' (mean=2.27) as mentioned above, it does not make any difference taking a paper-based or computer-based test. The question RE2 'I think that on-line tests are fairer because there is less room for human error than a paper-based one' (mean=2.39) indicates that students do not think that just by automatizing an exam/test it would make a better exam. The question SU5 'I think that I'm using cutting edge technology in my on-line tests' (mean=2.50) indicates that using cutting edge technology is not important to students. This question was included because it was thought it could affect the students' feelings toward using on-line test, based on the fact that young people are often fascinated with new technology.

*Initial measurement model evaluation.* As we have defined an empirical scale, which was developed considering non-validated items from prior research, the first step was to check their appropriateness and reliability. This was checked by calculating Cronbach's alpha coefficients for the entire dataset and each construct separately as shown in the Table 6.5.

The results show that the internal consistency of the whole scale is good ($\alpha$ = .81). The constructs suitability and reliability reach the cut-off value suggested for an exploratory study ($\alpha$ = .60) (Hair et al., 2010). The rest of the constructs show a limited internal consistency. Therefore, adjustments to the scale are required. Two approaches were followed: For Model A, we run an exploratory factor analysis to define the underlying structure among the items. For Model B, we reorganize the items conceptually.

*Table 6.5. Cronbach's alpha coefficients for original constructs*

| | **Affective factors α = .102** |
|---|---|
| AF1 | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) |
| AF2 | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) |
| AF3 | On-line assessments test my computer skills as well as the topic being studied. |
| | **Utility α = .572** |
| UT1 | I am used to reading lecture notes on-line. |
| UT2 | Electronic assessments help me to get a deeper understanding of the subject. |
| UT3 | I find it useful that I can do on-line exercises at a time of my choosing. |
| | **Suitability α = .697** |
| SU1 | I think that on-line tests and exams are appropriate to test my ability in using mathematics. |
| SU2 | It is easier for me to take an on-line test because I can get information using the internet. |
| SU3 | It is easier to get support from others students in an on-line assessment than in a paper-based one. |
| SU4 | I can save time by answering a test on-line. |
| SU5 | I think that I'm using cutting edge technology in my on-line tests. |
| | **Feedback α = .691** |
| FE1 | On-line feedback encourages me to do better in my studies. |
| FE2 | I receive sufficient personalised feedback on my on-line tests. |
| FE3 | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. |
| FE4 | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. |
| | **Reliability α = .384** |
| RE1 | It is easier to cheat during an electronic assessment than when doing it on paper. (R) |
| RE2 | I think that on line tests are fairer because there is less room for human error than a paper-based one. |
| RE3 | On-line tests can often be passed by using flaws in the software. (R) |

The appropriateness of constructs by examining the entire correlation matrix was determined. A visual inspection revealed a substantial number of correlations greater than .30. Thus, factor analysis is appropriate.

*Model A. Exploratory Factor Analysis (EFA).* For this empirical scale, to identify logical combinations of variables and better understand the interrelationships among variables is necessary. It provides an empirical basis for judging the structure of the variables and the impact of this structure when interpreting the results from other multivariate techniques (Hair et al., 2010). This technique also helps us to examine the reliability and validity of the related constructs. Therefore, an EFA following the five-step procedure suggested in Hair's book, as shown in the Appendix D, was executed. After several iterations, the weaker items are detected and removed, thus obtaining an empirically validated four-factor solution (see Table 6.6). Based on a qualitative assessment of the loadings, two constructs were confirmed with the same indicators. A new construct emerges named “usefulness”.

For the initial model, without any adjustment the results of the KMO (Kaiser-Meyer-Olkin), Measure of Sampling Adequacy test was = .741, which indicates that it is appropriate to modify this model in order to obtain a better fit. To estimate the validity of each construct an extraction method running principal component analysis using a rotation Promax method with Kaiser normalization keeping values above .3 was made. The communality values were satisfactory with values at least .50 (Hair et al., 2010). Convergent validity and discriminant validity were also reached. For details of this analysis see Appendix D.

The reliability of the scale was also obtained with Cronbach's alpha coefficients shown in Table 6.6. The results show that the model has enough factors to meet the suggested percentage of explained variance (60%) (Hair et al., 2010) for a reliable model. The model explains the 70% variance.

*Table 6.6. Solution emerged from exploratory factor analysis*

| | |
|---|---|
| | **Usefulness α = .698** |
| FE2 | I receive sufficient personalised feedback on my on-line tests. |
| PU2 | Electronic assessments help me to get a deeper understanding of the subject. |
| PS5 | I think that I'm using cutting edge technology in my on-line tests. |
| | **Reliability α = .615** |
| PR1R | It is easier to cheat during an electronic assessment than when doing it on paper. (R) |
| PR3R | On-line tests can often be passed by using flaws in the software. |
| | **Feedback α = .733** |
| FE4 | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. |
| FE3 | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. |
| | **Affective factors α = .636** |
| AT1R | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) |
| AT2R | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) |
| PS3 | It is easier to get support from others students in an on-line assessment than in a paper-based one. |

After running EFA the results of the KMO test was = .679. Bartlett's Test of Sphericity (Approx. Chi-Square = 262.232, df = 45, Sig. = .000). A statistically significance Bartlett's Test of Sphericity (sig<.05) indicates that sufficient correlations exist among the variables. Loadings above the cut-off point by ± .40 for interpreting proposes were obtained. The reliability of each construct was better than the original scale: The reliability of the scale was α = .722, which is considered as acceptable to (rule of thumb: α ≥ .9 excellent, .9 > α ≥ .8 good, .8 > α ≥ .7 acceptable, .7 > α ≥ .6 questionable, .6 > α ≥ .5 poor, .5 > α unacceptable).

*Model B. Reorganizing items conceptually* In order to obtain an alternative model, the results of the EFA were taken into account. This helped identify and remove the weaker items (AF3, PR2) and also gave us insight in how to reorganize items according to the statistical structure of the scale. It is important to consider the statistical output, but it is essential to consider the theoretical foundation of each construct. Thus, the items were reorganized conceptually by clustering those items that shared similar features with those that loaded on a same construct. Four constructs were obtained as displayed in Table 6.7.

*Table 6.7. Solution emerged by reorganising constructs conceptually*

| | **Affective factors α = .636** |
|---|---|
| AF1 | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) |
| AF2 | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) |
| AF3 | It is easier to get support from others students in an on-line assessment than in a paper-based one. |
| | **Usefulness α = .737** |
| US1 | Electronic assessments help me to get a deeper understanding of the subject. |
| US2 | I find it useful that I can do on-line exercises at a time of my choosing. |
| US3 | I think that on-line tests and exams are appropriate to test my ability in using mathematics. |
| US4 | It is easier for me to take an on-line test because I can get information using the internet. |
| US5 | I can save time by answering a test on-line. |
| US6 | I think that I'm using cutting edge technology in my on-line tests. |
| | **Feedback α = .691** |
| FE1 | On-line feedback encourages me to do better in my studies. |
| FE2 | I receive sufficient personalised feedback on my on-line tests. |
| FE3 | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. |
| FE4 | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. |
| | **Reliability α = .615** |
| RE1 | It is easier to cheat during an electronic assessment than when doing it on paper. (R) |
| RE2 | On-line tests can often be passed by using flaws in the software. (R) |

The reliability of the scale α = .824. To validate both solutions and thus test the hypotheses, a SEM was executed as discussed in the following section.

## 6.7 Results and Conclusions

### 6.7.1 Evaluation of the Model Quality

In order to prove the quality of the models and to test the hypotheses the data was analysed using the PLS-SEM approach, using the SmartPLS version 2 software. Both models (proposals A and B) were tested to determine the best fit.

#### 6.7.1.1 Model A. Evaluation of Measurement Model Emerged from EFA

The first step in evaluating a PLS-SEM model is to examine the outer model in an effort to validate the measurement model (Hair et al., 2014). To this end, relationships between the constructs and their indicators are assessed. Confirmatory Factor Analysis (CFA) was undertaken to further assess the factor structure and validate the scale. CFA determines whether the loading patterns of the measurement items correspond to the theoretically anticipated factors. Reliability and validity of the measurement model is proved by internal consistency, convergent validity and discriminant validity (Barclay et al. 1995).

As shown in Table 6.8, reliability measures derived from CFA are a composite of reliability and the average variance extracted (AVE).

*Table 6.8. Model A. Construct and convergent validity coefficients. Discriminant validity coefficients: bivariate correlations. In the main diagonal, the squared-root AVE of each construct*

| | AVE | Composite Reliability | Cronbach's Alpha | AT | FE | RE | SU |
|---|---|---|---|---|---|---|---|
| **AT** | 0.58 | 0.8 | 0.64 | **0.76** | | | |
| **FE** | 0.79 | 0.88 | 0.73 | 0.35 | **0.89** | | |
| **RE** | 0.56 | 0.68 | 0.62 | 0.21 | -0.02 | **0.75** | |
| **SU** | 0.62 | 0.83 | 0.7 | 0.37 | 0.41 | 0.07 | **0.79** |

Table 6.8 shows composite reliabilities ranged from .68 to .88 exceeding the minimum requirement of .7 for all constructs except RE, with a value very close to the cut-off. The AVE values ranged from .56 to .79 while the recommend cut-off is .50 (Hair et al., 2012), confirming convergent validity and implicitly content validity in all constructs.

This table also shows the AVEs on the diagonal and the squared inter-construct correlations on the off-diagonal entries. The Fornell–Larcker criterion (Fornell & Larcker, 1981) demonstrates that all AVEs are higher than the squared inter-construct correlations. Therefore, overall discriminant validity was achieved. Thus, using both convergent and discriminant validity demonstrated that there is evidence of construct validity.

#### 6.7.1.2 Model B. Evaluation of Reorganized Measurement Model

The measurement model for the proposal B was also evaluated. As shown above, construct validity was proved by checking convergent and discriminant validity. Table 6.9 shows values for composite reliabilities ranging from .76 to .81, exceeding the minimum requirement. The AVE values ranged from .51 to .64 exceeding also the cut-off recommended. Likewise, values for Cronbach's alpha were considered as good. Therefore, all values obtained confirm convergent validity.

*Table 6.9. Model B. Construct and convergent validity coefficients. Discriminant validity coefficients: bivariate correlations. In the main diagonal, the squared-root AVE of each construct*

| | AVE | Composite Reliability | Cronbach's Alpha | AT | FE | RE | SU |
|---|---|---|---|---|---|---|---|
| **AT** | 0.53 | 0.81 | 0.7 | **0.73** | | | |
| **FE** | 0.53 | 0.81 | 0.7 | 0.4 | **0.73** | | |
| **RE** | 0.64 | 0.76 | 0.62 | -0.14 | 0.18 | **0.80** | |
| **SU** | 0.51 | 0.81 | 0.68 | 0.59 | 0.6 | 0.23 | **0.71** |

Discriminant validity by checking all AVEs higher than the squared inter-construct correlations were also proved. The Fornell–Larcker criterion (Fornell & Larcker, 1981) was demonstrated; all AVEs were higher than the squared inter-construct correlations. All values reached a good level of reliability, even for the indicator PR3R that has a value allowing for exploratory studies.

### 6.7.2 Model A. Evaluation of the Structural Equation Model Emerged from EFA

#### 6.7.2.1 Testing Hypotheses

We calculated the sizes and significance of the path coefficients that represent the hypothesized relationships. To obtain the significance levels a bootstrapping procedure was run using 800 sub-samples. Figure 6.3 shows the t-statistics for each relationship. Table 6.10 also shows the path coefficients, t-statistics, and summarizes the results of the hypotheses test.

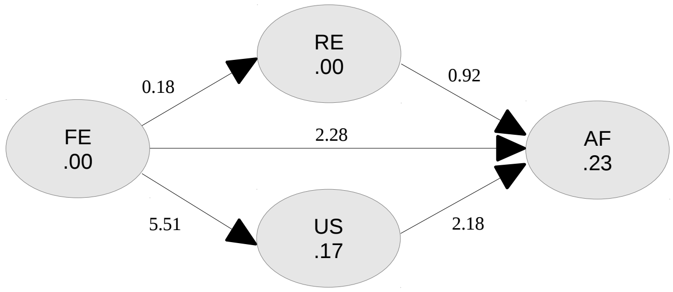


*Figure 6.3. The structural model used for hypotheses testing (path coefficients and R-squared values)*

*Table 6.10. Model A. Results of hypotheses test (summary of path coefficients and significance levels)*

| Hypotheses | Hypotheses path | Path coefficients | T-values significance | Accept/Reject |
|---|---|---|---|---|
| H3 | FE → AF | 0.26 | 2.28 | Accept |
| H1 | FE → RE | -0.02 | 0.18 | Reject |
| H2 | FE → US | 0.4 | 5.51 | Accept |
| H4 | RE → AF | 0.2 | 0.92 | Reject |
| H5 | US → AF | 0.25 | 2.18 | Accept |

*Critical t-values for a two-tailed test are: <1.96 (p>.05*), 1.96 (p=.05**), and 2.58 (p=.001***)*

## 6.7.3 Model B. Evaluation of the Structural Equation Model Emerged from the Reorganized Scale

### 6.7.3.1 Testing Hypotheses

The procedure shown above to obtain the path coefficients and t-statistics was followed. Figure 6.4 shows the t-statistics for each relationship. Table 6.11 shows the path coefficients, t-statistics and summarizes the results of the hypotheses test.

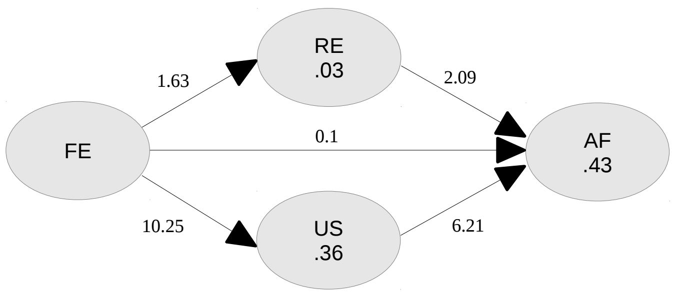


*Figure 6.4. The structural model used for hypotheses testing (t-values)*

*Table 6.11. Model B. Results of hypotheses test (summary of path coefficients and significance levels)*

| Hypotheses | Hypotheses path | Path coefficients | T Statistics (\|O/STERR\|) | Accept/Reject |
|---|---|---|---|---|
| H3 | FE → AF | 0.1 | 0.99 | Reject |
| H1 | FE → RE | 0.18 | 1.63 | Reject |
| H2 | FE → US | 0.6 | 10.25 | Accept |
| H4 | RE → AF | -0.3 | 2.09 | Reject |
| H5 | US → AF | 0.59 | 6.21 | Accept |

*Critical t-values for a two-tailed test are: <1.96 (p > .05*), 1.96 (p = .05**), and 2.58 (p = .001***).*

#### 6.7.3.2 Hypotheses Examination

The results for both empirical research models demonstrated supports all hypotheses except hypothesis H3.

H1 states there is a causal relationship between feedback and reliability; both models showed that this relationship is not significant. Therefore, the hypothesis is not supported. Both models showed that feedback had a significance impact on usefulness (H2) with paths coefficient of (β: 0.4, p = .001) and (β: 0.6, p = .001). Therefore, this hypothesis is supported. The causal relationship between feedback and affective factors (H3) was accepted (β: 0.26, p = .001) in case A whereas that in case B was rejected (β: 0.1, p = .001). H4 predicts a causal relationship between reliability and affective factors; both models showed that reliability does not impact affective factors with values at (β: 0.2, p = .001) and (β: -0.3, p = .001). This hypothesis is not supported. The causal relationship between usefulness and affective factors (H5) is strongly accepted for both models with values at (β: 0.25, p = .001) and (β: 0.59, p = .001).

The affective factors construct was predicted by feedback, reliability and usefulness; these constructs together explain 43% ($R^2$ = .43) of the variance in affective factors indicating an overall $R^2$ value. Reliability was predicted by feedback, this variable explains 3% ($R^2$ = .03) of the variance in reliability. Usefulness was predicted by feedback, this variable explains 36% ($R^2$ = .36) of the variance in usefulness.

## 6.8 Discussion

Determining students' attitude towards on-line assessments gives us an understanding of what really matters to students in their learning. Our awareness should be ‘‘a critical criterion in the evaluation of computer courses and in the development of computer-based curricula’’ (Woodrow 1991). These insights also help to determine to what extent students will or (will not) accept using on-line testing. As students, they have to accept the learning methods and academic assignments that lecturers establish. However, the degree to which they are willing to do so is a factor that should be considered. The same issue is also seen in commercial settings; an employee has to be willing use the information system that the enterprise has chosen. That is why information system researchers have underlined the importance of personal factors, such as attitudes, beliefs, cognitions, culture and behaviours regarding technology acceptance (Davis et al., 1992). There are important attitudinal factors that affect the usage of technology successfully such as instructors' and students' mind-set about on-line learning and level of collaboration in the course. These are crucial factors for successful applications of on-line course resources (Benson Soong et al., 2001).

Our findings show that usefulness is the major factor impacting students' attitude to use on-line testing environments. This result is consistent with the work of Terzis, Moridis, Economides & Mendez (2013). This means, when students perceive on-line assessment as useful and where it also increases their productivity, their intention to use it will be significant increased. It seems that the main driver for adoption of such a technology is the user's belief that technology will enrich his/her performance (Davis, 1989). Thus, in order to obtain a positive attitude towards on-line assessments, it is necessary to increase perceived usefulness. Educational managers may encourage lecturers to use on-line assessments by providing technological tools that raise their aspirations to use innovative teaching methods. In turn, lecturers directly influence students; they can persuade students to use on-line tools for achieving their learning goals. If that is so, they will be more willing to use technology.

The outcomes also reveal that reliability is not an important factor in students' attitude towards on-line testing. Literature on educational assessment asserts that a fundamental component of validity is the concept of reliability. This means to what extent students consider it important that an on-line test measures what it claims to measure (Gikandi et al., 2011). This fact is surprising since issues that are a common cause for complaint, such as the lack of fairness of assessments, play such a small role in this study. This outcome is contrary to the opinion of (Segers et al., 2003) who report that students' learning is affected by an assessment task, in how students perceive that task, its fairness, validity and the values it embodies. It is also supported by (Scouller, 1998) who points out that the student value perception and validity of assessment affects their learning. Perhaps, this result is due to the lack of accuracy of the items to determine this issue ($\alpha$ = .615). Moreover, the feedback construct was not a factor that predicts the adoption of on-line assessments. As happened with the reliability construct their items were not sharp enough to exert an impact. However, an item shown to be a crucial feature that students truly value.

The item highlights the fact that students can obtain their grades immediately after doing an on-line test. For this study it means immediately after completing a section of the test. In this way the on-line testing environment helps students to determine how they are doing. Therefore, they appreciate the on-line feedback practices. This matches with the aim of feedback mentioned earlier. It means allowing students to be aware of how effective they are as learners through reflection and regular feedback, namely it is the comparison between student's present state in relation with goals and standards to determine if they should continue as they are, or if a type of change is required.

Another important feature that students value is the academic support that an on-line assessment system brings to their learning. In terms of this study obtaining better understanding of the subject. Students already taking classes in a traditional way and adding different ways to learn, such as using an on-line testing environment represents innovative ways in which students can acquire and process new knowledge. This outcome shows that students genuinely value it as a form of enhancing their learning. Students acknowledge it as an important factor to obtain a good preparation. This matches with Gill & Greenhow (2008) who recognize that students are able to improve their performance in formative and summative assessments while they are engaged with the computer-assisted assessment assignments, especially by spending time studying the feedback.

The fact that technology can promote and facilitate communication was found important for students. They are very familiar in sharing information using social media. The fact that they can get support from other students is a feature that is also appreciated. Formative assessment has shown to be very useful for students when they can learn from peer comments and make further revisions before their final submission. Formative feedback implies a partnership and reciprocal relationships. It could be said that the relationship finishes when students produce a correct assignment, which shows that learning goals have been correctly addressed.

This can be summarized by citing Teo (2006), to include information relating to students' attitudes towards computers could be useful for various reasons ranging from curriculum design to teacher training. Given the monetary and opportunity costs involved, it is crucial for users and policy makers to be informed by research into the impact of computer attitudes on various aspects of IT use and implementation in the schools.

## 6.9 Conclusions

To incorporate useful electronic tools such as on-line testing environments to the teaching and learning processes has the aim of enriching the way in which students acquire knowledge. Therefore this demonstrate the usefulness of the technology and the form in which they would be more willing to use it. Technology enables them to obtain grades quicker, to support their learning, as well as enhancing on-line communication, and these are the characteristics that students value most. Thus, these insights should be taken into account to build punctual strategies to implement an on-line assessment environment. This is a matter for the educational managers who are able to create strategies to give an impulse. There is a need to build some effective schemes to convince instructors who have the power to persuade students to use technology in a broader manner. Students as digital natives are always more willing to use technological tools as long as they think that these are useful and enjoyable.

These findings will be very helpful for developing future tests. In the later chapters it is our belief that researching the factors allowing or hampering students the use of on-line assessment must continue in order to promote their use. It is even more relevant to investigate these particularly in the fields of mathematics, physics and engineering where technology can have a strong impact making its benefits even more evident.

## 6.10 Limitations and restrictions

Empirical results for both studies have been described. Performing these studies has provided us with the experience to make the next phase of this research. As lecturers' and students' scales were designed early on in this research, some items included did not sharply reflect the construct needed to be measured. This mainly occurred in the case of the first designed student's questionnaire.

For the next stage, making important changes to both scales bringing in finer details was considered. Also, the models and hypotheses have been restated in order to prove the most relevant factors implicated mainly in the context of our interest: mathematical subjects. This is particularly interesting, as at the moment, there have not been many previous studies.

Likewise, as part of the aims of this research was to extended this study to Mexico in order to compare results from both UK and Mexico. The new version of the survey was applied in the National Polytechnic Institute in Mexico which is the largest institute of public education in Mexico in fields of mathematics, physics and engineering. This is really interesting since information was obtained from very different contexts, cultures and economies. It is important to explore the major determinants in the use of on-line assessment and how these would be measured by their context. The following chapters will explain these issues.

# Chapter 7. Case Study: Teachers in the IPN-Mexico

## 7.1 Introduction

In order to understand the factors affecting the adoption of on-line assessment technologies in mathematical subjects from a different cultural perspective, we have extended our research to Mexico. We have decided to focus our research on the National Polytechnic Institute (IPN) particularly in the Interdisciplinary Professional Unity of Engineering, Social Sciences and Management (UPIICSA).

The IPN is the most important public technological higher education centre in Mexico. The Institute was founded on the premise of supporting the country's industrialization and national development process, as well as offering education opportunities to various social sectors, especially the less favoured financially. The IPN supports three educational sectors, technical high-school, undergraduate and postgraduate. For the period 2012-2013, the high-school level had an enrolment of 63,363. Comprising 61,513 students in 36 academic programs in a traditional setting and 1850 students in 5 academic programs in an on-line context.

In the IPN, a total of 103,712 students attended undergraduate education during 2012-2013 comprising 101,168 students in traditional settings in 26 schools offering 56 academics courses; 2,544 in on-line in 5 academic courses. Currently, the IPN has 149,409 technical high-school, undergraduate and graduate level students; 131,564 students at the virtual campus and in continuous education centres; 53,835 students in foreign-language units. In postgraduate education, the IPN has 20 scientific and technological research centres with 15,602 scholars and researchers (National Polytechnic Institute, 2015).

UPIICSA is one of the schools in the IPN. During 2013 the school had an academic staff of 845 teachers with 59% between 46 and 65 years old. The number of students enrolled was 12,274 in 5 academic courses: Industrial Management, Industrial

Engineering, Informatics, Computer Science and Transport Engineering.

UPIICSA does not yet have any strategies regarding the usage of technology for teaching and learning. Most of the teachers use technology for supporting their marking and management activities, but not as a tool for assessing students' academic performance in academic departments which include mathematics, and this represents an issue in performing this study.

A formal invitation to the Head of the school and the Head of the mathematics department was made in order to extend our research project. The latter suggested that the decision to take part in the project should be a unanimous decision taken by all the mathematics teaching staff. After some meetings, teachers explained that they were concerned with some issues. The first concern was regarding the bank of maths exercises. They mentioned that “the bank does not match their actual curricula”. In that case, they would prefer to develop their own bank. However, they explained that lack of time was their main barrier to cover a heavy academic assignment (teaching an average of 4 classes of 50 students every day).

Another concern was that the on-line platform does not consider the mathematical procedure that students have to perform in order to reach a result. They mentioned the platform only takes into account final typed results and that in their department, the written mathematical procedure is also considered in order to give a grade. On the other hand, they were also concerned whether the on-line assessment process would be consistent and fair. Technology is seen as a secondary teaching process that does not have the same validity as a paper-based assessment process. Another important concern is the belief that students cheat on paper-based exams; an on-line exam would foster this practice.

Summarizing their comments, they thought that it is not trivial to transfer traditional teaching learning processes into new ways, especially when it is a matter of teaching mathematics. Their main concern was the lack of time, either doing their own bank of mathematical exercises or spending time in learning how to use the on-line learning platform. The comments cited represent a sample of the common arguments that teachers had regarding the adoption of on-line teaching. These comments far, from favouring adoption, considered more the impediments.

We believe that one of the causes of these problems in this university is the lack of appropriate dissemination strategies that clearly show the benefits that teachers and students can obtain by using an on-line learning platform. It is worth mentioning that, although on-line teaching activities are already considered for academic rewards, these do not have enough value in order to be attractive for teachers.

It is clearly shown that as the IPN still does not have a culture of on-line based learning it was a challenge to involve teachers in our research. In fact, the project was achieved by the participation of 4 teachers motivated by the desire to support the project and not by the support of the department of mathematics.

Since the IPN has a different culture of on-line based learning it would be really interesting to compare the conclusions from the UK and from Mexico, especially for teachers.

To better understand how to propose some guidelines that can help to stimulate the usage of on-line tools for teaching and assessing mathematics we include the following factors into a causal model to help us understand what really matters to teachers when the on-line assessment technologies are involved.

