A Comparison Of Computer Aided Learning And Traditional Didactic Lectures For Teaching Clinical Decision Making Skills To Optometry Undergraduates

BHAVNA RICKESH PANCHOLI

Doctor of Philosophy

ASTON UNIVERSITY

June 2016

Aston University
A Comparison Of Computer Aided Learning And Traditional Didactic Lectures For Teaching Clinical Decision Making Skills To Optometry Undergraduates

Bhavna Rickesh Pancholi
Doctor of Philosophy
2016

Thesis abstract

This study was designed to compare computer aided learning (CAL), in the form of a Virtual Patient (VP), and traditional didactic lectures as methods of teaching clinical decision making (CDM) skills to second year Optometry undergraduates. Comparisons were based on performance in multiple-choice examinations testing CDM skills (actual mastery), student feedback relating to confidence in CDM skills (perceived mastery or self-efficacy) and student satisfaction. The influence of sex, learning style and academic ability was also investigated. This is the first time that these aspects of teaching pedagogy have been studied together.

Current literature informed development of didactic lectures and an online VP. Both teaching methods were designed to ensure that the same clinical content was included. This content was aimed at training students to perform problem-orientated eye examinations. A cohort of 102 students was taught using the traditional didactic lectures in academic year 2010-11 and 93 students using the online VP in academic year 2011-12.

An established Index of Learning Styles instrument was used to classify students according to their preference in four learning style dimensions. Both teaching methods were designed to cater for both poles of each learning style dimension. Most students had no strong learning preferences but those that did had a tendency towards the active-sensing-visual-sequential profile.

Actual and perceived mastery were scored for five key learning objectives; question selection, critical symptom recognition, test selection, critical sign recognition and referral urgency selection. The influence of academic ability and teaching method differed for each learning objective; didactic lectures favouring some, the VP others. Learning style and sex had no influence, indicating that both teaching methods catered equally for males and females with all learning styles. Comparisons between perceived and actual mastery revealed poor self-assessment accuracy.

Student satisfaction, rated on a five point Likert scale, was equally high for both teaching methods. Sex was the only influential variable, with males favouring one aspect of VP training.

Overall, the findings suggest that CAL should be used to supplement traditional teaching rather than replace it in order to ensure that all students benefit equally. Future research may wish to focus on self-assessment accuracy as a means of improving academic performance.

Keywords: learning styles, academic ability, sex, self-assessment accuracy, academic performance

For Shayna and Kaylen Pancholi

Acknowledgements

I would like to begin by thanking Dr Mark Dunne for the support and guidance he has shown me throughout the course of this thesis. I am sure it would have not been possible without his help.

Thanks to Abdullah Bhamji, Pardeep Chohan, Hardish Dhillon, Sharon Lamson, Hui Lok, Ebrahim Lorgat, Nashreen Mulla, Andrew Munson and Taha M Jalal, the focus group of Optometry undergraduate students at Aston University for their assistance in testing and evaluating the Virtual Patient software while it was in the developmental stages.

I am truly indebted and thankful for the incredible support and motivation from my loving parents. And finally I would like to thank my husband, Rickesh Pancholi, whose encouragement and support has been priceless.

List of contents

Abbreviations used

CAL	Computer aided learning
CDM	Clinical decision making
CHAID	Chi-squared automatic interaction detection
CPD	Clinical Practice Development
DTA	Decision tree analysis
ELT	Experiential Learning Theory
ILS	Index of learning styles
LSI	Learning styles inventory
LSQ	Learning style questionnaire
MBTI	Myers-Briggs type indicator
MCQ	Multiple choice question
MSD	Mind Style Delineator
PBL	Problem based learning
RFV	Reason for visit
VP	Virtual patient

List of tables

List of figures

Chapter 1 – Introduction

1.1 Introduction

Computer aided learning (CAL) and its pros and cons are discussed in this chapter. The key objectives of the present study are then outlined.

1.2 Computer aided learning

Improving student learning is an important issue in higher education (Dermo, 2009). There is continual research in how to improve teaching methods (Levinson-Rose & Menges, 1981; Kneebone et al., 2002) and thereby enhance the student learning experience.

There has been a considerable amount of research into how CAL or computerised simulators can be used to aid teaching. CAL and computerised simulators provide a number of benefits in comparison to traditional teaching methods, especially when they are used to train clinical skills. The development of new clinical skills relies on learning from experience and repeated practise with a variety of patients (Kneebone et al., 2002). Traditional teaching methods have three key problems. Firstly, the practise of clinical skills on real patients needs to be supervised and by doing so this imposes time constraints. Students can only practise skills at specific times when a supervisor is available. Secondly, universities normally have to recruit patients for students to practise on. This approach may risk reducing the variety of scenarios that the students have to experience and can be costly for the university. Thirdly, there is nearly always a lack of one to one supervision. This means that supervisors are unable to provide specific feedback as it is not possible for them to monitor each and every step taken by the student.

There are also other concerns with traditional teaching methods, especially when considering assessments. Different supervisors may have varying criteria for assessing the same clinical skills (Sharaf et al., 2007). Furthermore, grades given by the same supervisor can vary on different occasions (Sharaf et al., 2007).

