Intelligent Diagnostic Feedback In Virtual Learning Environment

Ruisheng Guo

A Thesis submitted to School of Computing and Digital Media,

London Metropolitan University

for the degree of Doctor of Philosophy in Computing

2019

Acknowledgement

First of all, I would like to extend my sincere gratitude to my supervisors, Prof. Dominic Palmer-Brown, Prof. Hassan Kazemian and Dr. Fangfang Cai, for their helpful support, serious response, instructive advice, and useful suggestions on my research. I am deeply grateful for their help in both studying and living. They have given me great help and support in learning, research, work, and life continually. Especially, Prof. Hassan Kazemain, he helped me when I was facing the most difficult time in my life. There was a serious car accident happened, my parents were serious injured and one of their best friend died. My family spend about 1 million RMB on this accident. I spend more than half year to help them move out from it. Prof. Hassan Kazemian helped me put. He gave me some valuable suggestion and support to keep me continuing my PhD.

I am also deeply indebted to all the other supervisors, Dr Sin Wee Lee, and Dr. Christina, Dragonova for their direct and indirect help to me. It would not be possible to carry out my research without their guiding, encouragement and supervision.I also would like to thank London Metropolitan University for providing me very strong support in both studying and living. It has been my pleasure to study at Londonmet University, and my favourite study experience of my life.

Ultimately, my thanks go to my beloved family for their loving considerations and great confidence in me all through these years. I also owe my sincere gratitude to my friends and my colleagues who gave me their help and time in listening to me and helping me work out my problems during the difficult times.

Abstract

This research investigates the current key issues in the area of e-learning in higher education and focuses on how to provide automatic, effective and convenient online feedback to students in order to support students' learning. In recent years, e-learning has becoming increasingly commonplace in higher education. On the other hand, according to the National Student Survey (NSS) reports (2007-2010), in England, about half of students and 35% of students (2011 -2014), and 30% of students (2015-2016) did not agreed with that: 1, feedback on their work has been prompt; 2, feedback on their work has helped them clarify things they did not understand; 3, they have received detailed comments on their work. These reports reveal that the feedback and its related fields are one of the weakest areas in higher education in England.

This research compares and contrasts several methods in order to investigate the effective use of intelligent feedback towards modelling the stages of students' learning. The work explores the potential benefits of integrating an artificial neural network (ANN) into a Virtual Learning Environment (VLE) system as a means of identifying grouping together of students who would benefit from the same feedback. It investigates the relative effectiveness of different types of feedback and how to optimize the feedback to maximize the facilitation of learning. It explores the ability of neural networks and data analysis techniques to model the stages of students' learning. The research also assesses the difference in the progress of students' learning with and without using intelligent diagnostic feedback. The E-learning Snap-Drift Neural Network (ESDNN) is evaluated as one of the potential tools for providing diagnostic, and effective feedback. The ESDNN is enhanced following the first trial, and the enhanced ESDNN system is introduced to the MCQs-Online Feedback System (M-OFS). Four hypothesis are formulated as follows: 1, during the trials, students improved

their understanding by reading given feedback; 2, after using M-OFS system, students get a higher mark in a separate paper test than before; 3, the students who used the system gained higher marks in the final examination than those who did not use the system; 4, students are satisfied with this system. Several trials are conducted in order to evaluate the approach and the system.

The findings are analyzed and lead to the conclusion that under certain conditions online diagnostic feedback is an effective means of enhancing student learning across a wide range of subject.

Table of Contents

Table of Figures

Table of Tables

Table of Charts

Table of Abbreviations

Artificial Neural Network (ANN)
Adaptive Resonance Theory (ART)
distributed Snap-Drift Neural Network (dSDNN)
E-learning Snip-Drift Neural Network (ESNN)
Hypothesis (H)
Intelligent Tutoring Systems (ITSs)
Java Server Faces (JSF)
Jinqiao University (JQU)
Kunming Technology University (KTU)
Learning Vector Quantization (LVQ)
Multiple Choice Questions (MCQs)
Multiple-Choice Questions Online Feedback System (M-OFS)
Model-View-Controller (MVC)
National Student Survey (NSS)
University of East London (UEL)
Virtual Learning Environment (VLE)

1. Introduction

1.1. Research Area

This research investigates the current key issues in the area of e-learning in higher education and focuses on how to provide an intelligent, diagnostic, automatic, instant, and effective feedback system to support students learning.

In recent years, e-learning has becoming increasingly commonplace in higher education. The involvement of intelligent e-learning system enhances the accessibility of higher education to more people with increasing convenience, efficiency and quality of learning. Moreover, the provision of effective feedback to students is the subject being studied.

1.2. Problem in this Area

On the other hand, according to the National Student Survey (NSS) reports (2007-2010), in England, only about half of students agreed with that: 1, feedback on their work has been prompt; 2, feedback on their work has helped them clarify things they did not understand; 3, they have received detailed comments on their work. And according to NSS (2011 to 2014), in England, only about 65% students thought that: 1, feedback on their work has been prompt; 2, feedback on their work has helped them clarify things they did not understand; 3, they have received detailed comments on their work. And according to the NSS (2015 to 2016), in England, there was still about 30% students did not think that: 1, feedback on their work has been prompt; 2, feedback on their work has helped them clarify things they did not understand; 3, they have received detailed comments on their work. These reports reveal that the feedback and its related fields are one of the weakest areas in higher education in England.

1.3. About this research

The previous work had been studied by Palmer-Brown, Lee and Draganova (2008), and Lee, S.W., Palmer-Brown, D., Draganova, C., Preston, D. and Kretsis, M. (2009). The E-learning Snip-Drift Neural Network (ESNN) was designed and developed, and few small trial was applied. The initial and simple work is to provide students intelligent feedback, and modeling the stages of students' knowledge.

One Journal paper and two conference paper had been published based. All of the experiments, trials, and findings are the original work of this research. Furthermore, the original work of this research is to continue and further investigate the relative effectiveness of different types of feedback, and how to optimize the feedback to maximize the facilitation of learning. Several trials are applied to a large cohort of students in different country and across different subjects. It will also further compare and contrast several methods in order to investigate the effectiveness of using intelligent feedback that involves modelling the stages of students' knowledge acquisition. A Multiple Choice Questions (MCQs) online feedback system is explored as one of the effective tools for providing significant diagnostic feedback to learners.

Moreover, this research also explores the potential benefits of integrating novel artificial neural network (ANN) models into a Virtual Learning Environment (VLE) system. A series of trials of this system are applied to a range of academic fields in order to investigate the adaptability of the system. This research also explores the potential benefits of understanding the learning behaviour.

Furthermore, a survey of learner satisfaction of using the system is designed and conducted after each trial for evaluating the system. It explores the ability of the novel VEL tool and related data analysis techniques to model the stages of students' knowledge. The research also assesses the difference in the progress of students' learning with and without using intelligent diagnostic feedback.

1.4.Structure of the Thesis

This thesis is structured as follows. Chapter 2 establish the importance of this research by reviewing the previous work and background. Chapter 3 Meteorology: introduces the Multiple-Choice Questions Online Feedback System (M-OFS), a neural network system that supports the learning process based on intelligent diagnostic feedback. And the approach of this research, the hypothesis assumed, and the instruments used, the data collected, and the case study are presented. Chapter 4 Summary of Multi-subject Experimental Trial: introduce the trials in details and show how many works had been carried out. Chapter 5 Findings: describes the experiments that are applied to students with using M-OFS, and discusses the results obtained from testing the hypothesis. Furthermore, analyzes and discusses the learning behaviours, behavioural groups, and knowledge state transactions. Finally, Chapter 6 Conclusion: summarizes this research and lists the contributions to knowledge, and outlines future work.

1.5.Research Objectives

The objectives of the research can be summarised into the following ways:

1. To investigate the different types of effective feedback in the context of online multiple-choice questions (MCQs); and to investigate the effectiveness of intelligent feedback for modelling the stages of students learning.

2. To investigate the learner's behaviours and to model the knowledge state transitions that occurs during learning. In particular to further our understanding of the kind of errors learners are making, the knowledge state transitions during learning, and of the learner's behaviours.

3. To apply and analyze the effects of providing an intelligent, automatic and diagnostic feedback in modelling the phases of students' learning and to make comparisons with the students' process without the feedback tool. To establish a mechanism that can optimise feedback to maximise the facilitation of learning.

4. To evaluate learner performance and learner satisfaction.

5. To produce an understanding of the potential of the on-line diagnostic feedback approach across different subject areas

2. Literature Review

2.1. Introduction

In this chapter, the current states of research in relevant aspects of E-learning, Artificial Intelligence, and Neural Networks is introduced. Then it further discusses the advantages and disadvantages of Multiple Choice Questions (MCQs) Feedback Systems, and the E-learning Snap-Drift Neural Network (ESDNN) including example and current researches. Finally, the potential advantages and limitations of the e-learning system ESDNN are summarised. Furthermore, the enhanced novel e-learning system M-OFS is introduced, and the two main aspects of this research are presented at the end of this chapter.

2.2. E-Learning

Since computer network came into ordinary people's life, and information and communication technologies were increasingly widely used, a number of scholars have been committing to research on how to use these technologies to help people to learn, which led to the raise of e-learning. Clark and Mayer (2007) defined e-learning as 'training delivered on a computer that is designed to support individual learning or organizational performance goals'. European Commission (2001) gives the definition of e-learning as 'the use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services, as well as remote exchange and collaboration'. According to Rossen and Hartley (2001), 'e-learning refers to anything delivered, enabled, or mediated by electronic technology for the explicit purpose of learning'. Many practitioners also use online learning, internet-based learning, or web-based learning as an alternative expression of e-learning (Capper, 2001).

2.2.1. Development and Application of E-learning

Recalling the development and application of e-learning in the most recent decades, e-learning technologies have become an indispensable teaching and training means, especially in higher education areas. As Kanuka and Kelland (2008) state, as the rapid development of internet and web-based communication technologies, 'higher education has moved into a third decade of change in how courses and programs are designed and delivered'.

A large number of education institutions across the word have been paying more attention on e-learning and start to adopt e-learning technologies (Goldberg, 1997b; Hsu and Backhouse, 2001 cited by Ifinedo, 2006). Moreover, lots of them have developed their own e-learning systems to facilitate their own teaching needs. For instance, the British Open University, whose development is regarded as the turning point of distance education (Bates, 2005), RMIT University, which spent AUS$50 million over the period 1999-2001 on aligning information technology to the needs of the core business of the university, and The University of Melbourne, which spent $12 million since 1997 for its development of multimedia enhanced teaching and learning (Alexander, 2001).

2.2.2. Strengths and Barriers of E-learning

Why has e-learning been used so widely and adopted so fast? It is basically because of some main strengths of e-learning. Firstly, e-learning makes the learning process more flexible, as students can study anytime anywhere they want and plan their own learning schedules (Malone, 2003). Many student surveys have pointed out that flexibility has become one of the best features of e-learning as well as one of the most important reasons for them to choose e-learning. For example, according to Bullen and Janes (2007), most respondents reported that the e-learning system was more flexible than the traditional learning model and the flexibility of e-learning system enhanced their learning experience. In addition, Pan, Zhang and Rhalibi et al. (2008) also stated that flexible learning time and self learning methods were reported by the students as the greatest advantages of their e-learning platform. Secondly, e-learning deliveries quality learning experience with efficient and cost-saving features (Honey, 2001). According to Li, Buhalis, Lockwood and Benzine (2007)'s survey, 7 out of 8 interviewees reported that they are very satisfied with the e-learning solutions since they have witnessed lower costs on travelling and training expenses in their organizations. Thirdly, e-learning improves access to education and training (Bates, 1997). As Archer, Garrison and Anderson (1999) said, e-learning technologies provide students an intense and immediate tool for learning. Rossen and Hartley (2001) also state that e-learning overcomes both time conflict and venue constraint by providing 24/7 access to training everywhere in the world. Other strengths of e-learning also include that e-learning can develop students' time management skills, research, writing, and computer skills, make students accountable for his or her learning (Hurt, 2008), and offer students more fun and motivation in learning (Pan, Zhang and Rhalibi et al, 2008), and so on.

However, despite of the strengths, e-learning also has its disadvantages or barriers. Hurt (2008) summarized its main disadvantages as lack of face-to-face contact, requiring technology skills for both trainers and trainees, and internet infrastructure support. Other barriers also include the lack of culture for e-learning and limited accessibility to computers (Li, Buhalis, Lockwood and Benzine, 2007).

Therefore, there are still many obstacles that need to be overcome to apply e-learning successfully and maximize its strengths. However, the most essential technology issue is how to develop an effective e-learning system, which can conquer the communication, personalization and interaction problems to develop learners' skills and knowledge.

2.2.3. Existing literature on e-learning and the role of feedback

In the past, students' evaluations were based on campus-based courses where instructors used to provide feedback directly to the students, but with the advancement of e-learning technology, universities are exploring the use of online feedback systems to cater to the needs of those students who prefer to take classes at a distance through e-learning (Bangert, 2004). In the internet-based teaching and learning scenarios, the old concept of passive students has been minimized with the notion that learners are able to construct the knowledge and meaning from their personal experiences (Svinicki, 1999). Guo (2018) corroborates the same that it is not the direct support from teachers through which the learner is able to acquire knowledge, but the learner can construct knowledge by using suitable sources of learning. According to Martinez-Arguelles (2013), the past literature exhibits that giving or receiving feedback in an e-learning environment both from the students as well as teacher's perspective is full of problems. The reason being, the virtual classrooms are mostly associated with high ratio students-teacher relationship which put a significant burden on teacher to cope with the follow-up task to the distance students (Ley, 1999; Buchanan, 2000). The e-learning is facing many difficulties because teachers do to always receive verbal or physical cues from the students. Levy (2007) highlights that in e-learning, the rate of attrition is comparatively high. The research study conducted by Boyd (2008) concludes that the role of feedback comes on the top-priority when students are receiving distance education as they would never meet other students and instructors. In the context of students learning, the role of formative feedback assessment is undeniably critical (Hounsell, 2003; Knight & Yorke, 2003), but the latest trend on summative assessments and a significant reduction in feedback have provided negative effects on students' learning (Yorke & Longden, 2004). Gibbs (2006) argues that this situation has instigated the students to get away with their learning desires and concentrate only on getting good marks. Therefore, Gibbs & Simpsons (2004) formulated eleven conditions wherein they underline the effective provision and importance of feedback which should be provided to the students. These conditions emphasize that the feedback must be sufficient in detail and frequency, it should be directed to learning and not merely on getting high marks, it should be comprehendible and associated to the criteria of assessment, and the main motive of feedback is to be utilizable by the students so that they improve their academic work. On the other hand, Nicol and Dick (2006) formulated a list of seven principles relevant to good feedback practice with a focus on self-regulation of students during their learning process. These seven principles were focused on clarifying the standards and objectives of good performance, facilitating the reflection and self-assessment in learning, high-quality information delivery to students so that they self-correct themselves, promoting peer dialog and teacher-student dialog, promoting self-esteem and positive motivational belief, providing opportunities so that the student is able to act on the provided feedback, and lastly, providing teachers with appropriate information so that they shape their teaching practices accordingly. As a matter of fact, despite the important role of feedback in students' learning, there are numerous weaknesses in the provision of feedback system to date. Draper (2007) highlights it that the present feedback systems are not giving self-regulated learning to the students, rather students are getting entangled into more confusion. However, Draper (2007) highlights that under Nicol principles, the feedback is used by the students to regulate the content of their learning, whereas Gibbs principles focus on regulating the time and effort of the students. Yang & Cornelius (2004) analyzed the issues related to online learning quality for students through document collection, observations, and interviews. They concluded that students had negative experiences due to delayed feedback and they had negative experiences because of lack of self-motivation and self-regulation. All these literature findings corroborate the undeniably important role of feedback in the context of e-learning and a great need of further research so as to refine the role of feedback in the e-learning scenarios.

2.2.4. Key Factors of Effective E-Learning

In the knowledge rapid change age, in order to achieve success e-learning, Omoda-Onyait G., Lubega J.T., Maiga G. (2013) point out that there is a need to broaden what, when, where, and how students learn, and the rate at which they progress in achieving learning outcomes. Futhermore, Omoda-Onyait G., Lubega J.T., Maiga G. (2013) also point out that in e-learning environment, it requires to create a more effective interaction between e-learning content and learners.

To develop an effective e-learning system, several key factors of effective e-learning need to be emphasized. As Barron (2006) presented, successful e-learning should have

10 factors, which are: shared learning centered vision; comprehensive course design process; customized scoring guides; group work strategies; characteristics of effective facilitators; faculty training and support; great expectations; meaningful feedback; monitoring and evaluation; and continuous improvement. HEFCE (2004, cited by Sharma and Mishra, 2007) identified 6 key factors of effective e-learning: connectivity, interactivity, motivation, flexibility, collaboration, and extended opportunities. In addition, Villiers (2007) suggested using the Hexa-C model to design and evaluate the e-learning system, which includes 6 factors: cognitive learning, constructivism, components, creativity, customization, and collaborative learning.

Mulqueeny, K., Kostyuk, V., Baker, R.S. et al. (2015) stated that in order to attract students attention, the effective e-learning system should include several factors as follows: 1, developing deep learning; 2, minimizing cognitive load leads to improved engagement; 3, meaningful and useful outcomes.

Reviewing all these factors by different authors, combining with the major barriers of e-learning, we can find that an intelligent e-learning technology, which can achieve better communication, personalization and interaction to a certain extent, is highly desirable and crucial to the success of e-learning.

2.3. Intelligent Tutoring Systems (ITSs)

2.3.1. Advanced Features of ITSs

In order to satisfy the key factors of effective e-learning presented in various papers and overcome the barriers to e-learning, intelligent e-learning systems have been becoming more and more widely used as an advanced e-learning system and deeply investigated by many researchers (Kuri-Morales and Simari, 2010; Kordaki and Daradoumis, 2009; Ma, 2006). Graesser, A.C. et al., (2012) state that Intelligent Tutoring Systems (ITSs) are computerized learning environments that incorporate computational models from the cognitive sciences, learning sciences, computational linguistics, artificial intelligence, mathematics, and other fields.

Generally speaking, as an application of the advanced e-learning technology closely relating to artificial intelligence, web-based information systems and cognitive science, web-based ITSs have many advanced features basically because they try to play the role of the teacher in traditional classroom learning model, so that they can monitor and evaluate the student's performance, and give the student meaningful, constructive and adaptive feedback to improve the student's learning progress. Ma (2006) described ITSs as the milestone of the advanced generation of computer-aided instruction systems, and concluded their key feature as 'the ability to provide a user-adapted presentation of the teaching material'.

Rane and Sasikumar (2007) pointed out that to overcome the lack of the presence of a teacher, intelligent tutoring systems attempt to simulate a teacher, who can guide the student's study based on the student's level of knowledge by giving intelligent instructional feedback. Furthermore, according to Blessing, Gilbert, Ourada and Ritter (2007), the intense interaction and feedback achieved by intelligent tutoring systems can significantly improve student learning gains. In addition, in Gheorghiu and Vanlehn (2008)'s paper, they also suggested that meaningful, constructive and adaptive feedback is the essential feature of ITSs, and it is such feedback that helps students achieve strong learning gains.

All of these evidences provide a clear indication that web-based ITSs provide an effective e-learning technology that can significantly improve student learning gains mainly by offering the student customized and constructive feedback.

2.3.2. Attributes of Effective Feedback

Since feedback is the essential feature of ITSs and plays a very important role on promoting the effectiveness of e-learning and optimizing student learning experience as discussed above, it has drawn our attention to a great extent. Fallows & Ahmet, (1999) point out that the positive and encourage guiding feedback is one of the important methods for the educator.

In addition, Cullen *et al* (2002, cited by Heinze et al, 2007) suggested that feedback is one of the most important indicators of good education. Little (2001) presented that providing feedback, which can continuously reinforce learning progress and promote learners' attention and engagement, is crucial to effective learning. Besides, Vasilyeva, Pechenizkiy and Bra (2008) also stated that 'the design of feedback is a critical issue of online assessment development within web-based learning systems'.

However, reviewing the education reality, we surprisingly found that as one of the most important communication methods between teachers and students, feedback was one of the weakest areas according to the National Student survey (2007-2017) in the UK. Thus, how to design an optimized intelligent feedback system becomes a major and critical problem to develop a successful e-learning system.

According to Hyland (2000, cited by Hatziapostolou and Paraskakis, 2010), feedback is an essential component in all learning contexts, which has various functions including evaluating students' performance, developing students' competences and knowledge, elevating students' motivation and confidence, and so on. Formative feedback enables students to monitor and evaluate their own learning process (Özdener & Satar, 2009), identify their strengths and weaknesses (Brown, 1997), and guide them to achieve the leaning outcomes (Sadler, 1989, cited by Hatziapostolou and Paraskakis, 2010). Felix (2003) pointed that computer based automated feedback can provide several benefits to student as follows: 1, equal and fair individualised feedback anytime; 2, immediate and more frequent for guiding; 3, more independently provide feedback at any time any place for anyone. McIntyre and Wolff (1998) stated that: "One of the powers of interactivity in a web environment is the capability to engage by providing rapid, compelling interaction and feedback to students."

To make the feedback effective and meaningful, a range of quality attributes need to be achieved. Hatziapostolou and Paraskakis (2010) summarized the work by Race (2006), Irons (2008) and Juwah et al (2004) and suggested that in order to improve learning gains, formative feedback should address as many as possible of the following attributes, including constructive, motivational, personal, manageable, timely and directly related to assessment criteria and learning outcomes. Jamie Barron (2006) also stated that leaner strongly agree one thing they extremely needed is detailed, meaningful, and timely feedback in both public and private.

a, Constructive. Effective feedback should be constructive, as constructive feedback can lead to more thinking and cognitive learning which in turn improve the student's learning (Bang, 2003 and Alessi et al, 2001). As Nelson and Schunn (2009) argued, effective feedback should be able to guide a learner 'to change performance in a particular direction rather than just towards or away from a prior behaviour'. In addition, the result of Özdener and Satar (2009)'s research is also consistent to this point of view, as it revealed that explanatory feedback was more effective than confirmation feedback. Similarly, Clark and Mayer (2009) also pointed out that explanatory feedback was more effective than corrective feedback.

b, Motivational. Effective feedback should be motivational to empower and encourage students to learn more, as feedback can affect students' feelings and attitudes towards study, which in turn affect their engagement in the learning process (Juwah et al, 2004, cited by Hatziapostolou and Paraskakis, 2010). Similarly, Saddler and Andrade (2004) also suggest including praise as a feature of feedback to motivate students.

c, Personal. Effective feedback should be personalized, customized and diagnostic to fit each student's learning level and style, so that it can reflect individual student's strengths and weaknesses, and give tailored guides and recommendations to the student. Garber (2004) also argued, 'the more personalized the feedback becomes, the more meaning it can have for the individual receiving it…and the more likely the individual will be receptive to the feedback…If a person does not believe in the reliability or validity of the feedback, it will have little or no benefit'.

d, Manageable. Effective feedback should be detailed enough to ensure that students clearly understand their strengths and weaknesses, and have enough materials to guide them to achieve the learning goals. At the same time, the feedback should not be over-detailed to avoid confusing, and make students can easily interpret it and get the point (Hatziapostolou and Paraskakis, 2010).

e, Timely. Feedback should be delivered timely, as students can more easily utilize feedback when they can still remember how they just processed the task (Race, 2006), and the reasoning that led to the error is still accessible (Reiser and Kimberg et al, 1992).

As Anderson, Boyle and Reiser (1985, cited by Reiser and Kimberg et al, 1992) argued, tutors should provide immediate feedback to students, as ‘the learning mechanism for adjusting a faulty rule or forming a new correct rule relies upon the problem situation being active in memory’. Sung, (2009) pointed out that both time and method for providing feedback are very important in any tutoring system.

f, Quality. Felix, (2000) stated that although students thanks for immediately automated feedback, they need more quality feedback. In addition, Lanny & Musumeci, (2000) pointed that some leaner complain the missing of useful online advise. Barron, (2006) also pointed that leaners expect meaningful feedback when they are absent.

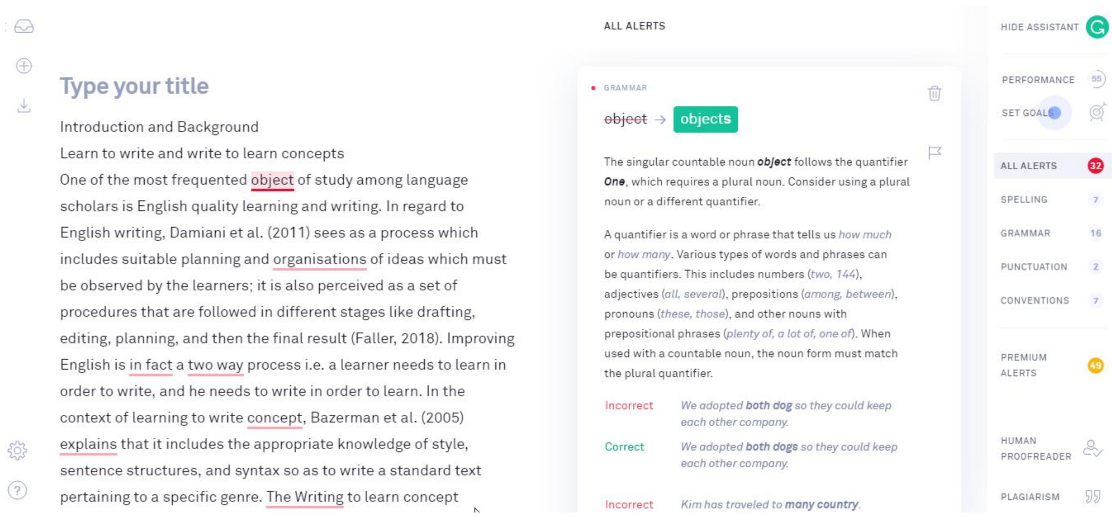
g, Directly related to assessment criteria/learning outcomes. Effective feedback should be directly related to assessment criteria/learning outcomes so that it can explain students’ achievement towards the intended learning outcomes, knowledge gaps and specific errors (Hatziapostolou and Paraskakis, 2010). Thus, the students can be guided and adjust their effort to achieve the intended leaning outcomes (Race and Brown, 2005). Additionally, Clark and Mayer (2009) further argued that learning goal oriented feedback is more effective than performance goal oriented feedback. In another word, feedback should be designed to inform the learners their progress toward achieving a learning goal rather than compare a learner’s performance with other learners’.

In addition to the required attributes above, many researchers also mentioned that various methods should be used in feedback to ensure better perception of feedback. For example, Özdener and Satar (2009) suggested using animation techniques to achieve better reception and perception of the feedback. Springgay and Clarke (2007) suggested including examples in feedback to achieve better perception of feedback. “Providing consistent, instant, and detailed feedback to students has been a big challenge in teaching Web based computing, given the complexity of project assignments and the comprehensive requirements on security, reliability, and robustness" (Fu, X. et al., 2008).

Agent software is a topic of growing interest to users and developers in the education field and learner feedback system (L. Teresita, and S. Raymund, 2003). The idea of intelligent software that performs the role of a human assistant is being explored in a wide range of education field. The intelligent system can monitor the activities of the learner, provide feedback, and find regularities (Buabeng-Andoh Charles & Asirvatham David 2002). Intelligent feedback system can provide learners with individualized, dedicated feedback based partly upon an analysis of the procedures followed by the user, which may provide some assistance on how the user should progress (Joung-Souk Sung, 2009). Intelligent feedback system provides continuous, intelligent feedback, guiding students into the learning style that's best for them (J.S. Sung and D.H. Lim, 2006). Intelligent feedback system is become a more capable feedback system so that learner can get learner's test done more efficiently (Sung, 2009).

2.3.3. Critical analysis of e-learning and the role of feedback

The above literature findings give credence to a clear knowledge-based action, but the researcher deems it important to critically analyze the research problem at hand with a presumption "is it necessarily needed to have an e-learning based intelligent feedback system or an ordinary feedback directly provided by the online teacher would be enough"? Secondly "why self-regulation of students learning is an important factor to consider while designing e-learning feedback system for them"? When critically analyzing these factors, the researcher came across the fact that there is a noticeable difference between today's online learning and the traditional learning scenarios in the past. Previously, all the learning was based inside a classroom setting where teacher was able to observe many physical and verbal cues of the students. But when technology came into play so as to make distance learning easier, the role of teacher has become limited and students have full responsibility to learn with a sense of autonomy. But the new format of learning instigates students to participate fully in their own learning and this desire becomes manifold in the context of e-learning when students do not have to wait for their virtual teacher to provide feedback against their work, but the feedback should be provided by a robust intelligent system which correctly provides assessment to each individual student without any delay. To analyze it critically, the researcher examined the Grammarly-based English language feedback to analyze the built-in advantages of it in the context of language learning. The researcher realised that the instant feedback provided by Grammarly software intends to engage him more through a different experiential world where Grammarly provides a long list of English language errors, punctuation flaws, typos, and a relevant explanation for each and every single error as shown in the following figure:



This implies how much effective, robust, and fast a simple language learning feedback tool is for a student against the traditional feedback provided by a teacher where teacher needs to give attention to each and every student, and it is not uncommon to have delays in providing feedback as well. Hence, this gives credence to the view that in e-learning scenarios, an intelligent feedback system is critically required not only because of the robustness and accurate feedback, but also because of increased student-base and a limited capacity of the teacher in terms of dealing equally with all the students. Secondly, self-regulation of student's learning is an important factor because constructivism theory also supports this type of learning since it is student-centered, and it states that the individual is the person who constructs knowledge actively through his experiential world which makes "knowing" and "knowledge" nothing shorter than an adaptive process (Sjoberg, 2007; Husain, 2010). This is itself a depiction of self-regulation in students where teacher is not responsible to transfer the entire knowledge into the student's braincells, but it the student who has to become the main body to construct the knowledge and absorb it in a self-regulated manner. This kind of self-regulated learning is necessary in the modern times and an intelligent feedback in e-learning environment develops the same type of accurate and self-regulated learning without any delays.

2.3.4. Overview of Neural Networks

The dictionary-based definition of Neural Networks is described in the following words:

> A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain

operates. Neural networks can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria (Investopedia, 2019).

There are many models proposed in the context of e-learning environment. For example, for monitoring and feedback in e-learning scenarios, a machine learning based framework is proposed by Joseph (2014). The main aim was how the presentation of the content could be utilized by utilizing the algorithms of machine learning. The process was to be conducted as per the student's provided feedback. In this process, neural networks were utilized so that the interactions could be mapped and the learner's interest and disinterest against the material could be forecasted. This is the simple example of utilization of neural networks in e-learning situations where different algorithms are used so that the data could be processed just like the intelligent human brain, and based on the changing input, the networks could be adapted accordingly thereby without needing to redesign the output criteria, the best possible results could be obtained. Moubayed et al. (2018) highlight different types of neural networks such as Artificial neural networks (ANN) which mimic the functioning of our brain whenever we have many features-based much training data and we desire a hypothesis function that is non-linear. The main components of these neural networks are input layer, hidden layer, and the output layer. According to Naim et al. (2011), a diverse range of neurons are found in each layer together with weighting functions. The flow of information is unidirectional i.e. through the hidden layers, the information flows from input to the output layer. It is of note that neural networks are used in e-learning on the score that rule-based programming is not utilized in these networks, rather their performance is based on using the learning algorithms which work in terms of tuning outputs to inputs (Ahmed, 2016). These neural networks have been widely utilized in the domain learning, assessment, and feedback. For example, Lee et al. (2009) highlight the utilization of the Snap-Drift neural network (SDNN) that are adept with the past students' responses, to online e-learning system so as to provide feedback and relevant guidance to the students. This neural network has the capability to take students responses and then categorise their responses based on its algorithms relevant to the lecturer-composed feedback for a certain level of students who have a certain level of understanding about the subject. Of note that any specific question is not attached to provide the feedback, instead, concept-based mapping is used in SDNN feedback. This implies that if the learner repeats the same test, he would not be provided the same feedback, but a different one, based on his concepts (state of knowledge). This type of diagnostic feedback could be effectively used to provide intelligent analysis of the real data to the students about their learning progress in particular subject.

2.3.5. Concept-based Feedback, student's knowledge state, and neural networks

The approach of concept-based feedback, as explained earlier, is different from the other types of feedbacks. The reason being, it is not based or oriented upon questions, but it is grounded upon the concept. This implies that this feedback would help the students to review those concepts that they did not understand earlier, and then they would be able to retake the test one more time accordingly. That is to say, it would not be like if the student had given wrong answers to a particular set of questions, then he would only try to learn those answers and get high marks in the retake test again. It would be like, at every retake, the concept-based feedback program would again check the student's concepts and grade it again based on his concepts, not upon his correct answers delivered in the previous test. This would result in receiving new feedback by the system when the student retake the test which would be specifically based on his prevalent misunderstandings of that student group category based on his level of understanding. Therefore, when the student is going to receive a new feedback each time based on his understanding level, and not a robotic type of question-oriented feedback, the student is likely to learn more and the whole process will lead to more self-learning as well (Payne et al., 2007). Furthermore, the concept-based feedback system also assists the student that they do not utilise fake guessing during the tests. Guo et al. (2014) corroborate the same that if the student would try to make fake guesses so as to click the right answers and does not pay attention to the feedback, the student might not even get to know what answers he clicked incorrectly.