## 7.2 Proposed Constructs

Latent variables such as perceived usefulness, perceived ease of use, computer-self efficacy, social influence, feedback, attitude and behavioural intention have already been explained in chapter 5. This chapter includes a new variable, availability of information technology (AIT) services. We have also extended the feedback construct to obtain finer grading of teachers' opinions.

In order to examine the impact of how external factors such as the accessibility of IT services may affect the attitude and usage intentions we have included the availability IT (AIT) construct. It is similar to what Thompson, Higgins & Howell (1991) and Venkatesh et al. (2003) call “technology and resource facilitating conditions” and Ajzen (1991) mention as “perceived behavioural control”.

In the context of PC use, Thompson et al. (1991) define facilitating conditions as “the provision of support for users may be one type of facilitating condition that can influence system utilization. By training users and assisting them when they encounter difficulties, some of the potential barriers to use are reduced or eliminated”. This is consistent with what Venkatesh (2000) explains “in the context of workplace technology use, specific issues such as the availability of support staff, which is an organizational response to help users overcome barriers and hurdles to technology use, especially in the early stages of learning and use”. In other words, the facilitating conditions include the factors in the environment that shape a person’s perception of ease or difficulty of performing a task (Teo, 2012). It embraces factors such as technical support (the provision of help-desks and on-line support services). Technical support has been cited as one of the important factors in the acceptance of technology for teaching and in user satisfaction (Williams, 2002), (Teo, 2012).

In fact, for teachers, technical support was ranked highly on the list of factors that affect teachers’ implementation of technology (Teo, 2009). Lim & Khine (2006) reveal in their study that teachers cited poor facilitating conditions (e.g. lack of access to computers, inadequate technical support given teachers) as barriers to ICT integration in the classroom (Teo, 2009).

The constructor “feedback” (FE) has been included in chapter 5. We defined this construct as “the lecturers' belief that they can enhance their teaching by using on-line feedback”, consequently, enriching students learning. We have explored the effects of this construct on lecturers' attitude and usage intention. The hypothesis was that feedback would have a direct positive effect on lecturers' attitude and behavioural intentions to use web-based testing. The results show that effects of feedback on attitude and behavioural intentions were significantly more influential in the latter. However, we have made important changes in this construct. It has been strengthened by including specific questions asking the context of our interest in teaching mathematics. In this model we also want to explore the effects of feedback on usefulness, attitude and usage intentions. All the constructs proposed were built taking into account the nomological network of TAM. Table 7.1 shows the constructs and indicators included in this model.

*Table 7.1. Distribution of the variables into each construct*

| Indicators | | Construct |
|---|---|---|
| PU1, PU2, PU3, PU4, PU5 | belongs to | PU |
| PEU1, PEU2, PEU3, PEU4 | belongs to | PEU |
| CSE1, CSE2, CSE3, CSE4 | belongs to | CSE |
| SI1, SI2, SI3, SI4 | belongs to | SI |
| AIT1, AIT2, AIT3 | belongs to | AIT |
| FE1, FE2, FE4, FE5, FE6, FE7, FE8 | belongs to | FE |
| AT1, AT2, AT3, AT4 | belongs to | AT |
| BI1, BI2, BI3, BI4, BI5, BI6, BI7 | belongs to | BI |

Eight constructs were included in the model. In turn, each construct is supported by their corresponding indicators. In order to bring more finer details, two new constructs were included.

## 7.3 Designing research models

*Proposing hypotheses*

Following the theoretical frameworks explained in chapter 5 and accomplishing with the requirements to build a valid model, the following hypotheses were established:

- *H1: states there is a causal relationship between perceived usefulness and attitude to use web-based assessment.*
- *H2: states there is a causal relationship between feedback and attitude to use web-based assessment.*
- *H3: states there is a causal relationship between feedback and behaviour intention to use web-based assessment.*
- *H4: states there is a causal relationship between feedback and perceived*

*usefulness to use web-based assessment.*

- *H5: states there is a causal relationship between perceived ease of use and perceived usefulness to use web-based assessment.*
- *H6: states there is a causal relationship between computer self-efficacy and perceived ease of use web-based assessment.*
- *H7: states there is a causal relationship between availability of IT services and perceived ease of use web-based assessment.*
- *H8: states there is a causal relationship between attitude and behavioural intention to use web-based assessment.*
- *H9: states there is a causal relationship between social influence and behavioural intention to use web-based assessment.*
- *H10: states there is a causal relationship between social influence and attitude to use web-based assessment.*

Figure 7.1 shows schematically the causal relationships proposed by the hypotheses within the teacher's model.

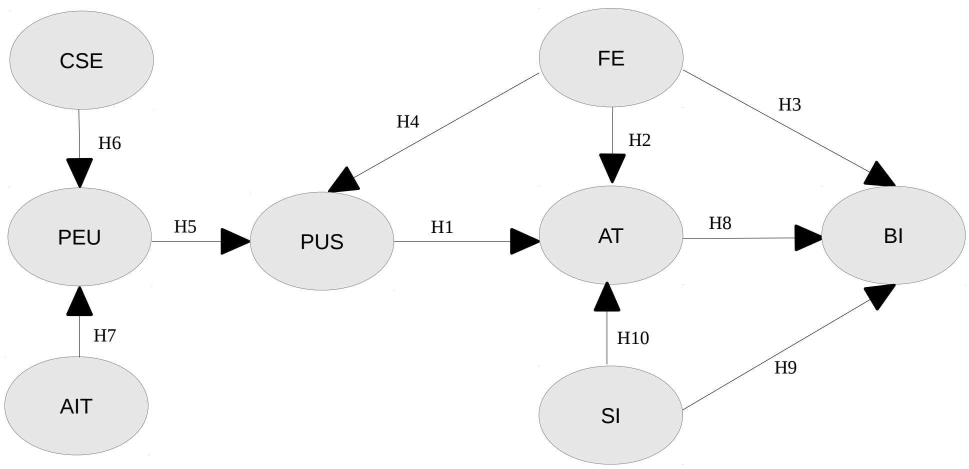


*Figure 7.1. A visual representation of the path model for PLS-SEM teacher's model showing the hypothetical relationships between latent variables*

# 7.4 Instrumentation

## 7.4.1 Developing the scale

Table 7.2 shows the new scale including the new constructs. All the constructs included are either new items or rewritten items which fit the context of this research, the educational on-line assessment for teaching mathematics. All items are assessed by 5-point-Likert scales ranging from 1 (strongly disagree) to 5 (strongly agree). The constructs and their respective indicators are listed below.

*Table 7.2. Empirical scale used to measure teachers' perceptions of mathematical on-line assessments*

| Construct/Indicator | Representa-tive study |
|---|---|
| **Perceived Usefulness (PUS)** | (Davis, 1989) |
| PUS1. I believe that on-line assessments can help support teaching in mathematical subjects. | |
| PUS2. I think that on-line assessments tools can help to enhance the teaching of mathematics. | |
| PUS3. I find on-line tests useful to assess my students in mathematical subjects. | |
| PUS4. On-line assessments are useful because they save time and cost. | |
| PUS5. I think that on-line assessments are a helpful part of teaching mathematics. | |
| **Perceived Ease of Use (PEU)** | (Davis, 1989) |
| PEU1. I find it easy to author on-line assessments containing mathematical material. | |
| PEU2. I find it straightforward to use on-line mathematical assessments to support my teaching. | |
| PEU3. It is straightforward to become skilful at using on-line assessments containing mathematics. | |
| PEU4. I think that electronic assessments for mathematical subjects are easy to use. | |
| **Computer Self-Efficacy (CSE)** | (Bandura, 1986) |
| CSE1. I find it easy to enhance my teaching by providing mathematical activities on computers. | |
| CSE2. I am able to use on-line assessments even if there is no one around to explain me how to use the system. | |
| CSE3. I feel confident that I have adequate skills to author on-line mathematical assessments. | |
| CSE4. I feel confident using electronic assessments in my teaching. | |
| **Social Influence (SI)** | (Fishbein & Ajzen, 1975) |
| SI1. Colleagues are positive about using on-line tests for mathematical subjects. | |

| | |
|---|---|
| SI2. My students want to do mathematical homework and tests using a web-based platform. | |
| SI3. In general, my university provides support when using on-line assessments. | |
| SI4. If I need support authoring on-line tests including mathematics, a colleague is available to help me | |
| **Availability IT Services (AIT)** | (Thompson et al., 1991) |
| AIT1. When I need help to learn to use on-line assessment specialised university staff is there to support me. | |
| AIT2. Internet speed at my university is fast enough to use an on-line learning environment. | |
| AIT3. My university has enough computing infrastructure to support on-line testing. | |
| **Feedback (FE)** | Self-developed |
| FE1. On-line feedback can provide students sufficient information to help them understand where they went wrong in a mathematical question. | |
| FE2. On-line feedback can help students to learn how to improve their work in mathematics based subjects. | |
| FE3. On-line feedback can help resolve doubts about mathematical material faster than traditional feedback. | |
| FE4. On-line assessments of mathematical subjects allows us to produce grades faster, so students know how well they are doing in the topic. | |
| FE5. I think that on-line feedback can be useful in understanding mathematical subjects. | |
| FE6. On-line feedback can help students determine their strengths and weaknesses in mathematics faster. | |
| FE7. I believe that timely on-line feedback can encourage students to do better in their studies. | |
| FE8. On-line assessment allows students to get on-line support from their peers more easily. | |
| **Attitude (AT)** | (Fishbein & Ajzen, 1975) |
| AT1. On-line tests that contain mathematics are useful. | |
| AT2. I would like to use on-line assessments with mathematical content. | |
| AT3. Mathematical on-line assessments are interesting. | |
| AT4. I find it more useful to do mathematical tests on-line than with pencil and paper. | |
| **Behaviour Intention (BI)** | (Davis, 1989) |
| BI1. I would like to be able to use on-line assessment to support my teaching of mathematics. | |
| BI2. I intend to use on-line assessments to assist my teaching of mathematics in the future. | |
| BI3. I would like to use on-line assessments in my mathematical subjects. | |
| BI4. I plan to use electronic tools for assessing mathematical problems in the future. | |

| | |
|---|---|
| BI5. I expect I will use on-line exams for assessing my students in the future. | |
| BI6. I plan to use electronic tools for assessing problems with mathematical content. | |
| BI7. I am committed to use on-line assignments to assess the mathematical skills of my students. | |

## 7.5 Data Collection Procedures

We gathered teachers' opinions form STEM subjects at UPIICSA by running an on-line survey. Three e-mails with an invitation to participate in the research were sent to teaching staff of mathematics, physics and chemistry departments. The survey was answered during the semester August to December 2014. A total of 31 full answers were gathered. 1 response was removed as it was considered as not an complete engaged response. 1 answer was included since it had more than 80% of the items answered following a case-wise replacement strategy.

Results of the experience as a teacher show that 58.1% of the teachers have been teaching for more than 15 years, 12.9% for 11-15 years,  16.1% for 6-10 years, 12.9% for 1-5 years, none of them has less than 1 year (0%). These results are similar to the results from lecturers at the University of Manchester, most of the teachers have a good experience teaching. Results of gender (29.3% female, 70.97% male) show similar results such as the lecturing staff being predominantly male which is also a common issue in STEM areas.

## 7.6 Data Analysis Techniques

The descriptive statistics (measures of mean, standard deviation and variance) for each item are shown in Table 7.3. All mean scores are in a range of 2.65 to 4.13. The standard deviations range from 0.70 to 1.10.

*Table 7.3. Descriptive statistics (mean, standard deviation and variance) per item for teacher's model*

| **Perceived Usefulness (PUS)** | **Mean ()** | **Std deviation ()** | **Variance ()** |
|---|---|---|---|
| PUS1. I believe that on-line assessments can help support teaching in mathematical subjects | 3.94 | 1.06 | 1.13 |
| PUS2. I think that on-line assessments tools can help to enhance the teaching of mathematics | 4.00 | 0.93 | 0.87 |
| PUS3. I find on-line tests useful to assess my students in mathematical subjects | 3.61 | 1.05 | 1.11 |
| PUS4. On-line assessments are useful because they save time and cost | 3.90 | 1.14 | 1.29 |
| PUS5. I think that on-line assessments are a helpful part of teaching mathematics | 3.81 | 1.05 | 1.10 |
| **Perceived Ease of Use (PEU)** | | | |
| PEU1. I find it easy to author on-line assessments containing mathematical material | 3.52 | 0.96 | 0.93 |
| PEU2. I find it straightforward to use on-line mathematical assessments to support my teaching | 3.42 | 1.12 | 1.25 |
| PEU3. It is straightforward to become skilful at using on-line assessments containing mathematics | 3.45 | 1.03 | 1.06 |
| PEU4. I think that electronic assessments for mathematical subjects are easy to use | 3.32 | 1.14 | 1.29 |
| **Computer Self-Efficacy (CSE)** | | | |
| CSE1. I find it easy to enhance my teaching by providing mathematical activities on computers | 3.97 | 0.98 | 0.97 |
| CSE2. I am able to use on-line assessments even if there is no one around to explain me how to use the system | 3.58 | 1.03 | 1.05 |
| CSE3. I feel confident that I have adequate skills to author on-line mathematical assessments | 3.55 | 1.09 | 1.19 |
| CSE4. I feel confident using electronic assessments in my teaching. | 3.52 | 1.26 | 1.59 |
| **Social Influence (SI)** | | | |
| SI1. Colleagues are positive about using on-line tests for mathematical subjects | 2.71 | 0.97 | 0.95 |
| SI2. My students want to do mathematical homework and tests using a web-based platform | 3.52 | 0.93 | 0.86 |
| SI3. In general, my university provides support when using on-line assessments | 2.87 | 1.06 | 1.12 |
| SI4. If I need support authoring on-line tests including mathematics, a colleague is available to help me | 2.65 | 1.11 | 1.24 |
| **Availability IT Services (AIT)** | | | |
| AIT1. When I need help to learn to use on-line assessment specialised university staff is there to support me | 2.81 | 1.05 | 1.10 |
| AIT2. Internet speed at my university is fast enough to use an on-line learning environment | 2.81 | 1.20 | 1.43 |
| AIT3. My university has enough computing infrastructure to support on-line testing | 2.97 | 1.17 | 1.37 |

| | | | |
|---|---|---|---|
| **Feedback (FE)** | | | |
| FE1. On-line feedback can provide students sufficient information to help them understand where they went wrong in a mathematical question | 3.90 | 1.08 | 1.16 |
| FE2. On-line feedback can help students to learn how to improve their work in mathematics based subjects | 4.06 | 0.93 | 0.86 |
| FE3. On-line feedback can help resolve doubts about mathematical material faster than traditional feedback | 3.19 | 1.30 | 1.70 |
| FE4. On-line assessments of mathematical subjects allows us to produce grades faster, so students know how well they are doing in the topic | 3.97 | 1.05 | 1.10 |
| FE5. I think that on-line feedback can be useful in understanding mathematical subjects | 4.13 | 0.96 | 0.92 |
| FE6. On-line feedback can help students determine their strengths and weaknesses in mathematics faster | 3.74 | 1.03 | 1.07 |
| FE7. I believe that timely on-line feedback can encourage students to do better in their studies | 3.94 | 0.93 | 0.86 |
| FE8. On-line assessment allows students to get on-line support from their peers more easily | 3.61 | 1.05 | 1.11 |
| **Attitude (AT)** | | | |
| AT1. On-line tests that contain mathematics are useful | 3.87 | 1.12 | 1.25 |
| AT2. I would like to use on-line assessments with mathematical content | 3.77 | 1.18 | 1.38 |
| AT3. Mathematical on-line assessments are interesting | 3.68 | 1.08 | 1.16 |
| AT4. I find it more useful to do mathematical tests on-line than with pencil and paper | 3.29 | 1.04 | 1.08 |
| **Behaviour Intention (BI)** | | | |
| BI1. I would like to be able to use on-line assessment to support my teaching of mathematics | 4.00 | 0.93 | 0.87 |
| BI2. I intend to use on-line assessments to assist my teaching of mathematics in the future | 3.84 | 1.00 | 1.01 |
| BI3. I would like to use on-line assessments in my mathematical subjects | 3.94 | 1.00 | 1.00 |
| BI4. I plan to use electronic tools for assessing mathematical problems in the future | 3.42 | 1.03 | 1.05 |
| BI5. I expect I will use on-line exams for assessing my students in the future | 3.71 | 0.97 | 0.95 |
| BI6. I plan to use electronic tools for assessing problems with mathematical content | 3.52 | 1.06 | 1.13 |
| BI7. I am committed to use on-line assignments to assess the mathematical skills of my students | 3.87 | 0.96 | 0.92 |

# 7.7 Findings and Results

## 7.7.1 Evaluation of Measurement Model (Model Quality)

We test the quality of the measurement model. Satisfactory results to ensure the reliability and validity of the constructs measured were obtained and therefore provide support for their inclusion in the path model.

Table 7.4 shows values for composite reliabilities ranged from .77 to .91 exceeding the minimum requirement. The AVE values ranged from .64 to .89 also exceeding the cut-off recommended. Values for Cronbach's alpha are as good. Therefore, all values obtained confirm convergent validity.

*Table 7.4. Construct and convergent validity coefficients. Discriminant validity coefficients: bivariate correlations. In the main diagonal, the squared-root AVE of each construct*

| | AVE | Composite Reliability | Cronbachs' Alpha | AIT | AT | BI | CSE | FE | PEU | PUS | SI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **AIT** | 0.84 | 0.91 | 0.82 | **0.92** | | | | | | | |
| **AT** | 0.79 | 0.92 | 0.87 | 0.04 | **0.89** | | | | | | |
| **BI** | 0.89 | 0.98 | 0.97 | 0.04 | 0.81 | **0.94** | | | | | |
| **CSE** | 0.77 | 0.93 | 0.9 | 0.26 | 0.72 | 0.71 | **0.88** | | | | |
| **FE** | 0.79 | 0.94 | 0.91 | 0.23 | 0.77 | 0.68 | 0.71 | **0.89** | | | |
| **PEU** | 0.86 | 0.95 | 0.92 | 0.33 | 0.58 | 0.58 | 0.85 | 0.64 | **0.93** | | |
| **PUS** | 0.78 | 0.94 | 0.91 | 0.1 | 0.84 | 0.73 | 0.74 | 0.88 | 0.67 | **0.88** | |
| **SI** | 0.64 | 0.77 | 0.48 | 0.4 | 0.64 | 0.59 | 0.49 | 0.58 | 0.42 | 0.6 | **0.80** |

Discriminant validity is also shown in Table 7.4, which displays that a construct shares more variance with its associated indicators than with any other construct. Therefore, Table 7.4 provides evidence for reliability and validity of the model estimators which serve as criteria to assess the model's predictive capabilities.

Figure 7.2 shows schematically the paths coefficients, $R^2$ values for the endogenous latent variables and factorial loadings (indicator validity coefficients).

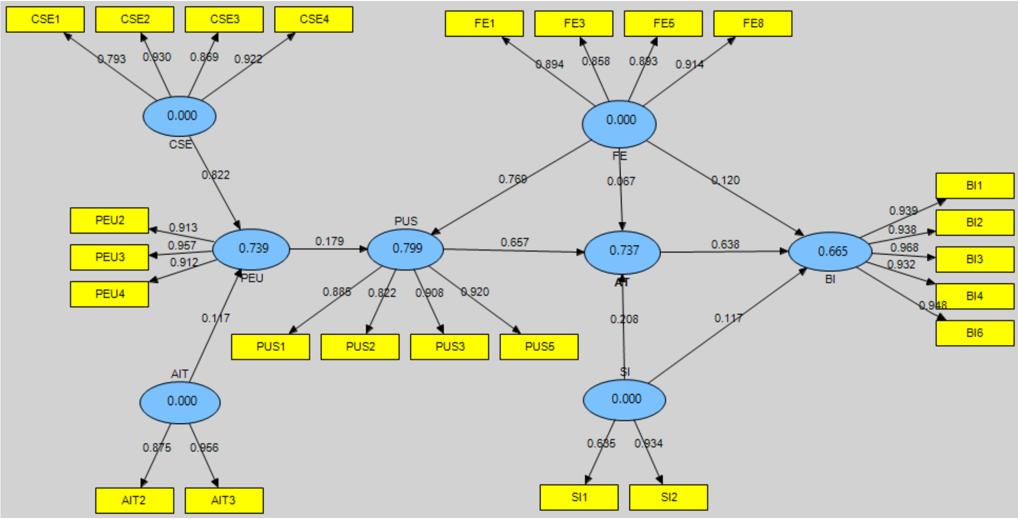

*Figure 7.2. Indicators loadings coefficients, path coefficients and R-squared values for endogenous constructs into the measurement model*

All loadings coefficients are higher than the cut-off value suggested to accomplish indicator reliability, including loading indicators of AT construct (AT1 0.89, AT2 0.89, AT3 0.89) not shown in Figure 7.2.

## 7.7.2 Evaluation of the Structural Equation Model

### 7.7.2.1 Testing Hypotheses

In order to test the statistical significance of the relationships in the model a bootstrap procedure with 200 re-samples was used. The results for the hypotheses are summarized in Table 7.5.

*Table 7.5. Results of hypotheses test (summary of path coefficients and significance levels)*

| Hypotheses | Hypotheses path | Path coefficients | T-values significance | Accept/reject |
|---|---|---|---|---|
| H7 | AIT → PEU | 0.12 | 0.94* | Reject |
| H8 | AT → BI | 0.64 | 2.76*** | Accept |
| H6 | CSE → PEU | 0.82 | 7.63*** | Accept |
| H2 | FE → AT | 0.07 | 0.31* | Reject |
| H3 | FE → BI | 0.12 | 0.53* | Reject |
| H4 | FE → PUS | 0.77 | 9.29*** | Accept |
| H5 | PEU → PUS | 0.18 | 1.67* | Reject |
| H1 | PUS → AT | 0.66 | 2.94*** | Accept |
| H10 | SI → AT | 0.21 | 1.55* | Reject |
| H9 | SI → BI | 0.12 | 0.8* | Reject |

*Critical t-values for a two-tailed test are: <1.96 (p > .05*), 1.96 (p = .05**), and 2.58 (p = .001***).*

Figure 7.3 represents schematically the significance of the paths of the measurement and structural models.

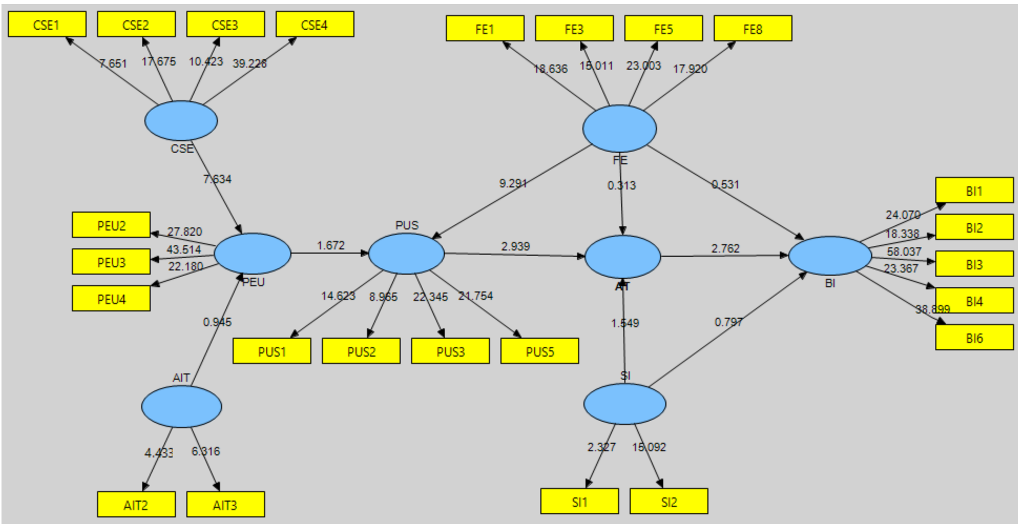

*Figure 7.3. Significance of indicators loadings coefficients and path coefficients for the structural model*

Figure 7.3 also shows the statistical significance of the indicator variables into the underlying construct displaying that the most significant items are (BI3 ← BI = 58.04), BI3. “I would like to use on-line assessments in my mathematical subjects”. (PEU3 ← PEU = 43.51) PEU3. “It is straightforward to become skilful at using on-line assessments containing mathematics”. (CSE4 ← CSE = 39.23) CSE4. “I feel confident using electronic assessments in my teaching”. (BI6 ← BI = 38.9) BI6. “I plan to use electronic tools for assessing problems with mathematical content". Taking into account these outcomes, we can conclude that teachers have a good attitude toward mathematical on-line exams (BI3) and are able to manage these activities easily (PEU3) and have confidence to perform the task (CSE4). In conclusion, they are willing to plan using mathematical on-line tests.

Looking into more details, we also analysed the statistical significance of indirect effects (analysis also known as total effects). Results of this analysis shows that the most significant path is perceived usefulness on behaviour intention (PUS → BI = 2.1). This indirect relationship confirms the strong impact of usefulness on usage intentions.

#### 7.7.2.2 Hypotheses examination

Hypothesis 1 predicts a causal relationship between perceived usefulness and attitude. The results shows that perceived usefulness had a significant impact on attitude with a path coefficient of (β: .66, $p = .001$) this hypothesis is strongly supported. The causal relationship between feedback and attitude (H2) is not accepted (β: .07, $p > .05$). H3 shows that the causal relationship between feedback and behaviour intention is not significant (β: .12, $p > .05$), therefore this hypotheses not supported. H4 predicts a causal relationship between feedback and perceived usefulness the path coefficient shows a strong relationship (β: .77, $p = .001$) therefore this hypotheses is accepted.

H5 that shows the causal relationship between perceived ease of use and perceived usefulness (β: .18, $p > .05$) is not significant, this hypothesis is not supported. Hypothesis H6 predicts the causal relationship between computer-self efficacy and perceived ease of use, the results (β: .82, $p = .001$) shows that this hypothesis is strongly supported. The causal relationship between availability of information technology and perceived ease of use (H7) is not supported (β: .12, $p > .05$). H8 predicts the causal relationship between attitude and behavioural intention (β: .64, $p = .001$) the result show that this hypothesis is strongly accepted. The relationship between social influence and behavioural intention (H9) is not accepted at (β: .12, $p > .05$). The causal relationship between social influence and attitude (H10) is not supported at (β: .21, $p > .05$).

## 7.8 Discussion

We see that a substantial amount of the variance in the endogenous construct "behavioural intention" is explained by the three constructs; attitude, feedback and social influence. All these constructs together explain 67% of the variance in behavioural intention, indicating a high overall $R^2$ value. These outcomes suggest that the attitude construct has the strongest influence on intention to use on-line assessments. On-line feedback and social influence have only moderate effects. In other words, teachers' feelings, opinions about performing the target behaviour (using technology) or object (technology) contribute significantly more to their usage intention than on-line feedback and opinions from colleagues and managers. Contrary to the conclusions of Davis, Bagozzi & Warshaw (1989) who found that attitude was only modest in predicting technology acceptance and that individuals may use a technology even if they did not have a positive attitude toward technology per se, as long as it is perceived to be useful and/or easy to use in ways that enhance their productivity (Teo & Noyes, 2011), we found that attitude matters. This finding supports the research of Teo (2006) who suggests that a positive feeling toward the use of technology is associated with factors that foster continued and sustained use of technology.

In order to obtain a favourable teachers' attitude it is important to take into account their opinions. This might mean that by the use of technology, teachers will expect to receive a gain (an estimated value or a utility) which will be perceived by evaluating its results (consequences). This suggests that showing clear evidence of the benefits of doing on-line assessment might convince teachers of its use. No matter how sophisticated and powerful the technology is, the extent to which it is implemented depends on teachers having a positive attitude toward it (Huang & Liaw, 2005), (Teo, 2012).

Findings also reveal that the on-line feedback construct has a strong influence on perceived usefulness, but only moderate effects on attitude and usage intentions. It clearly shows that teachers really value providing on-line feedback, they find it useful.

The fact that on-line feedback shows limited effects on attitude and usage intention reveal that feedback per se is not a factor that triggers teachers' positive feelings (willingness) towards adopting on-line assessment and therefore to the intention to use it, the usefulness is essential.

On the other hand, we see that the social influence construct does not have an effect on attitude and use intentions. This result is similar to findings in chapter 5. It may mean that teachers perceived the usage of technology in a voluntary way and do not feel pressure from managers or colleagues to use the on-line environment. This confirms that teachers in this study participated in a voluntary way. Social influence has a significant effect on intention in a mandatory setting, but not in a voluntary setting (Venkatesh & Davis 2000), (Venkatesh et al., 2003).

Results show that the attitude construct is strongly influenced by perceived usefulness. This can mean that in order to foster a positive teachers' attitude it is important to show tangible benefits from the usage of on-line testing. This result matches with previous studies from the TAM, showing positive effects of perceived usefulness on attitude (Ming Chi, 2010), (Gong et al., 2004), (Legris et al., 2003)

In turn, perceived usefulness is strongly determined by perceived ease of use. This explains 80% of the variance in perceived usefulness. Similar to the result for lecturers at the University of Manchester, when teachers consider it is easy to use an on-line assessment systems; it is more likely that they consider it also to be useful.

Perceived ease of use is predicted by computer self-efficacy and availability of information technology. Both constructs explain 74% of the variance in perceived ease of use. This can mean if teachers feel skilled enough to use an on-line environment, it is more likely that they think they are capable of using it. This means that a strong sense of teachers computer self-efficacy can affect the extent and the way technology can be used in everyday instructional practice, changing significantly both the teacher's and the student's roles. Teacher computer self-efficacy might determine to a considerable extent the ability to use on-line technologies as an important educational tool (Paraskeva, Bouta, & Papagianni, 2008). In this sense, self-efficacy has been shown to predict behavioural change with different types of participants in various settings.

Furthermore, the availability of information technology construct shows only a moderate effect meaning that the causal relationship between availability of technology and ease of use is not significant. This reveals that when teachers have IT resources, (technical support such as help-desks, on-line support services and guidance by the IT staff) that does not mean they perceive the IT as easy to use. Providing technical assistance does not make it easier to use the on-line environment. This may indicate that there is no technical assistance strategy or that this is not effective.

## 7.9 Conclusions

We see that attitude and on-line feedback (through attitude) are the most significant factors in predicting intention to use on-line testing for teachers in Mexico. The result from on-line feedback matches with that obtained in chapter 5. This confirms the significant role that feedback plays in on-line mathematical assessments. Benefits from on-line feedback such as the fact that it can provide students with sufficient information to help them understand where they went wrong in a mathematical question; can help resolve doubts about mathematical material faster than traditional feedback; that it can help students determine their strengths and weaknesses in mathematics faster and that it allows students to get on-line support from their peers easier, are important factors for teachers to enhance students' learning in mathematics.

The fact that attitude has a significant effect on usage intention reveals that teachers' opinions and feelings are important regarding the usage of technology for applying mathematical tests. Moreover, the results reveal that teachers feel very confident using computers, are highly capable of doing mathematical on-line testing which has a strong impact on predicting ease of use. Our findings also show that perceived usefulness has strong indirect effects on usage intention. This demonstrates that perceived usefulness plays a crucial role in predicting teachers' intentions to use mathematical on-line assessments.

Despite the difficulties encountered in performing this research, teachers recognize the usefulness of both on-line feedback and assessments which is reflected in their positive attitude and usage intention. These issues reveal an urgent necessity to propose helpful strategies to implement technology in teaching, learning and assessment. Therefore, we believe that these can really help to propose strategies to implement on-line assessment environments which enhances students' performance in mathematical subjects.

# Chapter 8. Case Study: Students in the IPN-Mexico

## 8.1 Introduction

This chapter reports our conclusions on students at IPN - UPIICSA in Mexico. In order to perform this research, the stack/moodle course “first-year students” including the bank of mathematics exercises was translated into Spanish. The web-based platform was also set up in Spanish. The on-line survey was also translated. 132 students from industrial management, industrial engineering and transport engineering were enrolled in the on-line mathematical course. The students had a window time of 45 days for practising each mathematics section three times; after the last trial the web-based platform used the highest score. After answering the on-line test the students were asked to answer an on-line survey. We obtained 35 answers.

## 8.2 Proposed Constructs

Similar to teachers' reported in the previous chapter, we include additional constructs to the model discussed in section 7.2 using our initial experience. Some new items were added in the scale, others were slightly rewritten. We include facilitating conditions and enjoyment constructs.

Facilitating conditions includes factors in the environment that shape a person’s perception of ease or difficulty of performing a task (Teo, 2012). It contains factors such as technical support (the provision of help-desks and on-line support services). This has been cited as one of the important factors in the acceptance of technology for teaching and in user satisfaction (Williams, 2002), (Teo, 2012). Similar to the chapter 7 we include it as availability of IT services (AIT).

We include enjoyment (playfulness) construct since it is considered as a significant reason to be interested in using a specific system. Although research on it has emerged in an organizational context, (Davis, 1989), (Davis, 1993) it looks to be generally relevant. In an educational setting, this construct has been shown to have a positive influence on usage intentions. Venkatesh (2000) explains that playfulness is related to intrinsic motivation or the "perceptions of pleasure and satisfaction from performing the behaviour". This also matches with the work of Davis et al. (1992) who argue that intrinsic motivation refers to the performance of an activity for no apparent reward other than the process of performing the activity per se. In his research Venkatesh (2000) explains that playfulness is a construct that is system-independent, supporting this by the fact that users who are more "playful" with computer technologies in general enjoy using a new system just for the sake of using it. Playfulness is still expected to be a relevant factor even when the systems perform a rather boring task, since it still involves exploration and discovery. Venkatesh states that, from a theoretical standpoint, higher levels of playfulness will lead to a lower perceptions of effort. In the context of this research, this would mean that students could perceive less time on effort invested if they perceive an electronic educational task (mathematical exercise) as enjoyable. If we consider that students are seen as "digital natives", accustomed to using technologies for fun, it should be interesting to explore the effects of this construct on attitude and intention to use. This could be a direct determinant to enhance learning of

On the other hand, Davis et al. (1992) found that perceived enjoyment and perceived usefulness mediate the influence of perceived ease of use on intention, explaining that "while usefulness will once again emerge as a major determinant of intentions to use a computer in the workplace, enjoyment will explain significant variance in usage intentions beyond that accounted by usefulness alone."

In the context of computer-based assessment, Terzis & Economides (2011) include perceived playfulness showing a positive effect of this construct on behavioural intention to use. Therefore, we will explore the effects of this construct on attitude and behavioural intentions.

Chapter 6 includes the constructor feedback. We have observed positive influence of feedback on attitude. Measuring this construct is important since it can help enhance students learning performance, which in turn, prompts students to achieve their goals. It is also considered for fostering their intrinsic motivation within the learning setting.

In this model we want to investigate students' opinion regarding on-line feedback in more detail. With this aim, we consider separating this construct into two. The first construct captures the perception of getting on-line feedback. We have asked for aspects such as the on-line feedback being precise, clear, helpful and timely. We have defined this construct before. Therefore, it is embedded as received feedback (RF) or “the students' belief that the support and feedback they receive from the on-line platform will enhance their learning”.

The second construct explores students' perception of getting on-line feedback versus face-to-face feedback. In others words, we want to obtain students' opinions regarding the differences of getting each type of feedback. Therefore, we have defined comparative feedback (CF) as “the students' perception of receiving on-line feedback when it is compared with traditional feedback”. The following Table 8.1 shows all constructs included in the model and the distribution of indicators into each construct.

*Table 8.1. Constructs and number of indicators proposed to the student's model*

| **Indicators** | | **Constructs** |
|---|---|---|
| PU1, PU2, PU3, PU4 | belongs to | Perceived Usefulness (PU) |
| PEU1, PEU2, PEU3 | belongs to | Perceived Ease of Use (PEU) |
| CSE1, CSE2, CSE3, CSE4 | belongs to | Computer-Self Efficacy (CSE) |
| SI1, SI2, SI3, SI4 | belongs to | Social Influence (SI) |
| AIT1, AIT2, AIT3 | belongs to | Availability IT Services (AIT) |
| EN1, EN2, EN3 | belongs to | Enjoyment (EN) |
| RF1, RF2, RF3 | belongs to | Received Feedback (RF) |
| CF1, CF2, CF3 | belongs to | Comparative Feedback (CF) |
| AT1, AT2, AT3, AT4 | belongs to | Attitude (AT) |
| BI1, BI2, BI3 | belongs to | Behavioural Intention (BI) |

## 8.3 Data collection procedures

### 8.3.1 Procedure for the questionnaire

Students were using the on-line platform during September 2014 and two weeks of October 2014. We obtained 35 responses to the questionnaire, 4 were removed as not fully complete, 1 was removed as it was considered as not an engaged response (st dev= 0), resulting in a dataset of 30 responses.

The gender balance was, 53.33% are female and 46.67% are male. The result shows most of the students surveyed as female. Regarding the distribution by school, Table 8.2 is shown below.

*Table 8.2. Distribution of students' responses showed by school*

| Course | Responses | Percentage |
|---|---|---|
| Industrial engineering | 12 | 40.00% |
| Informatics | 0 | 0.00% |
| Transport engineering | 2 | 6.67% |
| Computer science | 0 | 0.00% |
| Industrial management | 16 | 53.33% |

There are some issues worthy of mention. As discussed in the previous chapter four teachers were willing to participate in the research with a total of eight classes (two of transport engineering, four of industrial engineering, two of industrial management). However, only two teachers were fully engaged with the project resulting in a total of three groups (one of industrial engineering, one of transport engineering, one of industrial management). Then, we obtained two groups from STEM subjects and one from social sciences. Nevertheless, it seems that students from social sciences were more willing to participate in the project by answering the on-line survey. This is an important bias in our results.

## 8.4 Data analysis techniques

*Descriptive statistics.* The values shown in Table 8.3 gives insights about students' feelings; these show the favourable or unfavourable students' attitude. The values can help us to obtain preliminary results about predicting data.

*Table 8.3. Descriptive statistics (mean, standard deviation and variance) per item for student's model*

| **Perceived Usefulness (PUS)** | **Mean ()** | **Std. Deviation ()** | **Variance ()** |
|---|---|---|---|
| PUS1. I find on-line tests useful to support my learning of mathematical subjects | 3.90 | 0.89 | 0.78 |
| PUS2. Doing on-line tests enhance my mathematical knowledge | 3.80 | 0.96 | 0.92 |
| PUS3. On-line assessments help me to understand mathematical topics better | 3.20 | 0.96 | 0.92 |
| PUS4. I find it useful that I can answer mathematical on-line test at any time and at a place of my choice | 3.47 | 1.17 | 1.36 |
| **Perceived Ease of Use (PEU)** | | | |
| PEU1. My interaction with the systems providing mathematical on-line assessments is clear and understandable | 3.67 | 0.76 | 0.58 |
| PEU2. I find it straightforward to use on-line assessments to support my learning of mathematics | 3.27 | 1.02 | 1.03 |
| PEU3. It is straightforward to become skilful at using on-line assessments of mathematical subjects | 3.43 | 0.82 | 0.67 |
| **Computer Self-Efficacy (CSE)** | | | |
| CSE1. I feel confident doing mathematical on-line assessment | 3.33 | 0.80 | 0.64 |
| CSE2. I feel comfortable using mathematical on-line assessments on my own | 3.40 | 0.81 | 0.66 |
| CSE3. I am able to use on-line assessments of mathematics even if there is no one around to explain me how to use the system | 3.20 | 1.03 | 1.06 |
| CSE4. In general, I feel confident doing on-line assessment | 3.20 | 0.85 | 0.72 |
| **Social Influence (SI)** | | | |
| SI1. My lecturer expects me to do on-line tests that contain mathematics | 3.47 | 1.04 | 1.09 |
| SI2. Classmates are positive about the use of mathematical on-line assessments | 3.20 | 0.93 | 0.86 |
| SI3. Discussions in social media such as Facebook are supportive of using mathematical on-line assessment | 3.27 | 1.02 | 1.03 |
| SI4. In general, my university provides support when using web-based assessments for mathematical subjects | 3.20 | 0.93 | 0.86 |
| **Availability IT Services (AIT)** | | | |
| AIT1. When I need help to learn to use on-line assessment specialised university staff is there to support me | 3.20 | 0.85 | 0.72 |
| AIT2. Internet speed at my university is fast enough to use an on-line learning environment | 2.70 | 1.09 | 1.18 |
| AIT3. My university has enough computing infrastructure to support on-line testing | 3.07 | 1.05 | 1.10 |
| **Enjoyment (EN)** | | | |

| | | | |
|---|---|---|---|
| EN1. I enjoy using on-line assessments that require mathematical answers | 3.23 | 0.86 | 0.74 |
| EN2. Using on-line assessment of mathematics stimulates my curiosity | 3.57 | 1.01 | 1.01 |
| EN3. Doing mathematical on-line tests is enjoyable | 3.13 | 1.07 | 1.15 |
| **Received Feedback (RF)** | | | |
| RF1. The on-line feedback returned with my mathematical exercises and exams were fair & balanced | 3.50 | 1.11 | 1.22 |
| RF2. On-line feedback gave me enough information on where I went wrong in mathematical exercises and exams | 3.47 | 1.22 | 1.50 |
| RF3. From my on-line feedback, I learnt how to improve my work for mathematical subjects | 3.50 | 0.97 | 0.95 |
| **Comparative Feedback (CF)** | | | |
| CF1. On-line feedback helps me resolve faster doubts about the mathematical material than traditional feedback | 3.20 | 0.96 | 0.92 |
| CF2. Electronic assessments of my mathematical subjects allow me to get grades faster, so I know if I am doing well in my topic | 3.60 | 1.07 | 1.15 |
| CF3. On-line feedback helps me better understand mathematical subjects | 3.27 | 0.98 | 0.96 |
| **Attitude (AT)** | | | |
| AT1. I like doing on-line test and exams in subjects that require mathematical answers | 3.20 | 1.00 | 0.99 |
| AT2. I look forward to those aspects of my learning of mathematics that require me to use on-line assessment | 3.13 | 0.90 | 0.81 |
| AT3. On-line tests that contain mathematics are useful | 3.60 | 0.89 | 0.80 |
| AT4. I prefer answering mathematical tests on-line than with pencil and paper | 2.97 | 1.07 | 1.14 |
| **Behaviour Intention (BI)** | | | |
| BI1. I will use electronic tools to support my learning of mathematical subjects in the future | 3.87 | 0.90 | 0.81 |
| BI2. I would like to continue my use of on-line assessment to support my learning of mathematics | 3.47 | 1.07 | 1.15 |
| BI3. All things considered, I expect to continue doing on-line test or exams to assists my learning of mathematics | 3.63 | 1.03 | 1.07 |

## 8.5 Testing Models and Relationships

### 8.5.1 Evaluation of the Measurement Model (Model Quality)

This model deals with the same issues as the lecturer's model. As we obtained a limited sample size, a model can not be built that would test all hypotheses proposed in the beginning. Therefore, we test two simplified models. Model A tests the direct effects of both constructs perceived and comparative feedback on perceived usefulness, attitude and usage intentions. Model B tests the direct effects of perceived and comparative feedback on enjoyment, attitude and usage intentions. In order to obtain reliable measurement models, we test measures of quality: convergent, discriminant and indicator validity in both models.

From Model A and B we find that the effects of perceived and comparative feedback on perceived usefulness, attitude and usage intentions are limited, being only significant for the path of comparative feedback on attitude, which gave us a measure of the effects that feedback exerts on usefulness and enjoyment. As a result, we built a model that tests the impact of both feedback constructs together on perceived usefulness and enjoyment. This model measures the direct effects of perceived and comparative feedback on perceived usefulness and enjoyment and the indirect effects on attitude and usage intentions. This model is shown in Figure 8.7.

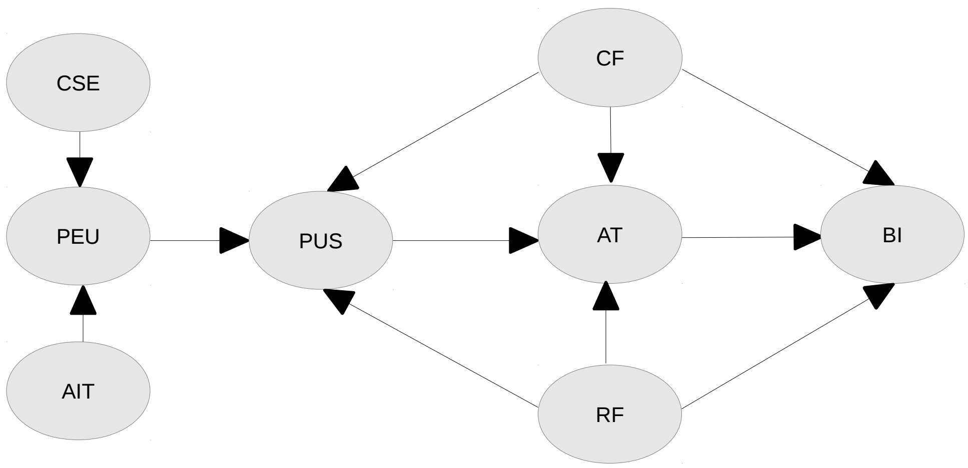


*Figure 8.1. Model A analyses the direct impact of comparative feedback and received on-line feedback on perceived usefulness, attitude and behaviour intentions*

In order to test measures of quality in Model A, the indicators CSE3, AT1 were removed. We include measures of quality (construct, discriminant and indicator validity) in Table 8.4. Figure 8.2 summarises results from the measurement model showing that all indicator loadings are higher than the common threshold criterion of .7 to reach indicator reliability including the values for AT (AT2 0.92, AT3 0.84, AT4 0.86) that are not shown. Figure 8.3 summarises measures for the structural model showing that relationships with t-values higher than 1.96 are considered as significant paths.

*Table 8.4. Model A. Construct and convergent validity coefficients. Discriminant validity coefficients: bivariate correlations. In the main diagonal, the squared-root AVE of each construct*

| | AVE | Composite Reliability | Cronbach's Alpha | AIT | AT | BI | CF | CSE | PEU | PUS | RF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **AIT** | 0.76 | 0.91 | 0.85 | **0.87** | | | | | | | |
| **AT** | 0.76 | 0.91 | 0.84 | 0.4 | **0.87** | | | | | | |
| **BI** | 0.76 | 0.9 | 0.83 | 0.3 | 0.86 | **0.87** | | | | | |
| **CF** | 0.71 | 0.88 | 0.8 | 0.48 | 0.84 | 0.74 | **0.84** | | | | |
| **CSE** | 0.63 | 0.83 | 0.71 | 0.4 | 0.56 | 0.34 | 0.49 | **0.79** | | | |
| **PEU** | 0.64 | 0.84 | 0.72 | 0.33 | 0.64 | 0.5 | 0.53 | 0.55 | **0.80** | | |
| **PUS** | 0.62 | 0.87 | 0.8 | 0.18 | 0.67 | 0.62 | 0.53 | 0.43 | 0.77 | **0.79** | |
| **RF** | 0.77 | 0.91 | 0.85 | 0.37 | 0.69 | 0.55 | 0.84 | 0.32 | 0.59 | 0.56 | **0.88** |

The AVE ranges from .62 to .77, confirming convergent validity and implicitly, content validity in all constructs. Composite reliabilities and Cronbach's alpha demonstrate reliability for all constructs.

All constructs accomplish the threshold required for empirical studies to reach indicator validity. In order to assess the structural model, we examine coefficients of determination showing that a large amount of variance (.75) in the latent construct

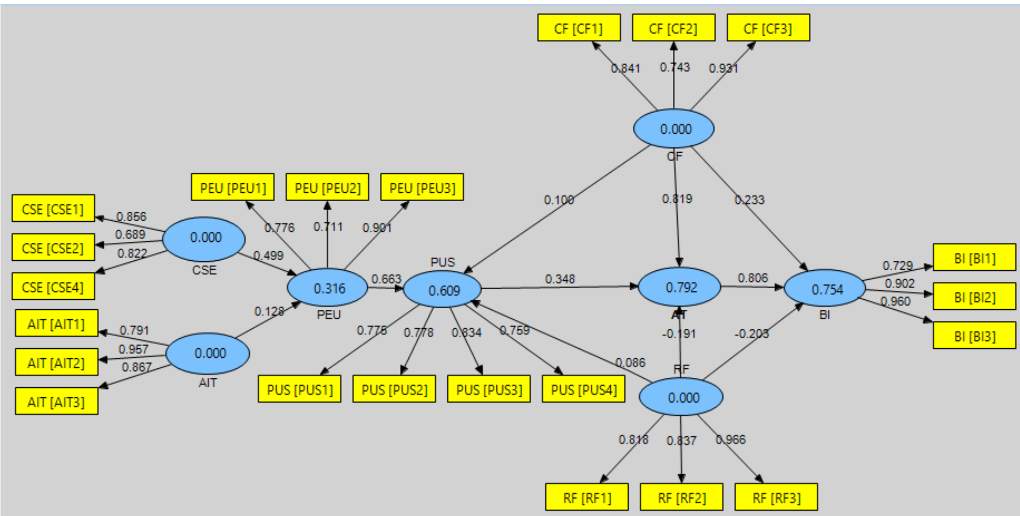


*Figure 8.2. Model A. Indicators loadings coefficients, path coefficients and R-squared values for endogenous constructs into the measurement model*

usage intention is explained by three constructs: comparative feedback, received feedback and attitude.

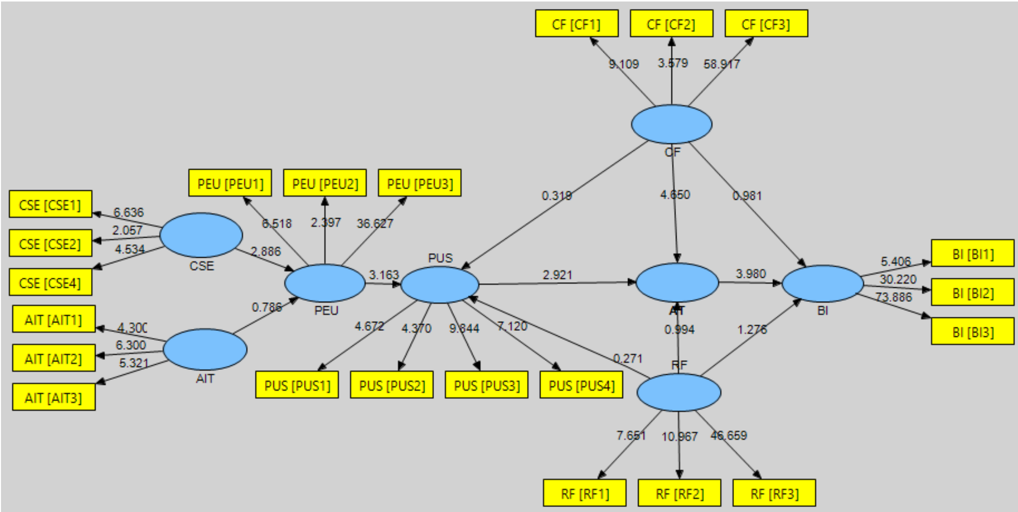


*Figure 8.3. Model A. Statistical significance of indicators loadings coefficients and path coefficients used for testing hypotheses in the structural model*

The results also show that the constructs BI3, CF3, RF3 and PEU3 are the most significant in this model. In others words, the items, CF3 "on-line feedback helps me better understand mathematical subjects", BI3 "all things considered, I expect to continue doing on-line test or exams to assists my learning of mathematics", RF3 "from my on-line feedback, I learnt how to improve my work for mathematical subjects" and PEU3 "It is straightforward to become skilful at using on-line assessments of mathematical subjects" are relevant according to students' opinion.