The concept-based feedback relates with neural networks in the sense that just like there is a standard approach of incorporating existing valid knowledge about a specific domain into a system, and then devise a new knowledge which more accurately provide a better solution, the neural networks are now being utilized to take advantage of the previous feedback designed by the academics in the shape of students' response templates. These templates are used by the NNs wherein they associate students responses in several groupings based on their level of understanding, and then accurate feedback texts are provided to the students which would identify their current state of knowledge, the misunderstood concepts that they have, and then system would store their responses in NNs accordingly. This way the student's knowledge level and understanding is stored in the concept-based feedback system and such knowledge states could also be used to refine or retrain the NNs so that it creates new refined groupings accordingly (Guo et al., 2014).

It is worth explaining that the behavior of students while answering the questions is termed as “knowledge state”. The patterns of answers selected by the students i.e. students’ responses form the knowledge state in particular. That is to say, some commonality is seen in a specific knowledge state in a set of responses against questions. For instance, if there are many students who click the identical answers such as incorrect or correct to, at the very least, two questions or more, then one output neuron will be formed by the snap-drift that will be linked to a specific group so that all similar cases are forwarded to that output neuron (Guo et al., 2014). This implies that group of similar answers are discovered by the neural network which signifies student’s different knowledge state. The NN would also associate students’ responses into particular “knowledge state transitions”. For example, when a student receives some feedback after they click a new set of answers, then the system places them or reclassify them into either the same state (based on their past feedback) or to a new state as they receive a new feedback. These are referred to as knowledge state transitions.

2.4. Multiple-Choice Questions (MCQs) Feedback System

2.4.1. Strengths and Barriers of MCQs

According to Little, J.L. & Bjork, E.L. Mem Cogn (2015), “Answering multiple-choice questions with competitive alternatives can enhance performance on a later test, not only on questions about the information previously tested, but also on questions about related information not previously tested—in particular, on questions about information pertaining to the previously incorrect alternatives.”

Multiple Choice Questions (MCQs) as one of Intelligent feedback system is an effective way to provide students with feedback. The use of multiple-choice questions has been widely studied. A number of advantages can be found in the Epstein, et al, (2002), Higgins & Tatham (2003), and Kuechler and Simkin (2003) researches: rapid feedback, automatic evaluation, perceived objectivity, easily computed statistical analysis of test results and the reuse of questions from databases as required, thus saving time for instructors. On the other hand there are also some researches (e.g. Paxton (2000)) shows that MCQs have some disadvantages: significant effort is required to construct MCQs, they only evaluate knowledge and recall, and they are unable to test literacy and creativity.

2.4.2. Current Research of MCQs

Although the MCQs have been primarily used for summative evaluation, they also serve formative assessment purposes. Formative assessment provides students with feedback that highlights areas for further study and indicates the degree of progress. There are many researches investigating the effect of different types of feedback in web-based assessments which show the positive results of using MCQs in online test for formative assessment (e.g. Higgins and Tatham (2003), Payne et al (2007), Dafoulas (2005), Fu et al (2008)). Higgins and Tatham (2003) researched the use of MCQs in formative assessment in a web-based format using WebCT for a unit on the level 1 of undergraduate law degree. They assume that they can forecast all the possible errors for a question and they can write a general feedback for this question. But in terms of this kind of feedback, predicting all the possible errors and write the general feedback for a combination of questions would be a very hard work and would be impossible for a large test banks (e.g. 3 questions with 5 answers need 125 possible answer combinations; 5 questions with 5 answers need 3125 combinations, etc.).

Fu et al (2008) presented an effective technique for providing instant and informative feedback to students in web computing classes. The major concern of teaching web based systems is: how do instructors provide consistent evaluation and detailed feedback to students, given an overwhelming number of student project submissions, each of which may consist of over 10 dynamic Web forms, 100 user controls, and possibly over 2000 lines of source code? Clearly, automated grading and feedback system is one good potential solution. The feedback system provides feedback information at two levels: (1) a summary report, and (2) detailed feedback for each requirement and the corresponding test cases. Overall, the web-based system can provide informative feedback to help students make reflective and iterative improvements in learning.

Payne et al (2007) assessed the effectiveness of three different types of feedback (corrective, corrective explanatory, and video feedback) that used in e-learning to support students' learning. In this kind of feedback, the feedback shows exactly which questions are answered correct or not, and further explanations of the corrective explanatory and video feedback. In our research, the feedback is different from above. The intelligent diagnostic feedback we present is concept-oriented instead of question-oriented. Furthermore, the learners are encouraged to review their misunderstood concepts in order to do the test again and study further. It is important that each category of answers is associated with carefully designed feedback based on the level of understanding and prevalent misconceptions of that category-group of students so that every individual student can get diagnostic feedback reflect his or her learning level and certain mistakes.

In addition, the feedback is concept-based rather than question-based so that the student is encouraged to retake the test and receive different feedback according to his or her knowledge state, which in turn leads to more self-learning. Moreover, concept-based feedback can also prevent the student from guessing the right answers, as if the student did not read the feedback carefully, he or she may even do not know which questions are answered incorrect. According to our current research, there is no other studies related to MCQs and formative web-based assessment have used any similar form of using intelligent agent to analyze the students' response in order to provide diagnostic feedback.

2.5. E-learning Snap-Drift Neural Network (ESDNN)

2.5.1. Artificial Neural Network (ANN)

Artificial neural network (ANN) is developed based on the way that the brain performs computations. According to Galushkin (2007), 'A neural network represents a highly parallelized dynamic system with a directed graph topology that can receive the output information by means of a reaction of its state on the input actions'. Every neural network is composed of a large number of interconnected neurons, each of which presents a nonlinear, parameterized function of its input variables (Dreyfus, 2005). Although these neurons are often quite simple, the network gains its computing power from the massive neurons being connected, with outputs from the neurons being input to others (Johnson and Picton, 1995). Thus, a neural network can solve very complex problems, as it breaks down the complex problem into many simple issues, and each of these issues can be solved by a certain neuron which has been defined for this particular issue (Bharath and Drosen, 1994).

According to Johnson and Picton (1995), ANNs have 6 main advantages: 1. they have the ability to learn from examples, so that they do not need to be programmed. This is the most important advantage, as well as a distinguishing feature of ANNs, since ANNs can be 'trained by feeding the inputs into the network and pairing them with corresponding known outputs' (Greenstein and Welsh, 2005). 2. They can generalize from their training data to other data. 3. They are fault tolerant. 4. On being damaged, they degrade in a progressive manner rather than fail catastrophically after isolate failures. 5. They are fast. 6. They are not very expensive to build and to train. Based on these advantages, ANNs have been widely used in many areas, including character recognition, speech processing, image processing, pattern classification and recognition, system control and robotics (Karayiannis and Venetsanopoulos, 1993; Johnson and Picton, 1995).

2.5.2. Snap-Drift Algorithm

Snap-drift algorithm is an advanced and novel ANN algorithm. It was first developed to try to overcome the limitations of Adaptive Resonance Theory (ART) learning in non-stationary environments, and then it has proved its invaluable role as an outstanding classifier in several applications, e.g. feature discovery in speech, classifying user requests in an active computer network simulation environment, grouping spatio-temporal variations associated with road traffic conditions and so on (Palmer Brown, et al. 2008).

According to Palmer Brown, et al. (2008), snap-drift algorithm combines fast, convergent, minimalist learning (snap) and more cautious learning (drift) together so that it can capture both precise sub-features in the data as well as more general holistic features. During the learning process, the input data patterns firstly enter into the distributed SDNN (dSDNN), which will learn to group them based on their features using snap-drift. The neurons whose weight prototypes result in them receiving the highest activations are adapted. In this process, weight vectors are normalised, which means only the angle of the weight vector is adapted. Then, the winning neurons as the output from dSDNN enter into the selection SDNN (sSDNN) for the purpose of feature grouping. Both dSDNN and sSDNN are subject to snap-drift learning.

Of note that Snap-Drift Neural networks (SDNNs) exhibit a strong algorithm method that is based on the categorisation of high-speed data (Lee et al., 2004; Lee et al., 2008). It is regarded as an approach based on modal learning wherein the snap and drift modes switch between each other and the main features of SDNN revolves around the switching its learning mode and appropriate utilization of components to categorise the data into different groups. By snap and drift, every weight vector is bounded where snapping provides minimum values angle in all directions while drifting provides average patterns angle that is grouped beneath the neuron. When in function, the highest activations to weight prototypes neurons are given when SDNN groups them into respective categories. The more detailed description of SDNN algorithm is described by Lee and Palmer-Brown (2006) and Lee et al. (2008). The slight difference between SDNN and dSDNN is that when an input data is presented through SDNN, the dSDNN takes the charge to check their features and subsequently categories them and group them using snap-drift. From the dSDNN, the winning output neurons perform as the input data to the SDNN selection group. On the other hand, the ESDNN provides training with regards to the responses of students for a course related particular topic wherein past cohorts of students' responses are used to obtain responses. Palmer-Brown and Jayne (2011) explain that a binary form holds the encoded responses from each of the students before training so as to prepare the ESDNN input patterns presentation. It is worth mentioning the role of dSDNN here which, at the input layer, upon presenting an input pattern during the training, starts learning to categorize the input into groups based on their general structures (Guo et al., 2014) .

In a word, the learning process of SDNN is to find a large number of features in the data, and group the data into categories based on these features (Alemán, et al. 2011). During the whole learning process, all of the weight vectors are bounded by both snap and drift. Snapping gives the angle of the minimum values and provides an anchor vector pointing at the bottom left hand corner of the pattern group for which the neuron wins. 'It represents a feature common to all the patterns in the group and gives a high probability of rapid convergence'; drifting, which gives the average angle of the grouped patterns, uses Learning Vector Quantization to make the vector point towards the centroid angle of the group and ensure that an average, generalised feature is included in the final vector (Alemán, et al. 2011).

2.5.3. Advantages of Integrate Snap-Drift Neural Network into a Web-based MCQs System

Multiple Choice Questions (MCQs) is an effective method of providing feedback to students. During students attempting such MCQs, they generate the invaluable data of understanding their learning behaviours. These data can provide a simple picture of their knowledge related to a given topic; in addition, these data is generally lost. In this research, the data is systemically collected and automatically analyzed; moreover its results are used to provide the customized, diagnostic feedback to support students' learning. Furthermore, the teachers can receive a picture of learning process of their students. The groups of answers are collected. The tutors can base on the features of groups to find out which concepts have been mastered and which have not. In addition, this information can be used to address any issue which students did not understand. This is achieved by developing a method which integrates a web-based system with the SDNN-based analysis of students' responses to the MCQs.

2.5.4. E-learning Snap-Drift Neural Network (ESDNN)

As an advanced and novel ANN algorithm, SDNN is very suitable to be applied to E-learning, as one of the significant advantages of snap-drift algorithm is its ability to adapt rapidly (Palmer Brown, et al. 2008). In addition, by the application of SDNN, the E-learning system can offer students immediate and frequent diagnostic feedback, which in turn will promote the student's learning gains.

In Palmer Brown, et al.'s research (2008), a ESDNN system was designed based on online multiple choice questions (MCQs), as MCQs is an effective way to provide students with immediate and frequent feedback and has been reported as having positive results in online tests for formative assessments (Palmer Brown, et al. 2008). The principle of the ESDNN system is as below:

1. The student answers the online MCQs on a chosen topic.

2. Their answers are input into a snap-drift neural network trained with answers of past students.

3. Snap-drift categorises the input answers as having a significant level of similarity with a subset of the students it has previously categorised.

4. The system gives the corresponding feedback link to the certain category of answers to the student.

5. The student learns from the feedback and retakes the MCQs till his or her answers to all the questions are correct or till timeout.

According to Guo, et al. (2014), it is important that each category of answers is associated with carefully designed feedback based on the level of understanding and prevalent misconceptions of that category-group of students so that every individual student can get diagnostic feedback reflect his or her learning level and certain mistakes. In addition, the feedback is concept-based rather than question-based so that the student is encouraged to retake the test and receive different feedback according to his or her knowledge state, which in turn leads to more self-learning. Moreover, concept-based feedback can also prevent the student from guessing the right answers, as if the student did not read the feedback carefully, he or she may even do not know which questions are answered incorrect.

In summary, according to Palmer Brown, et al. (2008), the features and strengths of this ESDNN system are as below:

1. It can provide immediate and frequent diagnostic feedback to a large number of students based on intelligent analysis of real data.

2. It encourages independent and deeper learning.

3. It provides a tool for self-assessment, which is accessible anywhere and anytime.

4. It is easy to install and run on a PC with network.

2.6. Summary

In chapter 1, the ESDNN was introduced as a potential advanced tool to offer students immediate and diagnostic feedback to improve their learning gains. However, in order to evaluate the performance and effectiveness of this novel e-learning system, abundant target-oriented testing of this system is required to be carried out in different fields. Furthermore, we also aim to enhance this system to overcome its deficiencies during practical applications.

Thus, this research is basically formed by three main parts. Firstly, we evaluate the existing ESDNN system by collecting and analysing a large number of testing data reflecting the students' learning gains by using this system as well as the survey data reflecting the students' satisfaction and attitudes towards this system. Four hypothesis are formulated for evaluating this system: 1, during the trials, students improved their understanding by reading given feedback; 2, after using M-OFS system, students get higher mark in a separate questionnaire than before; 3, the experimental students get higher marks than control group at the final examination; 4, most students are satisfied with this system.

Secondly, the investigation will lead to an understanding of the potential of the on-line diagnostic feedback approach across different subject areas. Moreover, this research should also produce guidelines for the design principles of on-line MCQs in the context of diagnostic feedback learning environments.

Thirdly, we will enhance the existing system according to the evaluation data, and rename it as MCQs-Online Feedback System (M-OFS).

3. Methodology

3.1. Introduction

In order to provide students intelligent diagnostic customizes feedback, several methods are used in this research. The snap-drift algorithm and e-learning snap-drift neural network (ESDNN) are the key methods. The ESDNN is an approach to integrating an e-learning system and the Snap-Drift Neural Network (SDNN) which can provide carefully targeted guidance and feedback at different steps/times of learning. Palmer-Brown, Lee and Draganova (2008) point out that the preliminary results indicates that integrating the SDNN into an online system of multiple choice questions (MCQs) for providing automatic diagnostic feedback is a viable solution. In addition, Guo, et al. (2012) also confirmed that the integrated system is a viable solution.

3.2 Multiple-Choice Questions Online Feedback Systems (M-OFS)

To analyze the students' answers, and integrate over a number of questions to gain insights into the students' learning needs, a snap-drift neural network (SDNN) approach is proposed. SDNN provides an efficient means of discovering a relatively small and therefore manageable number of groups of similar answers. In the following sections, the e-learning system based on SDNN is described.

3.2.1 Snap-Drift Neural Networks (SDNNs)

One of the strengths of the SDNN is the ability to adapt rapidly in a non-stationary environment where new patterns are introduced over time. The learning process utilises a novel algorithm that performs a combination of fast, convergent, minimalist learning (snap) and more cautious learning (drift) to capture both precise sub-features in the data and more general holistic features. Snap and drift learning phases are combined within a learning system that toggles its learning style between the two modes. On presentation of input data patterns at the input layer F1, the distributed SDNN (dSDNN) will learn to group them according to their features using snap-drift (Lee et al., 2004). The neurons whose weight prototypes result in them receiving the highest activations are adapted. Weights are normalised weights so that in effect only the angle of the weight vector is adapted, meaning that a recognised feature is based on a particular ratio of values, rather than absolute values. The output winning neurons from dSDNN act as input data to the selection SDNN (sSDNN) module for the purpose of feature grouping and this layer is also subject to snap-drift learning.

The learning process is unlike error minimisation and maximum likelihood methods in MLPs and other kinds of networks. These perform optimization for classification or equivalents by for example pushing features in the direction that minimizes error, without any requirement for the feature to be statistically significant within the input data. In contrast, SDNN toggles its learning mode to find a rich set of features in the data and uses them to group the data into categories. Each weight vector is bounded by snap and drift: snapping gives the angle of the minimum values (on all dimensions) and drifting gives the average angle of the patterns grouped under the neuron. Snapping essentially provides an anchor vector pointing at the 'bottom left hand corner' of the pattern group for which the neuron wins. This represents a feature common to all the patterns in the group and gives a high probability of rapid (in terms of epochs) convergence (both snap and drift are convergent, but snap is faster). Drifting, which uses Learning Vector Quantization (LVQ), tilts the vector towards the centroid angle of the group and ensures that an average, generalised feature is included in the final vector. The angular range of the pattern-group membership depends on the proximity of neighbouring groups (natural competition), but can also be controlled by adjusting a threshold on the weighted sum of inputs to the neurons. The output winning neurons from dSDNN act as input data to the selection SDNN (sSDNN) module for the purpose of feature grouping and this layer is also subject to snap-drift learning.

Codification	A:00001, B:00010, C:00100, D:01000, E:10000, N/A:00000
Response	Encoded response
[A, C, D, B, A]	[0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1]
[D, A, A]	[0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1]

Table 1: Example of input patterns for ESDNN

3.2.2 Training Neural Network

The ESDNN is trained with the students' responses to questions on a particular topic in a course. The responses are obtained from the previous cohorts of students. Before training, each of the responses from the students is encoded into binary form in preparation for presentation as input patterns for ESDNN. Table 1 shows examples of a possible format of questions for five possible answers and some encoded responses.

This version of ESDNN is a simplified unsupervised version of the snap-drift algorithm (Lee and Palmer-Brown, 2006) as shown in Figure 1.

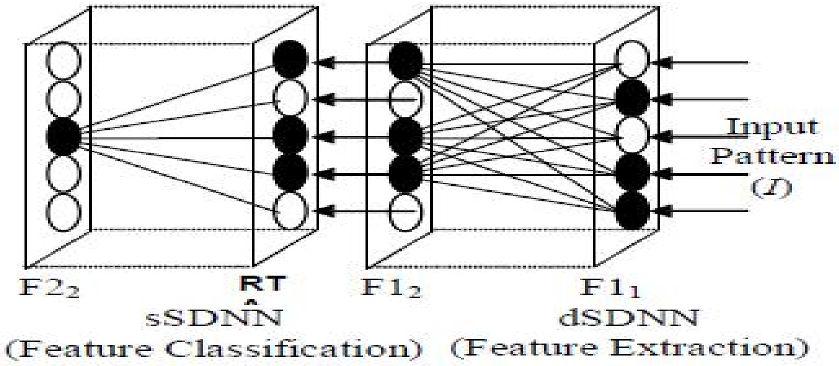


Figure 1: E-learning SDNN architecture

During training, on presentation of an input pattern at the input layer, the dSDNN will learn to group the input patterns according to their general features. In this case, 5 F12 nodes, whose weight prototypes best match the current input pattern, with the highest net input are used as the input data to the sSDNN module for feature classification. In the sSDNN module, a quality assurance threshold is introduced. If the net input of an sSDNN node is above the threshold, the output node is accepted as the winner; otherwise a new uncommitted output node will be selected as the new winner and initialised with the current input pattern. For example, for one group, every response might have in common the answer C to question 2, the answer D to question 3, the answer A to question 5, the answer A to question 6, the answer B to question 8, and the answer A to question 10. The other answers to the other questions will vary within the group, but the group is formed by the neural network based on the commonality between the answers to some of the questions (four of them in that case). From one group to another， the precise number of common responses varies in theory between 1 and X， where X is the number of questions. In this experiment， where there are 10 questions in 1st English trial (Section 5), the groups had between 5 and 8 (Trial 1) common answers. More details of the steps that occur in ESDNN and the ESDNN learning algorithm are given in (Lee et al., 2008).

The training relies upon having representative training data. The number of responses required to train the system so that it can generate the states of knowledge varies from one domain to another. When new responses create new groups, more training data is required. Once new responses stop creating new groups, it is because those new responses are similar to previous responses, and sufficient responses to train the system reliably are already available. The number of groups formed depends on the variation in student responses.

3.2.3 M-OFS Description

The MCQs-OFS tool is enhanced based on the ESDNN system for the following purposes: 1.) much easier to install and run in wider environment; 2.) provide more accurate, detailed, specific, customized, and targeted feedback to the learners; 3.) automatically record the data of studying process of each student; 4.) be able to apply to more area of study; 5.) Compare and contrast the learning performance between each individual student and the average of other students in each stages of knowledge. The MCQs-Online Feedback System (M-OFS) is developed using Java Server Faces (JSF) Technology, which is a component-based web application framework that enables rapid development. The JSF follows the Model-View-Controller (MVC) design pattern and its architecture defines clear separation of the user interface from the application data and logic. Furthermore, the (M-OFS) as a simple tool, it can be simply installed on a PC, and it also can be integrated into a Virtual Learning Environment.

This version of ESDNN is the unsupervised version of the snap-drift algorithm, as shown in figure 2. The working of ESDNN can be divided into two phases, training and deployment. Ultimately, these two phases will be mechanized and integrated into the e-learning system, which we call Multiple-Choice Questions Online Feedback System (M-OFS). The M-OFS has been designed and built using the Java-Server Faces Technology (JSF), which is a component-based web application framework that enables rapid development. The JSF follows the Model-View-Controller (MVC) design pattern and its architecture defines clear separation of the user interface from the application data and logic.

The M-OFS is integrated within the web application as part of the model layer. The M-OFS is trained for each set of questions offline with data available from previous years of students, and the respective weight text files are stored on the application server. The feedback for each set of questions and each possible set of answers is grouped according to the classification from the M-OFS and written in an XML file stored on the application server. The feedback includes several contents such as the basic concepts of the topic, examples of specific problem-solving steps, hints and some links of the website of further studies, but not providing the solution to students. It is Concept-Oriented instead of Question-Oriented. In a trial, the system will automatically provide the most suitable feedback to the students according to the different sets of student's responses to a set of questions on the given topic, and the given feedback is not dependent on a particular question. After then students are encouraged to review the questions, and then find out what mistakes they made and what the right method is. By doing so, the students will not get directly the right answers from the feedback, but guide the students through stages of learning until they reach the ultimate stage, i.e. getting the correct answer and understand how to do it. The principle is that the journey towards the answer (learning process) is much more important than the answer itself. In order to analyze the progress of the students in using the system they have to login into the system with their student id numbers. The set of answers, time and student id are recorded in the database after each student's submission of answers. After login into the system the students are prompt to select a module and a topic and this leads to the screen with a set of multiple choice questions specific for the selected module and topic. On submission of the answers the system converts these into a binary vector which is fed into the M-OFS. The M-OFS produces a group number; the system retrieves the corresponding feedback for this group from the XML feedback file and sends it to the student's browser. The student is prompted to go back and try the same questions again or select a different topic. A high level architectural view of the system is illustrated in Figure 2.

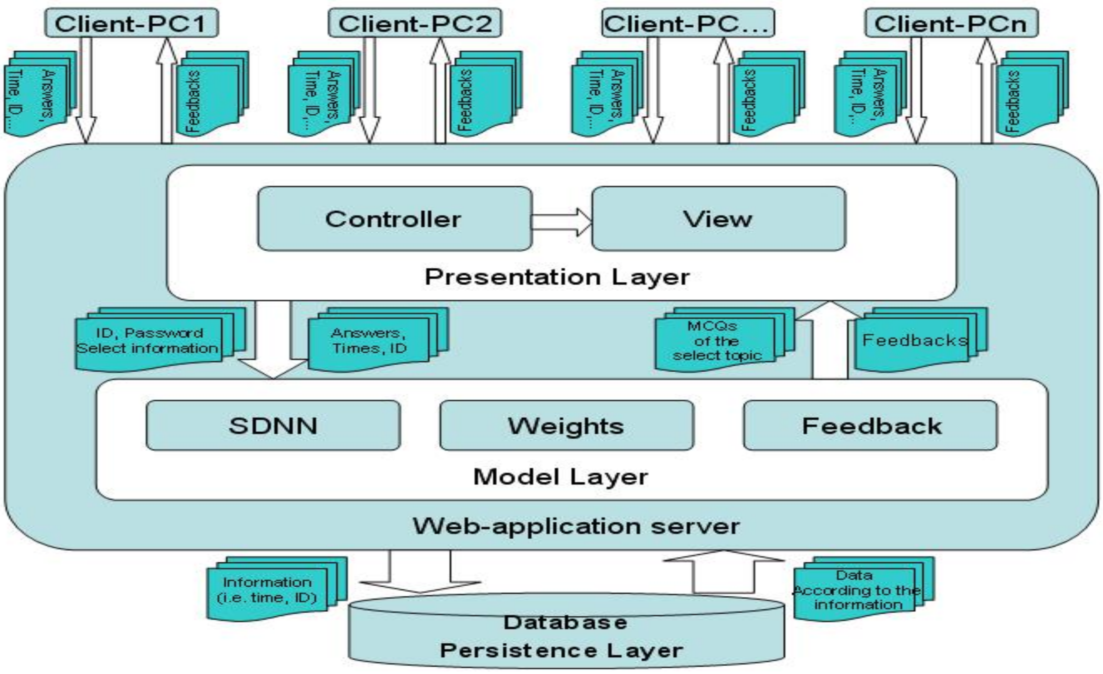


Figure 2: M-OFS system architecture and flow chart

3.2.4 How the System Guides Learning

The feedback is written by academics and designed so that it does not identify which questions were incorrectly answered. The academics are presented with the groups in the form of tem-plates of student responses. For example, "A/D B mix" represents a group characterized by all the students answering A or D to question 1, B to question 2, and mixed answers to question 3. Hence, the educator can easily see the common mistakes in the groups of the student answers highlighted by the tool. The feedback texts are associated with each of the pattern groupings and are composed to address misconceptions that may have caused the incorrect answers common to that pattern group.

The student responses, recorded in the database, can be used for monitoring the progress of the students and for identifying misunderstood concepts that can be addressed in subsequent face-to-face sessions. The collected data can be also used to analyze how the feedback influences the learning of individual students by following a particular student's progress over time and observing how that student's answers change after reading the feedback. In the trial, for example with 10 questions and 4 possible answers, there are more than 1 million possible combinations of answers, and students are not been told which is right answer. Therefore, random selection would be an unsuccessful strategy, moreover the students are unlikely to make improvement by guessing answers.

Student responses can also be used to retrain the neural network and see whether refined groupings are created, which can be used by the educator to improve the feedback. Once designed, MCQs and feedbacks can be reused for subsequent cohorts of students.

3.3 Experimental Environment

In order to evaluate the performance and effectiveness of this novel e-learning system, abundant target-oriented testing of this system is carried out in different fields across different subject. Furthermore, it investigates the effective use of intelligent feedback towards modelling the stages of students' learning. In addition, we also aim to enhance this system to overcome its deficiencies during practical applications. Thus, this study is composed of three main parts. Firstly, we evaluated the M-OFS system by collecting and analysing a large number of experimental data reflecting the students' learning gains by using this system as well as the survey and interview data reflecting the students' satisfaction and attitudes towards this system. Secondly, the investigation will lead to an understanding of the potential of the on-line diagnostic feedback approach across different subject areas. Furthermore, this research also produces guidelines for the design principles of on-line MCQs in the context of diagnostic feedback learning environments. Thirdly, in the future it will help to enhance the existing system according to the evaluation data.

3.3.1 hypothesis

Four hypothesis are formulated as following: H1, students are satisfied with using M-OFS; H2, students improved their understanding by reading given feedback; H3, in a separate MCQs paper test, students get higher mark in the first test than the second trial by learning from the M-OFS; H4, in the final examination, the average score of experimental group is higher than the average score of control group.

3.3.2 Trials

The trials are applied to several totally different subject areas, in order to evaluate the adaptability, feasibility, and ability of this novel system. The general steps of setting up a trial are shown as follow: 1.Collect and assess the data which include: examination questions, correct answers, students' ID and students' answers of previous years for given module; 2.Classify and filter the questions according to their relative topic; 3.Use the Snap-Drift Neural Network (SDNN) to train the initial data, in order to group the students' responses; 4.Identify the features of each group (or the sets of responses of students); 5.Design and write the targeted, customized and diagnostic feedback according to the feature of group; 6.Test the system; 7.Apply the trials; 8.Design and conduct a survey to evaluate students' experiment of using the system

Feedback as a core part of the system, it includes several contents such as the basic concepts of the topic, examples of specific problem-solving steps, hints, and some links of the website for deep learning, but it does not provide the answers to the questions to students. Furthermore, it is concept-oriented instead of questions-oriented. In a trial, the system will automatically provide the most suitable feedback to the students according to the different sets of student's responses to a set of questions on the given topic, and the given feedback is not solely dependent on any particular question. On receiving feedback students are encouraged to review the questions, and to consider what mistakes they made and what the right method is. By doing so, the students will not get directly the right answers from the feedback, but guide the students through stages of learning until they reach the ultimate stage, i.e. getting the correct answer and understand how to do it. The principle is that the journey towards the answer (learning process) is much more important than the answer itself.

According to Guo, et al. (2012), although a trial has been carried out on “Java Programming” module in March 2009, the amount of the recorded data is insufficient. The other trials that have been applied to the following modules: “English Language Study” and “Mathematics” in December 2010 were very successful. Especially “English Language Study” projects, 250 students participated the trials and filled the survey, and 20 students conducted the interview. A large number of data of learning process for each student had been automatically recorded by the M-OFS system during the trials. Furthermore, this data was analyzed and summarized to answer the questions following: 1.) How much time each student has spent on reading feedback? 2.) What is the average time all the students have spent on reading feedback? 3.) How many times each student has tried to answer questions during their online studying? 4.) What are the average times all the students have tried to answer question during their online studying? 5.) How long have each student spent on moving from one to another stage of knowledge? 6.) What is the average time all the students have spent on moving from one to another stage of knowledge? 7.) How long have each student taken to reach the final stage (all correct answers)? 8.) What is the average time the students have taken to reach the final stage? 9.) What is the percentage of students who have reached the final stage?

3.3.3 Evaluation

In this research, system evaluation consists of two parts:

Part 1: Subjective evaluation: In order to assess the student perception of the tests and feedbacks, a student survey and interview on the experiment of the system had been conducted after every trial.

Part 2: Objective evaluation: A large number of data had been recorded during the trials, in order to assess the performance and effectiveness of M-OFS. Furthermore, comparisons on the examinations and questionnaires results are carried out between students using and without using the MCQs-OFS and this is part of evidence in evaluating the system. In addition, separate paper questionnaires have been carried out three times (once for each test) before the "English trial" and three times after, in order to compare the results of the test before and after the system trial.

3.4 Experimental Environment and Process

3.4.1 Data Collection

The data collection in this research consists of two parts: 1.) the data are used to train the neural network (collect from previous examination before the trial); 2.) the data used to analyze learning behaviour (collect during the trial by the system).

Before the trial, the data of questionnaires (questions, options and correct answers) and students' answers of the given module of previous years' final examination are collected. Moreover, the students' response when they doing the trial were also collected. The data collection is precise, careful, and there was a large work of data collection. Thus the process of data collection is very important. The process of data collection is introduced as following: 1.) The target topic need to be identified first; 2.) The target students who can be conducted the trial need to be determined; 3.) Collect the appropriate data from previous final examination; 4.) Collect the response from students when they doing the trial. The collected data cannot be used directly.

The data of questions need to be classified and filtered, and grouped into several sets of questions according to the relative topic (see an example at figure 3). To classify and filtering the data questions according to relative topic based on the previous experimental experience, the number of MCQs in one topic is ideally between 5 and 12, and the key learning points of MCQs should be linked. Thus, if the initial set of data may include many questions and they are belong to many topics, the questions should be classified into several topics or sub-topics, and the number of questions in one topic should be reduced to a suitable size. The data of questions which collected before trial is used to train the neural network.

Furthermore, the data collected when a student doing the trial is used to analyze the learning behaviour of students and to group students learning behaviour. In addition, the data is also used to train the neural network.

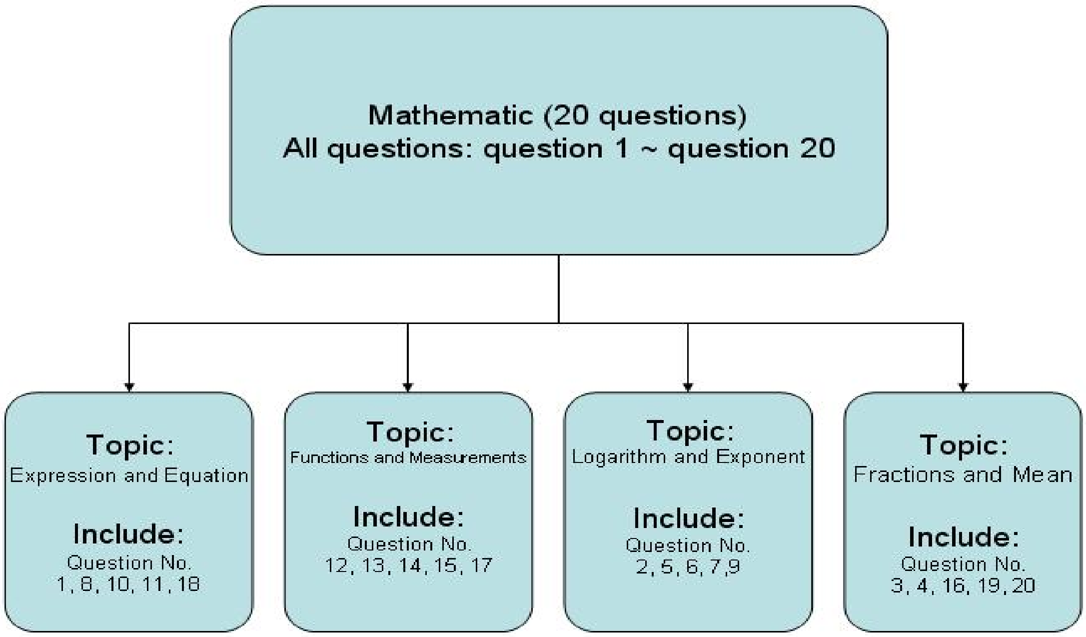


Figure 3: an example of classifying questions

3.4.2 Data Training

The data of each set of questions is used to train the snap-drift neural network (SDNN) several times until the output becoming to stable (see an example at figure 4). In this case, the SDNN is becoming stable after it was trained 3000 epochs (see an example at table 2). The SDNN is ready to use after training.