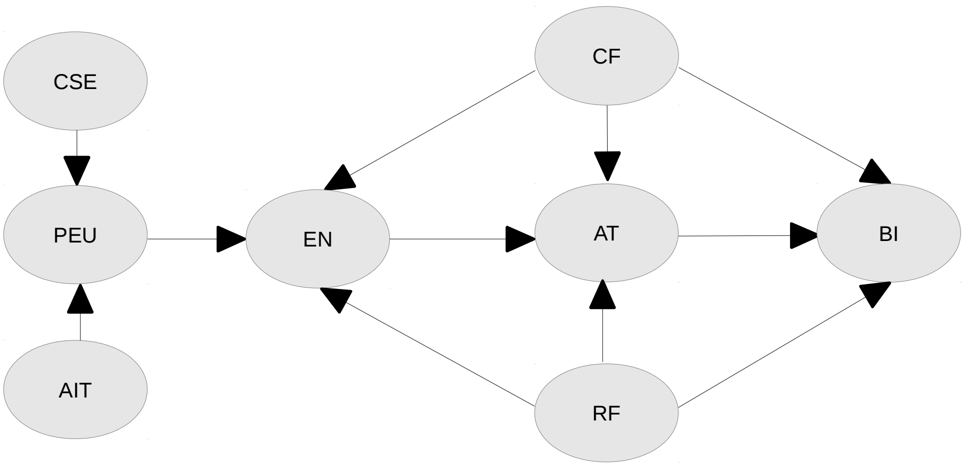


*Figure 8.4. Model B analyses the direct impact of comparative feedback and received on-line feedback on enjoyment, attitude and behaviour intentions*

In order to test measures of quality for Model B, we have followed the same procedure described for Model A. Table 8.5 summarises measures of quality. Figure

8.5 summarises the results from the measurement model showing that all indicator loadings are higher than the common threshold criterion of .7 to reach indicator reliability including the values for AT (AT2 0.92, AT3 0.84, AT4 0.85) that are not shown in the figure. Figure 8.6 summarises structural models emphasising the relevant paths.

*Table 8.5. Model B. Construct and convergent validity coefficients. Discriminant validity coefficients: bivariate correlations. In the main diagonal, the squared-root AVE of each construct*

| | AVE | Composite Reliability | Cronbach's Alpha | AIT | AT | BI | CF | CSE | EN | PEU | RF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **AIT** | 0.77 | 0.91 | 0.85 | **0.88** | | | | | | | |
| **AT** | 0.76 | 0.91 | 0.84 | 0.4 | **0.87** | | | | | | |
| **BI** | 0.76 | 0.9 | 0.83 | 0.3 | 0.86 | **0.87** | | | | | |
| **CF** | 0.71 | 0.88 | 0.8 | 0.49 | 0.84 | 0.72 | **0.84** | | | | |
| **CSE** | 0.63 | 0.83 | 0.71 | 0.41 | 0.57 | 0.35 | 0.5 | **0.79** | | | |
| **EN** | 0.81 | 0.93 | 0.88 | 0.42 | 0.84 | 0.63 | 0.82 | 0.65 | **0.90** | | |
| **PEU** | 0.63 | 0.84 | 0.72 | 0.35 | 0.65 | 0.52 | 0.55 | 0.57 | 0.69 | **0.79** | |
| **RF** | 0.77 | 0.91 | 0.85 | 0.37 | 0.69 | 0.55 | 0.84 | 0.33 | 0.7 | 0.6 | **0.88** |

The AVE ranged from .63 to .81, confirming convergent validity and implicitly, content validity in all constructs. Composite reliabilities and Cronbach's alpha demonstrated reliability for all constructs.

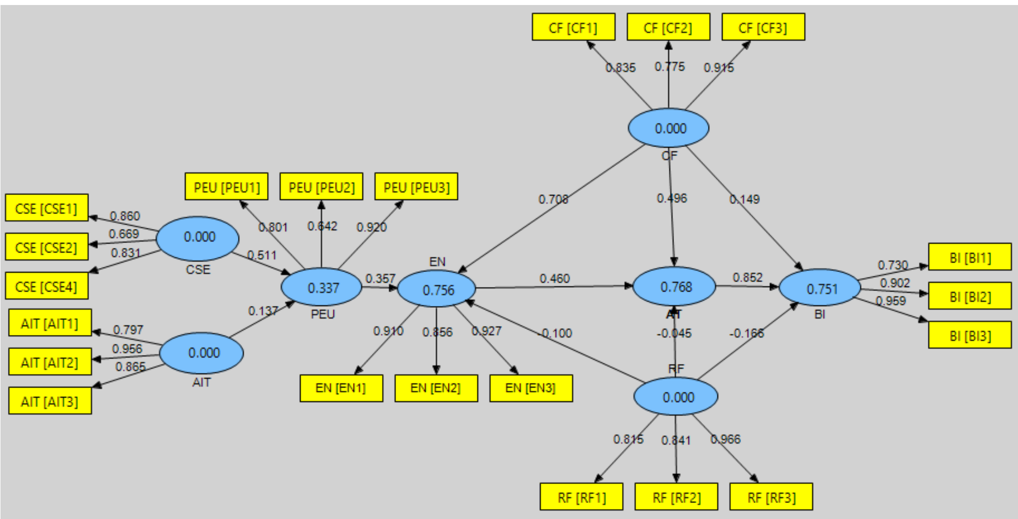


*Figure 8.5. Model B. Indicators loadings coefficients, path coefficients and R-squared values for endogenous constructs into the measurement model*

Figure 8.5 also shows $R^2$ values. These results show that substantial variance (.75) in the latent construct usage intention is explained by the three constructs: comparative feedback, received feedback and attitude. This shows that the research model predicts usage intentions efficiently.

Figure 8.6 shows that received feedback does not have influence on enjoyment, attitude or behaviour intentions. Comparative feedback has effects on enjoyment and attitude. It is clear that enjoyment mediates the effects of perceived ease of use on attitude. This effect is also reflected on behavioural intentions. Therefore, it clearly shows that enjoyment has an influence on attitude and usage intentions.

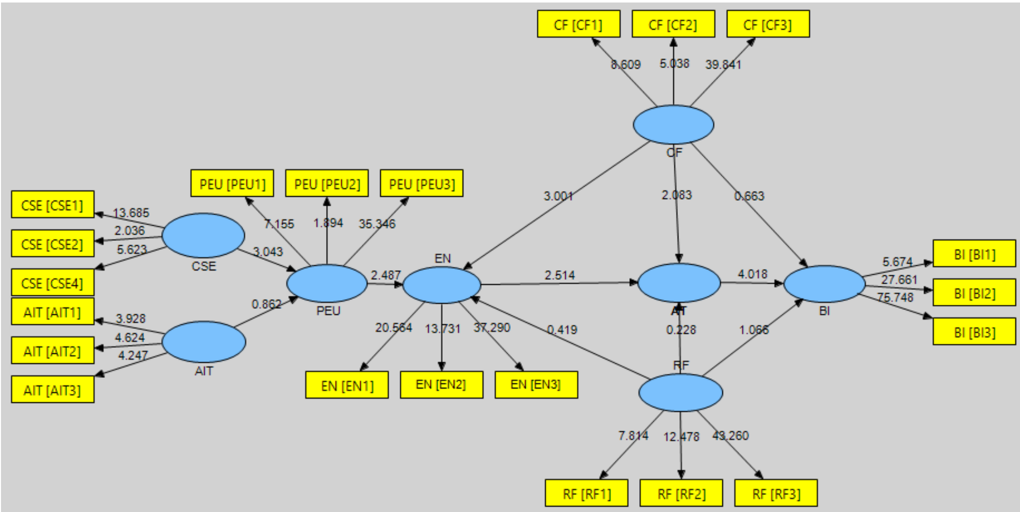


*Figure 8.6. Model B. Statistical significance of indicators loadings coefficients and path coefficients for the structural model*

The findings showed that the constructs BI3, CF3, EN3, RF3 and PEU3 were the most important for Model B. The constructs BI3, CF3, RF3 and PEU3 were also significant in Model A. This model shows the item EN3 "doing mathematical on-line tests is enjoyable" as being important for both models.

## 8.6 Designing research models

### 8.6.1 Proposing hypotheses

Both constructs related to feedback show limited effects on attitude and behaviour intention. In order to test further effects we decided to prove the direct effects of both constructs on perceived usefulness and enjoyment in the same model proving the effects of received feedback and comparative feedback on attitude and usage intentions through perceived usefulness and enjoyment, in other words, testing the indirect effects.

When an inner model is estimated, a causal path is not restricted to direct relationships. It can also examine total effects; that is, the sum of direct and indirect effects. Interpretation of total effects is particularly useful in studies with the objective of exploring the differential impact of different driver constructs on a criterion construct via several mediating variables (Albers, 2010), (Hair, Sarstedt, Pieper, & Ringle, 2012).

Based on the theoretical frameworks mentioned before, we proposed the following hypotheses:

- *H1: states there is a causal relationship between attitude and behavioural intention to use web-based assessment.*
- *H2: states there is a causal relationship between perceived usefulness behaviour intention to use on-line assessment.*
- *H3: states there is a causal relationship between perceived usefulness and attitude to use web-based assessment.*
- *H4: states there is a causal relationship between enjoyment and attitude to use on-line assessment.*
- *H5: states there is a causal relationship between enjoyment and behaviour intention to use web-based assessment.*
- *H6: states there is a causal relationship between comparative feedback and perceived usefulness to use on-line assessment.*
- *H7: states there is a causal relationship between comparative feedback and enjoyment to use web-based assessment.*
- *H8: states there is a causal relationship between received feedback and perceived usefulness to use on-line assessment.*
- *H9: states there is a causal relationship between received feedback and enjoyment to use web-based assessment.*
- *H10: states there is a causal relationship between perceived ease of use and perceived usefulness on-line assessment.*
- *H11: states there is a causal relationship between perceived ease of use and enjoyment web-based assessment.*
- *H12: states there is a causal relationship between availability of information*

*technology and perceived ease of use on-line assessment.*

- *H13: states there is a causal relationship between computer self-efficacy and perceived ease of use web-based assessment.*

Figure 8.7 represents schematic hypotheses proposed for the model.

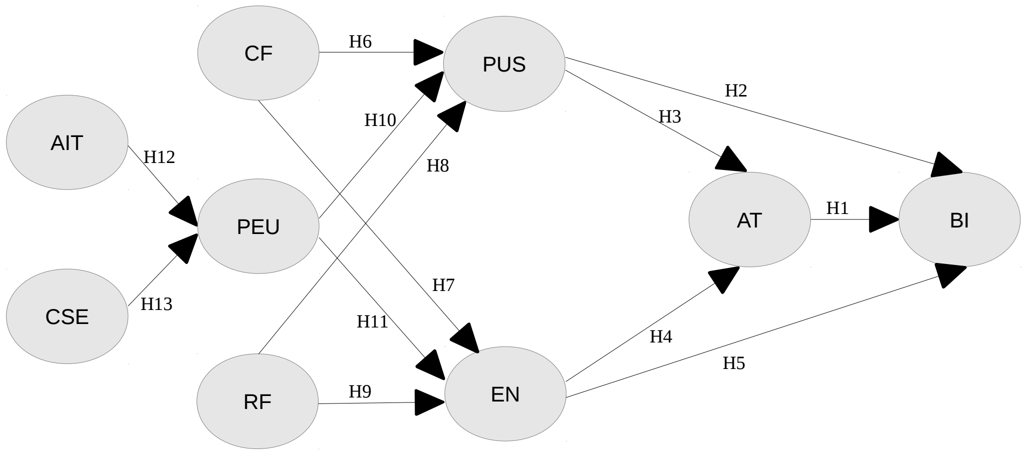


*Figure 8.7. A visual representation that display the hypotheses and variables relationships that are examined in the PLS-SEM model*

## 8.7 Findings and Results

### 8.7.1 Evaluation of the Measurement Model (Model Quality)

In order to test model quality we obtain measures of convergent, discriminant and indicator reliabilities. These values are shown in Table 8.6.

Figure 8.8 shows values from the measurement model demonstrating that all indicator loadings are higher than the common threshold criterion of .7 to reach indicator reliability including the values for AT (AT2 0.92, AT3 0.84, AT4 0.86) that are not shown. Figure 8.9 summarises measures for the structural model, t-values higher that 1.96 are considered as a significant path.

*Table 8.6. Construct and convergent validity coefficient. Discriminant validity coefficients: bivariate correlations. In the main diagonal, the squared-root AVE of each construct*

| | AVE | Composite Reliability | Cronbachs' Alpha | AIT | AT | BI | CF | CSE | EN | PEU | PUS | RF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **AIT** | 0.76 | 0.91 | 0.85 | **0.87** | | | | | | | | |
| **AT** | 0.76 | 0.91 | 0.84 | 0.4 | **0.87** | | | | | | | |
| **BI** | 0.76 | 0.9 | 0.83 | 0.3 | 0.86 | **0.87** | | | | | | |
| **CF** | 0.71 | 0.88 | 0.8 | 0.49 | 0.83 | 0.72 | **0.84** | | | | | |
| **CSE** | 0.63 | 0.83 | 0.71 | 0.41 | 0.57 | 0.35 | 0.5 | **0.79** | | | | |
| **EN** | 0.81 | 0.93 | 0.88 | 0.42 | 0.83 | 0.64 | 0.82 | 0.65 | **0.90** | | | |
| **PEU** | 0.64 | 0.84 | 0.72 | 0.33 | 0.64 | 0.51 | 0.54 | 0.55 | 0.67 | **0.80** | | |
| **PUS** | 0.62 | 0.87 | 0.8 | 0.18 | 0.67 | 0.62 | 0.5 | 0.43 | 0.49 | 0.76 | **0.79** | |
| **RF** | 0.77 | 0.91 | 0.85 | 0.37 | 0.7 | 0.55 | 0.84 | 0.33 | 0.71 | 0.6 | 0.56 | **0.88** |

The AVE values range from .62 to .81 exceeding the recommended cut-off. Likewise, values for Cronbach's alpha can be considered as good. Therefore, all values obtained confirm convergent validity.

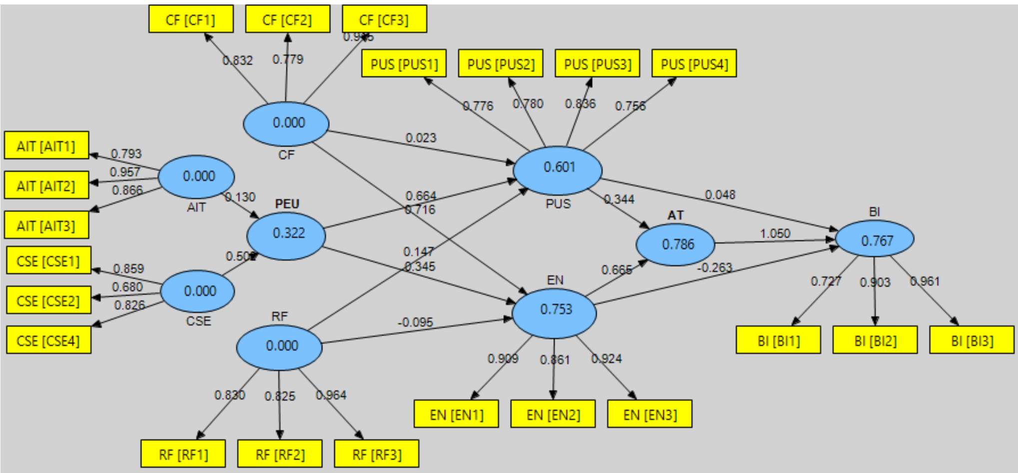

*Figure 8.8. Indicators loadings coefficients, path coefficients and R-squared values for endogenous constructs into the measurement model*

Figure 8.8 also shows the $R^2$ values for the endogenous constructs. The values for behavioural intention, $R^2$ = .77, attitude, $R^2$ = .79 and enjoyment, $R^2$ = .75, are considered as substantial. Perceived usefulness shows a moderate coefficient, $R^2$ = .60, while that for perceived ease of use displays a weak coefficient, $R^2$ = .32.

The three measurement models thus include indicators that appropriately measure their constructs to accomplish the threshold required to reach model quality. The three proposals proved to be reliable and consistent models. However, the real issue is the sample size, as it was not large enough to prove the hypotheses established from the beginning.

## 8.7.2 Evaluation of the Structural Equation Model

### 8.7.2.1 Testing Hypotheses

In order to test the statistical significance of the relationships in the model a bootstrap procedure with 200 re-samples was used. The results for the hypotheses are summarized in Table 8.7.

*Table 8.7. Results of hypotheses test (summary of path coefficients and significance levels)*

| Hypotheses | Hypotheses path | Path coefficients | T-values significance | Accept/reject |
|---|---|---|---|---|
| H12 | AIT → PEU | 0.13 | 0.81 | Reject |
| H1 | AT → BI | 1.05 | 5.46*** | Accept |
| H7 | CF → EN | 0.72 | 3.03*** | Accept |
| H6 | CF → PUS | 0.02 | 0.08 | Reject |
| H13 | CSE → PEU | 0.5 | 3.31*** | Accept |
| H4 | EN → AT | 0.67 | 7.32*** | Accept |
| H5 | EN → BI | -0.26 | 1.44 | Reject |
| H11 | PEU → EN | 0.35 | 2.48** | Accept |
| H10 | PEU → PUS | 0.66 | 3.28*** | Accept |
| H3 | PUS → AT | 0.34 | 3.95*** | Accept |
| H2 | PUS → BI | 0.05 | 0.34 | Reject |
| H8 | RF → EN | -0.1 | 0.4 | Reject |
| H9 | RF → PUS | 0.15 | 0.5 | Reject |

*Critical t-values for a two-tailed test are: <1.96 (p > .05*), 1.96 (p = .05**), and 2.58 (p = .001***).*

Figure 8.9 displays measures of statistical significance for the structural model showing that relationships with t-values higher than 1.96 are considered as significant paths.

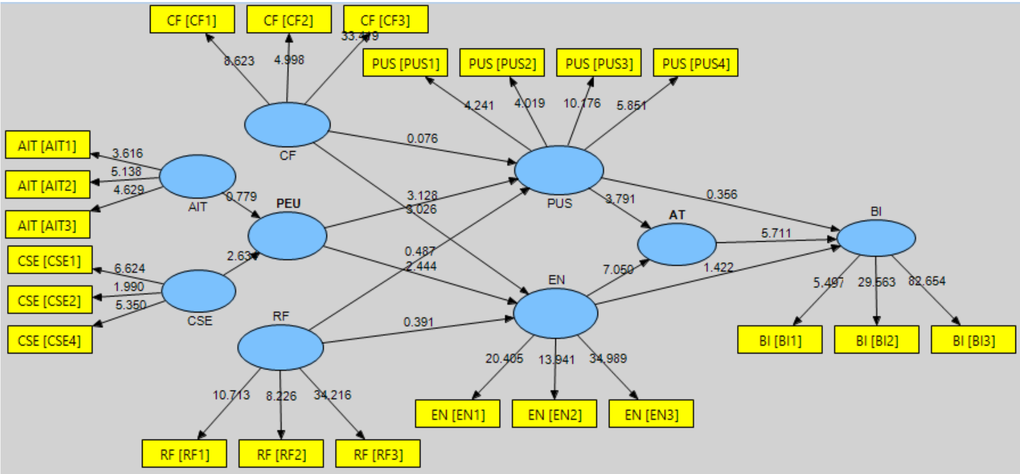

*Figure 8.9. Significance of indicator loading coefficients and path coefficients for the structural model*

In order to further examine significant relationships, we analyse the indirect effects of the constructs. This analysis serves to determine which are the most influential constructs of the model. Table 8.8 shows that the construct comparative feedback has strong impact on attitude and usage intentions. Computer self-efficacy has an important impact on attitude and perceived usefulness. Perceived ease of use influences in an important way attitude and behaviour intention. We can infer that these three constructs are relevant to students in Mexico.

*Table 8.8. Testing hypotheses: by examining direct and indirect effects*

| Path | Coefficients | T-Statistics (\|O/STERR\|) |
|---|---|---|
| AIT → AT | 0.06 | 0.82 |
| AIT → BI | 0.05 | 0.85 |
| AIT → EN | 0.04 | 0.66 |
| AIT → PEU | 0.13 | 0.81 |
| AIT → PUS | 0.09 | 0.88 |
| AT → BI | 1.05 | 5.46 |
| **CF → AT** | **0.48** | **2.87** |
| **CF → BI** | **0.32** | **2.12** |
| CF → EN | 0.72 | 3.03 |
| CF → PUS | 0.02 | 0.08 |
| **CSE → AT** | **0.23** | **2.14** |
| CSE → BI | 0.21 | 1.85 |
| CSE → EN | 0.17 | 1.77 |
| CSE → PEU | 0.5 | 3.31 |
| **CSE → PUS** | **0.33** | **1.98** |

| | | |
|---|---|---|
| EN → AT | 0.67 | 7.32 |
| EN → BI | 0.44 | 2.84 |
| **PEU → AT** | **0.46** | **3.68** |
| **PEU → BI** | **0.42** | **3.08** |
| PEU → EN | 0.34 | 2.48 |
| PEU → PUS | 0.66 | 3.28 |
| PUS → AT | 0.34 | 3.95 |
| PUS → BI | 0.41 | 2.71 |
| RF → AT | -0.01 | 0.08 |
| RF → BI | 0.02 | 0.14 |
| RF → EN | -0.09 | 0.4 |
| RF → PUS | 0.15 | 0.5 |

Results show that the indicator variables BI3, CF3, RF3, EN3, and PEU3 are the most significant for the model which also matches with the results from Models A and B. Furthermore, the item BI2 "I would like to continue my use of on-line assessment to support my learning of mathematics" is shown to be significant for the model. This item reveals that students in Mexico wish to continue taking mathematical on-line exams and they enjoy the learning experience.

#### 8.7.2.2 Hypotheses Examination

Hypothesis 1 predicts a causal relationship between attitude and behaviour intention. The results show that attitude has a significant impact on usage intentions with a path coefficient of (β: 1.05, $p = .001$), therefore this hypothesis is strongly supported. The causal relationship between perceived usefulness and behaviour intention (H2) is not accepted at (β: .05, $p = .01$). H3 shows that perceived usefulness is a significant determinant of attitude (β: .34 $p = .001$); therefore the hypotheses is strongly accepted. H4 predicts a causal relationship between enjoyment and attitude (β: .67, $p = .001$), this hypotheses is also accepted.

The causal relationship between enjoyment and behaviour intention (H5) is not significant at (β: -.26, $p = .001$). H6 states a causal relationship between comparative feedback and perceived usefulness, the outcome (β: .02, $p = .05$) shows that this hypothesis is not accepted. The causal relationship between comparative feedback and enjoyment (H7) is strongly accepted at (β: .72, $p = .001$). H8 states a causal relationship between received feedback and enjoyment, the result (β: -.01, p > .05) shows that this hypothesis is not accepted. H9 predicts a relationship between received feedback and perceived usefulness (β: .15, p > .05); this hypothesis was not supported by the data.

Hypothesis 10 predicts a causal relationship between perceived ease of use and perceived usefulness. The results show that perceived ease of use exerts a significance impact on perceived usefulness at (β: .66, p = .001); therefore this hypothesis is again supported. The causal relationship between perceived ease of use and enjoyment (H11) is accepted at (β: .35, p = .001). H12 shows that availability of information technology is not significantly determinant of perceived ease of use (β: .13 p = .05), therefore the hypotheses is not accepted. H13 predicts a causal relationship between computer self-efficacy and perceived ease of use, the result of (β: .05, p = .001) confirms that this hypotheses is readily accepted.

## 8.8 Discussion

The results indicate that the model explained 77% of the variance in behaviour intention. Similarly, 79% of the variance in attitude, 76% of the variance in enjoyment, 60% of the variance in perceived usefulness and 32% of the variance in perceived ease of use were explained by the related antecedent constructs.

Findings show that the direct effects of both constructs perceived usefulness and enjoyment on behaviour intention are not significant, only attitude predicts usage intention. It clearly shows that students' attitude strongly influences their usage intentions, even more than usefulness and enjoyment constructs.

On the other hand, results also show strong direct effects of perceived usefulness (PUS -> AT = 3.95) and enjoyment (EN -> AT = 7.32) on attitude. This result also contributes to the significant indirect effects on usage intentions (perceived usefulness through attitude on usage intentions) (EN -> BI 2.84), (PUS -> BI = 2.71) which is also proved by comparing the direct effects of both perceived usefulness and enjoyment on behaviour intention (not including attitude) (EN -> BI = 2.74), (PUS -> BI = 2.68). This clearly means that in order to obtain a positive students' willingness to do on-line assessment it is necessary to take into account their attitude. This shows that not just the usefulness of the technologies or the enjoyment matter in order to get a favourable usage intention. In other words, students' opinions weighs when favourable intentions to use an on-line assessments is the issue. This result is similar to that reported by Terzis & Economides (2011), who found that perceived playfulness has a positive direct effect on behavioural intention to use.

Comparative feedback has a positive effect on enjoyment but not on perceived usefulness. As this construct captures student's perception of getting on-line feedback versus face-to-face feedback, this result could mean that students find receiving on-line feedback more enjoyable than useful.

The outcome also shows that received feedback does not exert direct influence on either enjoyment or perceived usefulness, not even indirectly on attitude and usage intentions. Contrary to the findings shown in chapter 4 in which we have observed a positive influence of on-line feedback on attitude, this result shows that receiving on-line feedback does not make a strong impact to motivate students to use the system.

Perceived ease of use predicts positively enjoyment and perceived usefulness (PEU -> EN = 2.48), (PEU -> PUS = 3.28) which also contribute to the strong positive indirect effects on attitude and usage intentions (PEU -> AT = 3.68), (PEU -> BI = 3.08). This result is consistent with Davis et al. (1992) who found that perceived enjoyment and perceived usefulness mediated the influence of perceived ease of use on intention, explaining that “while usefulness will once again emerge as a major determinant of intentions to use a computer in the workplace, enjoyment will explain significant variance in usage intentions beyond that accounted for by usefulness alone.”

Computer self-efficacy and availability of information technology explains 32% of the variance in perceived ease of use. The fact that computer self-efficacy has a significant positive influence on perceived ease of use means that students who feel skilled doing an on-line testing are more likely to find it is easy to use which can mean that students with a strong sense of self-efficacy are more open to new ideas and they are more willing to experiment with new learning methods.

The “availability of information technology” construct has no effect on perceived ease of use which can mean the availability of IT resources, or the technical support provided (help-desks, on-line support services) and the guidance by the IT staff do not reflect positively; it does not help students to perceive that it is easier to use the on-line environment. This might reveal that there are no technical assistance strategies or that they are not effective.

## 8.9 Conclusions

Chapter 3 has already explained the advantages of integrating on-line feedback in teaching practices. Immediate on-line feedback is an essential feature of formative assessment having positive effects in enhancing students learning performance, activating their intrinsic motivation (Dreher et al., 2011) and therefore achieving their goals (Whitelock & Brasher, 2006). In this chapter we proved the effects (direct and indirect) of on-line feedback in students' attitude and usage intentions.

Our results show that receiving on-line feedback does not have a positive impact on students' perception of usefulness or enjoyment of doing on-line testing which in turn does not predict attitude or usage intentions. The students' experience of receiving on-line feedback in comparison with face-to-face feedback shows that this is more enjoyable than useful. It positively predicts the enjoyment factor but not the usefulness factor. The indirect effects show a positive influence on attitude and usage intentions meaning that comparative feedback influences positively on students' attitude and intentions to use on-line assessments.

Computer self-efficacy shows that students feel capable using on-line testing which has a positive impact on perceived ease of use. The problem of obtaining technological resources does not have any effect on predicting ease of use. Managers can invest in technological resources but results show that it does not mean that students perceive them as easy to use.

Attitude has the strongest influence on usage intentions which means that student feelings and opinions matter. The fact of weak effects of perceived usefulness and enjoyment on usage intention when these are compared with their effects through attitude demonstrates that attitude is a more important determinant for the intention to use on-line testing.

# Chapter 9. Case Study: Second Study for Students in the University of Manchester

## 9.1 Introduction

This chapter reports findings from a second study of students of the Faculty of Engineering & Physical Sciences at the University of Manchester. The aim of this research is to collect data which can be directly compared to that from students at the IPN in Mexico. We ran the same on-line test including the mathematics exercises using a similar learning strategy. Students also responded to the same on-line survey as students in Mexico. After that, we ran the same model testing the constructs. All the students in the case study came from STEM subjects.

## 9.2 Proposed Constructs

We include the same constructs from the student's model in Mexico as is shown in Table 9.1.

*Table 9.1. Constructs and number of indicators proposed to the student's model*

| Indicators | | Constructs |
|---|---|---|
| PU1, PU2, PU3, PU4 | belongs to | Perceived Usefulness (PU) |
| PEU1, PEU2, PEU3 | belongs to | Perceived Ease of Use (PEU) |
| CSE1, CSE2, CSE3, CSE4 | belongs to | Computer-Self Efficacy (CSE) |
| SI1, SI2, SI3, SI4 | belongs to | Social Influence (SI) |
| AIT1, AIT2, AIT3 | belongs to | Availability IT Services (AIT) |
| EN1, EN2, EN3 | belongs to | Enjoyment (EN) |
| RF1, RF2, RF3 | belongs to | Received Feedback (RF) |
| CF1, CF2, CF3 | belongs to | Comparative Feedback (CF) |
| AT1, AT2, AT3, AT4 | belongs to | Attitude (AT) |
| BI1, BI2, BI3 | belongs to | Behavioural Intention (BI) |

## 9.3 Data collection procedures

Students used the on-line platform during September and October 2014. Then, an on-line survey was used to gather students' responses. We obtained 121 full responses. The gender balance is, 25.62% (31) female, 74.38% (90) male. The school distribution is seen in Table 9.2 shown below.

*Table 9.2. Distribution of students' responses showed by school*

| School of | Responses | Percentage |
|---|---|---|
| Chemical Engineering & Analytical Science | 20 | 16.53% |
| Chemistry | 3 | 2.48% |
| Computer Science | 0 | 0.00% |
| Earth, Atmospheric & Environmental Sciences | 2 | 1.65% |
| Electrical and Electronic Engineering | 14 | 11.57% |
| Materials | 0 | 0.00% |
| Mathematics | 6 | 4.96% |
| Mechanical, Aerospace and Civil Engineering | 11 | 9.09% |
| Physics & Astronomy | 65 | 53.72% |

## 9.4 Data analysis techniques

Table 9.3 includes values of (mean, standard deviation and variance) of all items in the test. The values of the items use a Likert scale from (5 = strongly agree) to (1 = strongly disagree).

*Table 9.3. Descriptive statistics (mean, standard deviation and variance) of the items for student's model*

| **Perceived Usefulness (PUS)** | **Mean ()** | **Std. Deviati-on ()** | **Varian-ce ()** |
|---|---|---|---|
| PUS1. I find on-line tests useful to support my learning of mathematical subjects | 3.95 | 0.87 | 0.76 |
| PUS2. Doing on-line tests enhance my mathematical knowledge | 3.53 | 0.98 | 0.97 |
| PUS3. On-line assessments help me to understand mathematical topics better | 3.50 | 1.00 | 1.00 |
| PUS4. I find it useful that I can answer mathematical on-line test at any time and at a place of my choice | 4.43 | 0.81 | 0.66 |
| **Perceived Ease of Use (PEU)** | | | |
| PEU1. My interaction with the systems providing mathematical on-line assessments is clear and understandable | 3.58 | 1.00 | 1.00 |

| | | | |
|---|---|---|---|
| PEU2. I find it straightforward to use on-line assessments to support my learning of mathematics | 3.76 | 0.95 | 0.90 |
| PEU3. It is straightforward to become skilful at using on-line assessments of mathematical subjects | 3.83 | 0.84 | 0.71 |
| **Computer Self-Efficacy (CSE)** | | | |
| CSE1. I feel confident doing mathematical on-line assessment | 4.03 | 0.95 | 0.90 |
| CSE2. I feel comfortable using mathematical on-line assessments on my own | 4.23 | 0.80 | 0.65 |
| CSE3. I am able to use on-line assessments of mathematics even if there is no one around to explain me how to use the system | 4.18 | 0.77 | 0.60 |
| CSE4. In general, I feel confident doing on-line assessment | 4.17 | 0.78 | 0.61 |
| **Social Influence (SI)** | | | |
| SI1. My lecturer expects me to do on-line tests that contain mathematics | 3.83 | 1.23 | 1.52 |
| SI2. Classmates are positive about the use of mathematical on-line assessments | 3.29 | 0.99 | 0.97 |
| SI3. Discussions in social media such as Facebook are supportive of using mathematical on-line assessment | 2.96 | 0.97 | 0.94 |
| SI4. In general, my university provides support when using web-based assessments for mathematical subjects | 3.48 | 0.95 | 0.90 |
| **Availability IT Services (AIT)** | | | |
| AIT1. When I need help to learn to use on-line assessment specialised university staff is there to support me | 3.11 | 0.96 | 0.91 |
| AIT2. Internet speed at my university is fast enough to use an on-line learning environment | 4.40 | 0.79 | 0.62 |
| AIT3. My university has enough computing infrastructure to support on-line testing | 4.22 | 0.94 | 0.89 |
| **Enjoyment (EN)** | | | |
| EN1. I enjoy using on-line assessments that require mathematical answers | 3.35 | 1.16 | 1.35 |
| EN2. Using on-line assessment of mathematics stimulates my curiosity | 3.04 | 1.08 | 1.16 |
| EN3. Doing mathematical on-line tests is enjoyable | 2.99 | 1.11 | 1.24 |
| **Received Feedback (RF)** | | | |
| RF1. The on-line feedback returned with my mathematical exercises and exams were fair & balanced | 3.77 | 0.83 | 0.70 |
| RF2. On-line feedback gave me enough information on where I went wrong in mathematical exercises and exams | 3.40 | 1.11 | 1.22 |
| RF3. From my on-line feedback, I learnt how to improve my work for mathematical subjects | 3.60 | 1.06 | 1.13 |
| **Comparative Feedback (CF)** | | | |
| CF1. On-line feedback helps me resolve faster doubts about the mathematical material than traditional feedback | 3.21 | 1.22 | 1.49 |

| | | | |
|---|---|---|---|
| CF2. Electronic assessments of my mathematical subjects allow me to get grades faster, so I know if I am doing well in my topic | 4.11 | 0.91 | 0.83 |
| CF3. On-line feedback helps me better understand mathematical subjects | 3.49 | 0.91 | 0.84 |
| **Attitude (AT)** | | | |
| AT1. I like doing on-line test and exams in subjects that require mathematical answers | 3.31 | 1.20 | 1.45 |
| AT2. I look forward to those aspects of my learning of mathematics that require me to use on-line assessment | 3.05 | 1.12 | 1.26 |
| AT3. On-line tests that contain mathematics are useful | 3.82 | 0.86 | 0.73 |
| AT4. I prefer answering mathematical tests on-line than with pencil and paper | 2.26 | 1.20 | 1.44 |
| **Behaviour Intention (BI)** | | | |
| BI1. I will use electronic tools to support my learning of mathematical subjects in the future | 3.92 | 1.00 | 0.99 |
| BI2. I would like to continue my use of on-line assessment to support my learning of mathematics | 3.60 | 1.10 | 1.21 |
| BI3. All things considered, I expect to continue doing on-line test or exams to assists my learning of mathematics | 3.74 | 1.01 | 1.03 |

## 9.5 Designing research models

*Proposing hypotheses.*

Based on the theoretical frameworks mentioned in the previous chapter, we test the same hypotheses as for students in Mexico:

- *H1: states there is a causal relationship between attitude and behavioural intention to use web-based assessment.*
- *H2: states there is a causal relationship between perceived usefulness behaviour intention to use web-based assessment.*
- *H3: states there is a causal relationship between perceived usefulness and attitude to use web-based assessment.*
- *H4: states there is a causal relationship between enjoyment and attitude to use web-based assessment.*
- *H5: states there is a causal relationship between enjoyment and behaviour intention to use web-based assessment.*

- *H6: states there is a causal relationship between comparative feedback and perceived usefulness to use web-based assessment.*
- *H7: states there is a causal relationship between comparative feedback and enjoyment to use web-based assessment.*
- *H8: states there is a causal relationship between relative feedback and perceived usefulness to use web-based assessment.*
- *H9: states there is a causal relationship between relative feedback and enjoyment to use web-based assessment.*
- *H10: states there is a causal relationship between perceived ease of use and perceived usefulness web-based assessment.*
- *H11: states there is a causal relationship between perceived ease of use and enjoyment web-based assessment.*
- *H12: states there is a causal relationship between availability of information technology and perceived ease of use web-based assessment.*
- *H13: states there is a causal relationship between computer self-efficacy and perceived ease of use web-based assessment.*
- *H14: states there is a causal relationship between social influence and behaviour intention to use web-based assessment.*

The following Figure 9.1 schematically depicts the hypotheses proposed on the research model in this study. This model again integrates enjoyment as a determinant of student intention to use on-line assessments. The research model also includes on-line feedback (comparative and received) as exogenous constructs mediating their effects through enjoyment and perceived usefulness to attitude and usage intentions.

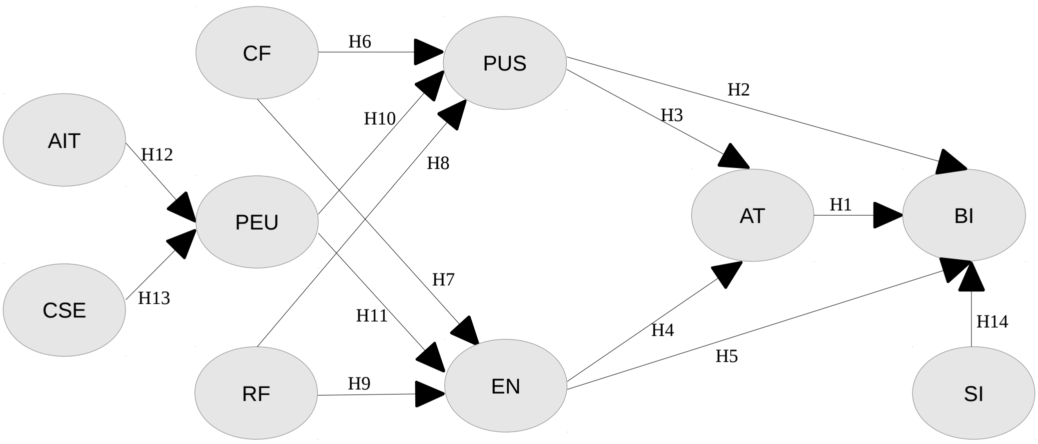


*Figure 9.1. A visual representation of the path model for PLS-SEM student's model showing the hypothetical relationships between the latent variables*

## 9.6 Findings and results

### 9.6.1 Evaluation of the Measurement Model (Model Quality)

Table 9.4 shows measures of construct and composite reliabilities. Results from Cronbach's alpha show that all constructs meet a good internal reliability except comparative feedback construct (.66) considered as acceptable (0.6 ≤ α < 0.7). The availability of information technology construct (.59) should be considered as poor (0.5 ≤ α < 0.6). However, analysing the values for composite reliability show they are all larger than the acceptable cut-off of .6 (Bagozzi & Yi, 1988). Therefore, the model meets the criterion for construct and scale reliabilities.

*Table 9.4. Construct and convergent validity coefficients for the measurement model. Discriminant validity coefficients: bivariate correlations. In the main diagonal, the squared-root AVE of each construct*

| | AVE | Composite Reliability | Cronbachs Alpha | AIT | AT | BI | CF | CSE | EN | PEU | PUS | RF | SI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **AIT** | 0.54 | 0.78 | 0.59 | **0.73** | | | | | | | | | |
| **AT** | 0.79 | 0.92 | 0.86 | 0.2 | **0.89** | | | | | | | | |
| **BI** | 0.81 | 0.93 | 0.89 | 0.15 | 0.66 | **0.90** | | | | | | | |
| **CF** | 0.6 | 0.82 | 0.66 | 0.28 | 0.63 | 0.46 | **0.77** | | | | | | |
| **CSE** | 0.63 | 0.87 | 0.8 | 0.4 | 0.41 | 0.3 | 0.38 | **0.79** | | | | | |
| **EN** | 0.75 | 0.9 | 0.83 | 0.29 | 0.8 | 0.67 | 0.64 | 0.39 | **0.87** | | | | |
| **PEU** | 0.69 | 0.87 | 0.78 | 0.51 | 0.41 | 0.34 | 0.35 | 0.51 | 0.42 | **0.83** | | | |
| **PUS** | 0.7 | 0.9 | 0.85 | 0.27 | 0.67 | 0.57 | 0.58 | 0.3 | 0.66 | 0.43 | **0.84** | | |
| **RF** | 0.68 | 0.86 | 0.76 | 0.38 | 0.53 | 0.43 | 0.56 | 0.25 | 0.53 | 0.31 | 0.39 | **0.82** | |
| **SI** | 0.67 | 0.8 | 0.51 | 0.29 | 0.37 | 0.43 | 0.42 | 0.36 | 0.39 | 0.41 | 0.4 | 0.33 | **0.82** |

Figure 9.2 shows schematic values for the measurement model (indicator loadings coefficients, path coefficients and $R^2$ values). In order to reach indicator validity the items SI3, SI4, AT4 were deleted.

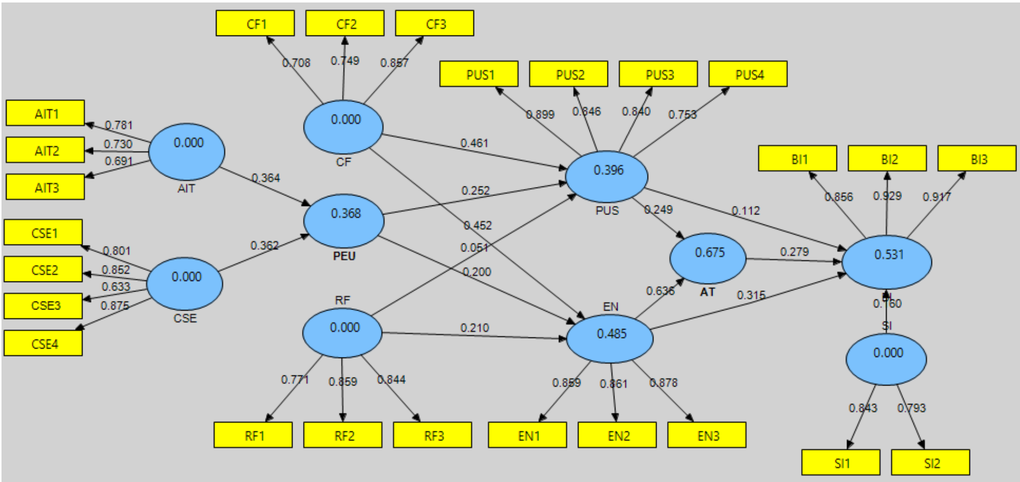

*Figure 9.2. Indicators loadings coefficients, path coefficients and R-squared values for endogenous constructs into the measurement model*

Figure 9.2 represents schematic outputs from the measurement model, showing values that meet the cut-off recommended to achieve indicator validity the values for the attitude construct AT (AT1 0.94, AT2 0.9, AT3 0.82) are not shown in the figure. $R^2$ values of endogenous constructs are also displayed. Behaviour intention construct only has a moderate value for $R^2$, (.53).

## 9.7 Evaluation of the Structural Equation Model

### 9.7.1 Testing Hypotheses

Figure 9.3 shows the relationships between latent variables. It shows the significant levels for each construct, which we have summarised in Table 9.5.

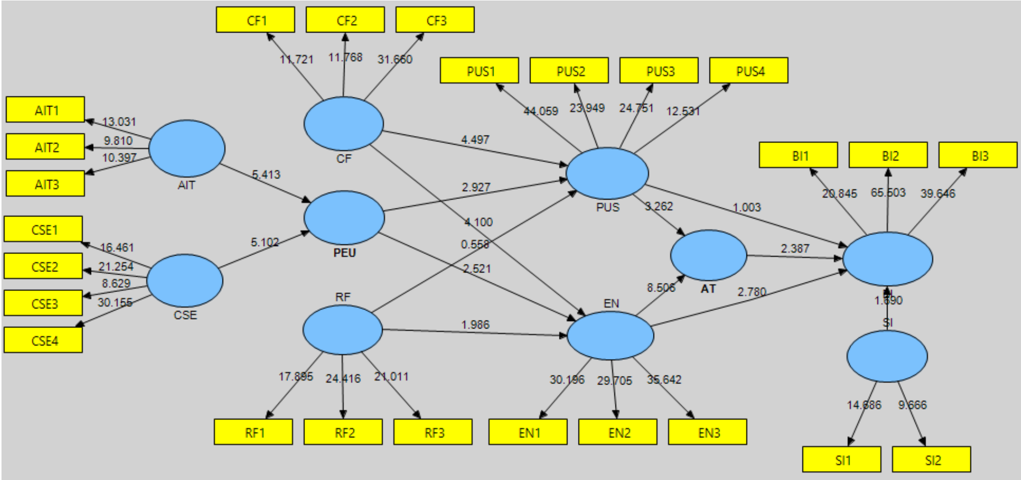

*Figure 9.3. Significance of indicators loadings coefficients and path coefficients for the structural model*

*Table 9.5. Testing hypotheses*

| Hypotheses | Hypotheses path | Path coefficients | T-values significance | Accept/reject |
|---|---|---|---|---|
| H12 | AIT → PEU | 0.36 | 5.41*** | Accept |
| H1 | AT → BI | 0.28 | 2.39*** | Accept |
| H7 | CF → EN | 0.45 | 4.1*** | Accept |
| H6 | CF → PUS | 0.46 | 4.5*** | Accept |
| H13 | CSE → PEU | 0.36 | 5.1*** | Accept |
| H4 | EN → AT | 0.64 | 8.51*** | Accept |
| H5 | EN → BI | 0.32 | 2.78*** | Accept |
| H11 | PEU → EN | 0.2 | 2.52*** | Accept |
| H10 | PEU → PUS | 0.25 | 2.93*** | Accept |
| H3 | PUS → AT | 0.25 | 3.26*** | Accept |
| H2 | PUS → BI | 0.11 | 1 | Reject |
| H8 | RF → EN | 0.21 | 1.99 | Accept |
| H9 | RF → PUS | 0.05 | 0.56 | Reject |
| H14 | SI → BI | 0.16 | 1.69 | Reject |

*Critical t-values for a two-tailed test are: <1.96 (p > .05*), 1.96 (p = .05**), and 2.58 (p = .001***).*

We also include Table 9.6 that shows the total effect, the sum of the direct and all indirect effects linking two constructs.

*Table 9.6. Testing hypotheses: by examining direct and indirect effects*

| Path | Sample Mean (M) | Standard Error (STERR) | T Statistics (\|O/STERR\|) |
|---|---|---|---|
| **AIT → AT** | **0.07** | **0.03** | **2.66** |
| **AIT → BI** | **0.05** | **0.02** | **2.6** |
| **AIT → EN** | **0.08** | **0.03** | **2.3** |
| AIT → PEU | 0.37 | 0.07 | 5.41 |
| **AIT → PUS** | **0.09** | **0.04** | **2.55** |
| AT → BI | 0.29 | 0.12 | 2.39 |
| **CF → AT** | **0.4** | **0.08** | **4.75** |
| **CF → BI** | **0.3** | **0.08** | **4.07** |
| CF → EN | 0.46 | 0.11 | 4.1 |
| CF → PUS | 0.46 | 0.1 | 4.5 |
| **CSE → AT** | **0.07** | **0.03** | **2.37** |
| **CSE → BI** | **0.05** | **0.02** | **2.32** |
| **CSE → EN** | **0.08** | **0.04** | **2.06** |
| CSE → PEU | 0.37 | 0.07 | 5.1 |
| **CSE → PUS** | **0.09** | **0.04** | **2.44** |
| EN → AT | 0.63 | 0.07 | 8.51 |
| EN → BI | 0.49 | 0.11 | 4.5 |
| **PEU → AT** | **0.19** | **0.06** | **3** |
| **PEU → BI** | **0.14** | **0.05** | **2.9** |
| PEU → EN | 0.21 | 0.08 | 2.52 |
| PEU → PUS | 0.25 | 0.09 | 2.93 |
| PUS → AT | 0.25 | 0.08 | 3.26 |
| PUS → BI | 0.19 | 0.11 | 1.71 |
| RF → AT | 0.14 | 0.08 | 1.85 |
| RF → BI | 0.12 | 0.07 | 1.64 |
| RF → EN | 0.2 | 0.11 | 1.99 |
| RF → PUS | 0.05 | 0.09 | 0.56 |
| SI → BI | 0.16 | 0.09 | 1.69 |

These results show that the indirect relationships (AIT → AT = 2.66), (AIT → BI = 2.6), (AIT → EN = 2.3), (AIT → PUS = 2.55), (CF → AT = 4.75), (CF →BI = 4.07), (CSE → AT = 2.37), (CSE → BI = 2.32), (CSE → EN = 2.06), (CSE → PUS = 2.44), (PEU → AT = 3), (PEU → BI = 2.9) are all significant for the model. Furthermore, significance of indicator variables shows that items AT1, AT2, BI2, BI3, CF3, EN3, and PUS1 are the most important in the model. Therefore, items AT1 "I like doing on-line test and exams in subjects that require mathematical answers", AT2 "I look forward to those aspects of my learning of mathematics that require me to use on-line assessment", BI2 "I would like to continue my use of on-line assessment to support my learning of mathematics", BI3 "all things considered, I expect to continue doing on-line test or exams to assist my learning of mathematics", CF3 "on-line feedback helps me better understand mathematical subjects", EN3 "doing mathematical on-line tests is enjoyable" and PUS1 "I find on-line tests useful to support my learning of mathematical subjects" were shown to be the significant factors for students' attitude.

### 9.7.2 Hypotheses Examination

Hypothesis 1 predicts a causal relationship between attitude and behaviour intention. The results show that attitude had a significance impact on usage intentions with a path coefficient of (β: 0.28, $p = .001$) therefore this hypothesis is strongly supported. The causal relationship between perceived usefulness and behaviour intention (H2) is not accepted at (β: 0.11, $p > .05$). H3 shows that perceived usefulness is a significant determinant of attitude (β: .25, $p = .001$); thus, the hypotheses is accepted. H4 predicts a causal relationship between enjoyment and attitude (β: 0.64, $p = .001$), this hypotheses is strongly accepted.

The causal relationship between enjoyment and behaviour intention (H5) is significant (β: .35, $p = .001$). H6 states a causal relationship between comparative feedback and perceived usefulness (β: .46, $p = .05$); this hypotheses is strongly accepted. The causal relationship between comparative feedback and enjoyment (H7) is strongly accepted at (β: .45, $p = .001$). H8 states a causal relationship between received feedback and enjoyment, the result (β: .21, $p = .05$) shows that this hypothesis is accepted. H9 predicts a relationship between received feedback and perceived usefulness (β: .05, $p > .05$); this hypothesis is not supported by the data.

Hypothesis 10 predicts a causal relationship between perceived ease of use and perceived usefulness. The results show that perceived ease of use exerts a significant impact on perceived usefulness at (β: .25, $p = .001$); therefore this hypothesis is strongly supported. The causal relationship between perceived ease of use and enjoyment (H11) is accepted at (β: .2, $p = .001$). H12 shows that availability of information technology is a significant determinant of perceived ease of use is strongly supported (β: .36, $p = .001$). H13 predicts a causal relationship between computer self-efficacy and perceived ease of use, the result of (β: .36, $p = .001$) confirms that this hypotheses is strongly accepted. The causal relationship between social influence and behavioural intentions (H14) is not significant at (β: .16, $p > .05$).

## 9.8 Discussion

The results show that the model explains 53% of the variance in behaviour intention. Similarly, 68% of the variance in attitude, 49% of the variance in enjoyment, 40% of the variance in perceived usefulness and 37% of the variance in perceived ease of use were explained by the antecedent constructs.

Outputs show that attitude and enjoyment have effects on predicting usage intentions. This illustrates that it is important for students to find using technology fun. It also shows that their opinions and feelings matter. Perceived usefulness does not have an effect; students are not influenced by perceived usefulness to use the system. This conclusion matches those of Terzis & Economides (2011) who also find that perceived usefulness has no direct effect on behavioural intention to use a computer-based assessment.

Attitude is predicted by perceived usefulness and enjoyment, but enjoyment has stronger positive effects than perceived usefulness. This could indicate that an on-line test has to include fun activities, challenging activities that make students find them enjoyable while they are learning mathematics. In other words, when an on-line assessment is fun, it will be more likely that students will use it. Furthermore, creating teaching strategies that include enjoyable activities is essential to enhance student's learning in mathematics. Making good use of games, quizzes, and other creative approaches to create more enjoyment and interest in learning mathematics is essential.