Number of epochs	1000	1500	2000	2500	3000	3500	4000
Percentage of grouping changes	35%	30%	20%	12%	6%	15%	23%

Table 2: an example of grouping changes by different training epochs

The training principle is described as following: 1.) the data was trained start with a small number of epochs. E.g. in the first trial, 800 epochs is used as start number of epoch; 2.) The outcomes were compare and contrast in order to find out whether the results were stable or not. From previous experimental results it shows that the grouping changing is become more stable if the changes less than 10%, and if the changing is more than 10% is unstable. E.g. in the training results of 800 epochs, the changes are more than 40%, and this results were not stable; 3.) Then the larger number of epochs were used to train the data; (e.g. 1000, 1500, 2000, 2500, 3000, 3500, and 4000 epoch were used in the first trial) 4.) The training was completed once the data getting stable after training. (e.g. 3000 epochs in the first trial)

The definition of stable result is defined as following: after training, the last two epochs were used to compare and contrast. If most students' response (more than 90%) were grouped into same group in both last two epochs, on the other word, if only a few students' response (less than 10%) were grouped into different group at the last two epochs, then the result was stable. It means the main group feature was found and a few students' response did not with a strong feature might be grouped into different group, since its feature were at the threshold between two main groups.

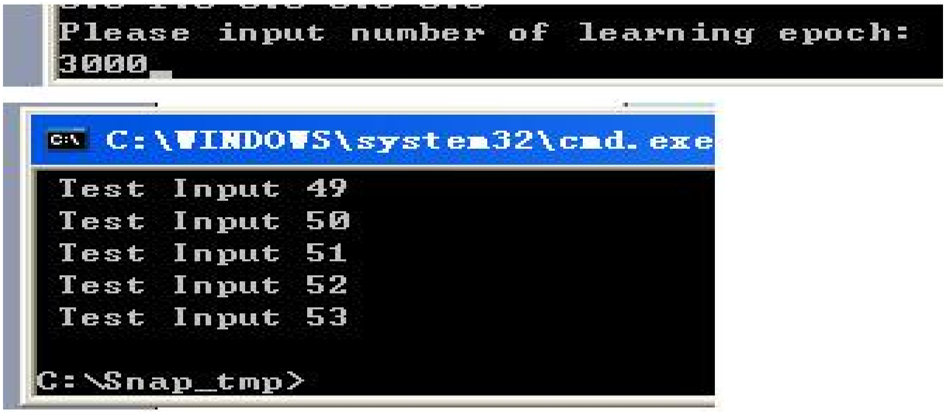


Figure 4: an example of training data

3.4.3 Grouping results and feature analysis

The data was group by neural network into the same/similar answers and same/similar answer order. Some of the groupings are not stable. In order to find out the stable results, compare and contrast the winning nodes of last 3 epochs after training. If the results only have a few changes and the changes are stable, then this result is acceptable. Then compare and contrast all the acceptable results in order to find out the best result. Once the best result was found, its respective weights will be stored on the application server.

Then identify the features of each group according to the final results (see an example at figure 5). In the stable result, the students with the same or similar answer of each MCQs were grouped into the same group (grouped by the neural network). For example, the students 26, 38, 39, 51, 55, 56, 59 were grouped into same group which the group number is 5. The strongest feature was defined as the group feature. For example, the feature of group 5 is (C, B/A, C, A, C). The group feature is summarized from the answer of each MCQs of most students' answer. For example, the small group 7, it includes two students, and its feature are (E, A, D, D/B, D).

User number	Given answer by students					grouping number
26	C	A	C	A	A	5
38	C	B	C	E	C	5
39	C	B	C	A	C	5
51	E	B	C	A	C	5
55	E	A	C	A	C	5
56	C	A	C	A	C	5
59	C	B	C	A	B	5
	C/E	B/A	C	A	C	
12	E	A	D	D	D	7
40	E	A	D	B	D	7
	E	A	D	D/B	D	
13	E	B	D	C	A	8
71	C	B	D	D	A	8
	E/C	B	D	C/D	A	
22	C	D	C	B	B	10
25	C	C	B	B	A	10
41	C	D	D	B	A	10
46	C	E	D	C	D	10
60	C	D	D	C	A	10
65	C	D	B	C	C	10
	C	D/c/e	D/b/c	B/C	A/mix	

Figure 5: an example of group's features

Feedback

G 2

It is very important to understand how the exponent laws work. When you are given a term with a form of $\mathbf{a^m}$, it is defined that: a is the ***base***, m is the ***exponent.*** Generally, more than one exponent law should be used to evaluate one exponent expression. The exponent laws will be given as following:

$$a^m \times a^n = a^{m+n}$$

$$\frac{a^m}{a^n} = a^{m-n}$$

$$(a^m)^n = a^{m \times n}$$

$$(ab)^n = (a^n) \times (b^n)$$

$$a^0 = 1$$

$$a^{-n} = \frac{1}{a^n}$$

$$\sqrt[m]{a^n} = a^{\frac{n}{m}} = (\sqrt[m]{a})^n = (a^{\frac{1}{m}})^n$$

Example 1: evaluate 1024^0

using law $\mathbf{a^0 = 1}$, (a=1024)

1024^0 => 1 (answer)

Hints:

1. Remember: $a^m \times a^n \neq a^{m \times n}$, $\frac{a^m}{a^n} \neq a^{\frac{m}{n}}$

2. It is also important to remember and understand the exponent laws before using it

Click the following link to find out more details about the exponents.

Exponent laws: http://www.mathexpression.com/exponent-rules.html

Exponents(Exponents and Calculations): http://www.aaastudy.com/exp.htm

Figure 6: an example of feedback

3.4.4 Feedback Design

After training, appropriate feedback text was written by academics for each of the group of students' responses that address the conceptual errors implicit in combinations of incorrect answers. During the trial, a current cohort of students was asked to provide responses on the same questions, they were given the feedback on the combination of incorrect answers and their responses recorded in the database. The feedback texts are composed around the pattern groupings and are aimed at misconceptions that may have caused the incorrect answers common within the pattern group.

An example of a typical response of to the questions of English Grammar Trial is shown as following: (1<D>2<B>3<A>4<B>5<B/D>6<A>7<A>8<C>9<A/C>10<B/D>)

1. no cause for alarm, the old man went back to his bedroom.

A. There was B. Since C. Being D. There being

2. Even as a girl, to be her life, and theater audiences were to be her best teachers.

A. performing by Melissa were.

B. it was known that Melissa's performances were

C. knowing that Melissa's performances were

D. Melissa knew that performing was

3. Agriculture is the country's chief source of wealth, wheat by far the biggest cereal crop.

A. is B. been C. be D. being

4. This company has now introduced a policy pay rises are related performance at *work.*

A. which B. where C. whether D. What

5. She managed to save she could out of her wages to help her brother.

A. how little money B. so little money

C. such little money D. what little money

6. He left orders that nothing touched until the police arrived here.

A. should be B. ought to be C. must be D. would be

7. As it turned out to be a small house party, we so formally.

A. need not have dressed up B. must not have dressed up

C. did not need to dress up D. must not dress up

8. I the party much more if there hadn't been quite such a crowd of people there.

A. would enjoy B. will have enjoyed

C. would have enjoyed D. will be enjoying

9. There ought to be less anxiety over the perceived risk of mountain climbing than _____ in the public mind today.

A. exists B. exist C. existing D. to exist

10. Fat cannot change into muscle muscle changes into fat.

A. any more than B. no more than C. no less than D. much more than

This is classified into Group 6, which generates the following feedback:

Group 6 Feedback: *Four points should be stressed. First, the logical subject of the adverbial phrase should agree with that of the main clause. Second, two verbs in a sentence need a conjunction. E. G. I am a teacher but you are a student. Third, the usage of various noun clauses should be familiar with. E.G. The news came that he died. ("that" does not serve as any part of the clause.) Fourth, some fixed structures in the comparative form should be memorized. E. G. not so…as…*

The diagnostic and effective feedback of each group can be designed and written according to the group's features and its related questionnaires, and it is saved in an XML file which is stored on the application server (see an example at figure 6).

The process of design the feedback is follows: 1.) The feature of each group is analyzed by the system; 2.) supervisor finds out the mistakes/misconceptions were made by most students in the group based on the feature of group; 3.) supervisor writes the key learning point/concept, related examples, and some hints to aim at prompt and guide students to improve their understanding of the knowledge. In another word, the supervisor writes the key learning point/concept according to the mistake/misconception which was made by students; 4.) supervisor may re-design the feedback according to the results from training of the new data which collected when students studying from the system.

Designer/teacher writes the feedback very carefully, and the feedback is suitable for every one made the same mistake or had any same misunderstood. In another word, the feedback is tailored for the learners in the group. All of the feedbacks are mostly suitable for the students who are trying to learn from the system, but except someone who guess the answers.

The feedback includes three main parts: 1.) Several key learning points will be introduced at the beginning of the feedback; 2.) The Feedback will also include some example related to the key learning point to help student to further understand the key learning point/concept; 3.) The related hints are also pointed out at the end, in order to provide students more information about key learning point/concept. An example of details of a feedback is shown in the figure 6.

3.4.5 System Set Up and Test

After the feedback is designed, the database, files, and system need to be set up according to each trial requirement.

The system was set up and tested in the virtual environment before the start of the real trials. Several jobs have to be done such as debugging program, checking threshold, testing results, testing stability in order to make sure the system will run smoothly during the trials. Two steps are used to test the system: 1.) Some students are randomly selected to use the system with the MCQs of the basic knowledge such as fundamental mathematics; 2.) These students were asked to use the system again with the MCQs which related to their topic.

3.4.6 Trial Preparation

The environment of the trial is prepared. The target students are given a user instruction of how to use this system, and it includes web site address, user name, password, and simple introduction to the trial and so on. Students are given 10 minutes to read the instruction and 10 minutes to try the system before the start of the trial. The students are then given predefined time to conduct the trial. Students are expected to try their best to answer all the MCQs within the given time. Students are also expected to improve their understanding of knowledge by reading the diagnostic feedback, and keep trying until either giving up or getting the all correct answer at the end. Survey and interview is applied after trial.

3.4.7 System Trial Application

3.4.7.1 Preparation

Before applying a trial, the system needs to be set up. It includes: software installation, collecting initial data, classify initial data, input and transfer data, to train the data and analyze the training results, according to the analysing results, identify the number of groups and the feature of each group, design and write feedback for each group, set up a trial and test it, apply a trial.

1. _____ no cause for alarm, the old man went back to his bedroom.	A. There was B. Since C. Being D. There being

2. Even as a girl,_____ to be her life, and theatre audiences were to be her best teachers.	A. performing by Melissa were. B. it was known that Melissa's performances were C. knowing that Melissa's performances were D. Melissa knew that performing was
3. Agriculture is the country's chief source of wealth, wheat______by far the biggest cereal crop.	A. is B. been C. be D. being
4. This company has now introduced a policy______ pay rises are related performance at work.	A. which B. where C. whether D. What
5. She managed to save__________she could out of her wages to help her brother.	A. how little money B. so little money C. such little money D. what little money
6. He left orders that nothing____ touched until the police arrived here.	A. should be B. ought to be C. must be D. would be

7. As it turned out to be a small house party, we _____ so formally.	A. need not have dressed up B. must not have dressed up C. did not need to dress up D. must not dress up
8. I ____ the party much more if there hadn't been quite such a crowd of people there.	A. would enjoy B. will have enjoyed C. would have enjoyed D. will be enjoying
9. There ought to be less anxiety over the perceived risk of mountain climbing than _____ in the public mind today.	A. exists B. exist C. existing D. to exist
10. Fat cannot change into muscle muscle changes into fat.	A. any more than B. no more than C. no less than D. much more than

Table 3: An example of English MCQs

3.4.7.2 Software and Running Environment Preparation

In order to use the system, some software has to be installed on the hosting computer such as: Java jre1.6.0_02, Tomcat5.5, apache-ant-1.7.0, MySQL, some drivers, QAUEL, and Snap-Drift Neural Network

It is very easy to set-up of this system, and the process is as following:

1. Use Ant to build a web-application
2. Place it on the application server
3. Start up the server and test it (see figure 7)

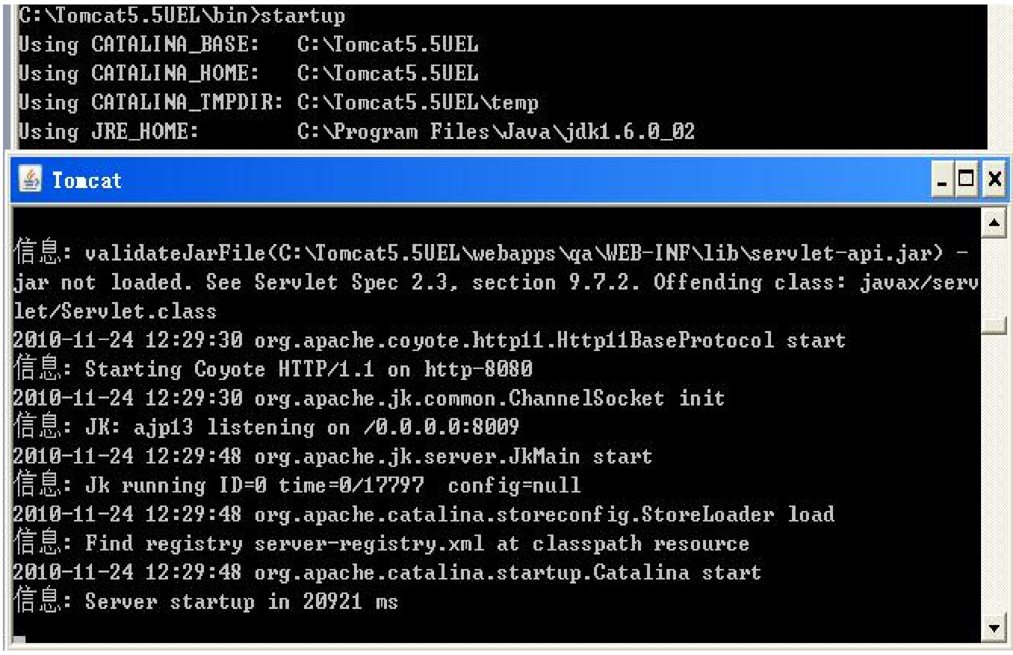


Figure 7: start up the Tomcat server

Module	**Summary**	**Details**
English	Five sets of data of final examinations have been collected (School of Language, JQ University and School of Foreign Literature, Kunming Technology University)	Date: 6th December 2010 Module title: Professional English Study Module Code: 20080305 Level of students: level 3 School: Foreign Language School Number of questions: 10

		Number of students: 386 Number of answers: 386 × 10=3860
Mathematic	Four sets of data of final examination have been collected School of Computing, UEL)	Date: 12th January 2010 Module title: Mathematics Module Code: SD0002 Level of students: level 1 School: CITE, UEL Number of questions: 20 Number of students: 74 Number of answers: 74 × 20=1480
Java Programming	Four sets of data of in class test have been collected (School of Computing, UEL)	Date: Semester A 2008-2009 Module title: Introduction to Software Development Level of students: level 1 School: CITE, UEL Number of students: 101 Set 1: Number of questions: 9 Number of answers:808 × 9=7272 Set 2: Number of questions: 11 Number of answers: 759 × 11=8349 Set 3: Number of questions: 5 Number of answers: 454 × 5=2270 Set 4: Number of questions: 12 Number of answers: 535 × 12=6420

Plagiarism	Two sets of data of questionnaires have been collected (School of Computing, UEL)	Module title: Skills for Academic Learning in Civil Engineering Module Code: CE1201 Level of students: level 1 School: CITE, UEL Set 1: Date: Summer, 2009 Number of questions: 5 Number of students: 53 Number of answers: 53×5=265 Set 2: Date: October, 2010 Number of questions: 5 Number of students: 103 Number of answers: 103×5=535

Information Security Management	Four sets of data of questionnaires have been collected (School of Computing, UEL)	Date: Semester B, 2009-2010 Module title: Information Security Management School: CITE, UEL Set 1 Number of questions: 5 Number of students: 15 Number of answers: $15\times5=75$ Set 2 Number of questions: 14 Number of student: 8 Number of answers: $8\times14=112$ Set 3 Number of questions: 17 Number of students: 8 Number of answers: $8\times17=136$ Set 4 Number of questions: 13 Number of students: 10 Number of answers: $10\times13=130$

Table 4: A summary of data collection (training data)

3.4.7.3 Data Collection

First of all, several suitable modules need to be selected in order to conduct a trial. The principles of module selection are: 1, modules should belong to different subject area; 2, modules should be able to apply trial; 3, there should be enough potential students to participate trials in the module. Secondly, the suitable data of previous examination of each selected module need to be collected in order to set up the system. The details of data collected are shown in table 4. In total data of six main trials were collected (3 English trials, 2 Math trials, 1 Plagiarism trial). As an example, The data of English Grammar Trial for training is collected from three previous year's MCQs tests. For these three tests, 94 students' answers were used to training. The trials data were collected during academic year 2010-2011. The data of two separate MCQ paper tests and final examination results were gathered. 83 students entered the survey and 16 students were randomly selected for interview. The states of knowledge of students were achieved by using ESDNN.

3.4.7.4 Data Classify and Filter

To classify and filtering the data questions according to relative topic based on the previous experimental experience, the number of MCQs in one topic is ideally between 5 and 12, and the key learning points of MCQs should be linked. Thus, if the initial set of data may include many questions and they are belong to many topics, the questions should be classified into several topics or sub-topics, and the number of questions in one topic should be reduced to a suitable size. These data is used to design and create MCQs.

Mathematics: In mathematics, in total 20 questions are classified into 4 sets/topics as follow: Set 1: Expression and Equation; Set 2: Functions and Measurements; Set 3: Logarithm and Exponent; Set 4: Fractions and Mean.

English: In English, in total 260 questions are classified into 3 topics and filtered into 10 questions for each topic as follow: Set 1: English Grammar; Set 2: English Vocabulary; Set 3: English Reading Comprehension

3.4.7.5 System Trial Application

In order to evaluate the ability, feasibility, and the adaptability of the M-OFS, several system tests are conducted. For example, Six different tests were applied to "English", "Mathematics", and "Java Programming" module with 20, 5, and 5 students. For the "English" tests, 3 trials, 20 surveys, 3 interviews and 3 questionnaires were conducted. During the tests, students can make as many attempts at the MCQs as they want. On average they attempted 12 times (12 sets of answers) over 50 minutes in the "English" test, and 3 attempts (3 sets of answers) over 15 minutes in the "Mathematic" test, and 3 attempts (3 sets of answers) over 18 minutes in the "Java Programming" test.

3.4.8 Experiment Results Analysis

In order to understand the relationship between learning behaviour and feedbacks, and evaluate if the feedback students received helped them to understand the topic. The results of system trials are analyzed by two steps:

1.) Basic data analysis: a.) Average number of attempts in each topic. b.) Number of students who got all correct answers in each topic. c.) How many students give the all correct answers at first attempt, and at the end? d.) The learning duration for each student/the time from first attempt to last attempt. e.) How long did the fastest student spend on getting all correct answers?

2.) Further data analysis: a.) It includes: How many students improve their understanding? b.) What is the highest percentage the student improved his answer? c.) What is the average percentage the students improved their answers? d.) Group/knowledge Transitions. e.) Learner's behaviours: it is to category the behaviour into: Rapid – Many attempt, Rapid – Few attempt, Slow – Many attempts, Slow – Few attempts, Rapid – Many – Long learning duration, and so on.

3.5 System Evaluation

The system evaluation includes two aspects. Two aspects of evaluation are conducted to assess the system: 1.) Subjective evaluation: In order to assess the student perception of the tests and feedbacks, a student survey and interview on the experiment of the system had been conducted after every trial. 2.) Objective evaluation: A large number of data had been recorded during the trials, in order to assess the performance and effectiveness of M-OFS. Furthermore, comparisons on the examinations and questionnaires results are carried out between students using and without using the MCQs-OFS and this is part of evidence in evaluating the system. In addition, three separate questionnaires test has been done three times before "English trial" and three times after, in order to compare the results of the test between before trial and after trial.

Survey and interview as a method to be applied in this research. Survey and Interview are conducted after each trial to assess learner's satisfaction and motivation in M-OFS. The students were asked to do the survey every time after they learning from the system, and some of the students were randomly selected to participate the interview. It can help to understand learners' satisfaction and motivation in using the system.

The survey includes 1.) students think the feedback is what they need; 2.) how do the students think about the system; 3.) students would like to use the system again; 3.) students would like to recommend the system to a friend or classmate in the future; 4.) students have never used similar system before.

For interviews, it includes: 1.) do the students feel this system is useful and helps them to improve their knowledge; 2.) what are the students expecting contents in the feedback. 3.) will the students lose patient by guessing the answers. 4.) Do the students want a picture of their learning process which can point out their weakness and a suggestion of how to improve their English.

3.6 System Enhancement

The system used in this research is based on ESDNN system, and it is not significant changes. Due to some limitations of the ESDNN system it is not feasible to apply it in some subject areas (e.g. feedback is very simple; system cannot display some mathematics symbols). It is necessary to enhance the system in order to achieve following objectives: 1, much easier to install and run in wider environment; 2, provide more accurate, detailed, specific, customized, and targeted feedback to the learners; 3, be able to apply to more area of study; 4, Compare and contrast the learning performance between each individual student and the average of other students in each stages of knowledge. The work of enhancing the system is shown in this section

Although there is no significant limitation of using system, some limitations are explored during testing the system. According to investigate and study the current ESDNN system, as an alternative database system to Oracle, MySQL is free, widely used, open source and much easier to install. There is a limitation of applying area (i.e. Mathematics area). The system cannot provide more than one topic's MCQs to user at a same time. The system cannot display mathematics symbol and diagram. The content of feedback is not good enough (i.e. it is tied on every particular question).

According to the limitations, the system enchantment work was carried out, and it is shown as following: 1.) Develop MySQL versions; 2.) Explore the feasible way to convert database from Oracle to MySQL; 3.) Data transfer; 4.) Test new version.

An example of coding is shown as figure 8: Modify the related programming code and system setting to adapt MySQL database (see figure 8)

```
<Context path="/qa" debug="5" reloadable="true" crossContext="true" docBase="C:/Tomcat5.5/webapps/qa">
  <Manager pathname="" />

  <Resource name="jdbc/qaDb"
            type="javax.sql.DataSource"  auth="Container"
            description="GMI database"
            maxActive="50" maxIdle="5" maxWait="10000"
            username="USERID" password="PASSWORD"
            driverClassName="oracle.jdbc.driver.OracleDriver"
            url="jdbc:oracle:thin:@DATABASE NAME"/>



  <Realm className="org.apache.catalina.realm.DataSourceRealm" debug="99"
     dataSourceName="jdbc/qaDb"
     localDataSource="true"
     userTable="qa_users" userNameCol="user_id" userCredCol="password"
     userRoleTable="qa_users" roleNameCol="user_role"/>
```

Figure 8: an example of JDBC setting of MySQL

3.7 Summary

In this chapter, the detailed methodology work of this research is mainly introduced, it includes four main parts: 1. the working principle of M-OFS system; 2. experimental environment and progress; 3. System Trials; 4. Evaluation. In these four parts, the second part is the most time consuming part with the most workload, which requires abundant careful work, while the third part are the most difficult parts to be completed, as it not only requires the support of the second part but also needs sufficient volunteer to take part in the testing to ensure the sample size.

In order to better assess the performance, effectiveness and adaptability, several trials have been applied to 3 totally different subject areas and in three universities of two countries. In next chapter, it will introduce the trials in details to better understanding how the system working across different subject.

4. Summary of Experimental Trial

4.1 Introduction

In the previous chapter, it introduced the principle of M-OFS and the method used to carry out target-oriented testing. In order to better assess the performance, effectiveness and adaptability, several trials have been applied to three totally different subject areas and in three universities of two countries. And the trials will be introduced in details in this chapter.

In this chapter, the detailed work of the trial is introduced and it includes three main parts: 1. preparation before testing (data collection, data input, data transfer, data training, groupings and grouping feature analysis, feedback design) ; 2. testing (trial apply); 3. Evaluation (survey, interview, and separate questionnaire). In these three parts, the first part is the most time consuming part with the most workload, which requires abundant careful work, while the second part and the third part are the most difficult parts to be completed, as it not only requires the support of the first part but also needs sufficient project students to take part in the testing to ensure the sample size and trial number.

Heretofore, we have already carried out six main trials: English trials (3 times, 250 participants in total), Mathematics (2 times. 21 participants in total), and Plagiarism (1 time, 156 participants), and two small trials: Information Security Management (1 time,), and Java Programming (1 time, 21 participants). All of the trials will be discussed in detail in this chapter**.** All eight trials are conducted. Java Programming trial, Plagiarism trial and Information Security Management trial are basically analyzed without survey and interview. Mathematics and English trials are further analyzed and discussed, especially English trials. Moreover, after introduced the system, the dean at JQU and KTU is very interested at this system, and he want the students in his faculty improve their knowledge, The students are invited to the experiment by dean of the faculty. In China, the dean are very powerful and students are pleasure to try new things to improve their learning in the university. The learning is the most important thing in the university, and all of the students try their best to get higher mark, and they always try to be a good student. Therefore, the students are happy to take the opportunity to learn from the system.

As the core part of the experiments, in order to understand the trial and how it be applied, the trials are introduced in details and with some real examples. This section is focused on the English trial. For example, the first trial of “English” the M-OFS was trained with the responses for 10 questions on a particular topic (English grammar) of the module English Language Study obtained from previous cohort of students. After training, the prepared appropriate feedback text was written by a supervisor of Foreign Language School of Kunming Technology University for each of the group of students' responses that address the conceptual errors implicit in combinations of incorrect answers. During the trial, a current cohort of students was asked to provide responses on the same MCQs, they were given the feedback on the combination of incorrect answers and their responses recorded in the database. The feedback texts are composed around the pattern groupings and are aimed at misconceptions that may have caused the incorrect answers common within the pattern group. An example of a typical response to the questions in table 3 is (D, C, D, D, D, D, A, C, A, A).

According to this set of answers, the system will automatically select an appropriate feedback (group 4's feedback), which is shown as below:

Group 4 Feedback:

Three points should be stressed. First, the logical subject of the adverbial phrase should agree with that of the main clause. E.G. Singing, I came. Second, the usage of various noun clauses should be familiar with. E.G. The news came that he died. ("that" does not serve as any part of the clause.) Third, subjunctive mood should be employed in the clause after some special verbs or adjectives. E. G. It is necessary that the door be closed. (Reference to Zhengjing Ting)

By reading the feedback of group 13 (a feature group created by system grouping), the students should be able, either immediately or after some reflection, to improve their understanding of this key knowledge point, and then improve their answer to question 2 to D. "Melissa knew that performing was". According to this change the system will automatically provide another suitable feedback which is group 8's feedback as below:

Group 8 Feedback:

Two points should be stressed. First, the usage of various noun clauses should be familiar with. E.G. The news came that he died. ("that" does not serve as any part of the clause.) Second, subjunctive mood could be employed with the auxiliary verbs. E. G. You should have handed in your assignment. (It is wrong that you did not hand in your assignment.)

At the end, the students perhaps can improve their answers to all correct answers after another couple of attempts, and then they will get the feedback of group 20:

Group 20 Feedback:

Well Done! All correct!

When student using the system to learn, the students are guided by the system towards achieving all of the learning outcome from the system. Furthermore, the system will also record all of their behaviour by each individual student, in order to analyze their learning behaviour. Moreover, after using the system to learn, students are invited to do the survey. In addition, students are randomly selected to do the interview and separate questionnaire, in order to assess the the performance, effectiveness and adaptability of the system, and the satisfaction of students using the system.

4.2 System Trial

4.2.1 Data Input

Because the initial data needs to be inputted and stored in a standardized format in text files in order to train them by SDNN (see figure 9), due to most of the initial data are on the paper work. The size of each subject data is shown as following: 1.) English subject: the data were collected from the final examination of English and it consists of 120 MCQs. There were 250 students who attended the examination. There were 30000 answers collected in total. 2.) Mathematic subject: the data were collected from the final examination of English and it consists of 20 MCQs. There were 74 students who attended the examination. There were 370 answers collected in total. 3.) Plagiarism subject: the data were collected from the final examination of English and it consists of 5 MCQs. There were 156 students attended the examination. There were 780 answers were collected in total. 4.) Java Programming subject: the data were collected from the final examination of English and it consists of 37 MCQs. There were 250 students attended the examination. There were 24311 answers were collected in total. 5.) Information Security Management subject: the data were collected from the final examination of English and it consists of 49 MCQs. There were 41 students attended the examination. There were 453 answers were collected in total.

A	B	C	D	E	F	G	H	I	J	K	L
	Student no.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	u0943342	E	B	A	E	B	A	E	D	C	A
2	U0834729	B	B	A	E	B	A	C	D	C	C
3	U0915704	E	E	B	A	B	A	C	A	B	C
4	U0917632	B	A	A	E	B	D	B	A	B	C
5	U0835864	A	D	B	E	D	C	B	A	B	C
6	U0834046	E	A	A	E	D	C	B	A	C	C
7	U0918009	C	E	A	E	D	B	B	A	B	C
8	U0922241	C	E	A	E	D	E	B	A	D	C
9	U0910051	E	E	A	E	B	C	B	A	B	C
10	U0944886	C	E	A	E	D	C	B	A	A	C
11	u0911442	C	E	A	E	D	C	B	A	B	C
12	U0837477	E	E	A	B	D	A	B	A	E	D
13	U0843508	E	B	E	E	D	E	A	B	C	D
14	U0843704	C	E	A	C	D	C	B	A	B	C
15	U0849721	E	C	A	E	B	A	C	B	C	A
16	U0853303	B	C	C	E	B	A	C	D	A	A
17	U0853614	C	E	A	E	D	C	B	A	B	C
18	U0902345	E	C	C	C	B	A	C	D	A	C
19	U0903064	D	D	C	B	E	E	C	B	A	D
20	U0903865	C	E	A	E	D	A	B	A	C	C
21	U0904124	E	A	A	E	E	B	B	A	B	C
22	U0904258	C	B	B	E	B	E	C	D	A	C
23	U0904455	E	A	A	E	E	B	B	D	B	C
24	U0904543	E	D	B	E	B	D	B	D	A	NA
25	U0904548	C	A	C	E	C	D	A	C	D	B
26	U0907821	C	E	A	E	D	C	B	A	B	C
27	U0912086	B	A	D	A	D	D	B	D	D	D

Figure 9: example of a set of questions

4.2.2 Data Transfer

Before the data training by SDNN, all of the data (answers e.g. a,b,c,d) need to be transferred into binary code by following rules (“A=0,0,0,0,1,”, “B=0,0,0,1,0,”, “C=0,0,1,0,0,”, “D=0,1,0,0,0,” “E= 1,0,0,0,0,”). It is an example of before data transfer (see figure 9) and after data transfer (see figure 10). Moreover, all of the input data need to be transferred before training with SDNN.

	Student no.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
1	u0943342	1,0,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,0,1,	0,1,0,0,0,
2	U0834729	0,0,0,1,0,	0,0,0,1,0,	0,0,0,0,1,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,1,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,1,0,0,	1,0,0,0,0,
3	U0915704	1,0,0,0,0,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,0,0,1,	0,0,1,0,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	1,0,0,0,0,
4	U0917632	0,0,0,1,0,	0,0,0,0,1,	0,0,0,0,1,	1,0,0,0,0,	0,0,0,1,0,	0,1,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
5	U0835864	0,0,0,0,1,	0,1,0,0,0,	0,0,0,1,0,	1,0,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
6	U0834046	1,0,0,0,0,	0,0,0,0,1,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,1,0,0,	0,0,1,0,0,	0,0,1,0,0,
7	U0918009	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,0,1,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
8	U0922241	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,
9	U0910051	1,0,0,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,0,0,1,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
10	U0944886	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,0,1,	0,0,1,0,0,	0,0,1,0,0,
11	u0911442	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
12	U0837477	1,0,0,0,0,	1,0,0,0,0,	0,0,0,0,1,	0,0,0,1,0,	0,1,0,0,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,1,0,0,0,
13	U0843508	1,0,0,0,0,	0,0,0,1,0,	1,0,0,0,0,	1,0,0,0,0,	0,1,0,0,0,	1,0,0,0,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,1,0,0,0,	0,0,1,0,0,
14	U0843704	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	0,0,1,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,0,1,0,
15	U0849721	1,0,0,0,0,	0,0,1,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,1,0,0,	0,0,0,1,0,	0,0,1,0,0,	0,0,0,0,1,	1,0,0,0,0,
16	U0853303	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,1,0,0,	0,1,0,0,0,	0,0,0,0,1,	0,0,0,0,1,	0,0,0,0,1,
17	U0853614	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
18	U0902345	1,0,0,0,0,	0,0,1,0,0,	0,0,1,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,1,0,0,	0,1,0,0,0,	0,0,0,0,1,	0,0,1,0,0,	1,0,0,0,0,
19	U0903064	0,1,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	1,0,0,0,0,	1,0,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,1,0,0,0,	1,0,0,0,0,
20	U0903865	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,0,0,1,	0,0,1,0,0,	0,0,1,0,0,	0,0,1,0,0,
21	U0904124	1,0,0,0,0,	0,0,0,0,1,	0,0,0,0,1,	1,0,0,0,0,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
22	U0904258	0,0,1,0,0,	0,0,0,1,0,	0,0,0,1,0,	1,0,0,0,0,	0,0,0,1,0,	1,0,0,0,0,	0,0,1,0,0,	0,1,0,0,0,	0,0,0,0,1,	0,0,1,0,0,	0,0,0,1,0,
23	U0904455	1,0,0,0,0,	0,0,0,0,1,	0,0,0,0,1,	1,0,0,0,0,	1,0,0,0,0,	0,0,0,1,0,	0,0,0,1,0,	0,1,0,0,0,	0,0,0,1,0,	0,0,1,0,0,	0,0,1,0,0,
24	U0904543	1,0,0,0,0,	0,1,0,0,0,	0,0,0,1,0,	1,0,0,0,0,	0,0,0,1,0,	0,1,0,0,0,	0,0,0,1,0,	0,1,0,0,0,	0,0,0,0,1,	0,0,0,0,0,	1,0,0,0,0,
25	U0904548	0,0,1,0,0,	0,0,0,0,1,	0,0,1,0,0,	1,0,0,0,0,	0,0,1,0,0,	0,1,0,0,0,	0,0,0,0,1,	0,0,1,0,0,	0,1,0,0,0,	0,0,0,1,0,	0,0,0,1,0,
26	U0907821	0,0,1,0,0,	1,0,0,0,0,	0,0,0,0,1,	1,0,0,0,0,	0,1,0,0,0,	0,0,1,0,0,	0,0,0,1,0,	0,0,0,0,1,	0,0,0,1,0,	0,0,1,0,0,	0,0,0,0,1,
27	U0912086	0,0,0,1,0,	0,0,0,0,1,	0,1,0,0,0,	0,0,0,0,1,	0,1,0,0,0,	0,1,0,0,0,	0,0,0,1,0,	0,1,0,0,0,	0,1,0,0,0,	0,1,0,0,0,	0,0,0,1,0,

Figure 10: example of binary code of a set of questions

4.2.3 Data Training

As described in chapter 3.4.2 Data Training. Use SDNN to train every set of data many times, in order to find out the best results. (It requires training the data many times by changing settings until the output becomes to stable. Setting example: “2000epoch, threshold>0.5, dSDNN_layer1_node = 50, hidden note=100, sSDNN_layer2_node = 25”). The work of training of each subject are show as follows: English: (3 sets of data) × (16 training/set)=48 times training; Mathematic: (4 sets of data) × (9 training/set)=36 times training; Plagiarism: (2 sets of data) × (10 training/set)=20 times training; Java Programming: (4 sets of data) × (16 training/set)=64 times training; ISM: (4 sets of data) × (6 training/set)=24 times training

4.2.4 Results of Groupings

As described in chapter 3.4.3 Grouping results and feature analysis. The minimum stability requirement is 80% stability. Training is continued until the number of pattern grouping is the same from one epoch to the next one, e.g. in figure 11, the stability is 95.9%. After all analysing completed, to compare and contrast several analysing results in order to find out the best one (the most stable grouping result/the least changes group result). The works of data analysis are show as follows: English: 48 times; Mathematic: 36 times training; Plagiarism: 20 times training; Java Programming: 64 times training;

Information Security Management: 24 times training.

Epoch 999	Epoch 1000	Epoch 1001
1	1	1
...	...	...
5	5	1
11	11	1
11	11	1
17	11	0
5	5	1
5	5	1
20	20	1
17	10	0
5	5	1
10	10	1
11	11	1
17	17	1
12	12	1
14	14	1
10	10	1
4	4	1
19	19	1
11	11	1
19	19	1
11	11	1
8	8	1
17	17	1
11	11	1
14	14	1
		Changes 4.1%

Figure 11: Compare and contrast among the last 3 epochs

4.2.5 Group Features analysis

As described in chapter 3.4.3 Grouping results and feature analysis. In order to write the feedback, it requires identifying the number of groups and the feature of each group according to the analysing results. For example, there are 10 groups in figure 12. In another word, there are 10 different strong features in this set of data. Furthermore, the feature of group 2 is: A, D, D/mix, B, D/mix. This feature means that most students in this group answered on MCQs were: A on question 1, D on question 2, D and mix answer on question 3, B on question4, D and mix answers on question 5. (mix answer means it consist of different answers)

After identifying the features of the group, designers/teachers can write the diagnostic feedback for this group according to the results of comparison between the feature and the correct answer. The features of each subjects and topic are show in table 5.

Group	Features
1	<E/B><D/A><A><D/A><A>
2	<B> <D> <C> <E><D/A>
3	<E><D><C/NA><E><B/D>
4	<A><A><A/E><BA><BD>
5	<C/E><B/A><C><A><C>
7	<E><A><D><D/B><D>
14	<E/C><B><D><C/D><A>
17	<C><D/c/e><D/b/c><B/C><A/mix>
18	<C/b/e><A><C><C><C/d>
19	<B><D><D><B/E><B/D>

Figure 12: an example of group's features

Subject	Topics	Content
English	English Grammar	17 groups, 10 answers/group
	English Vocabulary	19 groups, 10 answers/group
	English Reading Comprehension	20 groups, 10 answers/group
Mathematic	Expression and Equation	17 groups, 5 answers/group
	Functions and Measurements	12 groups, 5 answers/group
	Logarithm and Exponent	15 groups, 5 answers/group
	Fractions and Mean	17 groups, 5 answers/group
Plagiarism	Set 1	19 groups, 5 answers/group
	Set 2	11 groups, 5 answers/group
Java Programming	Set 1	21 groups, 9 answers/group
	Set 2	23 groups, 11 answers/group
	Set 3	19 groups, 5 answers/group
	Set 4	25 groups, 12 answers/group
Information Security	Set 1	5 groups, 5 answers/group
	Set 2	8 groups, 14 answers/group

Management	Set 3	7 groups, 17 answers/group
	Set 4	9 groups, 13 answers/group

Table 5: Features of each group

The number of groups is the number of strong features of the students group. The number of answers is how many answers are in one group.

4.2.5 Feedback Design

As described in chapter 3.4.4 Feedback Design. This is the most important and difficult part of work. Several issues have to be taken into account. For example: what kind of content of feedback can attract student, and what elements should be included in the feedback? In term of design the feedback, it requires the supervisor who supposed to write the feedback have sufficient knowledge in the selected subject area. The number of feedback designed for this research is shown in table 6.

4.2.6 System Set Up and Test

As described in chapter 3.4.5 System Set Up and Test. To set up a trial for a selected module, it requires changing some settings. All settings must be correct; otherwise the system cannot work correctly. For example: the number of input patterns, the module number, the topic number, the correct answers and so on. A test needs to be conducted after set up each trial. Nine trials were set up. Six of them are applied. The details of each subject are shown as follows: English: three trials in total; English Grammar: 1 trial; English Vocabulary: 1 trial; English Reading Comprehension: 1 trial; Mathematic:

four trials in total; Expression and Equation: 1 trial; Functions and Measurements: 1 trial; Logarithm and Exponent: 1 trial; Fractions and Mean: 1 trial; Plagiarism: 1 trial in total; Java Programming: 1 trial in total

4.2.7 Trial Implementation

The summary of six main trials are introduced in this section. Three trials have been applied at Kunming Technology University and Jinqiao University in China, and two trials were conducted in University of East London, and one trial was conducted in University of East London. The details of each trial are shown in the table 7.

Subject	Topic	Number of Feedback
English	English Grammar	17 feedbacks
	English Vocabulary	19 feedbacks
	English Reading Comprehension	20 feedbacks
Mathematic	Expression and Equation	17 feedbacks
	Functions and Measurements	12 feedbacks
	Logarithm and Exponent	15 feedbacks
	Fractions and Mean	17 feedbacks
Plagiarism	Set 1	19 feedbacks

	Set 2	11 feedbacks
Java Programming	Set 1	21 feedbacks
	Set 2	23 feedbacks
	Set 3	19 feedbacks
	Set 4	25 feedbacks
Information Security Management	Set 1	5 feedbacks
	Set 2	8 feedbacks
	Set 3	7 feedbacks
	Set 4	9 feedbacks

Table 6: Group Feedback

The Feedback is designed for each group, and the number of feedback is how many feedbacks under one topic.

Subject	**Trial**	**Detail**
English	First trial	Date: 10-Dec-2010 Location: Jinqiao University, China Participants: 84 students Duration: 50 minutes

	Second trial	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 84 students Duration: 1 hour
	Third trial	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 82 students Duration: 50 minutes
Mathematic	First trial	Date: 29-Nov-2010 Location: Library, University of East London, UK Participants: 13 students Duration: 20 minutes
	Second trial	Date: 30-Nov-2010 Location: Library, University of East London, UK Participants: 7 students Duration: 15 minutes
Java Programming	One trial	Date: 16-Mar-2010 Location: Library, University of East London, UK Participants: 12 students Duration: 20 minutes

Table 7: Details of Each Trial

4.3 System Evaluation

As described in 3.6 System Evaluation. Two aspects of evaluation are conducted to assess the system: 1.) Subjective evaluation 2.) Objective evaluation. Survey and Interview, and separate questionnaire methods are used to evaluate this system. All of the data of each trial was summarized and analyzed and the results are discussed in chapter 5 Findings.

To investigate and evaluate how the M-OFS guide and support students to learn, three English experiments were under taken by level 2 and level 3 students at JinQiao University (JQU) and Kunming Technology University (KTU) in China during the academic year 2010-2011. An example of 1st experiment is introduced as following. In the first experiment, data was collected from 148 students taking English language courses whom were randomly separated into two groups. The experimental group of 83 students used M-OFS, and the control group of 65 students received the same training but without using M-OFS. The system trial includes 10 MCQs with 4 potential answers, related to English grammar. The duration of this trial is flexible. When students were using M-OFS, they were encouraged to answer the MCQs (submit their answers) as many times as they wish until they got all the correct answers or gave up (students were not given answers or how many answers were correct in their feedback, except that they answered all correct answers). Two MCQ paper tests with different questions from system trials were applied to 116 students, and 83 students participated in both paper test and system trial. 83 students completed survey after second paper test. This system trial, paper test and survey were completed in practice lessons in a computer room at JQU.

The survey, interview, and separate questionnaire is not used in every trials, and the details work is introduced in this chapter.

4.3.1 Survey

Two surveys are designed for "Mathematics trial" and "English trial", and they are applied after each trial. The content of mathematics survey is designed first and the English survey are improved from the mathematics survey. All of the students are invited to do the survey. 250 students conducted surveys of "English trial" and 20 conducted surveys of "Mathematics trial".

English: survey of "English trial" includes 7 questions (see figure 13). 250 students conducted this survey: The details of survey is shown in table 8.

Survey	**Detail**
First Survey	Date: 10-Dec-2010 Location: Jinqiao University, China Participants: 84 students
Second Survey	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 84 students

Third Survey	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 82 students

Table 8: Details of English Survey

Survey	**Detail**
First Survey	Date: 29-Nov-2010 Location: Library, University of East London, UK Participants: 13 students
Second Survey	Date: 30-Nov-2010 Location: Library, University of East London, UK Participants: 7 students

Table 9: Details of Mathematics Survey

UEL E-MCQs Survey 2010
English 6 Questions

Student number: ___________

Ticket the box

How satisfied are you with our E-learning MCQs system?

1. ☐ Very satisfied
2. ☐ Somewhat satisfied
3. ☐ Somewhat dissatisfied
4. ☐ Very dissatisfied

Which of the following are true? Select all that apply.

1. ☐ The feedbacks were exactly what you need
2. ☐ The feedbacks were a part of what you need
3. ☐ The feedbacks were not what you need

Please rate the following attributes of our E-MCQs system.

	Excellent	Good	Fair	Poor
Studies support	☐	☐	☐	☐
Usefulness of feedback information	☐	☐	☐	☐
Quantity of feedback content	☐	☐	☐	☐
Facility of learning	☐	☐	☐	☐
Design of feedback	☐	☐	☐	☐
User support	☐	☐	☐	☐

How likely are you to use our E-MCQs system again?

1. ☐ Very likely
2. ☐ Somewhat likely
3. ☐ Very unlikely

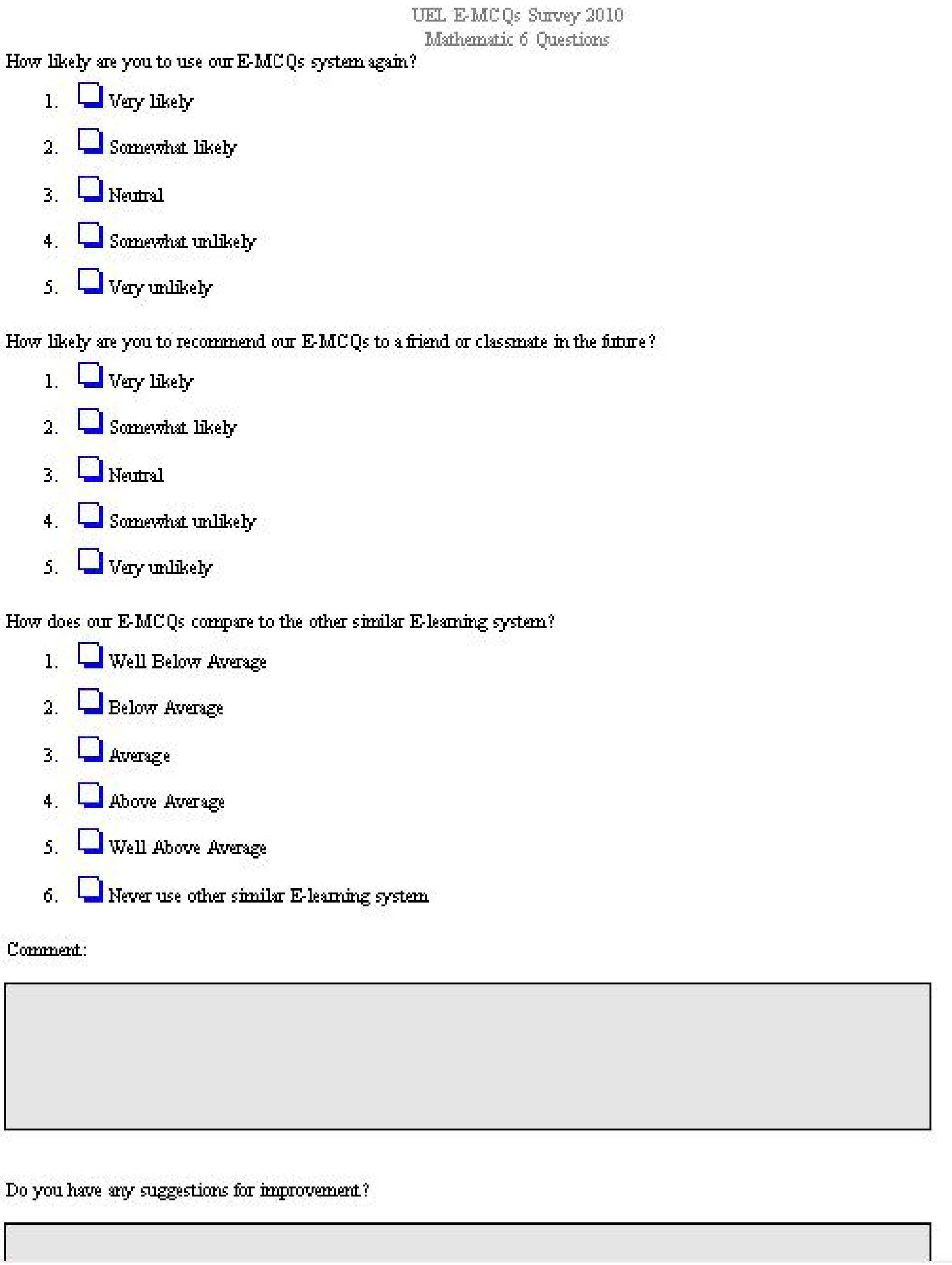
How likely are you to recommend our E-MCQs to a friend or classmate in the future?

1. ☐ Very likely
2. ☐ Somewhat likely
3. ☐ Very unlikely

Figure 13: a part of English survey

Mathematics: survey of “Mathematics trial” includes 9 questions (see figure 14).

20 students conducted this survey:



UEL E-MCQs Survey 2010
Mathematic 6 Questions

How likely are you to use our E-MCQs system again?

1. Very likely
2. Somewhat likely
3. Neutral
4. Somewhat unlikely
5. Very unlikely

How likely are you to recommend our E-MCQs to a friend or classmate in the future?

1. Very likely
2. Somewhat likely
3. Neutral
4. Somewhat unlikely
5. Very unlikely

How does our E-MCQs compare to the other similar E-learning system?

1. Well Below Average
2. Below Average
3. Average
4. Above Average
5. Well Above Average
6. Never use other similar E-learning system

Comment:

Do you have any suggestions for improvement?

Figure 14: a part of Mathematics survey

4.3.2 Interview

The interview includes several questions in order to investigate how the system guiding students deep learning. Students are randomly selected to participate to the interview, and the interview last about 10 minutes. The interview is face-to-face formal interview and it is conducted by teacher. In total 20 students conducted the interview after English trial, and 8 students participate interview after Mathematics trial. They provided valuable advice. The analysis of results will be shown in the data analysis section. The details of interviews is shown in table 10.

Subject	Interview	Detail
English	First Interview	Date: 10-Dec-2010 Location: Jinqiao University, China Participants: 8 students
	Second Interview	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 6 students
	Third Interview	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 6 students

Mathematics	First Interview	Date: 29-Nov-2010 Location: Jinqiao University, China Participants: 5 students
	Second Interview	Date: 30-Nov-2010 Location: Kunming Technology University, China Participants: 3 students

Table 10: Details of Interview

4.3.3 Separate Questionnaires Test

English: Six times of questionnaires test have been designed and applied.

Separate Questionnaire Test	**Detail**
Test before first trial	Date: 3-Dec-2010 Location: Jinqiao University, China Participants: 80 students
Test after first trial	Date: 17-Dec-2010 Location: Jinqiao University, China Participants: 22 students

Test before second trial	Date: 21-Dec-2010 Location: Kunming Technology University, China Participants: 80 students
Test after second trial	Date: 25-Dec-2010 Location: Kunming Technology University, China Participants: 80 students
Test before third trial	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 80 students
Test after third trial	Date: 23-Dec-2010 Location: Kunming Technology University, China Participants: 80 students

Table 11: Details of Separate Questionnaire Test

In order to better assess the effectiveness of using the system, three separate questionnaires has been designed and conducted for their corresponding trial. Each questionnaire is designed based on the topic of the trial, and it includes 20 MCQs. Furthermore, each questionnaire was applied twice for its corresponding trial, the first one before trial, the second one after trial. Any feedback or answers of the questionnaires were not given to student before or after every test. The details of each test are shown in table 9.

4.4 Summary

In this chapter, it introduced the trials in details with some example. Furthermore, it shows how many works had been done in this research. Moreover, the trials were designed and conducted step by step: first step: initial system test was applied to a small amount of students (e.g. Java Programming, Information security System, and Plagirism); second step: enhanced system with better better feedback was applied to a bigger amount of students (e.g. Mathematics trial); third step: stronger system with intelligent diagnostic feedback was applied to a large cohort of students (e.g. English trial). Survey and interview were applied to Mathematics and English trial. Furthermore, separate questionnaire test was applied to English trial.

The trial helps to understand learner and their learning behaviour more. More understanding of learner and their learning behaviour helps to build better trial. Better trial helps to understand leaner and their learning behaviour more. Based on the experimental results, in next chapter it will discuss the findings of understanding of learners and their learning behaviour.

5. Findings

5.1 Introduction

In last chapter, it introduced sufficient target-oriented experiments were carried out in different fields. In this chapter, it will discuss the findings from experiment such as: the performance and effectiveness of using this novel e-learning system; how the system support students deep learning; students' learning behaviour; knowledge state transactions. It will also discuss the results refer to the four hypothesis: 1, during the trials, students improved their understanding by reading given feedback (H1); 2, after using M-OFS system, students get higher mark in a separate questionnaire than before (H2); 3, the experimental students get higher marks than control group at the final examination (H3); 4, most students are satisfied with this system (H4).

The findings of experiments are composed of three main parts. Firstly, results from evaluating the system by collecting and analysing a large number of testing data reflecting the students' learning gains. Secondly, the results from survey and interview data reflecting the students' satisfaction and attitudes. Thirdly, the results of learning behaviour and knowledge transaction analysis.

The findings will lead to an understanding of the potential of the on-line diagnostic feedback approach across different subject areas. Furthermore, the findings should also produce guidelines for the design principles of on-line MCQs in the context of diagnostic feedback learning environments. In the further work, the findings will also be used to enhance the existing system. The details of the experiments results, result data analysis, learning behaviour, knowledge state transaction and further findings are shown in this chapter. It includes two parts: Results of Trials and Findings of Learning

Behaviour.

5.2 Results of Trials

5.2.1 Mathematics Trial:

5.2.1.1 Trial Result Analysis

A Trial was applied to students in University of East London (UEL) for Mathematics subject. There is no separated questionnaire test of mathematics (separated paper test) is applied. No students got the all correct answers at the first attempt. Two students achieve the all correct answers at the end after several attempts. On average the students attempted 3 times over 15 minutes. 55% students improved their answer during the trial, and it is one of the evidence to confirm hypothesis H2.

5.2.1.2 Survey and Interview Result Analysis

According to the data analysis of Mathematics trial, 87.5% students satisfied with M-OFS, and one student was dissatisfied, , and it is one of the evidence to confirm hypothesis H1. 56.25% students believe that the feedbacks were exactly what they need, and 31.25% students feel the feedbacks were a part of what they need, and 12.5% students are not sure whether the system is useful or not. 43.75% students were very likely to use M-OFS again, 37.5% students were somewhat likely to use M-OFS again, 12.5% were neutral, and 6.25% students were unlikely to use M-OFS again. 43.75% students are very likely to recommend M-OFS to a friend or classmate in the future, the same percent students are somewhat likely to do that, 6.25% students are neutral, and 6.25% students are unlikely to do that. 81.25% students feel the following attributes of

M-OFS are excellent and good: ease of navigation, accuracy of feedback information, quality of feedback content, quantity of feedback content, , and this is the evidence to confirm hypothesis H1. 31.25% students thought the layout/interface of M-OFS is fair and need to be improved.

Some of these students also gave some feedback of using M-OFS from interviews. For example: some of students reflect that the feedback received was hard to attribute to the overall questions, and the feedback should have a better overview, and there are too many choices for each question (A-E).

5.2.2 English Grammar Trial in JQU

The English Grammar Trial applied to students in Jinqiao University (JQU) for English grammar subject. The English Grammar Trial includes 5 components as follows: 1) Separate MCQs paper test (20 questions); 2) System trial (English grammar, 10 MCQs); 3) Survey with 6 questions; 4) Interviews; 5) Compare and contrast final examination results between control group and experimental group. Throughout there are 148 students involved in this experiment. 116 students completed the separate MCQs paper test before and after the system trial without any feedback. 83 of 116 students participated fully in the system trial doing both the system and paper test. The survey was conducted after the system trial, and all 83 students filled the survey form. 16 random students entered the interview. Details are presented in this chapter.

5.2.2.1 Trial Result Analysis

The purposes of analysing the results are to test the system, and help evaluate the research hypothesis. Furthermore, to analyze and assess learner's behaviours. The summary of using feedback system for English grammar trial is shown in table 12.

The methods of analysing the results are shown as follows: 1) to apply the system trial to students, and then the students can learn from the system; 2) the system will automatically record the data of students' learning process; 3) to collect and analyze the data in order to achieve the goals.

Subject	English grammar
Place	Computer Lab, Jinqiao University, Yunnan, China
Date	10-Dec-2010
Time	12:50 pm
Duration	flexible
Participants	116 students of School of Foreign Language-Professional English (Second year and Third year students in Bachelor degree).
Total attempts	1118
Total minutes	2143 minutes
Ave attempts	13.5
Ave time (per student)	25.8 minutes
Ave time (per attempt)	1.92 minutes

Ave score increased	12.77%
Min attempts	1
Max attempts	106
Increased their score	55/83 = 66.3%
Best improvement	70%
Decreased	9/83 = 10.8%
Unchanged	14/83 = 16.9%
Get all correct answer at the end	5/83 = 6%
Get all correct answer in first attempt	0/83 = 0%

Table 12: summary of English trial – feedback system

Students are not given the correct answer in the feedback. Consequently, guess work is discouraged. There are 1,048,676 possible answer combinations, and they are not been told which is right answer. Random selection would be an unsuccessful strategy. They must try to learn from the feedback, and they can only learn from the feedback in order to achieve all correct answer. In the English Grammar Trial, 148 students are involved in the first trial. 116 students completed the separate MCQs paper test before and after using the system. 83 students participated in system trial, and separate MCQs paper tests. For system trials, a total of 1118 answers/attempts were submitted and a total 2143 minutes were spent by 83 participants. All of the students submitted their answers at least once. The maximum number of attempts was 106 times and the minimum was 1. The average attempts for each student is 13.5 times. The average time spent by each student is 25.8 minutes and the average time of each attempt is 1.92 minutes. 2 students (2.4%) spent more than 60 minutes. 35 (42.2%) students spent more than the average time. No students achieved the all correct answers at the beginning. 55 (66.3%) students increased their scores by an average of 12.8%, whilst 1 student increased his score by 70%, and this is the evidence to confirm hypothesis H2. In this trial, with 10 questions and 4 possible answers, there are more than 1 million possible combinations of answers, thus the students are unlikely to make improvement by guessing answers; hence, the results show the feedback had a positive impact which is the evidence confirm hypothesis H2. For separate MCQs paper tests, the average score before system trial is 51.6%, and the average score after system trial is 59.1%. One student increased his score by 40%. 74% students increased their scores. In this test, the students were not given any answers or feedback between first (before system trial) and second (after system trial) test; furthermore, the first trial were applied 3 hours before the system trial and the second test were conducted 30 minutes after system trial; hence, the students are only learnt by using M-OFS but not any other ways; thus the results above are confident, therefore, it is the evidence to confirm hypothesis H3. In addition, this result is also the evidence to support hypothesis H2. For final examination, both the experimental group and the control group enter the same 4 days final examination. The experimental group got 79.52% and control group got 71.3% in English grammar module. This result confirms the hypothesis H4; furthermore, it also supports hypothesis H2.

5.2.2.2 Survey and Interview Result Analysis

Survey Result Analysis

The purposes of this part are to assess students' satisfaction of using the system, and to support the hypothesis. The methods are shown as follows: 1) to design and apply surveys to students; 2) to analyze and summary the results. Table 13 shows the summary of survey for English trial. Table 14 shows the results of survey for English trial.

Subject	English grammar
Location	Computer Lab, Jinqiao University, Yunnan, China
Date	10-Dec-2010
Time	14:00 pm
Duration	flexible
Participants	83 students of School of Foreign Language-Professional English (Level 2 and Level 3).

Table 13: summary of survey

	Excellent	Good	Fair	Poor
Studies support	❏15. 7%	❏57. 8%	❏25. 3%	❏1. 2%
Usefulness of feedback information	❏14. 5	❏56. 6%	❏24. 1%	❏4. 8%
Quantity of feedback content	❏10%	❏59%	❏28%	❏3%
Facility of learning	❏15. 7%	❏51. 8%	❏28. 9%	❏3. 6%
Design of feedback	❏15. 7%	❏44. 6%	❏37. 3%	❏2. 4%
User support	❏12. 1%	❏55. 4%	❏27. 7%	❏4. 8%

Table 14: result of survey

According to the results, there are 71.1% students were satisfied with using system, and there are 28.9% students are dissatisfied. 84.4% students think the feedback is what they need, and 12% students disagree that the feedback is what they need. Furthermore, 90.4% students would like to use the system again, and 92.8% students would like to recommend the system to a friend or classmate in the future. 81.9% students never used similar system before.

Interview Result Analysis

The purposes of this part are to assess students' satisfaction of using the system, and find out the deficiencies of system. The methods are shown as follows: 1) to design and apply interview to students; 2) to analyze and summary suggestion of students. Table 15 shows the summary of interview for English grammar trial.

Subject	English grammar
Location	Computer Lab, Jinqiao University, Yunnan, China
Date	10-Dec-2010
Time	afternoon
Duration	flexible
Participants	16 students of School of Foreign Language-Professional English (Level 2 and Level 3).

Table 15: summary of interview

There are several findings through the interview, and they are: 1) most students feel this system is useful and help them to improve their knowledge; 2) most students want exactly answer in the feedback at the end; 3) students also want a picture of their learning process which can point out their weakness and a suggestion of how to improve their English future; 4) some students feel that if they tried many time but cannot find the correct answer, they will lose patience at the end.

The survey and interview were conducted after the system trial. 83 (100%) students conducted the survey. 16 (19%) students were randomly chosen for interview. For the survey, 71.1% students are satisfied with using system. 84.4% students think the feedback is what they need. Using M-OFS to learn were positively evaluated by students, illustrate that it is the evidence to confirm hypothesis H1. 90.4% students would like to use the system again. 92.8% students would like to recommend the system to a friend or classmate in the future. 81.9% students have never used similar system before. For interviews, most students (94%) feel this system is useful and helps them to improve their knowledge, it is the evidence to confirm hypothesis H1; moreover, 69% students want the exact answers in the feedback in the end. Students also want a picture of their learning process which can point out their weakness and a suggestion of how to improve their English. Some students feel that if they tried many times but cannot find the correct answer, they will lose patience in the end. The details of survey and interview are introduced as following.

5.2.2.3 Separate MCQs Test Results Analysis

The purposes of this part are to test and evaluate the system, and support the research hypothesis. The details of separate MCQs test are shown in table 16. Furthermore, analyze and assess learner's behaviours. The methods are shown as follows: 1) to conduct two tests with the same MCQs which include the same topic as MCQs of the system trial, but totally different questions. The first test is applied 3 hours before the system trial, and the second test is applied immediately after the system trial, and the feedback is given to students their answer; 2) to collect the results of two test and compare them, in order to find out if the students learn from the system.

Subject:	English grammar
Place:	First test: classroom, Jinqiao University, Yunnan, China Second test: Computer Lab, Jinqiao University, Yunnan, China
Date:	10-Dec-2010
Time:	In the morning
Duration:	Flexible
Participant:	116 students of School of Foreign Language-Professional English (Level 2 and Level 3)
Students' score:	Before system trial: average score 51.6% After system trial: average score 59.15%
Average score increased	7.55%
Most increased:	40% (Student no. 200916031222)

Score changed	74% students increased their scores

Table 16: summary of English trial – separate MCQs test

5.2.2.4 Comparison and Contrast of Control Group and Experimental Group in Final Examination

The purpose of this part is to evaluate system and to support the hypothesis. And the methods are shown as follows: 1) Select random students to experimental group, and others to be the control group; 2) To collect the final examination grads of two groups; 3) To compare and contrast final examination results between control group and experimental group.

Subject: English grammar

Location: Jinqiao University, Yunnan, China

Date: Feb-2011

Participants: 115 students of School of Foreign Language-Professional English (Second year and Third year students in Bachelor degree).

5.2.3 English Grammar Trial in KTU

As a English Grammar Trial already described in details from last section. In this section, it will introduce a summary of English Grammar Trial in KTU instead of detailed introduction.

134 students are involved in the first experiment. 85 students participated in system trial. 101 students completed separate MCQs paper tests, and 60 students completed the separate MCQs paper tests before and after using the system. For system trials, a total of 563 answers/attempts were submitted and a total 1465 minutes were spent by 85 participants. All of the students submitted their answers at least once. The maximum number of attempts was 21 times and the minimum was 1. The average attempts for each student is 6.6 times. The average time spent by each student is 17.2 minutes and the average time of each attempt is 2.6 minutes. 44 (51.8%) students spent more than the average time. 3 students achieved the all correct answers at the beginning. 63 (74.1%) students increased their scores, whilst 1 student increased his score by 80%.

In this trial, with 10 questions and 4 possible answers, there are more than 1 million possible combinations of answers, thus the students are unlikely to make improvement by guessing answers; hence, the results show the feedback had a positive impact which is the evidence confirms hypothesis H2. For separate MCQs paper tests, the average score before system trial is 53.5%, and the average score after system trial is 62.65%. One student increased his score by 40%. 74% students increased their scores. In this test, the students were not given any answers or feedback between first (before system trial) and second (after system trial) test; furthermore, the first trial were applied 9 hours before the system trial and the second test were conducted 30 minutes after system trial; hence, the students are only learnt by using M-OFS but not any other ways; therefore the results above are confident, and it they are the evidences to confirm hypothesis H3. In addition, the results are also the evidences to confirm hypothesis H2. For final examination, both the experimental group and the control group enter the same 6 days final examination. The experimental group got 81.2% and control group got 73.7% in English grammar module. This result is the evidence to confirms the hypothesis H4 and also hypothesis H2. The survey and interview were conducted after the system trial. 83 (100%) students conducted the survey. 16 (19%) students were randomly chosen for interview. For survey, 71.1% students are satisfied with using system. 84.4% students think the feedback is what they need. Using M-OFS to learn were positively evaluated by students, illustrate that the hypothesis H1 is supported. 90.4% students would like to use the system again. 92.8% students would like to recommend the system to a friend or classmate in the future. 81.9% students have never used similar system before. For interview, that most students (94%) feel this system is useful and help them to improve their knowledge, and it is the evidence to confirm the hypothesis H1; moreover, 69% students want exactly answer in the feedback at the end.