On the other hand, comparative feedback reveals strong effects on predicting perceived usefulness and enjoyment. It is indicated by the strong direct effects (CF → PUS = 4.5), (CF → EN = 4.1). This illustrates that students value the fact of getting on-line feedback. They find that it is useful and enjoyable. This is also demonstrated by looking at the indirect effect on attitude (CF → AT = 4.75) and on usage intentions (CF → BI = 4.07), which reveal that this factor is very important in predicting positive students' attitude and usage intentions.

Received feedback has positive effects on enjoyment but not on perceived usefulness. This is illustrated by analysing the direct effects on enjoyment (RF → EN = 1.99) and perceived usefulness (RF → PUS = 0.56). This could mean that students find the experience of receiving mathematical on-line feedback more playful than useful. The indirect effects on attitude (RF → AT = 1.85) and usage intentions (RF → BI = 1.64) demonstrate that this factor does not trigger a favourable behavioural intention. All these outputs are similar to the results from the students in Mexico, except that students in the University of Manchester find that on-line feedback is enjoyable. This matches with the study of Moon & Kim (2001) who find that enjoyment has a positive impact on behavioural intentions.

Perceived ease of use has a strong influence on perceived usefulness (PEU → PUS = 2.93) and enjoyment (PEU → EN = 2.52). This also contributes to the strong indirect effect on attitude (PEU → AT = 3) and usage intentions (PEU → BI = 2.9). These match what we mentioned earlier: perceived enjoyment and perceived usefulness mediate the effects of perceived ease of use on intention. This means that this factor is essential to obtain positive attitude and usage intentions. This is proven by following the causal chain (PEU → EN → AT → BI) that could indicate that if students perceive it as easy to use, they are more likely to have an enjoyable experience, and they are more willing to use it. The causal links (PEU → PUS → AT → BI) shows that when technology is easier to operate, it is more useful, and therefore, students are more willing to apply it. Enjoyment and perceived usefulness mediate the impact of three constructs (perceived feedback, comparative feedback, perceived ease of use) on attitude this has a direct effect on usage behaviour. This reveals that enjoyment and perceived ease of use are powerful factors for predicting usage intentions. This is consistent with what Davis et al. (1992) point out "enjoyment will explain significant variance in usage intentions beyond that accounted for by usefulness alone.” On the other hand, perceived ease of use has been hypothesized as an important factor influencing usage behaviour (Davis, 1989).

Availability of information technology and computer self-efficacy have a strong effect on perceived ease of use. This could indicate that students who encounter some difficulties during an on-line test (regarding system's operation or questions' content) need technical support such as help-desks, on-line support services, guidance by the IT staff and faculty to overcome these situations. Therefore, when technical assistance is provided, it is more likely to find using the on-line environment easier. This is supported by Terzis & Economides (2011) who explain that in the context of computer-based assessment, the availability of information technology determines perceived ease of use. We also see that students who feel comfortable using computers, will find it easier for mathematical on-line assessments.

## 9.9 Conclusions

Our results reveal that attitude and enjoyment are important factors influencing usage intentions. Usefulness, however, does not have any effect. Usefulness and enjoyment, in turn, strongly predict attitude.

Comparative feedback has a strong effect on predicting perceived usefulness (CF → PUS = 4.5) and enjoyment (CF → EN = 4.1). This illustrates that students place real value on on-line feedback. They find it both useful and enjoyable. This is also demonstrated by looking at the strong positive indirect effect on attitude (CF → AT = 4.75) and on usage intentions (CF → BI = 4.07), which reveal that this factor is very important in predicting positive students' attitude and usage intentions.

Received feedback has positive effects on enjoyment but not on perceived usefulness. This is shown by the direct effects on enjoyment (RF → EN = 1.99) and perceived usefulness (RF → PUS = 0.56). This could mean that students find the experience of receiving mathematical feedback on-line more enjoyable than useful. The indirect effects on attitude (RF → AT = 1.85) and usage intentions (RF → BI = 1.64) demonstrate that this factor does not give a favourable behavioural intention.

Perceived ease of use is an important factor influencing usefulness and enjoyment. This also contributes to the strong indirect effect on attitude and usage intentions. Therefore, enjoyment and perceived ease of use are the most powerful factors for predicting usage intentions. Availability of information technology and computer self-efficacy have a strong effect on perceived ease of use. Taking into account these factors can be the best way to design a maths e-assessment activity for UK students.

# Chapter 10. Comparing Results for Mexico and the UK

As we have seen in the previous chapters we have results for both Mexico and UK students. We procured a sample size of 30 cases in Mexico (16 female and 14 male students). As mentioned in chapter 6, most of the students that answered the on-line survey are studying a social subject (53.33%) and (46.67%) are studying a STEM subject. In Manchester we obtained a sample size of 121 students all studying a STEM subject.

Table 10.1 shows Cronbach's alpha coefficients for both models demonstrating that almost all constructs have similar values, except the constructs "availability IT services" and "comparative feedback". This means that the internal consistency (inter-correlations) among items were perceived in a different way.

Our findings show that items belonging to the availability of IT services construct are not strongly correlated for students in the UK, whereas they are for students in Mexico. This suggests that items AIT1. "when I need help to learn to use on-line assessment specialised university staff is there to support me", AIT2. "internet speed at my university is fast enough to use an on-line learning environment", AIT3. "my university has enough computing infrastructure to support on-line testing" are not strongly linked. This could mean that students in Manchester find there is no link between that getting technical support and the availability of computing infrastructure (including internet speed).

*Table 10.1: Evaluating differences between the alpha coefficients of reliability for both models*

| | **Cronbach's Alpha values for student's model in Manchester** | **Cronbach's Alpha values for student's model in Mexico** |
|---|---|---|
| **AIT** | **0.59** | **0.85** |
| AT | 0.86 | 0.84 |
| BI | 0.89 | 0.83 |
| **CF** | **0.66** | **0.8** |
| CSE | 0.8 | 0.71 |
| EN | 0.83 | 0.88 |
| PEU | 0.78 | 0.72 |
| PUS | 0.85 | 0.8 |
| RF | 0.76 | 0.85 |
| SI | 0.51 | n/a |

Likewise, the Cronbach's alpha coefficients for the comparative feedback construct also show differences for students in Manchester (.66) with respect to students in Mexico (.88). Further analysis, suggests that: CF1. “on-line feedback helps me resolve faster doubts about the mathematical material than traditional feedback”, CF2. “electronic assessments of my mathematical subjects allow me to get grades faster, so I know if I am doing well in my topic”, CF3. “on-line feedback helps me better understand mathematical subjects” are correlated, but not strongly (although the value is considered as acceptable). These results suggest that for students in Manchester the aspects such as “it helps a better understanding of the topic”, “it helps to resolve doubts faster” and “it helps to get grades faster” are not strongly linked (or they do not measure the same characteristic).

Comparing both models we find that the causal relationships comparative feedback and perceived usefulness (CF -> PUS); received feedback and enjoyment (RF -> EN); availability of information technology and perceived ease of use (AIT -> PEU) are significant for students in Manchester, but not for students in Mexico.

These outcomes reveal that for students in Manchester find that on-line feedback is more useful when it is compared with traditional feedback, that receiving mathematical on-line feedback is enjoyable. Moreover, students find that the availability of information technology construct contributes to positively predicting

*Table 10.2. Percentage of variance explained for both models*

| | $R^2$ explained in student's model in Manchester | $R^2$ explained in student's model in Mexico |
|---|---|---|
| AT | 67% | 79% |
| BI | 53% | 77% |
| EN | 49% | 75% |
| PEU | 37% | 32% |
| PUS | 40% | 60% |

ease of use. In other words, to get technical computer support is important to facilitate using the on-line environment.

These differences show that students in Manchester value the experience of doing mathematical on-line assessments. They find it is enjoyable. This also means that students in Mexico do not find a true relationship between availability of information technology and perceived ease of use, they do not know that this relationship can help. As we use enriched formative on-line feedback designed to enhance teaching mathematics, as half of the sample of the Mexican students came from a social science background, this could affect the result (they do not find it is enjoyable and as useful as students in Manchester). It is possible that gathering a sample from Mexican students in STEM subjects only could reveal similar results.

On the other hand, we have also found differences in the coefficient of determination ($R^2$) as is shown in Table 10.2. Findings from the $R^2$ coefficients show that the model in Mexico has a better fit than the one in Manchester. This can also explain and predict future outcomes. Since 77% of the variance is explained in usage intentions indicating a large value, this also shows that the model predicts usage intention. The model in Manchester explains 53% of the variance in usage intentions which shows that there are other factors involved in students' usage intentions.

We think that some of the differences described are due to there not being a robust technological culture in educational institutions in Mexico. The lack of strategies that encourage the implementation of on-line learning and testing is one of the barriers. The lack of interest in creating such initiatives from authorities is another issue, which is also reflected in teachers' apathy to use on-line assessments that in turn, affects students' attitude and usage intentions.

Mexico has not recognized the essential role that educational technology plays in institutions as other countries have done. The UK has invested substantial effort and resources (time and money) in making sure that educational technologies work. Unfortunately, Mexico does not have the same level of awareness. This can be seen in the level of engagement that teachers gave to the project. It would be very convenient for Mexico to move in that direction.

Students, considered as “digital natives” have the technical skills and abilities to efficiently manage any kind of task that involves technology. There will always be students with the ability to handle technology efficiently.

Moreover, we find that the important items for the student's model in Mexico are BI3, CF3, RF3, EN3, and PEU3, but for the student's model in Manchester are AT1, AT2, BI2, BI3, CF3, EN3, and PUS1. Therefore, only items BI3, CF3, EN3 are important for both models. These state that:

BI3. “all things considered, I expect to continue doing on-line test or exams to assist my learning of mathematics”,
CF3. “on-line feedback helps me better understand mathematical subjects”,
EN3. “doing mathematical on-line tests is enjoyable”.

This shows that students in both countries find that it is enjoyable, it helps them to enrich their learning and they expect to continue doing on-line testing. This reveals that immediate on-line feedback in mathematics can help students to understand better the learning material and enhance their usage experience.

## 10.1 Final Conclusions

There are several factors that make teachers and students willing to use technology for teaching and learning. We have thoroughly investigated these factors in the context of teaching mathematics and have explored factors such as perceived usefulness, social influence and usage intentions. We also created specific variables such as comparative feedback and received on-line feedback to investigate their effects. Students' feelings have been analysed regarding on-line feedback in comparison with traditional feedback, in other words, to what extent students believe that this is useful and how these feelings may influence the acceptance of technology for learning mathematics.

Our results suggest that for students in the UK attitude and enjoyment are important factors influencing usage intentions. Usefulness does not demonstrate any effect; in turn, usefulness and enjoyment strongly predict attitude.

Comparative feedback reveals strong effects on predicting usefulness and enjoyment. This illustrates that students really value the fact of receiving on-line feedback. They find it is useful and enjoyable. This result is also corroborated by looking at their strong positive indirect effects on attitude and usage intentions which reveal that this factor is very important in predicting positively both students' attitude and usage intentions.

Received feedback has positive effects on predicting enjoyment but not on usefulness. This result could mean that students find the experience of receiving mathematical on-line feedback more diverting than useful. The indirect effects of this factor on attitude and usage intentions demonstrate that it does not trigger a favourable behavioural intention.

Ease of use is an important factor influencing usefulness and enjoyment. This also contributes to the strong indirect effect on attitude and usage intentions. Therefore, enjoyment and ease of use are powerful factors for predicting usage intentions. The factors of availability of information technology and computer self-efficacy have a strong effect on ease of use.

For students in Mexico attitude strongly influences their usage intentions. This result is also confirmed by looking at the strong direct effects of usefulness and enjoyment on attitude. It clearly means that in order to obtain a positive students' willingness to do on-line assessments it is necessary to take into account their attitude.

Comparative feedback has a positive effect on enjoyment. This result could mean that students find on-line feedback enjoyable. Received feedback does not exert direct influence on either enjoyment or usefulness. This result shows that receiving on-line feedback does not have a strong effect on students' intention to use on-line assessments. Ease of use positively predicts enjoyment and usefulness. This result is consistent with Davis et al. (1992) who found that perceived enjoyment and perceived usefulness mediated the influence of perceived ease of use on intention, explaining that "while usefulness will once again emerge as a major determinant of intentions to use a computer in the workplace, enjoyment will explain significant variance in usage intentions beyond that accounted for by usefulness alone."

Computer self-efficacy has a significant positive influence on ease of use, meaning that students who are proficient in on-line testing are more likely to find it easy to use. Attitude has the strongest influence on usage intentions which could mean that student feelings and opinions matter. The fact of weak effects of usefulness and enjoyment on usage intention when these are compared through attitude demonstrates that attitude is an important determinant of the intention to use on-line testing.

In the case of lecturers in the UK the results demonstrate that on-line feedback is an important determinant for predicting positive attitude and intention to use on-line assessments. They agree that on-line assessments are useful tools to enrich instructional strategies. Teachers in Mexico have a similar opinion (on-line feedback has a strong influence on perceived usefulness), which means that they really value providing on-line feedback. Both findings reveal that on-line feedback plays a crucial role in predicting instructors' intentions to use mathematical on-line assessments.

This also makes it evident that instructors recognize the positive relationship between usefulness and attitude, which might mean that they think that on-line assessments are practical and helpful tools that can enhance their productivity. For both countries, our findings also confirm that teachers feel very confident using computers, which probably means that they are capable of doing mathematical on-line testing. All these outcomes suggest that providing clear evidence of the benefits of on-line assessments for students' learning and performance might convince teachers to adopt such technologies.

These findings can help to propose better strategies to implement educational technology for teaching and learning mathematics. First of all, there is a need to build some effective schemes to convince instructors who have the power to persuade students to use technology in a broader manner. Students as "digital natives" are usually willing to use technological tools as long as they think that these are useful and enjoyable. The usage of ICT in teaching mathematics can help to introduce innovative ways of learning making the experience of learning more enjoyable for students.

Institutions should enhance the teaching and learning processes, building a favourable reputation in a competitive sector. These conclusions are important for the IPN in Mexico to learn the factors that ensure a successful adoption. The use of ICT opens the door to more opportunities for on-line higher education. Teaching mathematics should be an innovative process.

We are particularly interested in enriching mathematical activities and tasks involved in the assessment process by making them more efficient and effective by the use of technology. We consider that the use of technology can greatly enhance the students' performance by making the learning activities more effective and consequently in the way in which they are assessed. Furthermore, instructors can considerably boost their teaching practices with the use of helpful technological

We have gained some interesting insights regarding the adoption of technology considering perceptions and opinions of instructors and students for the assessment of mathematical subjects. We consider that these insights can be useful for all stakeholders, policy-makers, instructors and students. Likewise, these results can be very helpful as a way to fulfil necessities of more and better educational processes in developing countries. In addition, it is believed that these insights can contribute to lead a successful implementation of web-based assessment systems in a context of higher education.

There must be a clear governmental support for the development of strategical initiatives and policies that promote public and private investments in the development of learning, innovation and research. In turn, these create strategies for a better quality of life in society where technology can be the key to open the door to the future for universities.

# Appendix A

*Table A.1. The main factors that affect the uptake of e-assessment*

| | Dimension | Factors | Literature |
|---|---|---|---|
| 1 | Faculty concerns | Lack of faculty time | Dermo (2007); Mc Cann (2010); Whitelock (2006); Warburton (2003) |
| 2 | Faculty concerns | Lack of confidence in systems | Mc Cann (2010); Warburton (2008) |
| 3 | Faculty concerns | Lack of computer literacy | Warburton (2008); Purvis (2011); Dermo (2007); Ashton (2008) |
| 4 | Administrative structures | Incompatibility (technology – faculty culture); computer systems do not fit into neat departmental or faculty units. | Mc Cann (2010); Sieber (2008) |
| 5 | Administrative structures | Lack of policy; Lack of an agreed and enforced institutional policy | Mc Cann (2010); Dermo (2007); Heinrich et al. (2009) |
| 6 | Administrative structures | Lack of computer training (groups too large, staff with different level of technical competence, no consider disciplinary difference, introduce a new pedagogical model, balance on feedback) | Dermo (2007); Mc Cann (2010); Chew (2010); Whitelock (2006) |
| 7 | Technological infrastructure and systems | Lack of infrastructure (equipment, resources) | Mc Cann (2010); Boyle (2011); |
| 8 | Technological infrastructure and systems | Lack of technical support (reliability and security) | Dermo (2007); Sieber (2008); |
| 9 | Technological infrastructure and systems | High risk of technical failure | Chew (2010); Sieber (2008); Dermo (2007) |

# Appendix B

*Strategic initiatives and regulators in the UK.* The four organisations responsible for regulating qualifications in the UK are the Qualifications and Curriculum Authority for England (QCA); the Department for Children, Education, Lifelong Learning and Skills (DCELLS) for Wales; the Council for the Curriculum, Examinations and Assessments (CCEA) for Northern Ireland and the Scottish Qualifications Authority (SQA) for Scotland.

The QCA was created as a public organization that "maintains and develops the national curriculum and associated assessments, tests and examinations; and accredits and monitors qualifications in colleges and at work" (Qualifications and Curriculum Authority, 2004). In 2004, the QCA published a five-year programme for the implementation of on-demand e-assessment expressed in the QCA Blueprint for e-Assessment document in 2004.

In 2007, the Qualifications and Curriculum Development Agency (QCDA) was a public organization of the Department for Education which carried out similar roles to QCA, although regulatory functions regarding examinations and assessment were transferred to the Office of Qualifications and Examinations Regulation (OFQUAL) in 2007 as an independent regulator of exams and tests in England. The remaining work of the QCA was transferred to the QCDA. The QCA was formally dissolved in 2010 when the QCDA and OFQUAL gained statutory status.

In 2010, the Secretary of State announced his intention to promote legislation that would abolish QCDA which was closed in 2012 and has been replaced by the Standards and Testing Agency (STA) since 2011 which is responsible for developing and delivering all statutory assessments for school pupils in England. The STA is regulated by the examinations regulator such as the Office for Standards in Education (OFSTED) (National STEM Centre, 2014).

OFQUAL is a non-ministerial government department that regulates qualifications, exams and tests in England and vocational qualifications in England and Northern

Ireland. It supervises GCSEs and A levels in England. They also regulate the National Curriculum Assessments in England.

OFQUAL have seen that as the use of technology for assessment has increased, (whether it is an end-to-end solution or blended learning/ assessment combining electronic and traditional methods), this will have a regulatory impact. As a measure to ensure the integrity reliability and validity of assessment they have proposed initiatives that encourage innovative practice in e-Assessment.

One of their initiatives is "the regulatory principles for e-Assessment" designed to facilitate, promote and encourage innovation across the whole range of e-Assessment. An important concern included in the principles is to ensure that regulation does not impede innovation. They have proposed moderning their qualification system through the use of technology to maximise the opportunities available to the learner in demonstrating their achievements.

Scotland has also proposed initiatives to support e-Assessment practices. The Scottish Qualifications Authority (SQA) is "the national accreditation and awarding body in Scotland". They accredit vocational qualifications that are offered across Scotland such as Scottish Vocational Qualifications. Since they propose supporting high quality teaching and assessment practices in Scotland, the SQA has clearly set out the e-Assessment vision a strategy (SQA) for the next five years and beyond that includes:

- Where it is appropriate, e-Assessment of SQA's qualifications will be routine
- Assessment will be one part of an integrated process of teaching and learning where the distinction between learning and assessment becoming blurred.
- Learners will expect 'personalised' assessment opportunities; being able to undertake assessment at a time and in a place of their choosing, subject to appropriate authenticity checks.
- Learners will expect to store evidence of their achievements in a personal e-portfolio.

Moreover, in the effort to impulse the electronic assessment practices, the four regulatory bodies mentioned above have collaborated to develop the publication: “regulatory principles for e-Assessment” (2007) which aims to encourage innovative practices to enhance the quality assurance of e-Assessment.

The qualification regulators for England, Wales and Northern Ireland developed initiatives such as “e-Assessment: guide to effective practice” (2007) that we have cited already. This document was developed with the aim of supporting effective practice and quality improvement in the assessment of qualifications. It gives practical information and advice to people involved in the management and delivery of e-Assessment within qualifications. It covers two key aspects of e-Assessment - the management and delivery of e-testing and the use of e-portfolios for assessment -. The Scottish Qualifications Authority (SQA) developed their own version of the guide for use in Scotland (e-Assessment: guide to effective practice - SQA).

In England the role of the Higher Education Funding Council (HEFCE) has been crucial in promoting and funding high quality cost-effective teaching and research. The aim of the council is to “distribute public money to higher education to universities and colleges in England, and ensure that this money is used to deliver the greatest benefit to students and the wider public” (HEFCE, 2014). Since 2004 this organisation has proposed several national initiatives to foment the e-Assessment practices such as HEFCE Strategy for e-Learning, 2005, QCA Blueprint for e-Assessment, 2004 and the SQA Guidelines on e-Assessment for Schools, 2005. The HEFCE Strategy for e-Learning, 2005 (HEFCE, 2005) had the objective to enable institutions to meet the needs of learners and their own aspirations for development to effectively help the sector use new technology to become embedded as a part of their activities.

The Joint Information Systems Committee (JISC) was founded in 1993 as a Committee, but over the last decade has been working as a company. It plays an important role in stimulating the use of digital technology in higher education and research in the UK. JISC has recognised the importance of e-learning practices in the UK higher education and research community, with several initiatives such as the JISC e-Learning programme (2012) aim to “enables the development and effective use of digital technologies to support learning and teaching in universities and colleges, so that staff benefit from e-learning and students enjoy a more flexible learning experience”.

JISC has identified e-Assessment practices as an important part of the e-learning activities and has recognized the important role to play through this area. The company has been working in support of technology-enhanced assessment for over a decade focusing on issues such as pedagogy and institutional context for the appropriate use of a wide range of technology to support the assessment and feedback process, as well as technologies and interoperability standards for the delivery of on-screen tests.

As a way for continuous support to assessment practice, JISC has founded the programme “assessment and feedback” running from 2011 to 2014, to “supporting large-scale changes in assessment and feedback practice, supported by technology, with the aim of enhancing the learning and teaching process and delivering efficiencies and quality improvements”.

New capabilities of ICT has led to its growing use to deliver, score and record responses of tests and assessments in a wide range of educational and other contexts. In order to recognize the importance of taking control of the e-Assessment practices, the British Standards Institution (BSI) (2002) developed the British Standard BS7988:2002 that is the code of practice for the use of information technology (UIT) in the delivery of assessments. Due to efficient practices in 2007 the BS7988:2002 became an international standard ISO/IEC 23988:2007. The standard UIT delivery offers advantages of speed and efficiency, better feedback and improvements in validity and reliability, but its increased use has raised issues about the security and fairness of UIT-delivered assessments, as well as resulting in a wide range of different practices.

BS ISO/IEC 23988:2007 gives recommendations for the use of UIT to deliver assessments to candidates and to record and score their responses. Its scope is defined in terms of three dimensions: the types of assessment to which it applies, the stages of the assessment life cycle to which it applies and the standard focus on specific UIT aspects.

# Appendix C

The following Figure C.1 shows the course "first-year students" provided each year to incoming students to the School of Physics and Astronomy of The University of Manchester. The course is based on STACK (system for teaching and assessment using a computer algebra kernel) which is an open-source system particularly designed to provide formative assessment.

STACK asks for mathematical expressions and evaluates these using computer algebra system (CAS). CAS provides a library of functions with which to manipulate students' answers and generate feedback. Using CAS can also help generate random yet structured problems, and corresponding worked solutions.

STACK allows teachers to author and manage their own questions. It also allows to create question versions randomly generated within structured templates. STACK is an attractive on-line tool as it can generate tailored feedback according to a decision tree (potential response tree). The feedback can include: textual comments for the student, a numerical mark, answer notes from which statistics for the teacher are compiled. Another advantage of STACK is that teachers can generate questions which may have any number of inputs and any number of potential response trees, therefore partial credit is possible when an expression only satisfies some of the required properties in the kind of questions (multi-part mathematical) (GitHub moodle-qtype_stack, 2014).

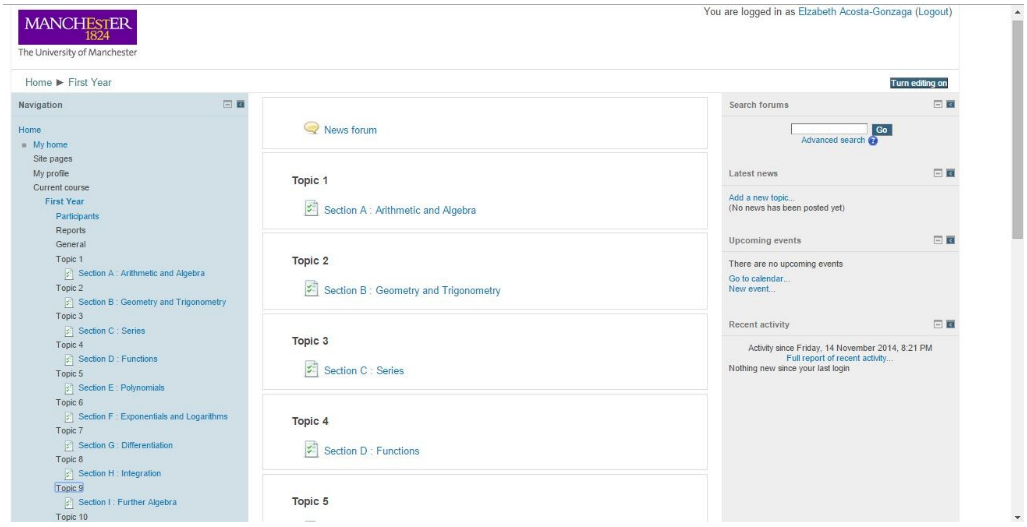


*Figure C.1. First-year course (main screen)*

As example, we have included question 1 of the section "differentiation" which is shown in Figure C.2 below.