Students also want a picture of their learning process which can point out their weakness and a suggestion of how to improve their English in the future. Some students feel that if they tried many time but cannot find the correct answer, they will lose patience at the end. In order to explore the characteristics of students, and relate these to student responses and performance in the tests, five behavioural variables were analyzed: the number of attempts (submissions), the average score changed between attempts, the average score at the end of trial, the amount of time spent to make each attempt, and the learning duration.

5.2.4 Summary

In Summary, there are several trials are conducted and analyzed successfully. According to the test results, in Mathematics and English subject, most students improved their MCQs test scores after using ESDNN system. Furthermore, the students increased their paper test scores after using the system. It means students improved their knowledge by learning from the system. The system helped students to learn and understand knowledge across different subjects.

It will explore the students' learning behaviour in next section in order to further understand how the system guide and support students' deep learning.

5.3 Findings of Learning Behaviours

5.3.1 Introduction of Behaviours of Groups

In order to explore the characteristics of students, and relate these to student responses and performance in the tests, five behavioural variables were analyzed: 1) the number of attempts (submissions), 2) the average score changed between attempts, 3) the average score at the end of trial, 4) the amount of time spent on each attempt, and 5) the learning duration. The study is focus on the English Grammar Trial, and the English Grammar Trial includes two trials: English Grammar Trial in JQU, and English Grammar Trial in KTU.

5.3.2 Learning Behavioural Group Analysis

In order to explore the characteristics of students, and relate these to student responses and performance in the tests, five behavioural variables were analyzed: the number of attempts (submissions), the average score changed between attempts, the average score at the end of trial, the amount of time spent to make each attempt, and the learning duration. Chart 2 illustrates a learning behaviour of this group of students by analysing the relationship between average scores increased and learning duration. Each blue point represents average scores increased of all students used the same learning time, and its coordinate of x-axis represents student's learning duration, and its coordinate of y-axis represents average scores increased. It can be achieved from this figure that average scores increased when students spent more time on studying from the system.

5.3.3 English Grammar Trial in JQU

Behavioural Group	Average score changed	Average score at the end	Number of students
Short learning duration	9.17%	48.33%	48
Long learning duration	17.14%	61.43%	35

Table 17: Short learning duration: time spent on learning < 25.6 minutes; Long learning duration: time spent on learning > 25.6 minutes

Behavioural Group	Average score changed	Average score at the end	Number of students
Many attempts	20.77%	63.07%	26
Few attempts	8.77%	49.65%	57

Table 18: Many attempts: number of attempt >13.4, Few attempts: number of attempt <13.4

Behavioural Group	Average score changed	Average score at the end	Number of students
Slow attempt	12.96%	53.52%	54
Rapid attempt	11.72%	54.48%	29

Table 19: Slow attempt: average time spent on each attempt >1.92 minutes; Rapid attempt: average time spent on each attempt <1.92 minutes

Behavioural Group	Average score changed	Average score at the end	Number of students
Rapid few attempts	2.22%	48.89%	9
Few slow attempts	10%	50%	48
Many rapid attempts	16%	57%	20
Slow many attempts	36.67%	81.67%	6

Table 20: Slow and many attempt are consistently associated with good score increases, and hence represent successful learning strategies amongst the students.

Behavioural Group	Average score changed	Average score at the end	Number of students
Slow, few attempt and short learning duration	9.35%	48.71%	31
Rapid, many attempts and short learning duration	15.7%	54.29%	7
Rapid, many attempt and long learning duration	16.15%	58.46%	13
Slow, few attempt and long learning duration	11.18%	52.35%	17

Table 21: Many attempts and long learning time are consistently associated with good score increases, and hence represent successful learning strategies amongst the students.

According to table 17-21 (time spent on learning is how long they spending on reading the feedback, and it is the time between each response to the system to submit their new answers), the group learning behaviour could be understood and studied. Chart 1 illustrates a learning behaviour of this group of students by analysing the relationship between average scores increased and learning duration. Each blue point represents average scores increased of all students used the same learning time, and its coordinate of x-axis represents student's learning duration, and its coordinate of y-axis represents average scores increased. It can be obtained from this char that average scores increased when students spent more time on studying from the system.

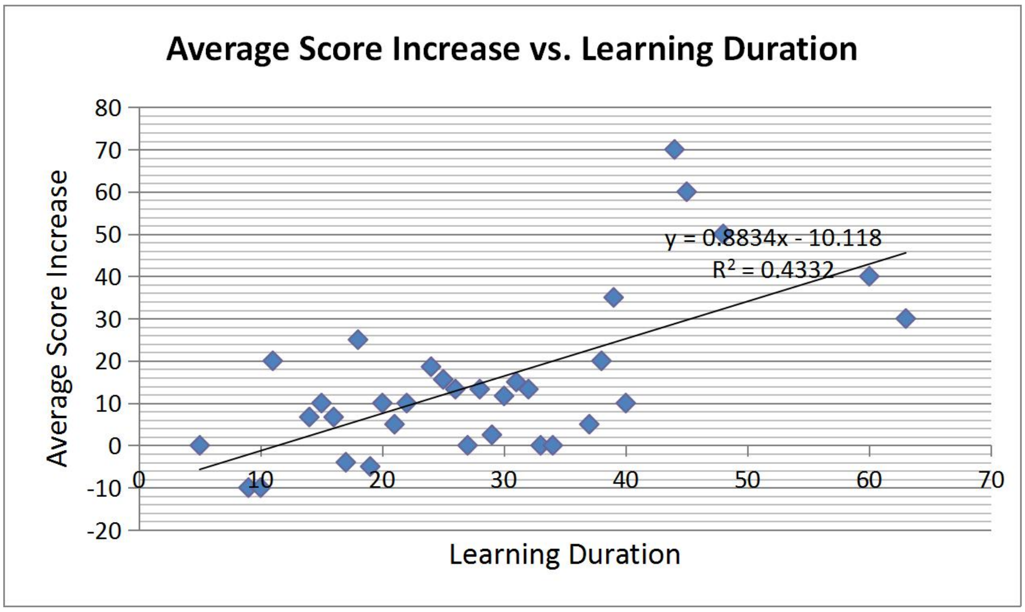


Chart 1: Average Score Increased – Learning Duration

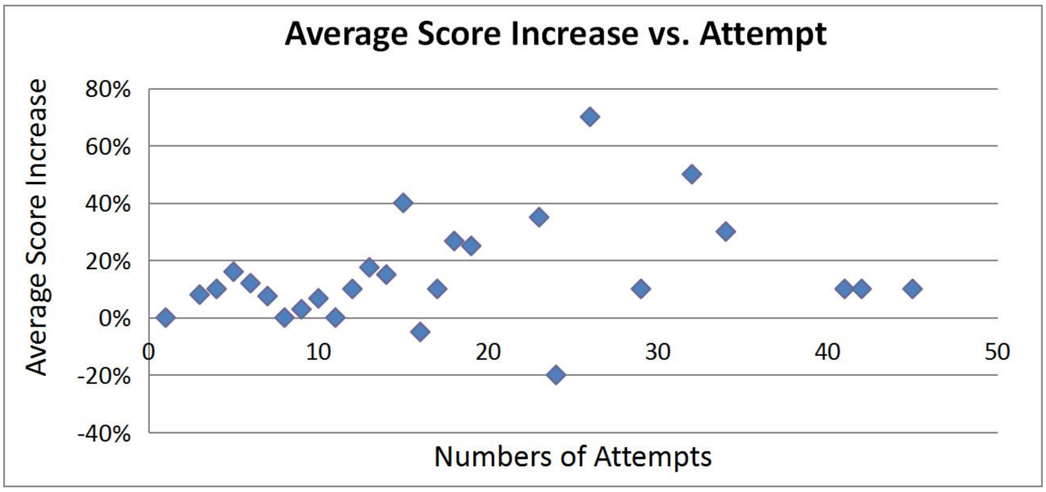


Chart 2: Average Score Increased – Attempt

Chart 2 illustrates a learning behaviour of this group of students by analysing the relationship between average scores increased and number of attempts. Each blue point represents average scores increased of all students did the same number of attempts, and its coordinate of x-axis represents number of students' attempts, and its coordinate of y-axis represents average scores increased. It can be obtained from this char that average scores increased when students did more attempts before peak, and average scores no longer increased when students did 26 attempts, and average scores decreased by doing more attempts after peak.

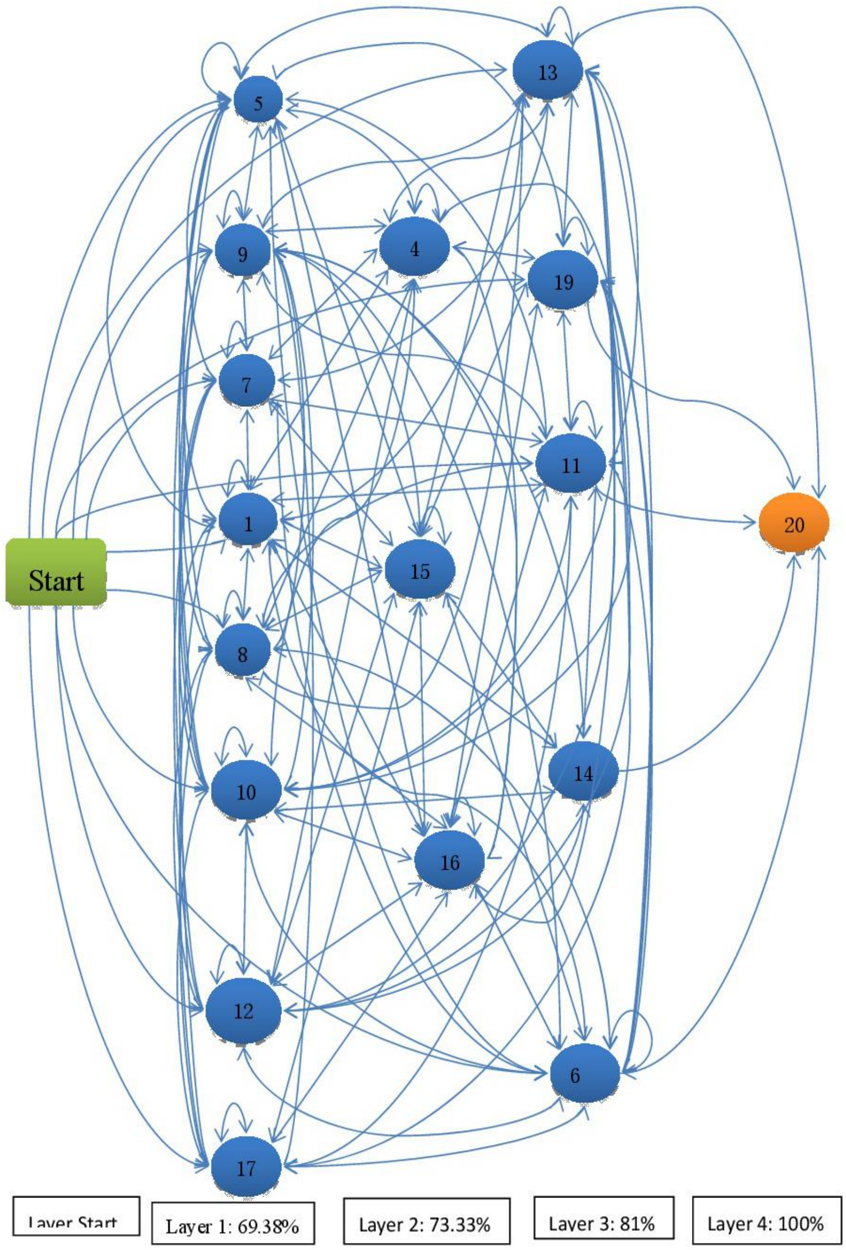


Figure 15: Knowledge State Transitions

Figure 15 illustrates the behaviour of students in terms of what might be called learning states (the details of group transitions see table 22 - 24). These states correspond to the output neurons that are triggered by patterns of question topic responses. In other words, the winning neuron represents a state of learning because it captures some commonality in a set of questions responses. For example, if there are several students who give the same answer (correct or incorrect) to two or more of the questions, snap-drift will form a group associated with one particular output neuron to include all such cases. That is an over simplification, because some of those cases may be pulled in to other 'stronger' groups, but that would also be characterized by a common feature amongst the group of responses.

It shows the knowledge state transitions. Each time a student gives a new set of answers, having received some feedback associated with their previous state, which in turn is based on their last answers, they are reclassified into a new (or the same) state, and thereby receive new (or the same) feedback. The tendency is to undergo a state transition immediately or after a second attempt or several attempts.

A justification for calling the states 'states of knowledge' is to be found in their self-organization into the layers of Figure 15. It classifies 5 layers such as: Layer Start, Layer 1, Layer 2, Layer 3 and Layer 4. A student on state 16, for example has to go via one of the states in the next layer such as state 13 before reaching the 'state of perfect knowledge' (state 20) which represents correct answers to all questions. On average, and unsurprisingly, the state-layer projecting onto state 20 (states 13, 19, 11, 14 and 6) is associated with more correct answers than the states in the previous layer. The states in the middle layer (layer 2) connect to neither start (layer star) nor state 20 (layer 4). The average score of each layer is shown below each layer. Students often circulate within layers before proceeding to the next layer. They may also return to a previous layer, but that is less common. The common knowledge state transitions (learning path ways) are shown in Figure 17. The average time spent on the questions was about 25.8

minutes, and the average increase in score was about 12.77% after an average of 13.4 attempts. Learning state transitions of students who got all correct answers are shown in Figure 16. The numbers of self-cycle of each knowledge state are shown in Figure 20, and numbers of learning points in feedback of each group are shown in Figure 21.

Average scores are 69.38% at beginning level (layer 1), 73.33% at connecting level (layer 2), and 81% advanced level (layer 3). Average numbers of self-cycling students did are 17.5 in layer 1, 8 in layer 2, and 52.2 in layer 3 (see chart 3). Average number of self-cycling in Layer 2 is less than Layer 1 and Layer 3. Students did not do many self-cycling in Layer 2 (connecting layer). It might mean students get more clearly feedback from this layer. Furthermore, there are many self-cycling in layer 3, and they are much more than layer 1 or layer 2. It might mean feedback in this layer made students confusing.

Group	**TtG**	**NoS**	**TNoS**	**NoA**	**TNoA**		Group	**TtG**	**NoS**	**TNoS**	**NoA**	**TNoA**
	13	**9**		**9**								
	19	**1**		**1**				5	**10**		**16**	
start	1	**8**	**83**	**8**	**83**			11	**7**		**9**	
	6	**11**		**11**			5	17	**1**	**43**	**2**	**57**
	5	**13**		**13**				16	**4**		**5**	
	11	**8**		**8**				15	**2**		**3**	

	7	**10**		**10**	
	9	**7**		**7**	
	8	**5**		**5**	
	10	**4**		**4**	
	12	**2**		**2**	
	17	**5**		**5**	
1	7	**5**	**28**	**7**	**34**
	11	**4**		**4**	
	1	**6**		**9**	
	9	**2**		**3**	
	5	**1**		**1**	
	13	**1**		**1**	
	10	**2**		**2**	
	8	**1**		**1**	
	6	**2**		**2**	
	15	**2**		**2**	
	16	**1**		**1**	
	14	**1**		**1**	
4	19	**1**	**10**	**1**	**11**
	6	**2**		**2**	
	4	**2**		**2**	

	13	**1**		**1**	
	9	**2**		**4**	
	1	**1**		**1**	
	12	**3**		**3**	
	6	**6**		**7**	
	4	**3**		**3**	
	19	**1**		**1**	
	8	**1**		**1**	
	10	**1**		**1**	
6	13	**3**	**52**	**3**	**71**
	9	**3**		**3**	
	11	**7**		**8**	
	16	**4**		**5**	
	12	**2**		**2**	
	5	**5**		**5**	
	6	**10**		**26**	
	8	**4**		**4**	
	1	**3**		**3**	

	1	**1**		**1**				15	**1**		**1**	
	9	**1**		**1**				10	**5**		**6**	
	13	**1**		**1**				7	**2**		**2**	
	7	**1**		**1**				20	**1**		**1**	
	12	**1**		**2**				19	**2**		**2**	

Table 22: Details of Group Transitions

TTG: Transfer to Group, NoA: Number of Attempts, NoS: Number of Students, TNoA: Total Number of Attempts, TNoS: Total Number of Students.

Group	**TTG**	**NoS**	**TNoS**	**NoA**	**TNoA**		Group	**TTG**	**TNoS**	**Total**	**NoA**	**TNoA**
7	9	**5**	**36**	**5**	**46**		10	13	**4**	**42**	**4**	**47**
	11	**1**		**2**				6	**5**		**5**	
	1	**3**		**6**				4	**1**		**1**	
	6	**4**		**4**				11	**4**		**4**	
	7	**5**		**7**				10	**10**		**13**	
	16	**5**		**5**				19	**1**		**1**	
	8	**6**		**8**				5	**3**		**3**	
	13	**2**		**2**				14	**1**		**1**	
	12	**1**		**1**				1	**2**		**2**	
	10	**1**		**2**				7	**3**		**3**	

	5	**1**		**1**	
	4	**1**		**2**	
	15	**1**		**1**	
8	8	**9**	**33**	**26**	**50**
	9	**2**		**2**	
	7	**4**		**4**	
	5	**1**		**1**	
	15	**3**		**3**	
	16	**2**		**2**	
	12	**3**		**3**	
	13	**1**		**1**	
	6	**1**		**1**	
	11	**3**		**3**	
	1	**1**		**1**	
	17	**1**		**1**	
	19	**1**		**1**	
	4	**1**		**1**	
9	9	**9**	**39**	**21**	**58**
	11	**9**		**12**	

	16	**2**		**2**	
	19	**1**		**1**	
	12	**2**		**3**	
	17	**1**		**1**	
	9	**2**		**3**	
11	9	**9**	**82**	**11**	**251**
	5	**9**		**13**	
	11	**24**		**178**	
	7	**1**		**2**	
	6	**5**		**6**	
	17	**10**		**14**	
	16	**4**		**4**	
	1	**2**		**2**	
	19	**13**		**15**	
	10	**2**		**3**	
	14	**2**		**2**	
	20	**1**		**1**	
12	5	**1**	**24**	**1**	**46**
	12	**7**		**28**	

	6	**1**		**1**			15	**1**		**1**	
	7	**4**		**6**			6	**3**		**3**	
	8	**3**		**3**			10	**4**		**5**	
	1	**1**		**2**			8	**3**		**3**	
	10	**4**		**4**			11	**1**		**1**	
	5	**4**		**5**			19	**2**		**2**	
	15	**1**		**1**			13	**1**		**1**	
	16	**2**		**2**			16	**1**		**1**	
	17	**1**		**1**							

Table 23: Details of Group Transitions

TTG: Transfer to Group, NoA: Number of Attempts, NoS: Number of Students, TNoA: Total Number of Attempts, TNoS: Total Number of Students.

Group	**TtG**	**NoS**	**TNoS**	**NoA**	**TNoA**	Group	**TtG**	**NoS**	**TNoS**	**NoA**	**TNoA**
13	10	**4**	**29**	**5**	**39**	16	11	**5**	**32**	**6**	**41**
	13	**6**		**15**			16	**8**		**14**	
	6	**2**		**2**			13	**1**		**1**	
	19	**2**		**2**			19	**1**		**1**	
	9	**2**		**2**			8	**3**		**3**	
	16	**3**		**3**			15	**1**		**1**	

	20	**1**		**1**	
	15	**1**		**1**	
	7	**3**		**3**	
	1	**2**		**2**	
	11	**2**		**2**	
	4	**1**		**1**	
14	12	**1**	**3**	**1**	**3**
	20	**1**		**1**	
	9	**1**		**1**	
15	19	**2**	**14**	**3**	**21**
	8	**1**		**1**	
	12	**1**		**1**	
	6	**2**		**2**	
	5	**2**		**2**	
	1	**2**		**2**	

	12	**3**		**3**	
	1	**1**		**1**	
	17	**2**		**2**	
	6	**1**		**2**	
	10	**2**		**2**	
	4	**2**		**2**	
	5	**1**		**2**	
	9	**1**		**1**	
17	19	**7**	**30**	**13**	**51**
	17	**6**		**20**	
	11	**8**		**9**	
	6	**2**		**2**	
	10	**1**		**1**	
	5	**3**		**3**	
	15	**1**		**1**	
	1	**1**		**1**	
	8	**1**		**1**	
19	11	**14**	**35**	**22**	**79**
	17	**5**		**7**	
	19	**9**		**42**	
	13	**1**		**2**	

	7	**1**		**1**	
	15	**2**		**8**	
	14	**1**		**1**	

	15	**1**		**1**	
	20	**1**		**1**	
	16	**2**		**2**	
	12	**2**		**2**	

Table 24: Details of Group Transitions

TTG: Transfer to Group, NoA: Number of Attempts, NoS: Number of Students, TNoA: Total Number of Attempts, TNoS: Total Number of Students.

According to compare and contrast feedback between each layer, it shows that average learning points in feedback of groups are 3 in layer 2, 3.25 in layer 1, and 2 in layer 3. Furthermore, the number of self-cycling of each knowledge transition layers are 18 in layer 1, 8 in layer 2, and 52 in layer 3 (see Chart 3).

It might mean there is optimal amount of information in each piece of feedback. For example, in the English trial, the most suitable amount of information is 2 key learning points in one feedback. It means that, there is a peak of improvement when students learning from the feedback that contain 2 key learning points.

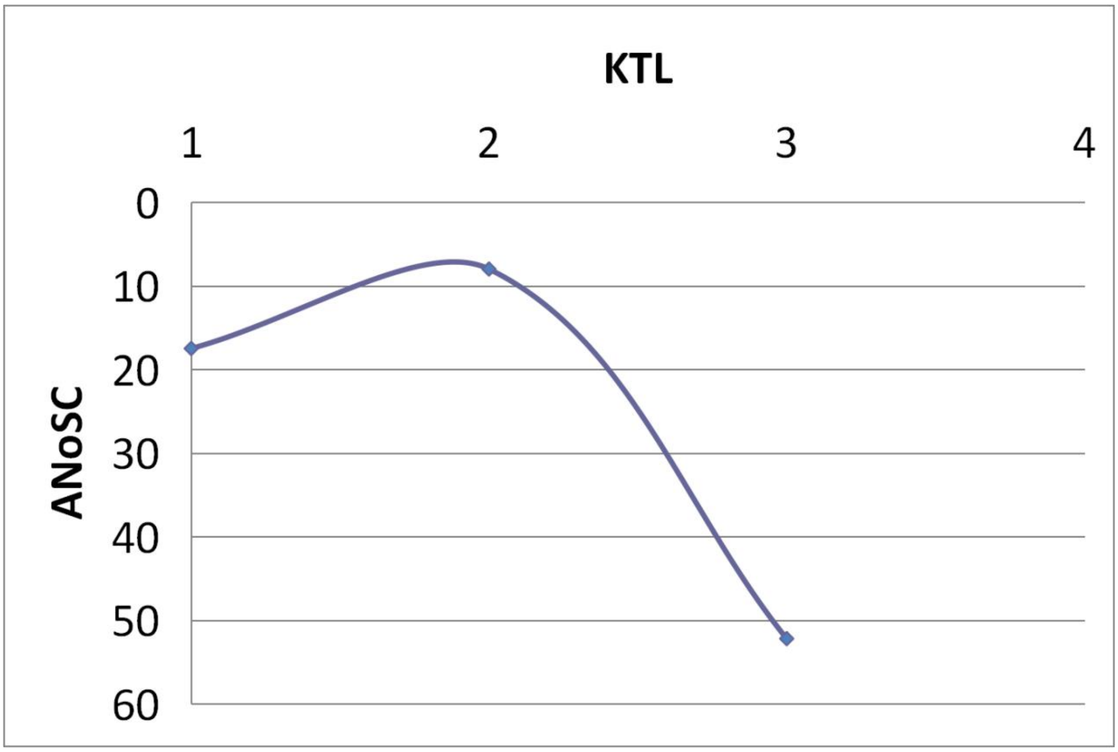


Chart 3: Average Number of Self-Cycling - Knowledge Transition Layer

ANoSC stands for Average Number of Self-Cycling, KTL stands for Knowledge Transition Layer

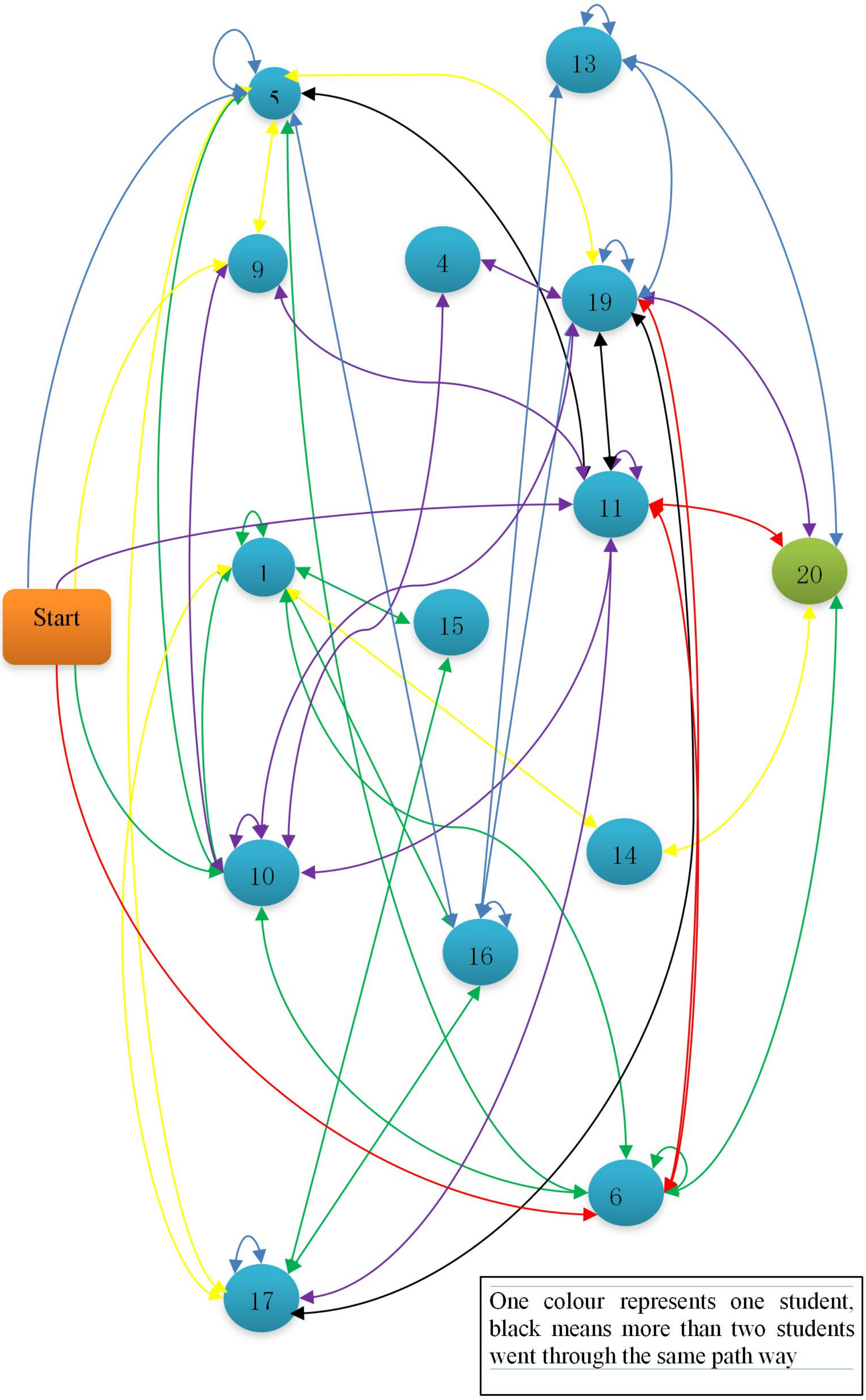


Figure 16: learning path ways of students who got all correct answers

Figure 16 illustrates knowledge transitions/learning path way of students who got all correct answers. Different colours represent different students, and black means more than two students went through the same path way. For example blue colour line represents a student went to group 5 from start, then went to group 16 and did a self-cycling, he went to group 13 and did a self-cycling, after then he went back to group 16, and then went back to group 5, and did some self-cycling, and then went to group 19 and did a self-cycling, then went to group 13 and then came back to group 19, then went to group 13, and finally went to group 20(all correct). This student's learning path way is start->group 5->group 16 (self-cycling) ->group 13 (self-cycling) ->group 16->group 19 (self-cycling) ->group 13->group 19 (self-cycling) -> group 13 (self-cycling) -> group 20.

Furthermore, for average, participants spent more time on moving to another group than doing self-cycling. It might mean students spent more time on thinking and reading feedback could be more helpful than they just tried simply changing answers to get all correct answer. There is no self-cycling on four groups (4, 9, 14, 15). According to comparing and contrasting feedback between self-cycling and non-self-cycling group, it shows that average learning points in the feedback for self-cycling group is 28.4% more than the non-self-cycling group**.** It might mean that feedback with succinct and clear information is better than feedback with much information on guiding students to make progress.

Figure 17 show that the common knowledge state transitions which has more than 9 attempts on each learning path ways except the ways to state 20 and self-cycling. Figure 18 shows the common knowledge state transitions which has more than 5 attempts on each learning path ways except the ways to state 20. Table 25 shows the details of Common Group Transitions. Figure 19 shows the group transitions flow, and table 26 shows the details of group Transition Flow (include self-cycling). According to the two figures and two tables, the picture of the most common learning path way and the group transitions flow could be draw up, in order to understand students' behaviours.

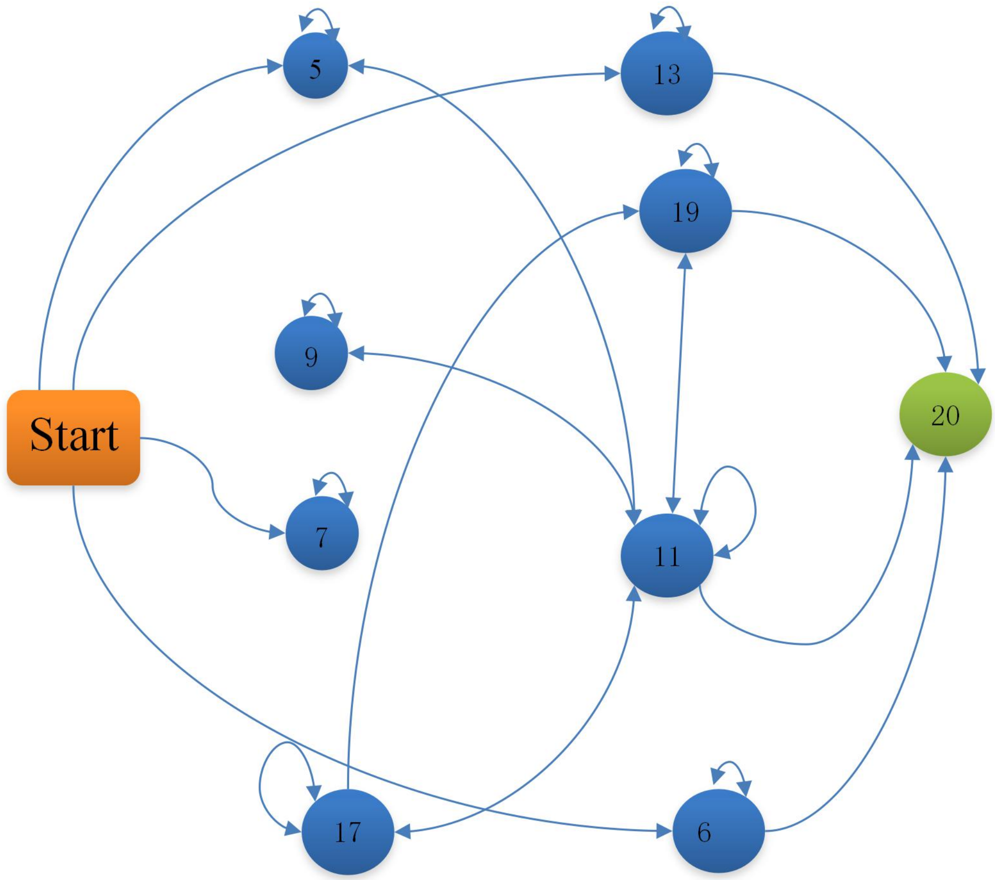


Figure 17: Common knowledge state transitions (9 attempts)

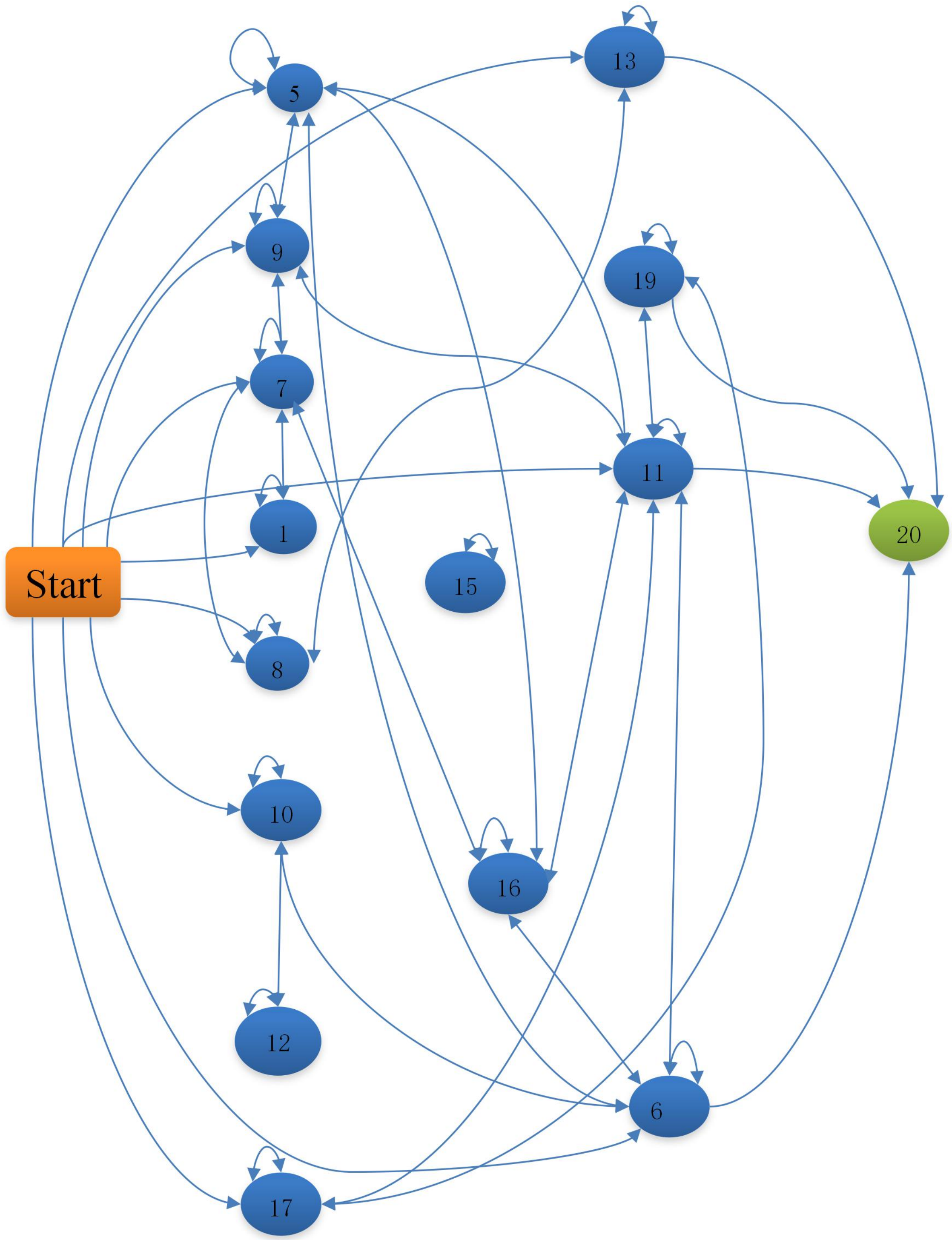


Figure 18: Common knowledge state transitions （5 attempts）

Group	**TtG**	**NoA**	**TNoA**
start	13	9	**43**
	6	11	
	5	13	
	7	10	
5	11	9	**25**
	5	16	
6	20	1	**27**
	6	26	
9	11	12	**33**
	9	21	
11	9	11	**232**
	5	13	
	17	14	
	19	15	
	20	1	
	11	178	
17	19	13	**42**
	11	9	
	17	20	
19	11	22	**65**
	20	1	
	19	42	

Table 25: Common Group Transitions

TtG stands for Transfer to Group, NoA stands for Number of Attempts, TNoA stands for Total Number of Attempts.

IN		**OUT**		IN		**OUT**
TNoA	Group	TNoA		TNoA	Group	TNoA
83	start	**83**		12	4	**11**
65	5	**57**		24	15	**21**
62	9	**58**		47	16	**41**
55	7	**46**		41	13	**39**
35	1	**34**		88	19	**79**
56	8	**50**		270	11	**251**
49	10	**46**		5	14	**3**
41	12	**46**		77	6	**73**
51	17	**51**				

Table 26: Group Transition Flow (include self-cycling)

TNoA stands for Total number of attempts

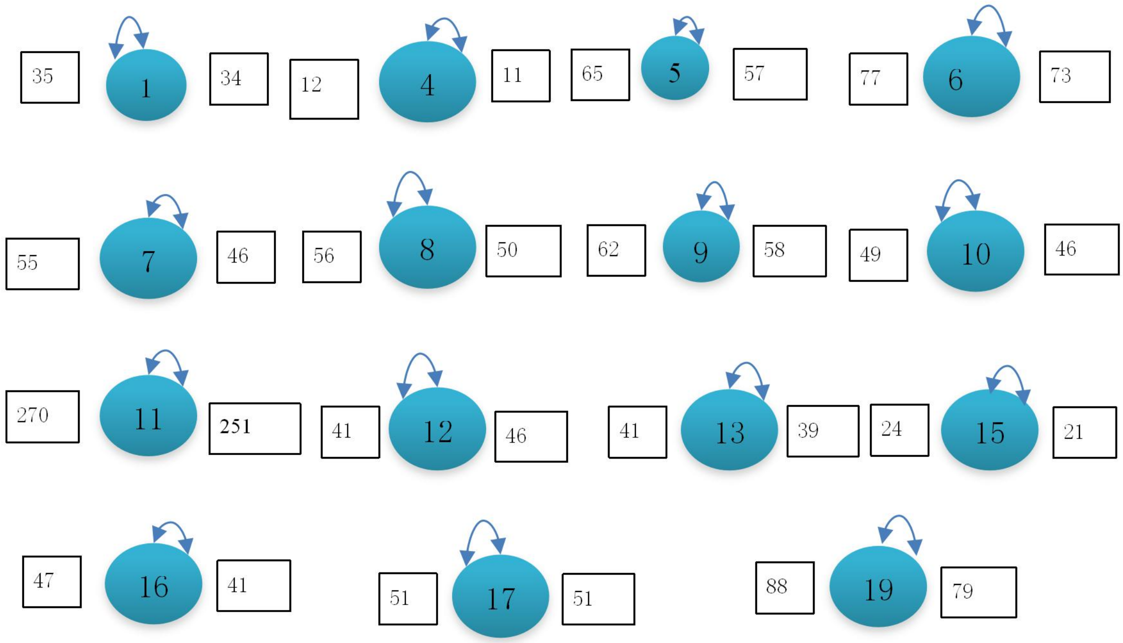


Figure 19: Group transition flow

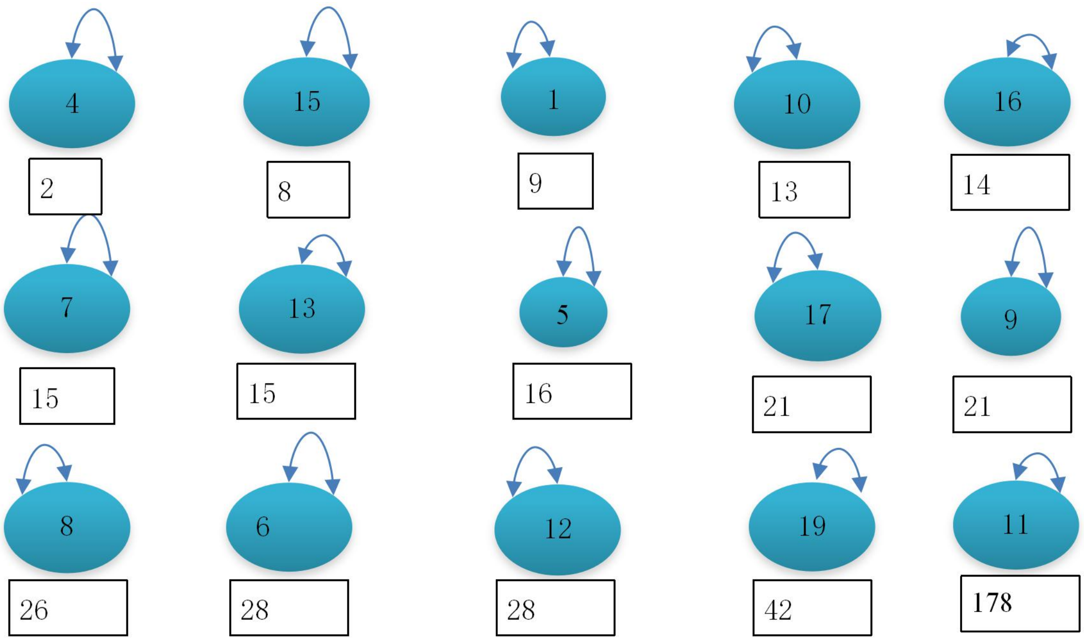


Figure 20: Numbers of self-cycling of each knowledge state

Figure 20 presents self-cycling of each knowledge state. Moreover it sorts by numbers of its loops. Students did average 17.5 times of self-cycling in layer 1, 8 times in layer 2, and 52.2 times in layer 3. Number of self-cycling in Layer 2 is less than Layer 1 and Layer 3. Students did not stay long in Layer 2 (connect layer). It might mean students get more clearly feedback from this layer. Furthermore, there are many self-cycling in layer 3 much more than layer 1 or layer 2. It might mean feedback in this layer made students confusing. It is no self-cycling on state 14. According to content of feedback 14, there is only 2 lines text in this feedback. Moreover, there are only a few words in feedback 11 and 19. Less content of feedback might make students confusing.

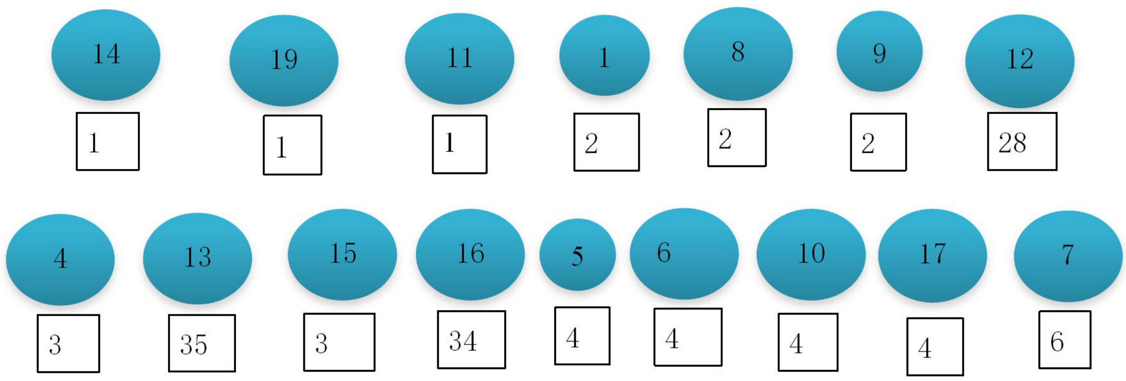


Figure 21: learning points within each knowledge state

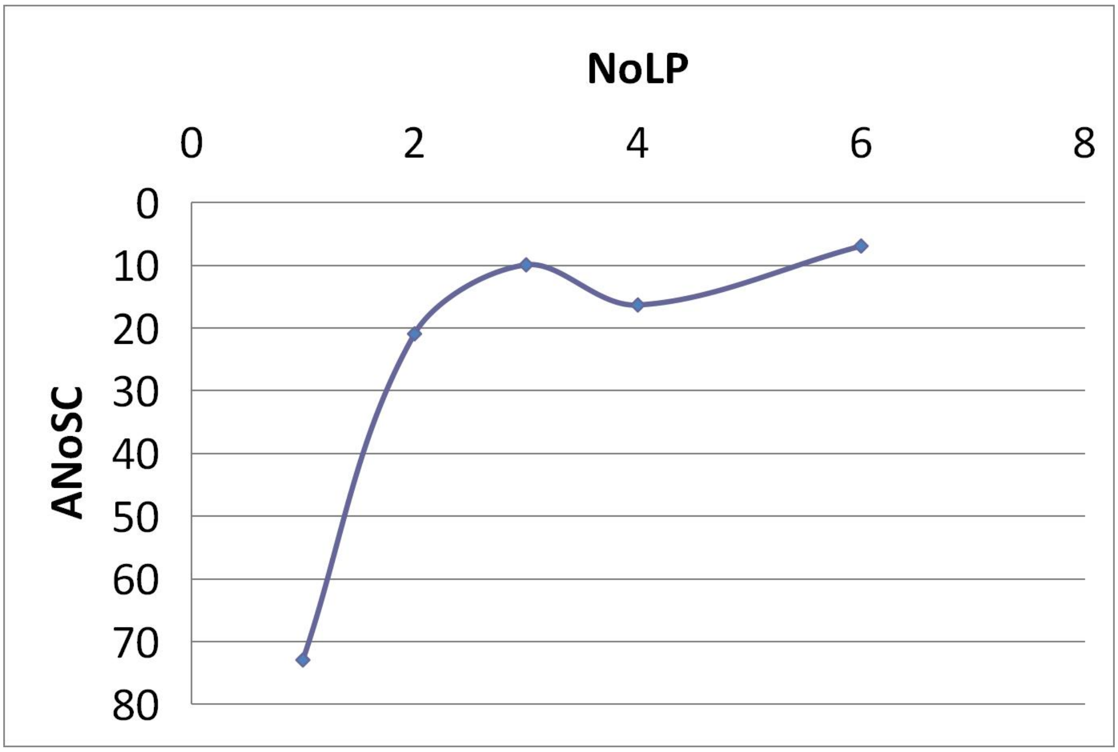


Chart 4: Average Number of Self-Cycling – Number of Learning Points

ANoSC stands for Average Number of Self-Cycling, NoLP stands for Number of Learning Points

Figure 21 presents number of learning points in each group feedback. Chart 4 illustrates that the number of self-cycling students did when they got the feedback with different number of leaning points. The results show that students did less average number of self-cycling when feedback with 3 and 6 learning points. The number of learning point is the learning point within each of the feedback, and it design by supervisor based on the group features. The learning point could be variable in different feedback.

5.3.4 English Grammar Trial in KTU

Chart 5 illustrates a learning behaviour of this group of students by analysing the relationship between average scores increased and learning duration. Each blue point represents average scores increased of all students used the same learning time, and its coordinate of x-axis represents student's learning duration, and its coordinate of y-axis represents average scores increased. It can be achieved from this figure that average scores increased when students spent more time on studying from the system.

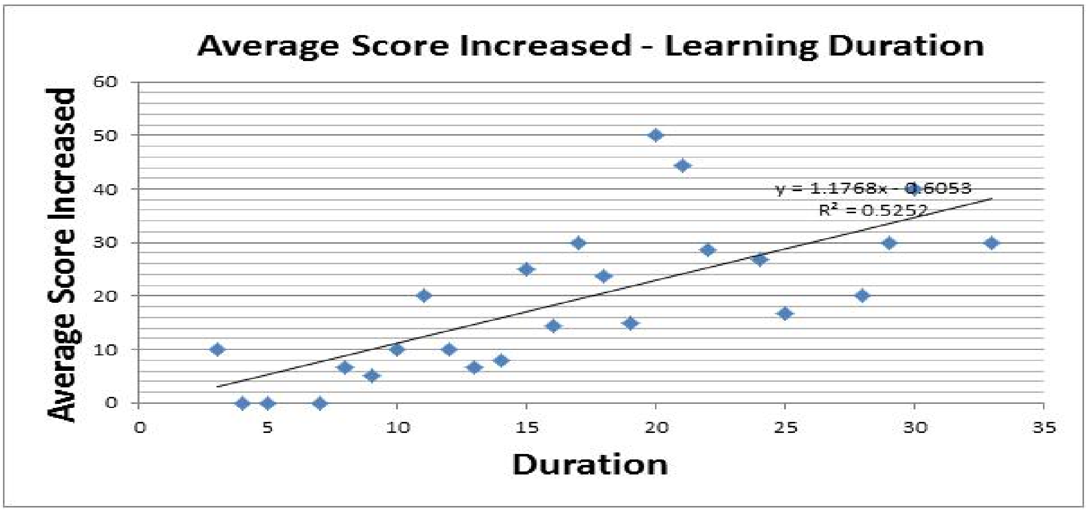


Chart 5: Average Score Increased vs. Learning Duration

Chart 6 illustrates a learning behaviour of this group of students by analysing the relationship between average scores increased and number of attempts. Each blue point represents average scores increased of all students did the same number of attempts, and its coordinate of x-axis represents number of students' attempts, and its coordinate of y-axis represents average scores increased. It can be deduced from this figure that average scores increased when students did more attempts before peak, and average scores no longer increased when students did 11 attempts, and average scores decreased by doing more attempts after peak.

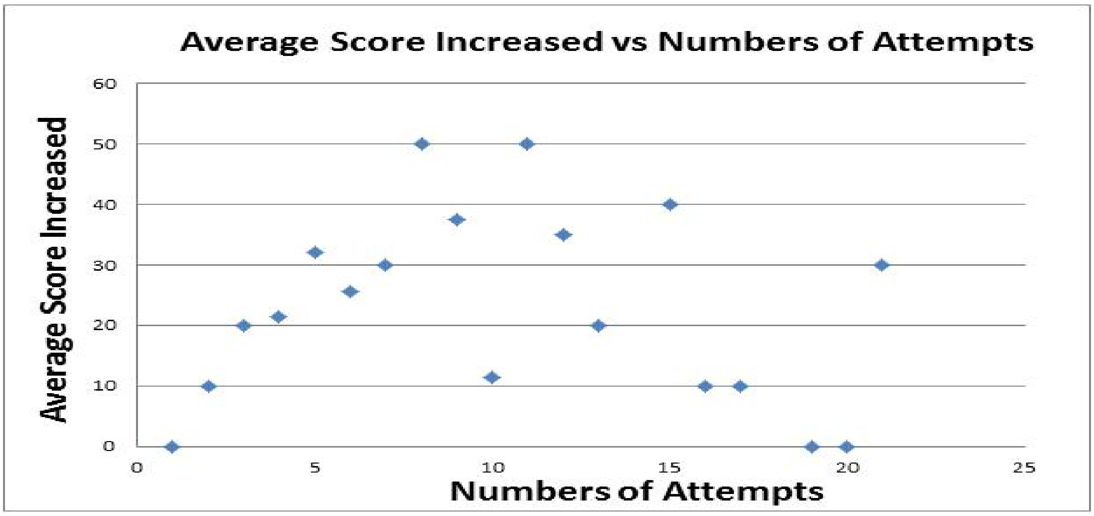


Chart 6: Average Score Increase vs. Numbers of Attempts

According to table 27-31, the behaviours of each learning group could be understood and studied. The results are similar to the first English experiment. Many attempts and long learning time are consistently associated with good score increases, and hence represent successful learning strategies amongst the students. Furthermore, in average, students cannot achieve more than 25% score increase with less than 15 minutes study duration from system, and students cannot achieve more than 30% score increase with less than 5 attempts or more than 21 attempts.

Behavioural Group	Average score changed	Average score at the end	Number of students

Short learning duration	13.41%	64.39%	41
Long learning duration	29.09%	84.09%	44

Table 27: Short learning duration: time spent on learning < 17.2 minutes; Long learning duration: time spent on learning > 17.2 minutes

Behavioural Group	Average score changed	Average score at the end	Number of students
Many attempts	26.47%	81.76%	34
Few attempts	18.24%	69.8%	51

Table 28: Many attempts: number of attempt >6.6, Few attempts: number of attempt <6.6

Behavioural Group	Average score changed	Average score at the end	Number of students
Slow attempt	21.57%	74.12%	51
Rapid attempt	21.47%	75.29%	34

Table 29: Slow attempt: average time spent on each attempt >2.6 minutes; Rapid attempt: average time spent on each attempt <2.6 minutes

Behavioural Group	Average score changed	Average score at the end	Number of students
Rapid few attempts	11.43%	62.86%	7
Few slow attempts	19.32%	70.91%	44
Many rapid attempts	24.07%	78.52%	27
Slow many attempts	35.71%	94.29%	7

Table 30: Rapid few attempt: average time spent on each attempt <2.6 minutes and number of attempt <6.6; Few slow attempts: number of attempt <6.6 and average time spent on each attempt >2.6 minutes; Many rapid attempts: number of attempt >6.6 and average time spent on each attempt <2.6 minutes; Slow many attempts: average time spent on each attempt >2.6 and number of attempt >6.6

Behavioural Group	Average score changed	Average score at the end	Number of students
Slow, few attempt and short learning duration	9.6%	60%	25
Rapid, many attempts and short learning duration	25.56%	77.78%	9

Rapid, many attempt and long learning duration	23.33%	78.89%	18
Slow, few attempt and long learning duration	32.11%	85.26%	19

Table 31: *Slow, few attempt and short learning duration*: **average time spent on each attempt >2.6 minutes, number of attempt <6.6, and time spent on learning <17.2 minutes;** *Rapid, many attempts and short learning duration*: **average time spent on each attempt <2.6 minutes, number of attempt >6.6, and time spent on learning <17.2 minutes;** *Rapid, many attempt and long learning duration*: **average time spent on each attempt <2.6 minutes, number of attempt >6.6, and time spent on learning >17.2 minutes;** *Slow, few attempt and long learning duration:* **average time spent on each attempt >2.6 minutes, number of attempt <6.6, and time spent on learning >17.2 minutes**

Figure 22 illustrates the behaviour of students in terms of what might be called knowledge states. These states correspond to the student responses triggered by patterns of student answer. In other words, a state of knowledge captures some commonality in a set of questions responses. For example, if there are several students who give the same answer (correct or incorrect) to two or more of the questions, snap-drift will form a group associated with one particular output neuron to include all such cases. That is an over simplification, because some of those cases may be pulled in to other 'stronger' groups, but that would also be characterized by a common feature amongst the group of responses. Figure 22 shows the knowledge state transitions. Each time a student gives a new set of answers, having received some feedback associated with their previous state, which in turn is based on their last answers, they are reclassified into a new (or the same) state, and thereby receive new (or the same) feedback. The tendency is to undergo a state transition immediately or after a second attempt or less commonly several attempts.

A justification for calling the states 'states of knowledge' is also to be found in their self-organization into the layers. There are 4 layers: Start, Layer 2, Layer 3, and Layer 4. A student on state 16, for example has to go via one of the states in the next layer such as state 12 before reaching the 'state of perfect knowledge' (state 20) which represents correct answers to all questions. On average, and unsurprisingly, the state-layer projecting onto state 20 (states 12) is associated with more correct answers than the states in the previous layer. This is true of state 12 which projects onto state 20, and it is also true of state 15 although this state does not project onto state 20. The states in the middle layer (layer 2) all connect to start layer, and layer 3 does not connect to start layer. The average score of each layer is increased from start to state 20 (Average scores are 69.7% at beginning level (layer 2), 77.5% at advanced level (layer 2), and 100% at state 20). Students often circulate within layers before proceeding to the next layer. They may also return to a previous layer, but that is less common.

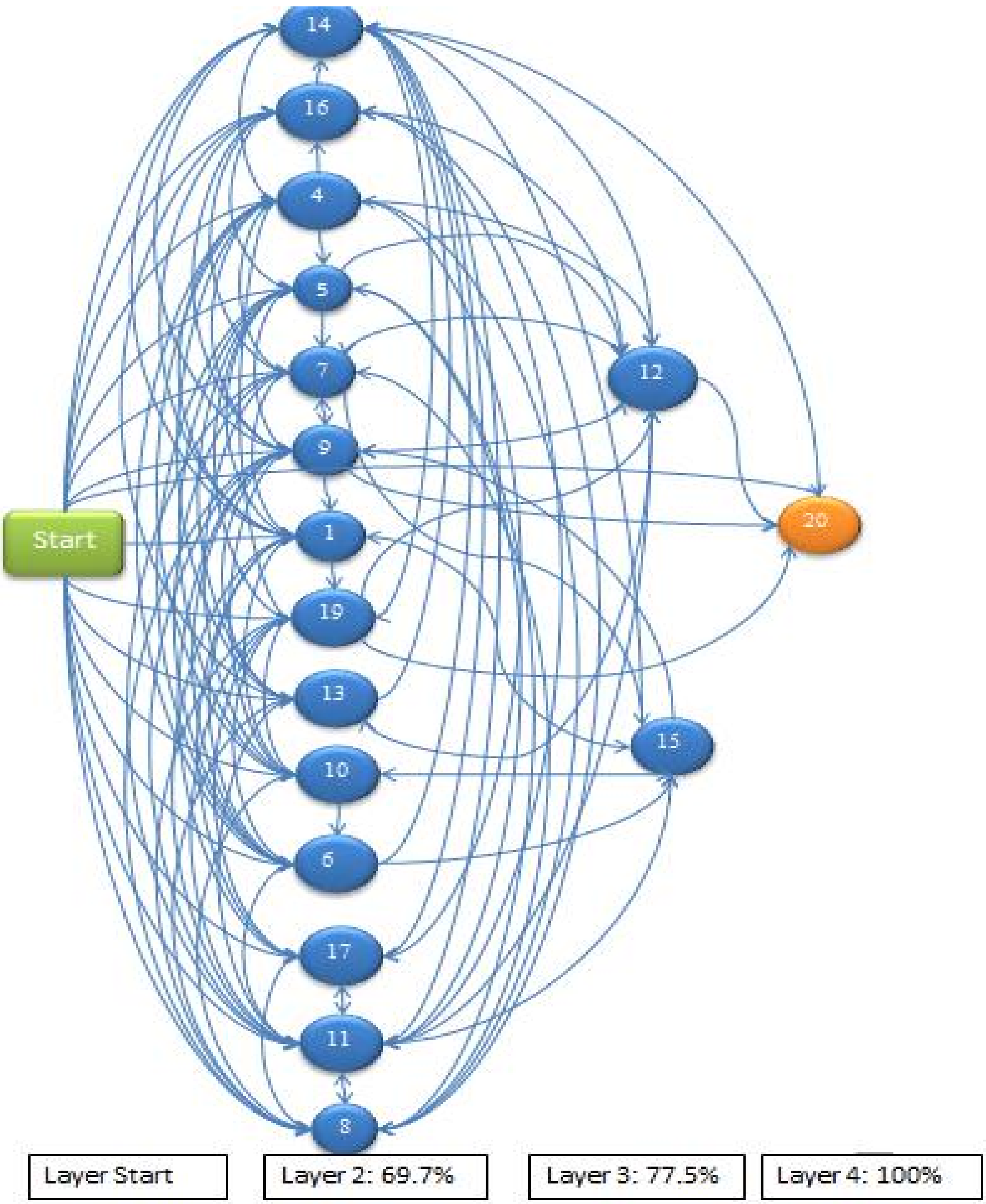


Figure 22: Knowledge State Transitions of 2nd trial

5.4 Summary

As amply described in chapter 4, chapter 5.1 and 5.2. Six main trials across three totally different subject areas have been carried out in three universities in two countries, the UK and China: 3 English trials, 2 Mathematics trials, and 1 Java Programming trial. The English trials are very successful with about 500 students participated. A abundant volume of data has been recorded when students using the system.

From the results (refer to chapter 5.2 Results of Trials) it shows that four hypothesis are supported. The results from system trial analysis show that the feedback had a positive impact which is the evidence to confirm hypothesis H2. The results from separate MCQs paper tests shows that the students are only learnt by using M-OFS but not any other ways which is the evidence to confirm hypothesis H3 and H2. From the results of compare and contrast of the experimental group and the control group, it shows that the experimental group got higher score than the control group which is the evidence confirms the hypothesis H4 and it also confirms hypothesis H2. Furthermore, according to the results of survey, the students are satisfied with using this system and they wish to recommend the system to others; Most students never use a similar system before and would like to us it again; Most student think the feedback from this system is what they need (illustrate that the hypothesis H1 is confirmed). From interviews, most students feel this system is useful and helps them to improve their knowledge (indicates the hypothesis H1 is confirmed); moreover, more than half students want the exact answers in the feedback at the end; Students also want a picture of their learning process which can point out their weakness and a suggestion of how to improve their English. Some students feel that if they tried many times but cannot find the correct answer, they will lose patience at the end.

As amply described in chapter 5.2.2 and 5.2.3. A large volume of data has been captured during the trials. The students participating in this second trial are from KTU that is one of the top universities in China. The entrance requirements of KTU are much higher than those of JQU, which means that students studying in KTU generally have a better academic background than students in JQU. The results of the second English system trials show that the average score is increased by 21.2% at the end. The results of the separate MCQ paper test indicated the average mark of the group is increased by 7.55%. In addition, the final examination, the average mark of the experimental group is 7.48% higher than the control group. Student surveys show that 87.8% students are satisfied with our e-learning system and 89.2% students feel that the intelligent diagnostic feedback is beneficial to enhance their learning experience.

According to all of the results, the M-OFS system demonstrates that it is working well for both less qualified and better qualified students at two universities JQU and KTU. In English Grammar Trial in KTU, the results of group behaviour shows that students in the group “Slow, few attempt and long learning duration” achieved much better learning effects than others. In comparison with the English Grammar Trial in JQU, the students in the group “Rapid, many attempt and long learning duration” achieved better learning effects than others. This has indicated that the students in the KTU who have better academic background prefer more reflective thinking and independent study from the feedbacks, whilst, the students in the JQU trial who have weaker academic background prefer more frequent and direct feedbacks as part of their learning process. Moreover, the feedback with more or less information make students confusing, and the feedback with succinct and clear information is better on guiding students to make progress. Furthermore, the average scores increased when students did more attempts before peak, and average scores no longer increased when students did peak attempts, and average scores decreased by doing more attempts after peak attempts. And the average scores increased when students spent more time on studying from the system. In addition, the knowledge state transaction analysis shows that: 1.) the knowledge of each layer is increased from start to final state (all correct answer); 2.) students commonly go through the knowledge layers from start to final state step by step; 3.) students often circulate within layers before proceeding to the next layer; 4.) they may also return to a previous layer, but that is less common. Moreover, the evidences are confirmed the four hypothesis: 1, during the trials, students improved their understanding by reading given feedback (H1); 2, after using M-OFS system, students get higher mark in a separate questionnaire than before (H2); 3, the experimental students get higher marks than control group at the final examination (H3); 4, most students are satisfied with this system (H4).

6. Conclusions and Future Work

6.1 Results and Benefit of this Research

In this research, a novel method for using snap-drift in a diagnostic tool to provide feedback has been presented. The most innovative aspect of the work is that: this is the first time a MCQ neural network diagnostic feedback approach has been systemically applied and evaluated across a range of different subject areas; with the neural network discovering groups of similar answers that represent different knowledge states of the students. The feedback targets the level of knowledge of individuals, and their misconceptions, guiding them toward a greater understanding of particular concepts. M-OFS has been applied for a large cohort of students in different subject areas. In total, six trials in three totally different subjects' area have been carried out in three Universities in two countries: three English trials, two Mathematics trials and one Java Programming trial. In terms of scale the English trials were the most successful with about five hundreds students participating in the experiments, and it consist of following part: re-test, trials, separate questionnaire test, experimental group and control group. Two of English trials, two mathematics trials have been analyzed, other data will be analyzed in the future work.

The results of six experiments show that an improvement in the learning process can be obtained by using M-OFS. The approach has been empirically evaluated across several subjects by carrying out the trials. These subjects included English Language, Mathematics, and Computer Programming (the data of Computer programming will analyzed in the future work) in the context of higher education in different countries. It was also evaluated on the topic of students' understanding of plagiarism avoidance. The trials were configured to assess the improvement in knowledge resulting from engagement with the online MCQ system that embodies the approach. A large amount of data has been captured during the trials and analyzed. The results of the English trials show that the average test score are significant increased. Moreover, the results of the separate questions paper test shows that the average mark of the group is increased as well. In addition, in terms of the results of final examination, the average mark of the experimental group is almost 10% higher than the control group. Furthermore, the results of analysis of surveys shows that most students are satisfied with our e-learning system and most students feel that the intelligent diagnostic feedback is what they need. In addition, the supervisors found that the system is helps them to discover the weak area of students' learning.

6.2 Research hypothesis

It summarized the evidence for four hypothesis as following: Hypothesis H1: more than 80% students feel the feedback is satisfied, about half of the students wish to use the system again, 90% students would like to recommend the system to their friend, and more than 90% students think the system is helpful and helps them to improve their knowledge. Hypothesis H2: there is million answer combinations of 10 MCQs with four choices, guessing answer is not make any improvement of students knowledge; more than half students improved their score in mathematics trial; in English Grammar Trial in JQU, more than half students improve their score by average 12.8%, and the best score improving is 70%; in English Trial in KTU, 74.1% students improved their score, and the best improvement is 80%; none of few students get all correct answer at the beginning; most students think the system is helpful and helps them to improve their knowledge; the experimental group (learning from the system) get higher mark than the control group in the final examination; students improve their scores in the separate questionnaire test after only learning from the system. H3: in the separate questionnaire test, most students improved their score after learning from the system, and best improvement is 40%, and the average score improved from 51.6% to 59.15% (in JQU) and 53.5% to 62.65% (in KTU). H4: in the final examination, the experimental group achieve much higher mark than the control group.

6.3 Group Learning Behaviour

This research also discovers the group learning behaviour is as follows: 1, average scores increased when students spent more time on studying from the system; 2, average scores increased when students made more attempts up to a threshold peak, and average scores no longer increased when students made more attempts than the threshold, and average scores decreased by doing more attempts; 3, many attempts and long learning time are consistently associated with good score increases, and hence represent successful learning strategies amongst the students. Furthermore, in average, students cannot achieve significant score increases with short study duration from system, and students cannot achieve significant score increase with very few attempts or too many attempts. Most students get the best learning effects from the feedback that with two learning point.

6.4 Knowledge State

The research find out the knowledge of each layer is increased from start to final state, students commonly go through the knowledge layers from start to final state step by step, students often circulate within layers before proceeding to the next layer, and they may also return to a previous layer, but that is less common.

This system can help students to discover their learning weakness not only on the key learning conception but the learning path. Students can learn and evaluate their learning by themselves. In addition, they can also improve their knowledge by themselves. On the other hand, the supervisors are also supported by this system. The system provides an intuitionistic summary of weak as part of students' learning. And this can save supervisors' time and help supervisors capture the misunderstanding of key learning points more accurately. Moreover, the system can also give supervisor a whole picture of both individual student's learning path and a group of students' learning path in order to help the supervisor analyze the group behaviour of students.