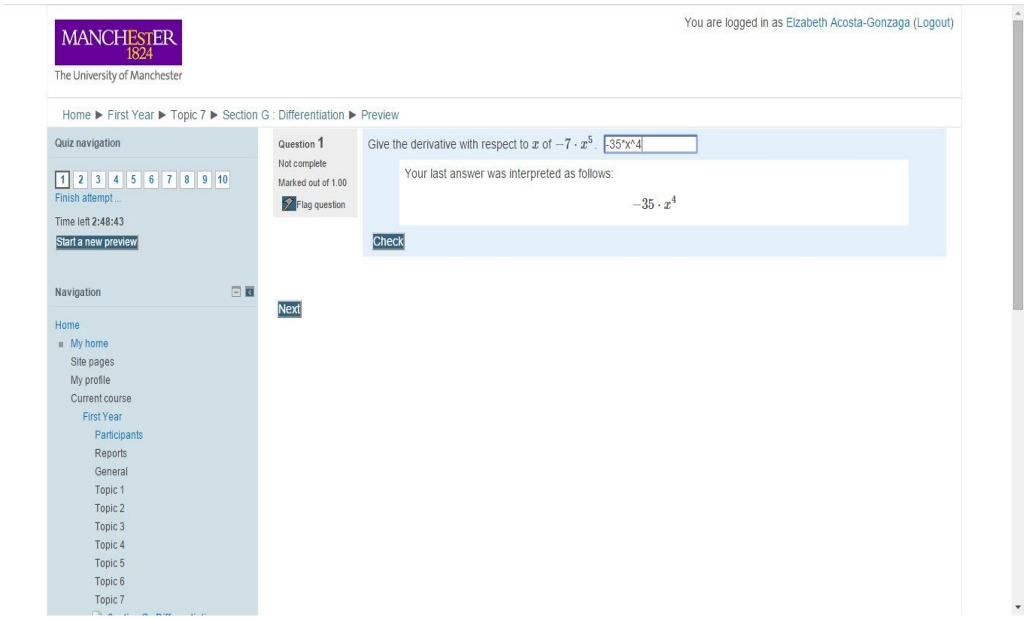


*Figure C.2. Introducing a mathematical expression*

When the student finishes typing, STACK syntactically validates the answer before being evaluated. In order to evaluate it, STACK implements a potential response tree algorithm to check whether the answer is algebraically equivalent to a model answer. In other words, the algorithm establishes the mathematical properties of the answer to provide feedback and a score. As the potential response tree is a non-empty acyclic directed graph of potential response nodes, each node compares two expressions using a specified test (answer test), and the result is either true or false. Then, a corresponding branch of the tree is performed making one of the following actions: adjusting the score, (e.g. assign a value, add or subtract a value); adding written feedback specifically for the student; leaving an answer note for reporting purposes; proposing the next potential response node, or ending the process. In this example, as is shown in Figure C.3 below, a correct response was introduced then a mark is obtained.

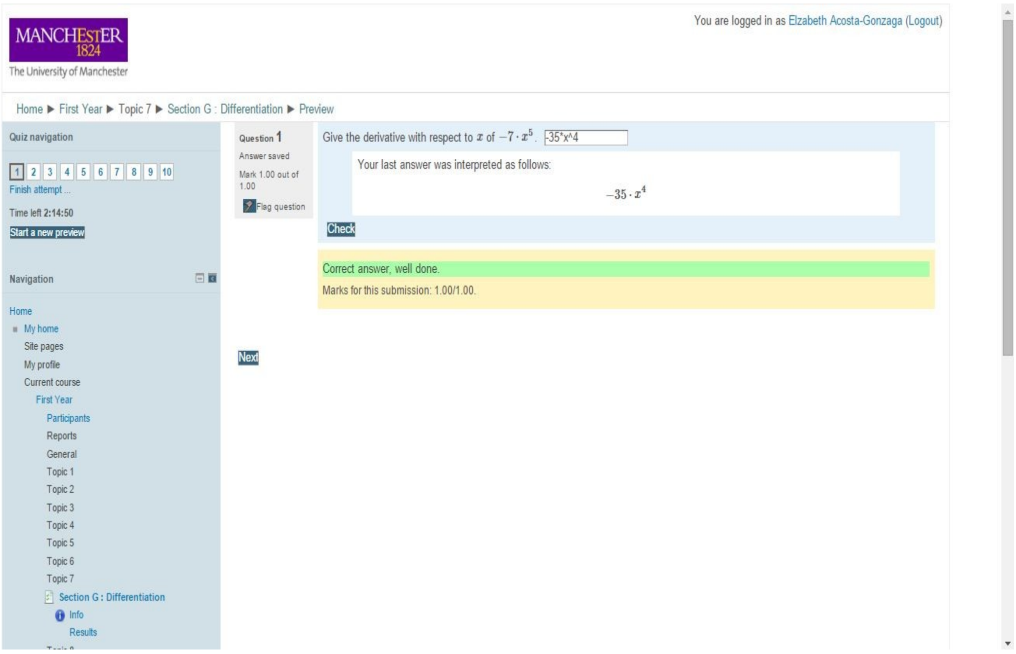


*Figure C.3. Feedback provided as a result of evaluating a correct answer*

In the case that the student having answered incorrectly, the web-based environment will display a penalty. However, after completing the section, formative feedback will be displayed for all questions not answered correctly. This represents a useful help to students, since they can learn from their own mistakes and do better in the next trial.

Furthermore, when the student's answer is partially correct, following our example of resolving a simple differential equation, STACK solves this issue by integrating the answer and compares it to the question. This is shown in Figure C.4.

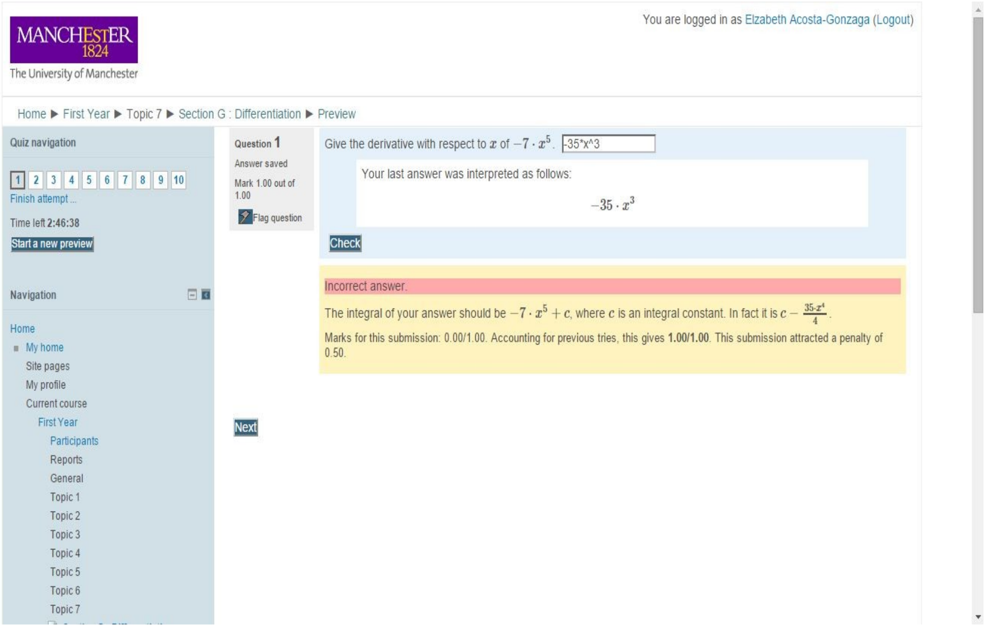


*Figure C.4. STACK identifies a possible misunderstanding*

This means that as the answer was not totally incorrect, the web-based environment gives formative feedback providing hints to students to help them find their mistake.

# Appendix D

**A. Executing Exploratory Factor Analysis (EFA)**

An EFA can be applied by following the six-stage model-building process suggested in (Hair et al., 2010).

*Stage 1. Objectives of factor analysis.*

An EFA is run to identify the structure of a set of variables. This analysis identifies the underlying factors (constructs) that explain the pattern of correlations within a set of measurement items. It calculates the relationships between all the measurement items, placing those most closely related (highly correlated) into factors, which are then matched to the researcher's theoretically positioned constructs (Gefen & Straub, 2005).

*Stage 2. Designing a factor analysis.*

The sample size required is determined by a general rule to contain at least five times as many observations as the number of variables to be analysed. The original scale is composed of 5 constructs. The sample size for this study is 133 respondents which falls within acceptable limits.

*Stage 3. Assumptions in factor analysis.*

To quantify the degree of inter-correlations among the variables, the following two values are applied.

1) Bartlett's Test of Sphericity, a statistical test to measure the presence of correlations among the variables. A statistically significance Bartlett's Test of Sphericity (sig<.05) indicates that sufficient correlations exist among the variables. A value Bartlett's Test of Sphericity = 624.189, significance = .000 was obtained.

Measure of sampling adequacy (MSA) values must exceed .50 for the overall test. This measure can be interpreted with the following guidelines: .80 or above, meritorious; .70 or above, middling; .60 or above, mediocre; .50 or above, miserable; and below .50 unacceptable.

*Stage 4. Deriving factors.*

Table D.1 shows the results for the extraction of component factors. It shows their explanatory power expressed by their eigenvalues. These values assist us in selecting the number of factors.

*Table D.1. Initial un-rotated factor solution*

| Component Matrix[a] | | | | | |
|---|---|---|---|---|---|
| | Component | | | | |
| | 1 | 2 | 3 | 4 | 5 |
| PU2 | .703 | | | | .305 |
| FE4 | .680 | | | -.504 | |
| PU3 | .655 | | | -.236 | |
| AT2R | .626 | | -.310 | .295 | |
| PS4 | .623 | | | | -.328 |
| PS1 | .596 | | | .368 | |
| PS5 | .582 | .320 | .370 | .250 | |
| FE1 | .576 | .260 | | -.292 | .216 |
| PS3 | .569 | -.414 | | .272 | |
| FE2 | .499 | | .466 | | |
| PR2 | .349 | | | | |
| PR1R | .212 | .815 | | | -.208 |
| PR3R | .223 | .612 | -.485 | | |
| PS2 | .500 | -.604 | | .222 | |
| AT3 | | | .705 | | |
| AT1R | .464 | -.253 | -.545 | | |
| FE3 | .518 | | | -.539 | -.320 |
| PU1 | .342 | | -.242 | -.256 | .721 |
| Extraction Method: Principal Component Analysis. | | | | | |
| a. 5 components extracted. | | | | | |

The factors of matrix loadings for the un-rotated factor matrix were examined. To identify significant loadings quickly, small coefficients below .02 were suppressed. Statistical significance employs the concept of statistical power to specify significant factor loadings values depending on differing sample sizes.

Having a sample size of 133 respondents, the factor loading suggested is .50 (Hair et al., 2010).

Therefore, based on the factor-loadings pattern sharing correlated values, EFA proposed an initial solution of 5 factors. Since it did not have a clean set of factor loadings, a rotation technique was required to improve the interpretation and obtain a simpler and theoretically more meaningful factor pattern.

*Stage 5. Interpreting the factors.*

Table D.2 shows the communalities values of each variable. This index is useful for assessing how much variance in a particular variable is accounted for by the factor solution. There are no statistical guides that indicate exact appropriate values. However, practical considerations state a lower level of .50.

*Table D.2. Communalities*

| Communalities | | |
|---|---|---|
| | **Initial** | **Extraction** |
| PS1 | 1.000 | .543 |
| PS2 | 1.000 | .708 |
| PS3 | 1.000 | .582 |
| PS4 | 1.000 | .523 |
| PS5 | 1.000 | .641 |
| PU1 | 1.000 | .764 |
| PU2 | 1.000 | .618 |
| PU3 | 1.000 | .537 |
| AT3 | 1.000 | .538 |
| FE1 | 1.000 | .541 |
| FE2 | 1.000 | .527 |
| FE3 | 1.000 | .698 |
| FE4 | 1.000 | .754 |
| PR2 | 1.000 | .176 |
| AT1R | 1.000 | .591 |
| AT2R | 1.000 | .587 |
| PR1R | 1.000 | .760 |
| PR3R | 1.000 | .694 |

Extraction Method: Principal Component Analysis.

*Table D.3. Un-rotated component analysis factor matrix*

| Total Variance Explained | | | | | | |
|---|---|---|---|---|---|---|
| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | |
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 4.835 | 26.861 | 26.861 | 4.835 | 26.861 | 26.861 |
| 2 | 1.969 | 10.940 | 37.801 | 1.969 | 10.940 | 37.801 |
| 3 | 1.669 | 9.270 | 47.072 | 1.669 | 9.270 | 47.072 |
| 4 | 1.224 | 6.799 | 53.871 | 1.224 | 6.799 | 53.871 |
| 5 | 1.084 | 6.022 | 59.893 | 1.084 | 6.022 | 59.893 |
| 6 | .990 | 5.499 | 65.393 | | | |
| 7 | .927 | 5.148 | 70.541 | | | |
| 8 | .839 | 4.663 | 75.204 | | | |
| 9 | .721 | 4.007 | 79.211 | | | |
| 10 | .660 | 3.666 | 82.876 | | | |
| 11 | .549 | 3.050 | 85.926 | | | |
| 12 | .520 | 2.888 | 88.814 | | | |
| 13 | .480 | 2.665 | 91.479 | | | |
| 14 | .399 | 2.218 | 93.697 | | | |
| 15 | .328 | 1.821 | 95.519 | | | |
| 16 | .298 | 1.653 | 97.172 | | | |
| 17 | .278 | 1.546 | 98.718 | | | |
| 18 | .231 | 1.282 | 100.000 | | | |
| Extraction Method: Principal Component Analysis. | | | | | | |

The index for the overall solution shows that 59.89 percent of the total variance is represented by the information contained in the factor matrix of the initial five-factor solution. The latent root criterion of retaining factors with eigenvalues (sum of squared factor-loadings) of greater than 1.0 was applied. The solution extracts the factors in order of their importance, with factor 1 accounting for the most variance, factor 2 slightly less and so on. The five factors proposed represent 59.8 percent of the variance of the variables. A (promax) rotation was applied.

Once it was determined that a rotated solution was required, a continuous cycle of the following steps were executed in order to specify the factor structure: 1. Examine the factor matrix of loadings. 2. Identify the significant loadings for each variable. 3. Assess the communalities of the variables. 4. Specify the factor model if needed.

Since cross-loadings are displayed in several interactions, the solution proposed was to decrease the number of factors to 4 as is shown in Table D.4.

*Table D.4. Promax-rotated component analysis matrix: full set of variables*

| Pattern Matrix[a] | | | | |
|---|---|---|---|---|
| | Component | | | |
| | 1 | 2 | 3 | 4 |
| FE2 | .845 | | | |
| PS5 | .836 | | | |
| PU2 | .636 | | | |
| AT1R | | .854 | | |
| PS3 | | .691 | | |
| AT2R | | .627 | | |
| FE3 | | | .937 | |
| FE4 | | | .815 | |
| PR1R | | | | .866 |
| PR3R | | | | .786 |
| Extraction Method: Principal Component Analysis. Rotation Method: Promax with Kaiser Normalization. | | | | |
| a. Rotation converged in 5 iterations. | | | | |

Loadings above the cut-off point with values ± .40 or above were obtained. Factor loadings above .70 mean that more than half of the variance is accounted for by the loadings of a single factor. All communalities were of sufficient size to warrant inclusion.

After rotating the pattern matrix the variance was redistributed. As is shown in Table D.5, the total amount of variance extracted was higher than the un-rotated solution. Thus, the explanatory power shifted slightly to a more even distribution because of the rotation. A simplified structure was obtained; the factor-loadings for each variable are maximized for each variable on one factor.

*Table D.5. Total variance explained*

| Total Variance Explained | | | | | | | |
|---|---|---|---|---|---|---|---|
| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadingsa |
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total |
| 1 | 3.043 | 30.430 | 30.430 | 3.043 | 30.430 | 30.430 | 2.379 |
| 2 | 1.577 | 15.766 | 46.196 | 1.577 | 15.766 | 46.196 | 2.152 |
| 3 | 1.338 | 13.376 | 59.572 | 1.338 | 13.376 | 59.572 | 2.069 |
| 4 | 1.026 | 10.259 | 69.831 | 1.026 | 10.259 | 69.831 | 1.618 |
| 5 | .680 | 6.802 | 76.633 | | | | |
| 6 | .596 | 5.961 | 82.594 | | | | |
| 7 | .527 | 5.268 | 87.863 | | | | |
| 8 | .515 | 5.145 | 93.008 | | | | |
| 9 | .377 | 3.775 | 96.783 | | | | |
| 10 | .322 | 3.217 | 100.000 | | | | |
| Extraction Method: Principal Component Analysis. | | | | | | | |
| a. When components are correlated, sums of squared loadings cannot be added to obtain a total variance. | | | | | | | |

Table D.6 shows the final scale and their Cronbach's alpha coefficients.

*Table D.6. Final solution emerged from EFA*

| | |
|---|---|
| | **Usefulness α = .698** |
| FE2 | I receive sufficient personalised feedback on my on-line tests. |
| PU2 | Electronic assessments help me to get a deeper understanding of the subject. |
| PS5 | I think that I'm using cutting edge technology in my on-line tests. |
| | **Reliability α = .615** |
| PR1R | It is easier to cheat during an electronic assessment than when doing it on paper. (R) |
| PR3R | On-line tests can often be passed by using flaws in the software. |
| | **Feedback α = .733** |
| FE4 | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. |
| FE3 | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. |
| | **Affective factors α = .636 personal reaction opinions, beliefs, sentiment** |
| AT1R | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) |
| AT2R | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) |
| PS3 | It is easier to get support from others students in an on-line assessment than in a paper-based one. |

*Stage 6: Validation of the factor analysis.*

This stage is shown in the section 6.7.1.1 evaluation of measurement model emerged from EFA.

**B. By reorganizing items conceptually**

*Table D.7. Removing weaker items*

| | |
|---|---|
| | **Affective factors α = .615** |
| AF1 | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) |
| AF2 | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) |
| | **Utility α = .572** |
| UT1 | I am used to reading lecture notes on-line |
| UT2 | Electronic assessments help me to get a deeper understanding of the subject. |
| UT3 | I find it useful that I can do on-line exercises at a time of my choosing. |
| | **Suitability α = .697** |
| SU1 | I think that on-line tests and exams are appropriate to test my ability in using mathematics. |
| SU2 | It is easier for me to take an on line test because I can get information using the internet. |
| SU3 | It is easier to get support from others students in an on-line assessment than in a paper-based one. |
| SU4 | I can save time by answering a test on-line. |
| SU5 | I think that I'm using cutting edge technology in my on-line tests. |
| | **Feedback α = .691** |
| FE1 | On-line feedback encourages me to do better in my studies. |
| FE2 | I receive sufficient personalised feedback on my on-line tests. |
| FE3 | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. |
| FE4 | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. |
| | **Reliability α = .615** |
| PR1 | It is easier to cheat during an electronic assessment than when doing it on paper. (R) |
| PR3 | On-line tests can often be passed by using flaws in the software. (R) |

*Table D.8. Final solution by reorganising constructs conceptually*

| | |
|---|---|
| | **Affective factors α = .636** |
| AF1 | I think that is more stressful doing an on-line test/exam than a paper-based one. (R) |
| AF2 | It is harder to concentrate on a question when taking an electronic assessment than in a paper-based one. (R) |
| AF3 | It is easier to get support from others students in an on-line assessment than in a paper-based one. |
| | **Usefulness α = .737** |
| US1 | Electronic assessments help me to get a deeper understanding of the subject. |
| US2 | I find it useful that I can do on-line exercises at a time of my choosing. |
| US3 | I think that on-line tests and exams are appropriate to test my ability in using mathematics. |
| US4 | It is easier for me to take an on-line test because I can get information using the internet. |
| US5 | I can save time by answering a test on-line. |
| US6 | I think that I'm using cutting edge technology in my on-line tests. |
| | **Feedback α = .691** |
| FE1 | On-line feedback encourages me to do better in my studies. |
| FE2 | I receive sufficient personalised feedback on my on-line tests. |
| FE3 | Immediate on-line feedback can help me resolve doubts about the material I am studying faster. |
| FE4 | Electronic assessments allow me to get grades faster, so I know if I am doing well in my topic. |
| | **Reliability α = .615** |
| RE1 | It is easier to cheat during an electronic assessment than when doing it on paper. (R) |
| RE2 | On-line tests can often be passed by using flaws in the software. (R) |

# References


Agudo-Peregrina, Á. F., Hernández-García, Á., & Pascual-Miguel, F. J. (2014). Behavioral intention, use behavior and the acceptance of electronic learning systems: Differences between higher education and lifelong learning. *Computers in Human Behavior*, *34*, 301–314. http://doi.org/10.1016/j.chb.2013.10.035

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. http://doi.org/10.1016/0749-5978(91)90020-T

Alexiou, A., & Paraskeva, F. (2010). Enhancing self-regulated learning skills through the implementation of an e-portfolio tool. *Procedia - Social and Behavioral Sciences*, *2*(2), 3048–3054. http://doi.org/10.1016/j.sbspro.2010.03.463

Angus, S. D., & Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology*, *40*(2), 255–272. http://doi.org/10.1111/j.1467-8535.2008.00916.x

Ashton, H. S., Beevers, C., & Thomas, R. (2008). Can e-assessment become mainstream? [Conference Contribution]. Retrieved April 30, 2012, from https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/4611

Astrachan, C. B., Patel, V. K., & Wanzenried, G. (2014). A comparative study of CB-SEM and PLS-SEM for theory development in family firm research. *Journal of Family Business Strategy*, *5*(1), 116–128. http://doi.org/10.1016/j.jfbs.2013.12.002

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning* (Vol. xiv). Oxford, England: Grune & Stratton.

Bagozzi, R. P., & Yi, Y. (1988). On the Evaluation of Structural Equation Models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94. http://doi.org/10.1177/009207038801600107

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall,.

Baylor, A. L., & Ritchie, D. (2002). What factors facilitate teacher skill, teacher morale, and perceived student learning in technology-using classrooms? *Computers & Education*, *39*(4), 395–414. http://doi.org/10.1016/S0360-1315(02)00075-1

Bennett, R. E. (1998). *Reinventing Assessment. Speculations on the Future of Large-Scale Educational Testing. A Policy Information Perspective.* Policy Information Center, Educational Testing Service, Princeton. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED424254

Bennett, R. E. (2001). How the Internet Will Help Large-Scale Assessment Reinvent Itself. *Education Policy Analysis Archives*, *9*, 5. http://doi.org/10.14507/epaa.v9n5.2001

Bennett, R. E. (2002). Inexorable and Inevitable: The Continuing Story of Technology and

Assessment. *The Journal of Technology, Learning and Assessment*, *1*(1). Retrieved from http://ejournals.bc.edu/ojs/index.php/jtla/article/view/1667

Benson Soong, M. ., Chuan Chan, H., Chai Chua, B., & Fong Loh, K. (2001). Critical success factors for on-line course resources. *Computers & Education*, *36*(2), 101–120. http://doi.org/10.1016/S0360-1315(00)00044-0

Bhuasiri, W., Xaymoungkhoun, O., Zo, H., Rho, J. J., & Ciganek, A. P. (2012). Critical success factors for e-learning in developing countries: A comparative analysis between ICT experts and faculty. *Computers & Education*, *58*(2), 843–855. http://doi.org/10.1016/j.compedu.2011.10.010

Black, P., & Wiliam, D. (1998a). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. http://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (1998b). Inside the Black Box: Raising Standards through Classroom Assessment. *The Phi Delta Kappan*, *80*(2), 139–148.

Bloom, B. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals.* London: Longman.

Bloxham, S., & Boyd, P. (2007). *Developing Effective Assessment In Higher Education: A Practical Guide: A Practical Guide*. McGraw-Hill International.

British Standards Institution. (2002). bsi. Retrieved March 7, 2014, from http://www.bsigroup.com/en-GB/

Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.

Bruner, J., Goodnow, J., & Austin, A. (1956). *A study of thinking*. New York: Wiley.

Bryceson, K. (2007). The online learning environment-A new model using social constructivism and the concept of “Ba” as a theoretical framework. *Learning Environments Research*, *10*(3), 189–206. http://doi.org/10.1007/s10984-007-9028-x

Bull, J. (2000). *Computer-assisted Asessment Centre - Annual Report* - (No. TLTP85/AR2/2000) (p. 12). University of Luton.

Buzetto-More, N. A., & Alade, A. J. (2006). Best Practices in e-Assessment. *Journal of Information Technology Education*, *5*, 251–269.

Byrnes, R., & Ellis, A. (2006). The prevalence and characteristics of online assessment in Australian universities. *Australasian Journal of Educational Technology*, *22*(1), 104.

Cassady, J. C., & Gridley, B. E. (2005). The Effects of Online Formative and Summative Assessment on Test Anxiety and Performance. *Journal of Technology, Learning, and Assessment*, *4*(1). Retrieved from http://eric.ed.gov/?id=EJ848518

Challis, D. (2005). Committing to quality learning through adaptive online assessment. *Assessment & Evaluation in Higher Education*, *30*(5), 519–527. http://doi.org/10.1080/02602930500187030

Cheung, R., & Vogel, D. (2013). Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning. *Computers & Education*, *63*, 160–175. http://doi.org/10.1016/j.compedu.2012.12.003

Chew, E., Jones, N., & Blackey, H. (2010). Implementing Institutional Online Assessment - Addressing the Challenges. In P. Tsang, S. Cheung, V. Lee, & R. Huang (Eds.), *Hybrid Learning* (Vol. 6248, pp. 453–464). Springer Berlin / Heidelberg. Retrieved from http://www.springerlink.com/content/0802v72064121055/abstract/

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In *Modern methods for business research* (pp. 295–336). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Chuttur, M. (2009). Overview of the Technology Acceptance Model: Origins, Developments and Future Directions. *Working Papers on Information Systems*, *9*(37). Retrieved from http://sprouts.aisnet.org/9-37

Clarke-Midura, J., & Dede, C. (2010). Assessment, Technology, and Change. *Journal of Research on Technology in Education*, *42*(3), 309–328.

Cooper, S., & Sahami, M. (2013). Reflections on Stanford's MOOCs. *Commun. ACM*, *56*(2), 28–30. http://doi.org/10.1145/2408776.2408787

Crippen, K. J., & Brooks, D. W. (2002). An Analysis of Student Learning at a Testing Web Site Emphasizing Descriptive Chemistry. *Journal of Computers in Mathematics and Science Teaching*, *21*(2), 183–201.