6.5 Contribution to Knowledge

In summary, this research investigates the current key issues in the area of e-learning in higher education and focuses on how to provide automatic, effective and convenient feedback to students in order to support students learning. The research compares and contrasts several methods in order to investigate the effective use of intelligent feedback towards modelling the stages of students' learning. The work explores the potential benefits of integrating an artificial neural network (ANN) into a Virtual Learning Environment (VLE) system as a means of identifying groups for the purposes of the getting feedback. It investigates the relative effectiveness of different types of feedback and how to optimize the feedback to maximize the facilitation of learning. It explores the ability of neural networks and data analysis techniques to model the stages of students' learning. The research also assesses the difference in the progress of students' learning with and without using intelligent diagnostic feedback. The E-learning Snap-Drift Neural Network (ESDNN) is assessed as one of the potential tool for providing diagnostic, and effective feedback. The ESDNN is enhanced following the first trial, and the enhanced ESDNN system is introduced to the MCQs-Online Feedback System (M-OFS).

6.6 Principles for Designing Diagnostic Feedback

The confirmation of the hypothesis demonstrates a significant level of effectiveness resulting from the deployment of diagnostic feedback in several subject areas and educational settings. This research establishes the potential for realising the benefits of diagnostic online feedback in a wide range of contexts. This research also leads to guidelines for the design principles of on-line MCQs and diagnostic feedback in learning environments. It includes 5 key characteristics that result in successful learning: 1, each MCQ should include 4 or 5 choices; 2, feedback should point out weakness in students' knowledge; 3, each feedback should include no more than 2 learning points; 4, the feedback should not include answers; 5, feedback should guide students in understanding the knowledge rather than simply towards getting the correct answers.

6.7 Future Work

The new data from student responses when using the system can also be used to retrain the neural network and see whether refined groupings are created, which can be used by the educator to improve the feedback. Moreover, the data of some trials will be analyzed in the future in order to assess the system. According to the results of data analysis, the system will be enhanced in the future. In future work, it is also intended to compare the effects of M-OFS provided feedback to the effects of other types of feedback. Another promising avenue for further inquiry is the extension of the tool to support knowledge state transition diagram construction and statistical data collection, which could help instructors to analyze the difficulty of the MCQs and the progression of the students during their learning process. In the future, the system can provide students a picture of their learning process that can point out their weakness and suggestions of how to improve their knowledge. The system should also encourage students by providing alternative avenues for learning when they lose patience after they tried many times but cannot find the correct answer.

Reference

1. Ahmed, R.K.A. (2016). Artificial Neural Networks in E-Learning Personalization: A Review. *International Journal of Intelligent Information Systems*, *5*(6), pp.104-108.

2. Alemán, J.L.F, Palmer-Brown, D. and Jayne, C. (2011). Effects of Response-Driven Feedback in Computer Science Learning. *IEEE Transactions on Education*, 54(3), pp.501-508.
3. Alessi, M. and Trollip, S. (2010). *Multimedia for learning: Methods and development*. 3rd ed. London: Allyn and Bacon.
4. Alexander, S. (2001). E-learning developments and experiences. *Education + Training*, 43(4/5), pp.240-248.
5. Archer, W., Garrison, R. and Anderson, T. (1999). Adopting disruptive technologies in traditional universities: Continuing education as an incubator for innovation. *Canadian Journal of University Continuing Education*, 25(1), pp.13-30.
6. Bang, P. (2003). *Engaging the learner – how to author for best feedback. Language Learning Online: Towards Best Practice*. 1st ed. The Netherlands: Swets & Zeitlinger

7. Bangert, A.W. (2004). The seven principles of good practice: A framework for evaluating on-line teaching. *The internet and higher education*, *7*(3), pp.217-232.

8. Barron, J. (2006). Top ten secrets of effective e-learning. *Industrial and Commercial Training*, 38(7), pp. 360-364.

9. Barron, J. (2006). Top ten secrets of effective e-learning. *Industrial and Commercial Training*, 38(7), pp.360-364.

10. Bates, A.W. (1997). Restructuring the university for technological change. The Carnegie Foundation for the Advancement of Teaching, London, http://bates.cstudies.ubc.ca/carnegie/carnegie.html, .

11. Bates, A.W. (2005). *Technology, e-learning and Distance Education (Routledge Studies in Distance Education)*. 2nd ed. New York: Routledge.

12. Bharath, R. (1994). *Neural Network Computing*. 1st ed. New York: Computing McGraw-Hill.

13. Blessing, S., Gilbert, S., Ourada, S. and Ritter, S. (2007). Lowering the bar for creating model-tracing intelligent tutoring systems. *Artificial intelligence in education*, pp.443-450.

14. Brown, G., Bull, J. and Pendlebury, M. (1997). *Assessing Student Learning in Higher Education*. 1st ed. London: Routledge.

15. Buabeng-Andoh Charles & Asirvatham David. (2002). Multimedia Intelligent System for Online Learning. [online] Available at: https://ieeexplore.ieee.org/document/1185857 [Accessed 1 Sep. 2019].

16. Buchanan, T. (2000). The efficacy of a World-Wide Web mediated formative assessment. *Journal of Computer Assisted Learning*, *16*(3), pp.193-200.

17. Bullen, M. and Janes, D.P. (2007). *Making the transition to E-learning: Strategies and issues*. 1ST ed. New York: Information Science Publishing.

18. Capper, J. (2001). The emerging market for online learning: insights from the corporate sector. *European Journal of Education*, 36(2), pp. 237_245.

19. Clark R. C. and Mayer, R. E. (2009). Instructional strategies for directive learning environments. In: *Handbook of Improving Performance in the Workplace: Instructional Design and Training Delivery*, 1st ed. San Franciscon: Pfeiffer, pp.329-360.

20. Clark, R.C. and Mayer, R.E. (2007). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning.* 2nd ed, San Franciscon: John Wiley & Sons.

21. Dafoulas G.A. (2005). The role of feedback in online learning communities. [online] Available at: https://ieeexplore.ieee.org/document/1508828 [Accessed 1 Sep. 2019].

22. Draper, S.W. (2007). *A Momentary Review of Assessment Principles*. [online] Ewds.strath.ac.uk. Available at: https://ewds.strath.ac.uk/REAP/reap07/Portals/2/CSL/keynotes/david%20nicol/A_momentary_review_of_assessment_principles.pdf [Accessed 1 Sep. 2019].

23. Drefus, G. (2005) Neural Networks: Methodology and Applications. 2nd ed. Paris: Springer.

24. Epstein, M.L., Lazarus, A.D., Calvano, T.B., Mathews, K.A., Hendel, R.A., Epstein, B.B., and Brosvic, G.M. (2002). Immediate feedback assessment technique promotes learning and corrects inaccurate first response. *The Psychological Record*, 52(2), pp. 187-201.

25. European Commission. (2001). *Communication from the commission to the council and the European parliament: the e-Learning action plan*. [online] Aic.lv. Available at: http://www.aic.lv/ace/ace_disk/Bologna/contrib/EU/e-learn_ACPL.pdf [Accessed 2 Sep. 2019].

26. Fallows, S. & Ahmet, K. (1999), *Inspiring students: Case studies in motivating the learner*. 1st ed. London: Kogan Page. pp. 1-5.

27. Felix, U. (2000). A multivariate analysis of students' experience of Web-based learning. *Australian Journal of Educational Technology*, 17 (1), pp.21-36.

28. Felix, U. (2003). Humanising automated online learning through intelligent feedback. *Proceedings of the 20th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE)*, [online] pp.178 - 186. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.475.7155&rep=rep1&type=pdf [Accessed 2 Sep. 2019].

29. Fu, X.; Peltsverger, B.; Qian, K.; Tao, L.; Liu, J. (2008). APOGEE: automated project grading and instant feedback system for web based computing. *Technical Symposium on Computer Science Education - SIGCSE*, pp.77-81, DOI: 10.1145/1352135.1352163

30. Fu, Xiang & Peltsverger, Boris & Qian, Kai & Tao, Lixin & Liu, Jigang. (2008). APOGEE-automated project grading and instant feedback system for web based computing. *SIGCSE'08 - Proceedings of the 39th ACM Technical Symposium on Computer Science Education*, pp.77-81. DOI: 10.1145/1352135.1352163.

31. G.Crisp, D.Thiele, I.Scholten, S.Barker and J.Baron (Eds), Interact, Integrate, Impact: Proceedings of the 20th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education. Adelaide,

32. Galushkin, A. I. (2007). *Neural networks theory*. 2nd ed. New York: Springer.

33. Garber, P. R. (2004). *Giving and Receiving Performance Feedback*. Canada: HRD Press.

34. Gheorghiu, R. and Vanlehn, K. (2008). *XTutor: an intelligent tutor system for science and Math based on Excel.* Intelligent tutoring systems: 9th International Conference, Paris: Springer. sd

35. Gibbs, G. (2006). Why assessment is changing. In *Innovative assessment in higher education.* 3rd ed. New York: Routledge, pp. 31-42.

36. Gibbs, G. and Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and teaching in higher education*, 8(1), pp.3-31.

37. Greenstein, M. M. and Welsh, M. J. (2005). Bankruptcy Prediction Using Ex Ante Neural Networks and Realistically Proportioned Testing Sets. *Artificial Intelligence in Accounting & Auditing: International Perspectives*. [online] Available at SSRN: https://ssrn.com/abstract=2850 [Accessed 2 Sep. 2019].

38. Guo R., Palmer-Brown D., Lee S.W., Cai F.F. (2012). Effective Diagnostic Feedback for Online Multiple-Choice Questions. *Artificial Intelligence Applications and Innovations*, 381(1), pp.316-326.

39. Guo, H., 2018. Application of a Computer-Assisted Instruction System Based on Constructivism. *International Journal of Emerging Technologies in Learning*, *13*(4), pp.33-44.

40. Guo, R., Palmer-Brown, D., Lee, S.W. and Cai, F.F. (2014). Intelligent diagnostic feedback for online multiple-choice questions. *Artificial Intelligence Review*, *42*(3), pp.369-383.

41. Guo, R., Palmer-Brown, D., Lee, S.W. et al. (2014). Intelligent diagnostic feedback for online multiple-choice questions, Artif Intell Rev. [online] Available at DOI: https://doi.org/10.1007/s10462-013-9419-6

42. Harris, K., Graham, S. and Urdan, T. (2012). *APA educational psychology handbook*. Washington, DC: American Psychological Association.
43. Hatziapostolou, T. and Paraskakis, I (2010). Enhancing the Impact of Formative Feedback on Student Learning through an Online Feedback System. *Electronic Journal of e-Learning*, 8 (2) , pp111-122
44. Heinze, A., Procter, C. and Scott, B. (2007). Use of conversation theory to underpin blended learning. *International Journal of Teaching and Case Studies*, 1 (2), pp. 108–120.
45. Higgins, E., and Tatham, L., (2003). Exploring the potential of multiple choice questions in assessment. *Learn. Teach. Action*, 2(1), pp. 35 – 48.
46. Honey, P. (2001). E-learning: a performance appraisal and some suggestions for improvement. *The Learning Organization*, 8(5), pp. 200-202.
47. Hounsell, D. (2003). Student feedback, learning and development. *Higher education and the lifecourse*, pp.67-78.
48. Hurt, J. (2008). The Advantages and Disadvantages of Teaching and Learning On-line. *Delta Kappa Gamma Bulletin*, 74 (4), pp5-11
49. Husain, Noushad. (2010). Computer Assisted Learning: Theory and Application. [online] Available at DOI: 10.13140/2.1.4591.4400.
50. Ifinedo, P. (2006). Acceptance and continuance intention of web-based learning technologies use among university students in a Baltic country. *the Electronic Journal on Information Systems in Developing Countries*, 23(6), pp.1-20.
51. Investopedia. (2019). *Neural Network Definition*. [online] Available at: https://www.investopedia.com/terms/n/neuralnetwork.asp [Accessed 1 Sep. 2019].

52. J.S. Sung and D.H. Lim, (2006). Adaptive Tutoring and Training System Based on Intelligent Agent. *International Journal of Multimedia and Ubiquitous Engineering*, 1(3), pp. 6-11.

53. Johnson, J. and Picton, P. (1995). *Designing Intelligent Machines: Concepts in artificial intelligence*. UK: Open University.

54. Joseph, H.R. (2014). September. Promoting education: A state of the art machine learning framework for feedback and monitoring E-Learning impact. In *2014 IEEE Global Humanitarian Technology Conference-South Asia Satellite (GHTC-SAS)* IEEE, pp. 251-254.

55. Kanuka, H. and Kelland, J. (2008). Has e-Learning Delivered on its Promises? Expert Opinion on the Impact of e-Learning in Higher Education. *Canadian Journal of Higher Education*, 38(1), pp.45-65

56. Karayiannis, N. B. and Venetsanopoulos, A. N. (1993). *Artificial neural networks: learning algorithms, performance evaluation, and applications*. 1st ed. New York: Kluwer Academic Publishers.

57. Knight, P. and Yorke, M. (2003). *Assessment, learning and employability*. 1st ed. London: McGraw-Hill Education.

58. Kordaki, M. and Daradoumis, T. (2009). Critical thinking as a framework for structuring synchronous and asynchronous communication within learning design-based e-learning systems. *Intelligent Collaborative E-Learning Systems and Applications*, 246(1), pp.83-698.

59. Kuechler, W.L., and Simkin, M.G. (2003). How well do multiple choice tests evaluate student understanding in computer programming classes?. *Journal of Information Systems Education*, 14(4), pp. 389-399.

60. Kuri-Morales, A. and Simari, G.R. (2010). *Advances in Artificial Intelligence - IBERAMIA 2010*, Peru: Springer.

61. Lanny, A. & Musumeci, D. (2000). Instructor attitudes within the SCALE efficiency projects. *Journal of Asynchronous Learning Network*, 3 (4), pp.215-223.

62. Lee, S., Palmer-Brown, D. and Draganova, C. (2008). Diagnostic feedback by snap-drift question response grouping. *World Scientific and Engineering Academy and Society*, 5(3), pp. 208-214.

63. Lee, S., Palmer-Brown, D., Draganova, C., Preston, D. and Kretsis, M. (2009). Question response grouping for online diagnostic feedback. *Proceedings of Advances in Computing and Technology*, pp.68-76.

64. Lee, S.W., Palmer-Brown, D. and Roadknight, C.M. (2004). Performance-guided neural network for rapidly self-organising active network management. *Neurocomputing*, *61*, pp.5-20.

65. Lee, S.W., Palmer-Brown, D., Draganova, C., Preston, D. and Kretsis, M. (2009). Question response grouping for online diagnostic feedback. *Proceedings of Advances in Computing and Technology*, pp.68-76.

66. Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & education*, *48*(2), pp.185-204.

67. Ley, K. (1999). Providing feedback to distant students. *Campus-Wide Information*

Systems, *16*(2), pp.63-69.

68. Li, L., Buhalis, D., Lockwood, A. and Benzine, K. (2007). *the use of e-learning in training in the UK hospitality industry: an exploratory study*. In: ECEL 2007: The 6th European Conference on e-Learning, Demark: Copenhagen Business School.
69. Little, B. (2001). Achieving high performance through e-learning. *Industrial and Commercial Training*, 33(6), pp. 203-206.
70. Little, J. and Bjork, E. (2014). Optimizing multiple-choice tests as tools for learning. *Memory & Cognition*, 43(1), pp.14-26.
71. Ma, Z. (2006). *Web-based intelligent e-learning systems*. 2nd ed. New York: Information Science Publishing.
72. Malone, S.A. (2003). *How to set up and manage a corporate learning centre*. 2nd ed. London: Gower.

73. Martínez-Argüelles, M.J., Plana-Erta, D., Hintzmann-Colominas, C., Badia-Miró, M. and Batalla-Busquets, J.M. (2013). Usefulness of feedback in e-learning from the students' perspective. In *European Conference on e-Learning (ECEL)* , 11(4), pp. 283-292.

74. McIntyre, D.R. and Wolff, F.G. (1998). Anexperiment with WWW interactive learning in university education", *Computers & Education*, 3(31), pp. 255-64.

75. Moubayed, A., Injadat, M., Nassif, A.B., Lutfiyya, H. and Shami, A. (2018). E-learning: Challenges and research opportunities using machine learning & Data analytics. *IEEE Access*, *6*, pp.39117-39138.

76. Mulqueeny, K., Kostyuk, V., Baker, R. and Ocumpaugh, J. (2015). Incorporating effective e-learning principles to improve student engagement in middle-school

mathematics. *International Journal of STEM Education*, 2(1), pp.1-14

77. Naim, N.F., Yassin, A.I.M., Zakaria, N.B. and Wahab, N.A. (2011). Classification of Thumbprint using Artificial Neural Network (ANN). In *2011 IEEE International Conference on System Engineering and Technology*, 42(3), pp. 231-234

78. National Student Survey (2017). *Hefce national student survey*. [online] Hefce.ac.uk. Available at: http://www.hefce.ac.uk/lt/nss/. [Accessed 1 Sep. 2019].

79. Nelson, M. M. and Schunn, C. D. (2009). The Nature of Feedback: How Different Types of Peer Feedback Affect Writing Performance. *Instructional Science*, 37(4), pp.375-401.

80. Nicol, D.J. and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, *31*(2), pp.199-218.

81. Omoda-Onyait G., Lubega J.T., Maiga G. (2013) Requirements Framework for Personalized Real-Time Feedback in Interactive Agent-Based E-Learning Systems. *Hybrid Learning and Continuing Education*, 8038, pp.290-300.

82. Özdener, N. & Satar, H. M. (2009). Effectiveness of various oral feedback techniques in CALL vocabulary learning materials. *Eurasian Journal of Educational Research (EJER)*, 34, pp.75-96.

83. Palmer-Brown, D., Lee, S. W., D. & Draganova, C. (2008). Diagnostic Feedback by Snap-drift Question Response Grouping. *Proceedings of the 9th WSEAS International Conference on Neural Networks (NN'08)*, pp 208-214.

84. Pan, Z., Zhang, X., and Rhalibi,A.E. et al. (2008). *Technologies for e-learning and digital entertainment*. 3rd ed. Germany: Springer.

85. Paxton, M. (2000). A linguistic perspective on multiple-choice questioning assessment and evaluation. *Assessment & Evaluation in Higher Education Assess.* Eval. Higher Educ., 25(2), pp. 109-110.

86. Payne, A., Brinkman, and Wilson, F. (2007). Towards effective feedback in e-learning packages: The design of a package to support literature searching, referencing and avoiding plagiarism. *Proceedings of HCI2007 workshop: Design and use and experience of e-learning systems*, pp.71-75.

87. Payne, A.M., Brinkman, W.P. and Wilson, F.C. (2007). Towards effective feedback in e-learning packages: the design of a package to support literature searching, referencing and avoiding plagiarism. *Proceedings of HCI2007 workshop: Design, use and experience of e-learning systems*, pp. 71-75.

88. Race, P. (2006). *The Lecturer's Toolkit – A Practical Guide to Assessment. Learning and Teaching*. 3rd ed. London: Routledge.

89. Race, P. and Brown, S. (2005). 500 Tips for Tutors. 2nd ed. London: Routledge

90. Rane, D. and Sasikumar, M. (2007). A constructive learning framework for language tutoring. *Innovations in e-learning, instruction technology, assessment, and engineering education,* pp.73-78.

91. Reiser, B. J. and Kimberg, D. Y. et al. (1992). Knowledge Representation and Explanation. *Computer Assisted Instruction and Intelligent Tutoring Systems*, pp.375-349.

92. Rossen, E. and Hartley, D. (2001) *Basics of E-Learning*. 1st ed. USA: ASTD Press.

93. Saddler, B. and Andrade, H. (2004). The writing rubric. *Educational Leadership*, 62(2), PP. 48–52.

94. Sharma, R.C. and Mishra, S. (2007). *Cases on global e-learning practices: successes and pitfalls.* 2nd ed. USA: published by Information Science Publishing.

95. Sjoberg, S. (2007). Constructivism and learning. *International encyclopaedia of education*, *3*(2), pp.118-126.

96. Springgay, S. and Clarke, A. (2007). *Mid-Course Feedback on Faculty Teaching: A Pilot Project, Collective improvisation in a teacher education community, Darling,* L. F. and Erickson, G. L. et al (Eds.), published by Springer, the Netherlands, Chapter 13, pp171-185.

97. Sung, J.S. (2009). Intelligent Feedback System (IFS) in Tele-Learning Environment. *International Journal of Advanced Science and Technology*, 9(4), pp.113-124.

98. Svinicki, M.D. (1999). New directions in learning and motivation. *New directions for teaching and learning*, *1999*(80), pp.5-27.

99. Teresita L. and Raymund S. (2003). Learner agents as learner modeling: Design and analysis. *IEEE International Conf. On Advanced Learning Technologies*.

100. Vasilyeva, E., Pechenizkiy, M. and Bra, P. D. (2008). Adaptation of elaborated feedback in e-learning. *Adaptive Hypermedia and Adaptive Web-Based Systems*, 5149, pp.235-244.

101. Villiers, R.D. (2007). The six C's framework for e-learning. *Advanced Principles of Effective E-Learning*, 3(4), pp.145-153.

102. Yorke, M. and Longden, B. (2004). *Retention and student success in higher education*. 2nd ed. London: McGraw-Hill Education.

Appendices

AIRE Journal Paper

Artif Intell Rev (2014) 42:369–383
DOI 10.1007/s10462-013-9419-6

Intelligent diagnostic feedback for online multiple-choice questions

R. Guo · D. Palmer-Brown · S. W. Lee · F. F. Cai

Published online: 9 August 2013

Abstract When students attempt multiple-choice questions (MCQs) they generate invaluable information which can form the basis for understanding their learning behaviours. In this research, the information is collected and automatically analysed to provide customized, diagnostic feedback to support students' learning. This is achieved within a web-based system, incorporating the snap-drift neural network based analysis of students' responses to MCQs. This paper presents the results of a large trial of the method and the system which demonstrates the effectiveness of the feedback in guiding students towards a better understanding of particular concepts.

Keywords Learning behaviour · Diagnostic feedback · Neural networks · On-line multiple-choice questions

1 Introduction

In recent years, e-learning has become commonplace in higher education. The involvement of intelligent e-learning systems has the potential to make higher education accessible with increasing convenience, efficiency and quality of study. According to the Hefce National Student Survey (2007–2010), in England, only about half of students believe that: (1) feedback on their work has been prompt; (2) feedback on their work has helped to clarify things

R. Guo (✉) · D. Palmer-Brown · S. W. Lee · F. F. Cai
Faculty of Life Sciences and Computing, London Metropolitan University,
Holloway Road, London N7 8DB, UK
e-mail: r.guo1@londonmet.ac.uk; ruisheng_g@hotmail.com

D. Palmer-Brown
e-mail: d.palmer-brown@londonmet.ac.uk

S. W. Lee
e-mail: s.w.lee@uel.ac.uk

F. F. Cai
e-mail: ff.cai@londonmet.ac.uk

they did not understand; (3) they have received detailed comments on their work. These reports reveal that the feedback and its related fields are one of the weakest areas in higher education in England. This research investigates the relative effectiveness of different types of feedback, and how to optimize feedback to facilitate deep learning. It compares and contrasts several methods in order to investigate the effectiveness of using intelligent feedback towards modeling the stages of students' knowledge. The investigation will lead to an understanding of the potential of the on-line diagnostic feedback across different subject areas.

The virtual learning environment (VLE) as a an e-learning system based on web which can model personal education by providing virtual access to classes, concepts, key learning point, tests, homework, grades, assessments, and other external resources. It can also provide a method to help students to enhance their learning experiences, and to help teachers manage the gap between teaching and learning. The VLE presented in this paper provides a generic method for intelligent analysis and grouping of student responses that applicable to any area of study. This tool offers important benefits: immediate feedback, significant time-saving evaluating assignments, and consistency in the learning process. The time taken to create the feedback is well spent not only because this feedback can be reused, but also it is made available through the system to large numbers of students.

2 Background and review of previous work

Ma (2006) described intelligent tutoring systems (ITSs) as the milestone of the advanced generation of computer-aided instruction systems, and concluded their key feature as 'the ability to provide a user-adapted presentation of the teaching material'. Rane and Sasikumar (2007) pointed out that to overcome the lack of the presence of a teacher, intelligent tutoring systems attempt to simulate a teacher, who can guide the student's study based on the student's level of knowledge by giving intelligent instructional feedback. Furthermore, according to Blessing et al. (2007), the intense interaction and feedback achieved by intelligent tutoring systems can significantly improve student learning gains. In addition, in Gheorghiu and Vanlehn (2008)'s paper, they also suggested that meaningful, constructive and adaptive feedback is the essential feature of ITSs, and it is such feedback that helps students achieve strong learning gains.

Cullen et al. (2002, cited by Heinze et al. 2007) suggest that feedback is one of the most important indicators of good education. Little (2001) also presents that providing feedback, which can continuously reinforce learning progress and promote learners' attention and engagement, is crucial to effective learning. Besides, Vasilyeva et al. (2008) also state that 'the design of feedback is a critical issue of online assessment development within web-based learning systems'. However, reviewing the education reality, it is surprising that as one of the most important communication methods between teachers and students, feedback was one of the weakest areas according to the Hefce National Student Survey (2007–2010) in the UK. Thus, how to design an optimized intelligent feedback system becomes a major and critical problem to develop a successful e-learning system.

To make the feedback effective and meaningful, a range of quality attributes need to be achieved. Hatziapostolou and Paraskakis (2010) summarized the work by Race (2006), Irons (2008) and Juwah et al. (2004) and suggested that in order to improve learning gains, formative feedback should address as many as possible of the following attributes, including constructive, motivational, personal, manageable, timely and directly related to assessment criteria and learning outcomes. (1) Constructive. Effective feedback should be constructive,

as constructive feedback can lead to more thinking and cognitive learning which in turn improve the student's learning (Bang 2003 and Alessi and Trollip 2001). As Nelson and Schunn (2009) argued, effective feedback should be able to guide a learner 'to change performance in a particular direction rather than just towards or away from a prior behavior'. (2) Motivational. Effective feedback should be motivational to empower and encourage students to learn more, as feedback can affect students' feelings and attitudes towards study, which in turn affect their engagement in the learning process, (Juwah et al. 2004, cited by Hatziapostolou and Paraskakis 2010). (3) Garber (2004) also argued, 'the more personalized the feedback becomes, the more meaning it can have for the individual receiving it. . .and the more likely the individual will be receptive to the feedback. . .If a person does not believe in the reliability or validity of the feedback, it will have little or no benefit'. (4) Manageable. Effective feedback should be detailed enough to ensure that students clearly understand their strengths and weaknesses, and have enough materials to guide them to achieve the learning goals. At the same time, the feedback should not be over-detailed to avoid confusing, and make students can easily interpret it and get the point (Hatziapostolou and Paraskakis 2010). (5) Timely. Feedback should be delivered timely, as students can more easily utilize feedback when they can still remember how they just processed the task (Race 2006), and the reasoning that led to the error is still accessible (Reiser and Kimberg 1992). As Anderson, Boyle and Reiser (1985, cited by Reiser and Kimberg 1992) argued, tutors should provide immediate feedback to students, as 'the learning mechanism for adjusting a faulty rule or forming a new correct rule relies upon the problem situation being active in memory'. (6) Directly related to assessment criteria/learning outcomes. Effective feedback should be directly related to assessment criteria/learning outcomes so that it can explain students' achievement towards the intended learning outcomes, knowledge gaps and specific errors (Hatziapostolou and Paraskakis 2010). Thus, the students can be guided and adjust their effort to achieve the intended leaning outcomes (Race and Brown 2005). Additionally, Clark and Mayer (2009) further argued that learning goal oriented feedback is more effective than performance goal oriented feedback. In another word, feedback should be designed to inform the learners their progress toward achieving a learning goal rather than compare a learner performance with other learners'.

In addition to the required attributes above, many researchers also mentioned that various methods should be used in feedback to ensure better perception of feedback. For example, Özdener and Satar (2009) suggested using animation techniques to achieve better reception and perception of the feedback.

Springgay and Clarke (2007) suggested including examples in feedback to achieve better perception of feedback. Multiple choice questions (MCQs) is an effective way to provide students with feedback. The use of multiple-choice questions has been widely studied. A number of advantages can be found in the Epstein et al. (2002), Higgins and Tatham (2003) and Kuechler and Simkin (2003) researches: rapid feedback, automatic evaluation, perceived objectivity, easily computed statistical analysis of test results and the reuse of questions from databases as required, thus saving time for instructors. On the other hand there are also some researches (e.g. Paxton 2000) shows that MCQs have some disadvantages: significant effort is required to construct MCQs, they only evaluate knowledge and recall, and they are unable to test literacy and creativity. Although the MCQs have been primarily used for summative evaluation, they also serve formative assessment purposes. Formative assessment provides students with feedback that highlights areas for further study and indicates the degree of progress. There are many studies investigating the role of different type of feedback in Web-based assessments that report positive results from the use of MCQs in online tests for formative assessments.

Many researches investigating the effect of different types of feedback in web-based assessments showed the positive results of using MCQs in online test for formative assessment (e.g. Epstein et al. 2002; Higgins and Tatham 2003; Kuechler and Simkin 2003; Payne et al. 2007). Higgins and Tatham (2003) studied the use of MCQs in formative assessment in a web-based environment using WebCT for a level one unit on undergraduate law degree. They summed that they could forecast all the possible errors for a question and write a general feedback for this question. However, using this type of feedback, it could be difficult to predict all the possible errors and produce the general feedback for a combination of questions, and it would be impossible for a large test banks (e.g. 3 questions with 5 answers would require 125 answer combinations; 5 questions with 5 answers require 3,125 combinations, etc.).

Payne et al. (2007) assessed the effectiveness of three different forms of feedback (corrective, corrective explanatory, and video feedback) used in e-learning to support students' learning. This type of feedback shows exactly which questions are answered correctly or not, with further corrective explanation and video feedback. Our approach to feedback is different from the above. The intelligent diagnostic feedback we present is concept-oriented instead of question-oriented. The learners are encouraged to review the concepts they misunderstood through the feedback in order to retake the test again and study further. It is important that each category of answers is associated with carefully designed feedback based on the level of understanding and prevalent misconceptions of that category-group of students so that every individual student can reflect on his or her learning level using this diagnostic feedback. In addition, when students retake the test they receive new feedback according to his or her knowledge state, which in turn leads to more self-learning. Moreover, concept-based feedback can also prevent the student from guessing the right answers; if the students do not read the diagnostic feedback carefully, they may not even know which questions were answered incorrectly. In conclusion the focus of our current research is to combine MCQs and formative online assessment using an intelligent agent to analyse the students' response in order to provide diagnostic feedback. To this end, we deploy a neural network

3 Multiple-choice questions online feedback systems (M-OFS)

To analyse the students' answers, and integrate over a number of questions to gain insights into the students' learning needs, a snap-drift neural network (SDNN) approach is proposed. SDNN provides an efficient means of discovering a relatively small and therefore manageable number of groups of similar answers. In the following sections, an e-learning system based on SDNN is described.

3.1 Snap-drift neural networks (SDNNs)

The learning process involves a combination of fast, convergent, minimalist learning (snap) and overall feature averaging (drift) to capture both precise sub-features in the data and more general holistic features. Snap and drift learning phases are combined within a learning system that toggles its learning style between the two modes. On presentation of input data patterns at the input layer, the distributed SDNN (dSDNN) learns to group them according to their features using snap-drift (Palmer-Brown and Jayne 2011). The neurons whose weight vectors result in them receiving the highest activations are adapted. Weights are normalised weights so that in effect only the angle of the weight vector is adapted, meaning that a recognised feature is based on a particular ratio of values, rather than absolute values. The output winning neurons from dSDNN act as input data to the selection SDNN (sSDNN) module within performs feature grouping and this layer is also subject to snap-drift learning.

The learning process is unlike error minimisation and maximum likelihood for SDNN toggles its learning mode to find a set of sub-features and average feature in the data and uses them to group the data into categories. Each weight vector is bounded by snap and drift: snapping gives the angle of the minimum values (on all dimensions) and drifting gives the average angle of the patterns grouped under the neuron. Snap creates a feature common to all the patterns in the group and gives a high probability of rapid (in terms of epochs) convergence (both snap and drift are convergent, but snap is faster). Drifting, which uses learning vector quantization (LVQ), tilts the vector towards the centroid angle of the group and ensures that an average, generalised feature is included in the final vector. The angular range of the pattern-group membership depends on the proximity of neighbouring groups (natural competition), and can be controlled by adjusting a threshold on the weighted sum of inputs to the neurons.

3.2 Training neural network

The e-learning snap-drift neural network (ESDNN) is trained with the students' responses to questions on a particular topic in a course. The responses are obtained from the previous cohorts of students. Before training, each of the responses from the students is encoded into binary form in preparation for presentation as input patterns for ESDNN. Table 1 shows examples of a possible format of questions for five possible answers and some encoded responses. This version of ESDNN is a simplified unsupervised version of the snap-drift algorithm (Palmer-Brown and Jayne 2011) as shown in Fig. 1.

During training, on presentation of an input pattern at the input layer, the dSDNN will learn to group the input patterns according to their general features. In this case, 5 F12 nodes,

Table 1 Example of input patterns for ESDNN

Codification	A:00001	B:00010	C:00100	D:01000	N/A:00000
Response	Recorded response				
[C, D, B, A]	[0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1]				
[E, B]	[1, 0, 0, 0, 0, 0, 0, 0, 1, 0]				
[D, A, A]	[0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1]				
[A, C, D, B, A]	[0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 10, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1]				

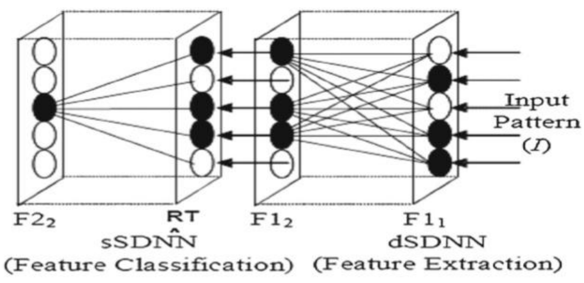


Fig. 1 E-learning SDNN architecture

whose weight prototypes best match the current input pattern, with the highest net input are used as the input data to the sSDNN module for feature classification. In the sSDNN module, a quality assurance threshold is introduced. If the net input of an sSDNN node is above the threshold, the output node is accepted as the winner; otherwise a new uncommitted output node will be selected as the new winner and initialised with the current input pattern. For example, for one group, every response might have in common the answer C to question 2, the answer D to question 3, the answer A to question 5, the answer A to question 6, the answer B to question 8, and the answer A to question 10. The other answers to the other questions will vary within the group, but the group is formed by the neural network based on the commonality between the answers to some of the questions (four of them in that case). From one group to another, the precise number of common responses varies in theory between 1 and X, where X is the number of questions. In this experiment, where there are 10 questions in 1st English trial, the groups had between 5 and 8 (Trial 1) common answers. More details of the steps that occur in ESDNN and the ESDNN learning algorithm are given in (Palmer-Brown and Jayne 2011). The training relies upon having representative training data. The number of responses required to train the system so that it can generate the states of knowledge varies from one domain to another. When new responses create new groups, more training data is required. Once new responses stop creating new groups, it is because those new responses are similar to previous responses, and sufficient responses to train the system reliably are already available. The number of groups formed depends on the variation in student responses.

3.3 How the system guides learning

The feedback is designed by academics so that it does not identify which questions were incorrectly answered. The academics are presented with the groups in the form of templates of student responses. Table 2 shows some examples of the group formed. For example, "A/D B mix" represents group 1 characterized by all the students answering A or D to question 1, B to question 2, and mixed answers to question 3. Hence, the educator can easily see the common mistakes in the groups of the student answers highlighted by the tool. The feedback texts are associated with each of the pattern groupings and are composed to address misconceptions that may have caused the incorrect answers common to that pattern group. The student responses, recorded in the database, can be used for monitoring the progress of the students and for identifying misunderstood concepts that can be addressed in subsequent face-to-face sessions. The collected data can be also used to analyze how the feedback influences the learning of individual students by following a particular student's progress over time and observing how that student's answers change after reading the feedback. Student responses can also be used to retrain the neural network and see whether refined groupings are created, which can be used by the educator to improve the feedback. Once designed, MCQs and feedbacks can be reused for subsequent cohorts of students.

Table 2 Example of answer groups

Group number	Group formed		
	Question 1	Question 2	Question 3
1	A/D	B	Mix
2	C	D	D
3	Mix	A	A/B/D

4 Experimental environment

In order to evaluate the performance and effectiveness of this novel e-learning system, abundant target-oriented testing needs to be carried out in different fields. Furthermore, we also aim to enhance this system to overcome its deficiencies during practical applications. Thus, this study is composed of three main parts. Firstly, we evaluated the system by collecting and analysing a large number of testing data reflecting the students' learning gains by using this system as well as the survey and interview data reflecting the students' satisfaction and attitudes towards this system. Secondly, the investigation will lead to an understanding of the potential of the on-line diagnostic feedback approach across different subject areas. Furthermore, this research should also produce guidelines for the design principles of on-line MCQs in the context of diagnostic feedback learning environments. Thirdly, we managed to enhance the existing system according to the evaluation data. The details of this experiment which are conducted to assess the use of M-OFS during academic year 2010–2011 are shown in next section.

4.1 Data collection and feedback generation

Meaning group have emerged and formed with small number of student with training data, not computational intensive scalability with large number of data. For the first trial presented in Sect. 5, the ESDNN was trained with the responses for 10 MCQs on English Grammar obtained from previous cohort of students. After training, appropriate feedback text was written by academics for each of the group of students' responses that address the conceptual errors implicit in combinations of incorrect answers. During the trial, a current cohort of students was asked to provide responses on the same questions, they were given the feedback on the combination of incorrect answers and their responses recorded in the database. The feedback texts are composed around the pattern groupings and are aimed at misconceptions that may have caused the incorrect answers common within the pattern group.

An example of a typical response to the questions below is $(1 < \mathrm{D} > 2 < \mathrm{B} > 3 < \mathrm{A} > 4 < \mathrm{B} > 5 < \mathrm{B/D} > 6 < \mathrm{A} > 7 < \mathrm{A} > 8 < \mathrm{C} > 9 < \mathrm{A/C} > 10 < \mathrm{B/D} >)$

1. ______ no cause for alarm, the old man went back to his bedroom.
A. There was *B. Since* *C. Being* ***D. There being***

2. Even as a girl,______ to be her life, and theater audiences were to be her best teachers.
A. performing by Melissa were.
B. it was known that Melissa's performances were
C. knowing that Melissa's performances were
D. Melissa knew that performing was

3. Agriculture is the country's chief source of wealth, wheat ______ by far the biggest cereal crop.
A. is *B. been* *C. be* *D. being*

4. This company has now introduced a policy_____ pay rises are related performance at work.
A. which ***B. where*** *C. whether* *D. What*

5. She managed to save ______she could out of her wages to help her brother.
A. how little money ***B. so little money***
C. such little money *D. what little money*

6. He left orders that nothing______ touched until the police arrived here.
A. should be *B. ought to be* *C. must be* *D. would be*

7. As it turned out to be a small house party, we _________ so formally.
*   **A. need not have dressed up**   B. must not have dressed up*
*   C. did not need to dress up   D. must not dress up*

8. I ______ the party much more if there hadn't been quite such a crowd of people there.
*   A. would enjoy   B. will have enjoyed*
*   **C. would have enjoyed**   D. will be enjoying*

9. There ought to be less anxiety over the perceived risk of mountain climbing than _____ in the public mind today.
*   A. exists   B. exist   **C. existing**   D. to exist*

10. Fat cannot change into muscle ______ muscle changes into fat.
*   A. any more than   **B. no more than**   C. no less than   D. much more than*

This is classified into Group 6, which generates the following feedback:

Group 6 Feedback 1 Four points should be stressed. First, the logical subject of the adverbial phrase should agree with that of the main clause. Second, two verbs in a sentence need a conjunction. E. G. I am a teacher but you are a student. Third, the usage of various noun clauses should be familiar with. E. G. The news came that he died. ("that" does not serve as any part of the clause.) Fourth, some fixed structures in the comparative form should be memorized. E. G. not so. . .as. . .

In total data of six trials were collected (3 English trials, 2 Math trials, 1 Plagiarism awareness trial). In this paper, it presents the details of data collection of 2nd English trial. The data for training is collected from three previous year's MCQs tests. For these three tests, 94 students' answers were used to training. The trials data were collected during academic year 2010–2011. The data of two separate MCQ paper tests and final examination results were gathered. 83 students entered the survey and 16 students were randomly selected for interview. The states of knowledge of students were achieved by using ESDNN.

4.2 English experiments

To investigate and evaluate how the M-OFS guide and support students to learn, three English experiments were under taken by level 2 and level 3 students at JinQiao University (JQU) and Kunming Technology University (KTU) in China during the academic year 2010–2011. The 1st experiment is introduced below.

In the first experiment, data was collected from 148 students taking English language courses whom were randomly separated into two groups. The experimental group of 83 students used M-OFS, and the control group of 65 students received the same training but without using M-OFS. The system trial includes 10 MCQs with 4 potential answers, related to English grammar. The duration of this trial is flexible. When students were using M-OFS, they were encouraged to answer the MCQs (submit their answers) as many times as they wish until they got all the correct answers or gave up (students were not given answers or how many answers were correct in their feedback, except that they answered all correct answers). Two MCQ paper tests with different questions from system trials were applied to 116 students, and 83 students participated in both paper test and system trial. 83 students completed survey after second paper test. This system trial, paper test and survey were completed in practice lessons in a computer room at JQU.

5 Empirical study

This section discusses the results from the first experiment in order to evaluate the effectiveness of using M-OFS to support students' deep learning.

The survey and interview were conducted after the system trial. 83 (100 %) students conducted the survey. 16 (19 %) students were randomly chosen for interview. For the survey, 71.1 % students are satisfied with using system. 84.4 % students think the feedback is what they need. Using M-OFS to learn were positively evaluated by students, illustrate that the hypotheses H1 is supported. 90.4 % students would like to use the system again. 92.8 % students would like to recommend the system to a friend or classmate in the future. 81.9 % students have never used similar system before. For interviews, most students (94 %) feel this system is useful and helps them to improve their knowledge, it indicates the hypotheses H1 is supported as well; moreover, 69 % students want the exact answers in the feedback in the end. Students also want a picture of their learning process which can point out their weakness and a suggestion of how to improve their English. Some students feel that if they tried many times but cannot find the correct answer, they will lose patience in the end.

5.1 Experiment and result

148 students are involved in the first experiment. 116 students completed the separate MCQs paper test before and after using the system. 83 students participated in system trial, and separate MCQs paper tests. For system trials, a total of 1,118 answers/attempts were submitted and a total 2,143 min were spent by 83 participants. All of the students submitted their answers at least once. The maximum number of attempts was 106 times and the minimum was 1. The average attempts for each student is 13.5 times. The average time spent by each student is 25.8 min and the average time of each attempt is 1.92 min. Two students (2.4 %) spent more than 60 min. 35 (42.2 %) students spent more than the average time. No students achieved the all correct answers at the beginning. 55 (66.3 %) students increased their scores by an average of 12.8 %, whilst one student increased his score by 70 %. In this trial, with 10 questions and 4 possible answers, there are more than one million possible combinations of answers, thus the students are unlikely to make improvement by guessing answers; hence, the results show the feedback had a positive impact which partially supports hypotheses H2. For separate MCQs paper tests, the average score before system trial is 51.6 %, and the average score after system trial is 59.1 %. One student increased his score by 40 %. 74 % students increased their scores. In this test, the students were not given any answers or feedback between first (before system trial) and second (after system trial) test; furthermore, the first trial were applied 3 h before the system trial and the second test were conducted 30 min after system trial; hence, the students are only learnt by using M-OFS but not any other ways; thus the results above are confident, therefore partially supporting hypotheses H3. In addition, this result also can partially support hypotheses H2. For final examination, both the experimental group and the control group enter the same 4 days final examination. The experimental group got 79.52 % and control group got 71.3 % in English grammar module. This result confirms the hypotheses H4; furthermore, it also supports hypotheses H2.

5.2 Some group behavioural characteristics

Previous work has made an initial investigation of the behavioural characteristics of students during their learning interaction with a diagnostic feedback system Alemán et al. (2011).

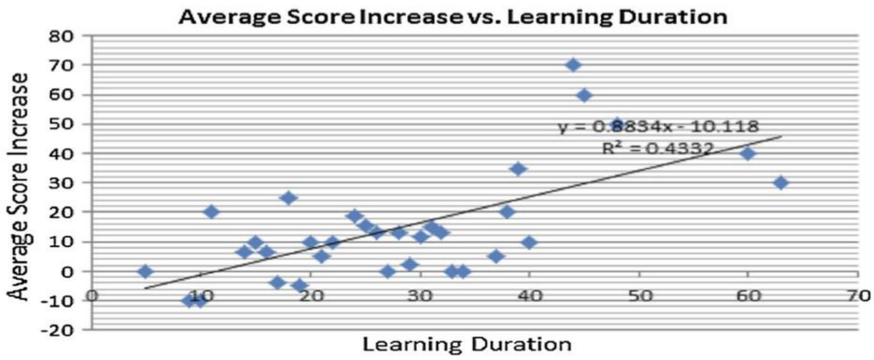


Fig. 2 Average score increased versus learning duration

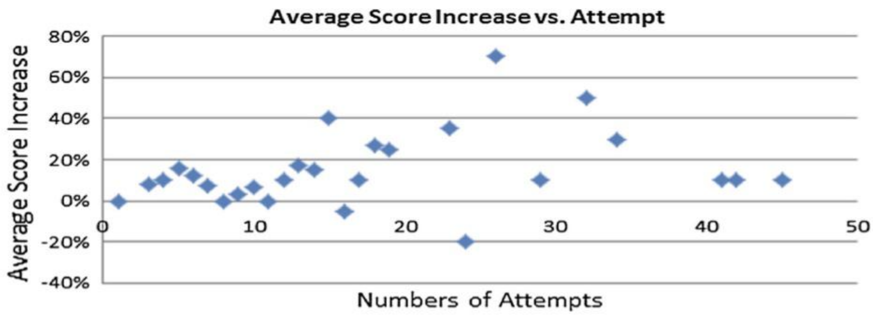


Fig. 3 Average score increase versus attempt

In order to explore the characteristics of students, and relate these to student responses and performance in the tests, five behavioural variables were analysed: the number of attempts (submissions), the average score changed between attempts, the average score at the end of trial, the amount of time spent to make each attempt, and the learning duration. Figure 2 illustrates a learning behaviour of this group of students by analysing the relationship between average scores increased and learning duration. Each blue point represents average scores increased of all students used the same learning time, and its coordinate of x-axis represents student's learning duration, and its coordinate of y-axis represents average scores increased. It can be achieved from this figure that average scores increased when students spent more time on studying from the system.

Figure 3 illustrates a learning behaviour of this group of students by analysing the relationship between average scores increased and number of attempts. Each blue point represents average scores increased of all students did the same number of attempts, and its coordinate of x-axis represents number of students' attempts, and its coordinate of y-axis represents average scores increased. It can be achieved from this figure that average scores increased when students did more attempts before peak, and average scores no longer increased when students did 26 attempts, and average scores decreased by doing more attempts after peak.

Long Learning Duration (time spent on learning $\geq$ 44 min) and Medium Attempts (35 $\geq$ number of attempt $\geq$ 18) are consistently associated with good score increases, especially the combination of Long Learning Duration and Medium Attempts, and hence represent successful learning strategies amongst the students (Tables 3, 4, 5, 6).

Table 3 *Short learning duration*: time spent on learning < 44 min; *Long learning duration*: time spent on learning ≥ 44 min

Behavioral group	Average score change (%)	Average score in the end (%)	Number of students
Short learning duration	9.2	50.6	77
Long learning duration	46.7	95	6

Table 4 *Many attempts*: number of attempt > 35, *Medium attempts*: $35 \geq$ number of attempt ≥ 18, *Few attempts*: number of attempt < 18

Behavioral group	Average score change (%)	Average score in the end (%)	Number of students
Many attempts	8.8	51.2	65
Medium attempts	31.5	72.3	13
Few attempts	15	57.5	4

Table 5 *Slow attempts*: average time spent on each attempt ≥ 4 min; *Rapid attempts*: average time spent on each attempt < 4 min

Behavioral group	Average score change (%)	Average score in the end (%)	Number of students
Slow attempt	12.96	53.52	54
Rapid attempt	11.72	54.48	29

Table 6 *Few rapid attempts*: average time spent on each attempt < 4 min and number of attempt < 18; *Many rapid attempts*: number of attempt >35 and average time spent on each attempt < 4 min; *Slow Few attempts*: average time spent on each attempt ≥ 4 and number of attempt < 18; *Rapid medium attempts*: $35 \geq$ number of attempt ≥ 18 and average time spent on each attempt < 4 min

Behavioral group	Average score change (%)	Average score in the end (%)	Number of students
Few rapid attempts	6	48.4	45
Many rapid attempts	15	57.5	4
Slow few attempts	16	54.5	20
Rapid medium attempts	31.5	72.3	13

Figure 4 illustrates the behaviour of students in terms of what might be called knowledge states. These states correspond to the student responses triggered by patterns of student answer. In other words, a state of knowledge captures some commonality in a set of questions responses. For example, if there are several students who give the same answer (correct or incorrect) to two or more of the questions, snap-drift will form a group associated with one particular output neuron to include all such cases. That is an over simplification, because some of those cases may be pulled in to other 'stronger' groups, but that would also be characterized by a common feature amongst the group of responses. Figure 3 shows the knowledge state transitions. Each time a student gives a new set of answers, having received some feedback associated with their previous state, which in turn is based on their last answers, they are

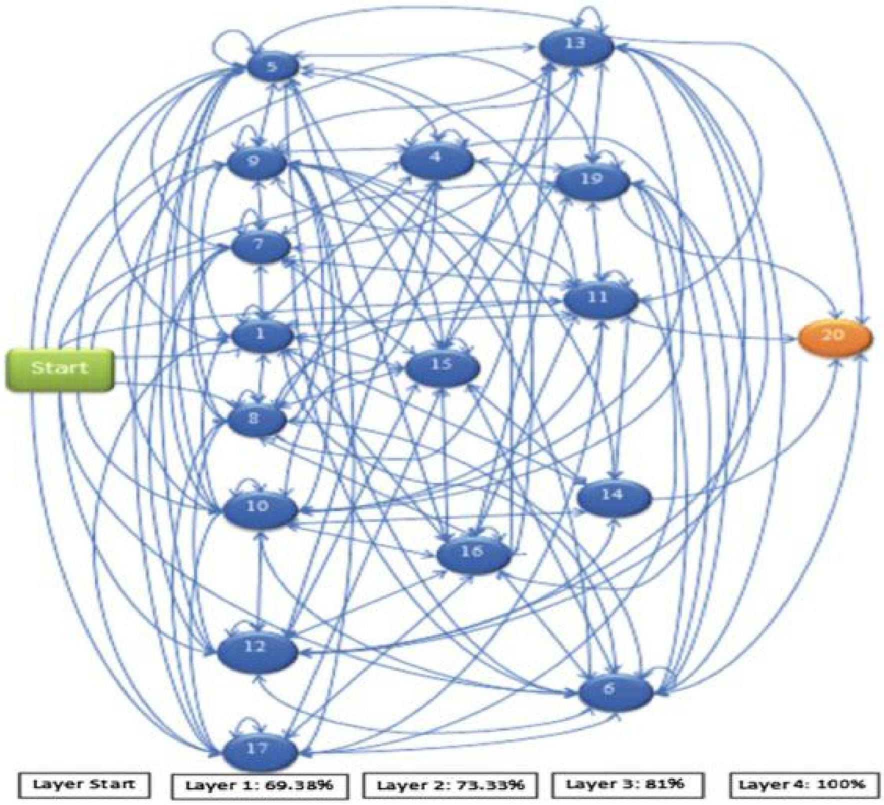


Fig. 4 Knowledge state transitions

reclassified into a new (or the same) state, and thereby receive new (or the same) feedback. The tendency is to undergo a state transition immediately or after a second attempt or several attempts. A justification for calling the states 'states of knowledge' is also to be found in their self-organization into the layers. There are 5 layers: Start, Layer 1, Layer 2, Layer 3, and Layer 4. A student on state 5, for example goes via one of the states in the further layer such as state 4 or 13, before reaching the 'state of perfect knowledge' (state 20) which represents correct answers to all questions. On average, and unsurprisingly, the state-layer projecting onto state 20 (states 12) is associated with more correct answers than the states in the previous layer. This is true of state 13 which projects onto state 20, and it is also true of state 4 although this state does not project onto state 20. The states in the middle layer (layer 2) do not connect to start layer or layer 4. The average score of each layer is increased from start to state 20 (Average scores are 69.38 % at beginning level (layer 1), 73.33 % at middle layer (layer 2), 81 % at advanced level (layer 3), and 100 % at state 20). Students often circulate within layers before proceeding to the next layer. They may also return to a previous layer, but that is less common.

6 Summary

Six trials across three totally different subject areas have been carried out in three universities in two countries, the UK and China: 3 English trials, 2 Mathematics trials, and 1 Java

Table 7 *Short learning duration many attempts*: time spent on learning < 44 min and number of attempt >35; *Short learning duration medium attempts*: time spent on learning < 44 min and 35 ≥ number of attempt ≥ 18; *Short learning duration few attempts*: time spent on learning < 44 min and number of attempt < 18; *Long learning duration few attempts*: time spent on learning ≥ 44 min and number of attempt < 18; *Long learning duration medium attempts*: time spent on learning ≥ 44 min and 35 ≥ number of attempt ≥ 18

Behavioral group	Average score change (%)	Average score in the end (%)	Number of students
Short learning duration few attempts	7.9	48.4	62
Short learning duration many attempts	16	57.5	4
Short learning duration medium attempts	23	64	10
Long learning duration few attempts	33.3	90	3
Long learning duration medium attempts	60	100	3

Programming trial. The English trials are very successful with 500 students participated. A large volume of data has been captured during the trials.

The students participating in this second trial are from KTU which is one of the top universities in China. The entrance requirements of KTU are much higher than those of JQU, which means that students studying in KTU generally have a better academic background than students in JQU. The results of the second English system trials show that the average score is increased by 21.2 % in the end. The results of the separate MCQ paper test indicated the average mark of the group is increased by 7.6 %. In addition, the final examination, the average mark of the experimental group is 7.5 % higher than the control group. Student surveys show that 87.8 % students are satisfied with our e-learning system and 89.2 % students feel that the intelligent diagnostic feedback is beneficial to enhance their learning experience. Furthermore, according to all of the results, the M-OFS system demonstrates that it is working well for both less qualified and better qualified students at two universities JQU and KTU. The results of group behaviours in Table 7 show that students in the group "Long learning duration medium attempts" achieved much better learning effects than others. In comparison with the 1st trial, the students in the group "Long learning duration medium attempts" achieved better learning effects than others. This has indicated that the students in the 2nd trial who have better academic background prefer more reflective thinking and independent study from the feedbacks, whilst, the students in the 1st trial who have weaker academic background prefer more frequent and direct feedbacks as part of their learning process.

7 Conclusion and future work

This research develops a novel method for using snap-drift in a diagnostic tool to provide feedback. The neural network diagnostic feedback approach in MCQ has been systemically applied to large cohorts of students and evaluated across a range of different subject areas. The neural network discovers groups of similar answers that represent different knowledge states of the students. The feedback targets the level of knowledge of individuals, and their misconceptions, guiding them toward a greater understanding of particular concepts. The results of the experiment show that an improvement in the learning process can be obtained by using M-OFS. In future work, it is intended to compare the effects of the feedback in new trials to the effects of other types of feedback already reported in the literature. Another promising avenue for further investigation is the extension of the tool to support knowledge state transition diagram construction and statistical data collection, which could help instructors to analyse the difficulty of the MCQs and to track students though the developmental stages of their learning.

References

Alemán JLF, Palmer-Brown D, Jayne C (2011) Effects of response-driven feedback in computer science learning. IEEE Trans Educ 54:501–508

Alessi M, Trollip S (2001) Multimedia for learning: methods and development, 3rd edn. Allyn and Bacon, London

Bang P (2003) Engaging the learner—how to author for best feedback. In: Felix U (ed) Language learning online: towards best practice. Swets & Zeitlinger, The Netherlands

Blessing S, Gilbert S, Ourada S, Ritter S (2007) Lowering the bar for creating model-tracing intelligent tutoring systems. In: Luckin R et al. (eds) Artificial intelligence in education. IOS Press, Amsterdam

Clark RC, Mayer RE (2009) Instructional strategies for directive learning environments. In: Silber KH, Foshay WR (eds) Handbook of improving performance in the workplace: instructional design and training delivery. Pfeiffer, San Francisco

Epstein ML, Lazarus AD, Calvano TB, Mathews KA, Hendel RA, Epstein BB, Brosvic GM (2002) Immediate feedback assessment technique promotes learning and corrects inaccurate first response. Psychol Record 52(2):187–201

Garber PR (2004) Giving and receiving performance feedback. HRD Press, Canada

Gheorghiu R, Vanlehn K (2008) XTutor: an intelligent tutor system for science and math based on excel. In: Woolf BP et al. (eds) Intelligent tutoring systems: 9th international conference, ITS 2008. Springer, Germany

Hatziapostolou T, Paraskakis I (2010) Enhancing the impact of formative feedback on student learning through an online feedback system. Electron J E-Learn 8(2):111–122. Available online at www.ejel.org

Hefce national student survey (2007–2010) HEFCE, London, UK. Online at: http://www.hefce.ac.uk/whatwedo/lt/publicinfo/nationalstudentsurvey/

Heinze A, Procter C, Scott B (2007) Use of conversation theory to underpin blended learning. Int J Teach Case Stud 1(1&2):108–120

Higgins E, Tatham L (2003) Exploring the potential of multiple-choice questions in assessment. Learn. Teach. Action 2(1). Assessment, Winter 2003

Kuechler WL, Simkin MG (2003) How well do multiple choice tests evaluate student understanding in computer programming classes? J Inf Syst Educ 14(4):389–399

Little B (2001) Achieving high performance through e-learning. Ind Commer Train 33(6):203–206

Ma Z (2006) Web-based intelligent e-learning systems. Information Science Publishing, USA

Nelson MM, Schunn CD (2009) The nature of feedback: how different types of peer feedback affect writing performance. Instr Sci 37(4):375–401

Özdener N, Satar HM (2009) Effectiveness of various oral feedback techniques in CALL vocabulary learning materials. Egitim Arastirmalari—Eurasian J Educ Res 34:75–96

Palmer-Brown D, Jayne C (2011) Snap-drift neural network for self-organisation and sequence learning. Neural Netw 24(8):897–905

Paxton M (2000) A linguistic perspective on multiple-choice questioning assessment and evaluation. Assess Eval High Educ 25(2):109–110

Payne A, Brinkman, Wilson F (2007) Towards effective feedback in e-learning packages: The design of a package to support literature searching, referencing and avoiding plagiarism. In: Proceedings of HCI2007 workshop: design and use and experience of e-learning systems, pp 71–75

Race P, Brown S (2005) 500 tips for tutors, 2nd edn. Routledge Falmer, London

Race P (2006) The lecturer's toolkit—a practical guide to assessment. Learning and teaching, 3rd edn. Routledge, London

Rane A, Sasikumar M (2007) A constructive learning framework for language tutoring. In: Iskander M (ed) Innovations in e-learning, instruction technology, assessment, and engineering education. Springer, The Netherlands

Reiser BJ, Kimberg DY et al (1992) Knowledge representation and explanation. In: Larkin JH, Chabay RW (eds) Computer assisted instruction and intelligent tutoring systems. Lawrence Erlbaum Associates, New Jersey

Springgay S, Clarke A (2007) Mid-course feedback on faculty teaching: a pilot project. In: Darling LF, Erickson GL, et al (eds) Collective improvisation in a teacher education community, Chapter 13. Springer, The Netherlands, pp.171–185

Vasilyeva E, Pechenizkiy M, Bra PD (2008) Adaptation of elaborated feedback in e-learning. In: Nejdl W, Kay J, Pu P (eds) Adaptive hypermedia and adaptive web-based systems. Springer, German

An Example Feedback of Mathematics

G 2	It is very important to understand how the exponent laws work. When you are given a term with a form of **a^m^**, it is defined that: a is the ***base***， m is the ***exponent***. Generally, more than one exponent law should be used to evaluate one exponent expression. The exponent laws will be given as following: $a^m \times a^n = a^{m+n}$ $\frac{a^m}{a^n} = a^{m-n}$ $(a^m)^n = a^{m\times n}$ $(ab)^n = (a^n)\times(b^n)$ $a^0 = 1$ $a^{-n} = \frac{1}{a^n}$ $\sqrt[m]{a^n} = a^{\frac{n}{m}} = (\sqrt[m]{a})^n = (a^{\frac{1}{m}})^n$ **Example 1: evaluate 1024^0** using law **a^0 = 1, (**a=1024) 1024^0 => 1 (answer) **Example 2: evaluate 3^{-2}** using law a^{-n} = $\frac{1}{a^n}$, (a=3, n=2) 3^{-2} => $\frac{1}{3^2}$ => $\frac{1}{9}$ **Hints:**

	1. Remember: $a^m \times a^n \neq a^{m \times n}$, $\frac{a^m}{a^n} \neq a^{\frac{m}{n}}$ 2. It is also important to remember and understand the exponent laws before using it Click the following link to find out more details about the exponents. Exponent laws: http://www.mathexpression.com/exponent-rules.html Exponents(Exponents and Calculations): http://www.aaastudy.com/exp.htm <C/mix><B><A><C><B/a/c>

G 5	You need a detailed study on the **Exponent** and the **Logarithm**. **Exponent:** It is very important to understand how the exponent laws work. When you are given a term with a form of $\mathbf{a^m}$, it is defined that: **a** is the ***base***, **m** is the ***exponent.*** Generally, more than one exponent law should be used to evaluate one exponent expression. The exponent laws will be given as following: $a^m \times a^n = a^{m+n}$ $\frac{a^m}{a^n} = a^{m-n}$ $(a^m)^n = a^{m \times n}$ $(ab)^n = (a^n) \times (b^n)$ $a^0 = 1$ $a^{-n} = \frac{1}{a^n}$ $\sqrt[m]{a^n} = a^{\frac{n}{m}} = (\sqrt[m]{a})^n = (a^{\frac{1}{m}})^n$ **Example: evaluate $\mathbf{2^{-3}}$** using law $a^{-n} = \frac{1}{a^n}$, (a=2, n=3) $2^{-3} \Rightarrow \frac{1}{2^3} \Rightarrow \frac{1}{8}$ **Hints:** 1. Remember: $a^m \times a^n \neq a^{m \times n}$, $\frac{a^m}{a^n} \neq a^{\frac{m}{n}}$ 2. It is also important to remember and understand the exponent laws before using it

G 7

It is very important to understand what the binary number and hexadecimal number are. They are two major numbering systems used in computer science. The binary number is the base 2 numbering system. The hexadecimal (hex) number is the base 16 numbering system. Generally, people use a decimal (base 10) numbering system in their daily life. Before converting between hex and binary numbers, you should to know how to convert the binary and hex number to decimal number. You must remember the table is given as below:

Hexadecimal	Decimal	Binary
0	0	0000
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
A	10	1010
B	11	1011
C	12	1100
D	13	1101
E	14	1110
F	15	1111

The best way to convert between binary and hex numbers is:

1. Consider that each four bits portion of binary number represent one digit

of hex number. On the other hand, each one digit of hex number represents four bits portion of binary number.

2. (binary-hex) Write down the hex number equivalent for each four bits of binary number. (less than 4 bits, complement 0)
3. (hex-binary) Write down the binary number equivalent for each one digit of hex number. (less than 4 bits, complement 0)

Example: (hex-binary)

3C2

3 -> 0011

C -> 1100

2 -> 0010

So: 3C2 -> 001111000010

You should remove the leading 00:

3C2 -> 1111000010 (answer)

Hints:

Base 2: there can only be two values for a specific digit; either 0 or 1.

Base 16: there can be sixteen values for a specific digit, from 0 to 9, then letter A to F.

Base 10: there can be ten values for a specific digit, from 0 to 9.

Click the following link to find out more details about the binary and hex number.

Binary number: http://www.computerhope.com/binhex.htm#01

Hexadecimal number: http://www.computerhope.com/binhex.htm#02

Convert between binary and hex numbers:

	http://www.ehow.com/how_4496769_convert-between-hexadecimal-binary-numbers.html <E><D><E><D><A>

An example Survey of Mathematics

The survey of mathematics is consist of five questions:
How likely are you to use the system again?
How likely are you to recommend the system to a friend or classmate in the future?
How does the system compare to the other E-learning system?
Please leave your comments
Do you have any suggestion for improvement?

UEL E-MCQs Survey 2010
Mathematic 6 Questions

How likely are you to use our E-MCQs system again?

1. ☐ Very likely
2. ☐ Somewhat likely
3. ☐ Neutral
4. ☐ Somewhat unlikely
5. ☐ Very unlikely

How likely are you to recommend our E-MCQs to a friend or classmate in the future?

1. ☐ Very likely
2. ☐ Somewhat likely
3. ☐ Neutral
4. ☐ Somewhat unlikely
5. ☐ Very unlikely

How does our E-MCQs compare to the other similar E-learning system?

1. ☐ Well Below Average
2. ☐ Below Average
3. ☐ Average
4. ☐ Above Average
5. ☐ Well Above Average
6. ☐ Never use other similar E-learning system

Comment:

Do you have any suggestions for improvement?

An example of English MCQs

1. _____ no cause for alarm, the old man went back to his bedroom.	A. There was B. Since C. Being D. There being
2. Even as a girl,_____ to be her life, and theatre audiences were to be her best teachers.	A. performing by Melissa were. B. it was known that Melissa's performances were C. knowing that Melissa's performances were D. Melissa knew that performing was
3. Agriculture is the country's chief source of wealth, wheat______by far the biggest cereal crop.	A. is B. been C. be D. being
4. This company has now introduced a policy_____ pay rises are related performance at work.	A. which B. where C. whether D. What
5. She managed to save__________she could out of her wages to help her brother.	A. how little money B. so little money C. such little money D. what little money
6. He left orders that nothing____ touched until the police arrived here.	A. should be B. ought to be C. must be D. would be
7. As it turned out to be a small house party, we _____ so formally.	A. need not have dressed up B. must not have dressed up C. did not need to dress up D. must not dress up
8. I ____ the party much more if there hadn't been quite such a crowd of people there.	A. would enjoy B. will have enjoyed C. would have enjoyed D. will be enjoying
9. There ought to be less anxiety over the perceived risk of mountain climbing than _____ in the public mind today.	A. exists B. exist C. existing D. to exist

10. Fat cannot change into muscle muscle changes into fat.	A. any more than B. no more than C. no less than D. much more than