Cross, K. P. (1996). Improving teaching and learning through classroom assessment and classroom research (pp. 3–10). Presented at the Improving Student Learning: using research to improve student learning, Oxford, Oxford Centre for Staff Development: Gibbs G. (Ed.).

Crumpacker, N. (2001). Faculty Pedagogical Approach, Skill, and Motivation in Today's Distance Education Milieu. *Online Journal of Distance Learning Administration*, *4*(4). Retrieved from http://www.westga.edu/~distance/ojdla/winter44/crumpacker44.html

Davis F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. http://doi.org/10.2307/249008

Davis F. D. (1993). User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies*, *38*(3), 475–487. http://doi.org/10.1006/imms.1993.1022

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003. http://doi.org/10.1287/mnsc.35.8.982

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology*, *22*(14), 1111–1132. http://doi.org/10.1111/j.1559-1816.1992.tb00945.x

Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: three experiments. *International Journal of Human-Computer Studies*, *45*(1), 19–45. http://doi.org/10.1006/ijhc.1996.0040

Dermo, J. (2009). e-Assessment and the student learning experience. A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, *40*(2), 203–214. http://doi.org/10.1111/j.1467-8535.2008.00915.x

Dettori, G., Garuti, R., & Lemut, E. (2002). From Arithmetic to Algebraic Thinking by Using a Spreadsheet. In R. Sutherland, T. Rojano, A. Bell, & R. Lins (Eds.), *Perspectives on School Algebra* (pp. 191–207). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/0-306-47223-6_11

Domino, G., & Domino, M. L. (2006). *Psychological testing: An introduction*. Cambridge University Press Cambridge.

Dreher, C., Reiners, T., & Dreher, H. (2011). Investigating Factors Affecting the Uptake of Automated Assessment Technology. *Journal of Information Technology Education*, *10*.

EDUCAUSE Learning Initiative. (2014). EDUCAUSE. Retrieved May 21, 2014, from http://www.educause.edu/

Ellis, R., & Goodyear, P. (2010). *Students' Experiences of e-Learning in Higher Education: The Ecology of Sustainable Innovation*. Routledge.

Fishbein M., & Ajzen, I. (1975). Belief, attitude, intension and behavior: An introduction to theory and research. MA: Addison Wesley, Reading,.

Fluck, A. (2010). eExams–high stakes summative assessment for the digital age. In *Cultures of Learning. Proceedings Teaching Matters 2010 Conference* (pp. 24–33).

Fornell, C., & Bookstein, F. L. (1982). Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory. *Journal of Marketing Research*, *19*(4), 440–452. http://doi.org/10.2307/3151718

Fornell, C., & Larcker, D. F. (1981). Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research*, *18*(3), 382–388. http://doi.org/10.2307/3150980

Galbraith, P., & Haines, C. (1998). Disentangling the Nexus: Attitudes to Mathematics and Technology in a Computer Learning Environment. *Educational Studies in Mathematics*, *36*(3), 275–290. http://doi.org/10.1023/A:1003198120666

Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information Systems*, *16*(1), 5.

Gibbs, G., & Lucas, L. (1996). "Using research to improve student learning in large classes" in using research to improve student learning. In *Proceedings of the 3rd International Improving Student Learning Symposium.* Edited by: Gibbs, G. Oxford: Oxford Centre for Staff Development.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, *1*(1), 3–31.

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, *57*(4), 2333–2351. http://doi.org/10.1016/j.compedu.2011.06.004

Gill, M., & Greenhow, M. (2008). How effective is feedback in Computer-Aided Assessments? *Learning, Media and Technology*, *33*, 207–220. http://doi.org/10.1080/17439880802324145

GitHub moodle-qtype_stack. (2014). Stack question type for Moodle. Retrieved November 19, 2014, from https://github.com/maths/moodle-qtype_stack

Gong, M., Xu, Y., & Yu, Y. (2004). An enhanced technology acceptance model for web-based learning. *Journal of Information Systems Education*, *15*(4), 365–374.

Götz, O., Liehr-Gobbers, K., & Krafft, M. (2010). Evaluation of Structural Equation Models

Using the Partial Least Squares (PLS) Approach. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of Partial Least Squares* (pp. 691–711). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/978-3-540-32827-8_30

Gray, L. (2008). Effective Practice with e-Portfolios - JISC. JISC. Retrieved from http://www.jisc.ac.uk/media/documents/publications/effectivepracticeeportfolios.pdf

Gunasekaran, A., McNeil, R. D., & Shaul, D. (2002). E-learning: research and applications. *Industrial and Commercial Training*, *34*(2), 44–53. http://doi.org/10.1108/00197850210417528

Gúzman, J. C., & Hernández, G. (1993). Implicaciones educativas de seis teorías psicológicas. Facultad de Psicología, UNAM.

Hadley, M., & Sheingold, K. (1993). Commonalities and Distinctive Patterns in Teachers' Integration of Computers. *American Journal of Education*, *101*(3), 261–315.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7 th ed.). Upper Saddle River, NJ: Prentice Hall.

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2014). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. SAGE Publications.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance. *Long Range Planning*, *46*(1–2), 1–12. http://doi.org/10.1016/j.lrp.2013.01.001

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The Use of Partial Least Squares Structural Equation Modeling in Strategic Management Research: A Review of Past Practices and Recommendations for Future Applications. *Long Range Planning*, *45*(5–6), 320–340. http://doi.org/10.1016/j.lrp.2012.09.008

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414–433. http://doi.org/10.1007/s11747-011-0261-6

HEFCE. (2005). Strategy for e-learning. Retrieved May 22, 2014, from http://webarchive.nationalarchives.gov.uk/20100202100434/http://www.hefce.ac.uk/pubs/hefce/2005/05_12/

HEFCE. (2014). Higher Education Funding Council for England. Retrieved May 21, 2014, from http://www.hefce.ac.uk/

Heinrich, E., Milne, J., & Moore, M. (2009). An Investigation into E-Tool Use for Formative Assignment Assessment--Status and Recommendations. *Educational Technology & Society*, *12*(4), 176–192.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (Eds.). (2009). The use of partial least squares path modeling in international marketing. In *Advances in International Marketing* (Vol. 20). Bingley: Emerald Group Publishing. Retrieved from https://www.escholar.manchester.ac.uk/uk-ac-man-scw:89555

Higher Education Academy. (2010). Assessment & Feedback. Retrieved June 11, 2014, from http://www.heacademy.ac.uk/assessment

Hodgson, P., & Pang, M. Y. C. (2012). Effective formative e-assessment of student learning:

a study on a statistics course. *Assessment & Evaluation in Higher Education*, *37*(2), 215–225. http://doi.org/10.1080/02602938.2010.523818

Howell, S. L., Saba, F., Lindsay, N. K., & Williams, P. B. (2004). Seven strategies for enabling faculty success in distance education. *The Internet and Higher Education*, *7*(1), 33–49. http://doi.org/10.1016/j.iheduc.2003.11.005

Huang, H.-M., & Liaw, S.-S. (2005). Exploring users' attitudes and intentions toward the web as a survey tool. *Computers in Human Behavior*, *21*(5), 729–743. http://doi.org/10.1016/j.chb.2004.02.020

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *Strategic Management Journal*, *20*(2), 195–204. http://doi.org/10.1002/(SICI)1097-0266(199902)20:2<195::AID-SMJ13>3.0.CO;2-7

Hung, D. (2001). Theories of Learning and Computer-Mediated Instructional Technologies. *Educational Media International*, *38*(4), 281–287. http://doi.org/10.1080/09523980110105114

Iannone, P., & Simpson, A. (2013). Students' perceptions of assessment in undergraduate mathematics. *Research in Mathematics Education*, *15*(1), 17–33. http://doi.org/10.1080/14794802.2012.756634

Ilias, A., Razak, M. Z. A., & Yasoa, M. R. (2009). Taxpayers' attitude in using e-Filing system: Is there any significant difference among demographic factors? *Journal of Internet Banking and Commerce*, *14*(1).

JISC. (2007). *Effective practice with e-assessment*. Retrieved from http://www.jisc.ac.uk/publications/programmerelated/2007/pub_eassesspracticeguide.aspx

JISC e-Learning programme. (2012). JISC. Retrieved May 27, 2014, from http://www.jisc.ac.uk/whatwedo/programmes/elearning.aspx

*Jöreskog, K. G., Sörbom, D., du Toit, S., du Toit, M.,*. (1999). [LISREL 8: New statistical features. Chicago, IL: Scientific Software].

Joyes, G., Gray, L., & Hartnell-Young, E. (2010). Effective practice with e-portfolios: How can the UK experience inform implementation? *Australasian Journal of Educational Technology*, *26*(1), 15–27.

Kaklauskas, A., Zavadskas, E. K., Pruskus, V., Vlasenko, A., Seniut, M., Kaklauskas, G., … Gribniak, V. (2010). Biometric and Intelligent Self-Assessment of Student Progress system. *Computers & Education*, *55*(2), 821–833. http://doi.org/10.1016/j.compedu.2010.03.014

Kellough, R. D., & Kellough, N. G. (1998). *Secondary School Teaching:a Guide to Methods and Resources, Planning for Competence: A Guide to Methods and Resources : Planning for Competence* (1 edition). Upper Saddle River, N.J: Pearson.

King, C. G., Guyette, R., & Piotrowski, C. (2009). Online exams and cheating: An empirical analysis of business students' views. *The Journal of Educators Online*, *6*(1), 1–11.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. http://doi.org/10.1207/s15430421tip4104_2

Larson, M. R., & Bruning, R. (1996). Participant perceptions of a collaborative satellite-based mathematics course. *American Journal of Distance Education*, *10*(1), 6–22.

http://doi.org/10.1080/08923649609526906

Lee, M. C. (2009). Factors influencing the adoption of internet banking: An integration of TAM and TPB with perceived risk and perceived benefit. *Electronic Commerce Research and Applications*, *8*(3), 130–141. http://doi.org/10.1016/j.elerap.2008.11.006

Lee, Y., Kozar, K. A., & Larsen, K. R. (2003). The technology acceptance model: past, present, and future. *Communications of the Association for Information Systems*, *12*(1), 50.

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management*, *40*(3), 191–204. http://doi.org/10.1016/S0378-7206(01)00143-4

Liaw, S.-S., & Huang, H.-M. (2013). Perceived satisfaction, perceived usefulness and interactive learning environments as predictors to self-regulation in e-learning environments. *Computers & Education*, *60*(1), 14–24. http://doi.org/10.1016/j.compedu.2012.07.015

Liaw, S.-S., Huang, H.-M., & Chen, G.-D. (2007). Surveying instructor and learner attitudes toward e-learning. *Computers & Education*, *49*(4), 1066–1080. http://doi.org/10.1016/j.compedu.2006.01.001

Lifelong Learning UK, (LLUK). (2010). Lifelong Learning UK (LLUK). Retrieved May 21, 2014, from http://www.lifelonglearning.co.uk/

Lim, C. P., & Khine, M. (2006). Managing Teachers' Barriers to ICT Integration in Singapore Schools. *Journal of Technology and Teacher Education*, *14*(1), 97–125.

Liu, S.-H., Liao, H.-L., & Pratt, J. A. (2009). Impact of media richness and flow on e-learning technology acceptance. *Computers & Education*, *52*(3), 599–607. http://doi.org/10.1016/j.compedu.2008.11.002

López, L. (2009). *Effects of delayed and immediate feedback in the computer-based testing environment* (Ph.D.). Indiana State University, United States -- Indiana. Retrieved from http://search.proquest.com/docview/304900802/abstract?accountid=12253

Maclellan, E. (2004). How convincing is alternative assessment for use in higher education? *Assessment & Evaluation in Higher Education*, *29*(3), 311–321. http://doi.org/10.1080/0260293042000188267

Martell, K. D., & Calderon, T. G. (2005). *Assessment of student learning in business schools: Best practices each step of the way* (Vol. 1). Association for Institutional Research (AIR).

McCann, A. L. (2010). Factors affecting the adoption of an e-assessment system. *Assessment & Evaluation in Higher Education*, *35*(7), 799–818. http://doi.org/10.1080/02602930902981139

McCollum, K. (1997). A Professor Divides His Class in Two to Test Value of On-Line Instruction. *Chronicle of Higher Education*, *43*(24). Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ541256

Miller, T. (2009). Formative computer-based assessment in higher education: the effectiveness of feedback in supporting student learning. *Assessment & Evaluation in Higher Education*, *34*(2), 181–192. http://doi.org/10.1080/02602930801956075

Ming Chi, L. (2010). Explaining and predicting users continuance intention toward e-learning:

An extension of the expectation–confirmation model. *Computers & Education*, *54*(2), 506–516. http://doi.org/10.1016/j.compedu.2009.09.002

Moon, J.-W., & Kim, Y.-G. (2001). Extending the TAM for a World-Wide-Web context. *Information & Management*, *38*(4), 217–230. http://doi.org/10.1016/S0378-7206(00)00061-6

National Polytechnic Institute, I. (2015). Instituto Politécnico Nacional. Retrieved March 28, 2015, from http://www.ipn.mx/english/Paginas/Inicio.aspx

National Student Forum-Annual Report. (2009). National Student Forum-Annual Report. Retrieved from http://webarchive.nationalarchives.gov.uk/20100519174430/nationalstudentforum.com/report2009/

Nickerson, R. S., Perkins, D. N., & Smith, E. E. (1985). *The Teaching of Thinking*. Hillsdale, N.J: Routledge.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*, 199–218. http://doi.org/10.1080/03075070600572090

NMC Horizon Report. (2014). Horizon Report-2014. Retrieved May 21, 2014, from http://www.nmc.org/publications/2014-horizon-report-higher-ed

Nunnally, J. C. (1978). *Psychometric theory* (2nd. ed.). McGraw-Hill.

Ocak, M. A. (2011). Why are faculty members not teaching blended courses? Insights from faculty members. *Computers & Education*, *56*(3), 689–699. http://doi.org/10.1016/j.compedu.2010.10.011

OCDE. (2012). *Perspectivas OCDE: México Reformas para el Cambio*. OCDE. Retrieved from www.oecd.org/centrodemexico

Paraskeva, F., Bouta, H., & Papagianni, A. (2008). Individual characteristics and computer self-efficacy in secondary education teachers to integrate technology in educational practice. *Computers & Education*, *50*(3), 1084–1091. http://doi.org/10.1016/j.compedu.2006.10.006

Peltenburg, M., Van Den Heuvel-Panhuizen, M., & Doig, B. (2009). Mathematical power of special-needs pupils: An ICT-based dynamic assessment format to reveal weak pupils' learning potential1. *British Journal of Educational Technology*, *40*(2), 273–284. http://doi.org/10.1111/j.1467-8535.2008.00917.x

Piaget, J. (1977). *The development of thought: Equilibrium of cognitive structures*. New York: Viking.

Piccoli, G., Ahmad, R., & Ives, B. (2001). Web-Based Virtual Learning Environments: A Research Framework and a Preliminary Assessment of Effectiveness in Basic IT Skills Training. *MIS Quarterly*, *25*(4), 401. http://doi.org/10.2307/3250989

Pintrich, P. R., & Zusho, A. (2002). Student Motivation and Self-Regulated Learning in the College Classroom. In J. C. Smart & W. G. Tierney (Eds.), *Higher Education: Handbook of Theory and Research* (pp. 55–128). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-010-0245-5_2

Purvis, A. J., Aspden, L. J., Bannister, P. W., & Helm, P. A. (2011). Assessment strategies to support higher level learning in blended delivery. *Innovations in Education and*

*Teaching International*, *48*, 91–100. http://doi.org/10.1080/14703297.2010.543767

Pynoo, B., Tondeur, J., van Braak, J., Duyck, W., Sijnave, B., & Duyck, P. (2012). Teachers' acceptance and use of an educational portal. *Computers & Education*, *58*(4), 1308–1317. http://doi.org/10.1016/j.compedu.2011.12.026

Reed, H. C., Drijvers, P., & Kirschner, P. A. (2010). Effects of attitudes and behaviours on learning mathematics with computer tools. *Computers & Education*, *55*(1), 1–15. http://doi.org/10.1016/j.compedu.2009.11.012

Reiners, T., Dreher, C., & Dreher, H. (2011). Six Key Topics for Automated Assessment Utilisation and Acceptance. *Informatics in Education*, *10*(1), 47–64.

Ringle, C. ., Wende, S., & Will, A. (2005). *SmartPLS software "SmartPLS 2.0," www.smartpls.de*. Hamburg. Retrieved from http://www.pls-sem.com/

Rosen, L. D., & Weil, M. M. (1995). Computer anxiety: A cross-cultural comparison of university students in ten countries. *Computers in Human Behavior*, *11*(1), 45–64. http://doi.org/10.1016/0747-5632(94)00021-9

Sánchez-Franco, M. J., Peral-Peral, B., & Villarejo-Ramos, Á. F. (2014). Users' intrinsic and extrinsic drivers to use a web-based educational environment. *Computers & Education*, *74*, 81–97. http://doi.org/10.1016/j.compedu.2014.02.001

Sangwin, C. (2013). *Computer Aided Assessment of Mathematics*. Oxford University Press.

Schunk, D. (2012). *Learning theories : an educational perspective* (6th ed.). Upper Saddle River  N.J.: Pearson/Merrill/Prentice Hall.

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, *35*(4), 453–472. http://doi.org/10.1023/A:1003196224280

Scriven, M. (1991). *Evaluation Thesaurus*. SAGE Publications.

Segers, M., Dochy, F., & Cascallar, E. (2003). The Era of Assessment Engineering: Changing Perspectives on Teaching and Learning and the Role of New Modes of Assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising New Modes of Assessment: In Search of Qualities and Standards* (pp. 1–12). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/0-306-48125-1_1

Selim, H. M. (2007). Critical success factors for e-learning acceptance: Confirmatory factor models. *Computers & Education*, *49*(2), 396–413. http://doi.org/10.1016/j.compedu.2005.09.004

Shook, C. L., Ketchen, D. J., Cycyota, C. S., & Crockett, D. (2003). Data analytic trends and training in strategic management. *Strategic Management Journal*, *24*(12), 1231–1237. http://doi.org/10.1002/smj.352

Sim, G., Holifield, P., & Brown, M. (2004). Implementation of Computer Assisted Assessment: Lessons from the Literature. *ALT-J: Research in Learning Technology*, *12*(3), 215–229.

Skinner, B. F. (1976). *About Behaviorism* (1 edition). New York: Vintage.

Smith, G., & Wood, L. (2000). Assessment of learning in university mathematics. *International Journal of Mathematical Education in Science and Technology*, *31*(1), 125–132. http://doi.org/10.1080/002073900287444

Stödberg, U. (2012). A research review of e-assessment. *Assessment & Evaluation in Higher*

*Education*, *37*(5), 591–604. http://doi.org/10.1080/02602938.2011.557496

Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & Education*, *50*(4), 1183–1202. http://doi.org/10.1016/j.compedu.2006.11.007

Teo, T. (2006). Attitudes toward computers: A study of post-secondary students in Singapore. *Interactive Learning Environments*, *14*(1), 17–24. http://doi.org/10.1080/10494820600616406

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, *52*(2), 302–312. http://doi.org/10.1016/j.compedu.2008.08.006

Teo, T. (2010). A path analysis of pre-service teachers' attitudes to computer use: applying and extending the technology acceptance model in an educational context. *Interactive Learning Environments*, *18*(1), 65–79. http://doi.org/10.1080/10494820802231327

Teo, T. (2012). Examining the intention to use technology among pre-service teachers: an integration of the Technology Acceptance Model and Theory of Planned Behavior. *Interactive Learning Environments*, *20*(1), 3–18. http://doi.org/10.1080/10494821003714632

Teo, T., Lee, C. b., & Chai, C. s. (2008). Understanding pre-service teachers' computer attitudes: applying and extending the technology acceptance model. *Journal of Computer Assisted Learning*, *24*(2), 128–143. http://doi.org/10.1111/j.1365-2729.2007.00247.x

Teo, T., & Noyes, J. (2011). An assessment of the influence of perceived enjoyment and attitude on the intention to use technology among pre-service teachers: A structural equation modeling approach. *Computers & Education*, *57*(2), 1645–1653. http://doi.org/10.1016/j.compedu.2011.03.002

Terzis, V., & Economides, A. A. (2011). The acceptance and use of computer based assessment. *Computers & Education*, *56*(4), 1032–1044. http://doi.org/10.1016/j.compedu.2010.11.017

Terzis, V., Moridis, C. N., & Economides, A. A. (2012). The effect of emotional feedback on behavioral intention to use computer based assessment. *Computers & Education*, *59*(2), 710–721. http://doi.org/10.1016/j.compedu.2012.03.003

Terzis, V., Moridis, C. N., Economides, A. A., & Mendez, G. R. (2013). Computer Based Assessment Acceptance: A Cross-cultural Study in Greece and Mexico. *Educational Technology & Society*, *16*(3), 411–424.

The New Media Consortium, NMC. (2014). The New Media Consortium. Retrieved May 21, 2014, from http://www.nmc.org/

Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal Computing: Toward a Conceptual Model of Utilization. *MIS Quarterly*, *15*(1), 125–143. http://doi.org/10.2307/249443

Ursini, S., Sánchez, G., & Orendain, M. (2004). Validación y confiabilidad de una escala de actitudes hacia las matemáticas y hacia las matemáticas enseñadas con computadora. *Educación Matemática*, *16*(3), 59–78.

Van Harmelen, M. (2012). *CETIS Analytics Series: Analytics for Understanding Research* (p. 41). (JISC CETIS Analytics Series). Retrieved from http://publications.cetis.ac.uk/2012/518

Venkatesh, V. (2000). Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Information Systems Research*, *11*(4), 342.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204. http://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478.

Volery, T., & Lord, D. (2000). Critical success factors in online education. *International Journal of Educational Management*, *14*(5), 216–223. http://doi.org/10.1108/09513540010344731

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Walet, N., & Birch, M. (2012). Using online assessment to provide instant feedback. In *Proceedings of the HEA STEM Learning and Teaching Conference (2012)* (Vols. 1–0). The Higher Education Academy. Retrieved from http://journals.heacademy.ac.uk/doi/abs/10.11120/stem.hea.2012.095

Walker, D. J., Topping, K., & Rodrigues, S. (2008). Student reflections on formative e-assessment: expectations and perceptions. *Learning, Media and Technology*, *33*(3), 221–234. http://doi.org/10.1080/17439880802324178

Wang, W.-T., & Wang, C.-C. (2009). An empirical study of instructor adoption of web-based learning systems. *Computers & Education*, *53*(3), 761–774. http://doi.org/10.1016/j.compedu.2009.02.021

Warburton, B. (2009). Quick win or slow burn: modelling UK HE CAA uptake. *Assessment & Evaluation in Higher Education*, *34*(3), 257–272. http://doi.org/10.1080/02602930802071080

Warkentin, M., Gefen, D., Pavlou, P. A., & Rose, G. M. (2002). Encouraging citizen adoption of e-government by building trust. *Electronic Markets*, *12*(3), 157–162.

Webster, J., & Hackley, P. (1997). Teaching Effectiveness in Technology-Mediated Distance Learning. *The Academy of Management Journal*, *40*(6), 1282–1309. http://doi.org/10.2307/257034

Whitelock, D. M. (2009). Editorial: e-assessment: developing new dialogues for the digital age. *British Journal of Educational Technology*, *40*(2), 199–202. http://doi.org/10.1111/j.1467-8535.2008.00932.x

Whitelock, D. M., & Brasher, A. (2006). Developing a Roadmap for e-Assessment: Which Way Now? In M. Danson (Ed.), *Proceedings of the 10th CAA International Computer Assisted Assessment Conference* (pp. 487–501). Loughborough, UK. Retrieved from http://oro.open.ac.uk/11950/

Whitelock, D. M., Mackenzie, D., Whitehouse, C., Ruedel, C., & Rae, S. (2006). Identifying innovative and effective practice in e-assessment: findings from seventeen UK case studies. In M. Danson (Ed.), *Proceedings of the 10th CAA International Computer Assisted Assessment Conference* (pp. 505–511). Loughborough, UK. Retrieved from http://oro.open.ac.uk/11949/

Whitelock, D. M., & Watt, S. (2008). Reframing e-assessment: adopting new media and adapting old frameworks. *Learning, Media and Technology*, *33*, 151–154. http://doi.org/10.1080/17439880802447391

Whitelock, D., & Raw, Y. (2003). Taking an electronic mathematics examination from home: what the students think. In *Computer Based Learning in Science, New Technologies and their Applications in Education* (Vol. 1, pp. pp. 701–713). Nicosia: Department of Educational Sciences, University of Cyprus, Cyprus,: In C.P. Constantinou and Z.C. Zacharia (eds.). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary;jsessionid=BB0BD94AD163923F458E4E8E255C6C17?doi=10.1.1.372.3458

Williams, L. J., Vandenberg, R. J., & Edwards, J. R. (2009). Structural Equation Modeling in Management Research: A Guide for Improved Analysis. *The Academy of Management Annals*, *3*(1), 543–604. http://doi.org/10.1080/19416520903065683

Williams, P. (2002). The Learning Web The Development, Implementation and Evaluation of Internet-Based Undergraduate Materials for the Teaching of Key Skills. *Active Learning in Higher Education*, *3*(1), 40–53. http://doi.org/10.1177/1469787402003001004

Williams, P. (2008). Assessing context-based learning: not only rigorous but also relevant. *Assessment & Evaluation in Higher Education*, *33*, 395–408. http://doi.org/10.1080/02602930701562890

Wilson, K., Boyd, C., Chen, L., & Jamal, S. (2011). Improving student performance in a first-year geography course: Examining the importance of computer-assisted formative assessment. *Computers & Education*, *57*(2), 1493–1500. http://doi.org/10.1016/j.compedu.2011.02.011

Wu, I.-L., & Chen, J.-L. (2005). An extension of Trust and TAM model with TPB in the initial adoption of on-line tax: An empirical study. *International Journal of Human-Computer Studies*, *62*(6), 784–808. http://doi.org/10.1016/j.ijhcs.2005.03.003