This is where computerised simulators can be beneficial (Zary et al., 2006). They can be pre-programmed to behave as consistent assessors. They can also monitor every step taken by the student and provide individual feedback on the performance. Immediate feedback is valued in learning (McGraw & O'connor, 1999; Kneebone et al., 2002). Computerised simulators can be used unsupervised and so allow self-paced learning at any time convenient to the student. It also provides a risk-free platform for students to practise clinical skills. They can make mistakes as they learn. Another benefit provided by computerised simulators is that they can easily simulate any clinical scenario. This is particularly useful for rare conditions that would otherwise be difficult to find in patients recruited by the university. In the long term, computerised simulators can be a cheaper alternative to recruiting patients. Once the software is developed, there are fewer costs associated with its maintenance.

While CAL has many of benefits, Greenhalgh (2001) has pointed out that without the correct support for staff and appropriate training for students, CAL can prove to be an ‘expensive disaster’. She highlighted that a computer can be a ‘temperamental and unforgiving beast’ if not programmed correctly and the initial programming of CAL material can be very labour intensive and time consuming. Despite this, she still acknowledged that it is becoming ‘a truth universally acknowledged’ that the education of undergraduate students will be enhanced through the use of CAL.

Use of computerised simulators has been proven to improve clinical performance in comparison to traditional teaching methods alone (Glittenberg & Binder, 2006). Glittenberg and Binder developed a computerised simulator of a three dimensional animation demonstrating the neuro-ophthalmological properties of the oculomotor system. Students who used the simulator performed significantly better on a multiple choice question (MCQ) examination in comparison to those who were taught with a conventional lecture. Students rated this type of learning material highly and expressed strong preference for more learning material in other subject areas to be presented in this manner.

Over the last couple of decades there has been a considerable growth in the use of CAL and computerised simulators to aid teaching in many healthcare related disciplines such as medicine, dentistry and nursing (Rogers et al., 1998; Teichman & Richards, 1999; Maleck et al., 2001; Mattheos et al., 2001; Parker & Seifter, 2001; Kneebone, 2005; Jenkins et al., 2008).

Until now, optometry has been falling behind in this area. Although a small number of computerised simulators designed to teach optometrists do exist (Findl, 1998; Lasslo et al., 2002; Bracewell & Anderson, 2003), none of these teaching materials have yet been evaluated.

1.3 Objectives

This study was designed to determine whether CAL in the form of a virtual patient (VP) was more efficient at training clinical decision making (CDM) skills of second year Optometry undergraduates in comparison to traditional didactic lectures. The key research questions were:

- Does examination performance (actual mastery) differ between the two teaching methods? (See chapter 5)
- Does self-efficacy (perceived mastery) differ between the two teaching methods? (See chapter 6)
- Does self-assessment accuracy (the relationship between perceived and actual mastery) differ between the two teaching methods? (See chapter 7)
- Does student satisfaction differ between the two teaching methods? (See chapter 8)

1.4 Summary

The key objectives of the present study have been outlined in this chapter. The research methodology and study sample will be described in the next chapter.

Chapter 2 – Overview of research methodology and study sample

2.1 Introduction

This chapter provides a review of the literature on teaching CDM skills prior to describing how these skills were taught at Aston University. An outline of the purpose, methodology and study sample then follows. Finally, the study sample is described.

2.2 Literature review and background to the development of the Clinical Practice Development (CPD) module

CDM in healthcare professionals has its challenges (Tanner et al., 1987) and requires a broad knowledgebase and critical thinking (Penn, 2008) to process patients' presenting symptoms and signs with the goal of making differential diagnoses and choosing an appropriate management plan. The complexity of CDM is often increased due to missing or incomplete information and also the fact that not all of the presenting information is relevant to the decision making process (Smith et al., 2008, p.89).

Traditionally in many healthcare disciplines, the development of clinical judgement and CDM has been left to on-going personal experiences and by healthcare professionals encountering a variety of patient episodes (Howe et al., 1984). However, the growing expectations that patients have of healthcare professionals has led to changes in teaching with the aim of developing CDM skills prior to having encounters with real patients (Jones et al., 2001). Furthermore, healthcare professionals are accountable for their decisions at a professional level; therefore, it is imperative that they are aware of how to make these important decisions.

Most students have difficulty understanding and grasping the processes involved in CDM (Conrick, 1996). CDM is a multidimensional process (Smith et al., 2008, p.90) and, therefore, teaching CDM skills is also challenging (Conrick, 1996). A number of studies have shown that problem-based learning (PBL) or student-centred learning is more efficient at developing CDM skills (Morales-Mann &

Kaitell, 2001; Wells et al., 2009; Nango & Tanaka, 2010; Carrió et al., 2011; Oja, 2011) (See Chapter 4 for more information regarding PBL).

The teaching and assessment of CDM skills needs to be balanced to create the right learning environment. Heavy workloads and high stress situations such as pressure of assignments or examinations lead to students adopting a rote learning method (Trigwell & Prosser, 1991; Biggs, 1993). Rote learning, often referred to as surface learning, is where lists or techniques are simply memorised by repetition. CDM relies on a deeper level of knowledge and on critical thinking (Penn, 2008, p. 386).

Miller (1990) developed a model of clinical competence (Figure 2.1). Although this model was produced primarily to develop assessment methods corresponding to the different levels of clinical competence identified, it also provides a clear insight into the different stages of learning that occur in the development of clinical competence. Miller's (1990) model has four levels. The lower two levels represent cognition, whereas the upper two levels represent behaviour. The "knows" level involves information and fact gathering. The "knows how" level develops further from this and involves the interpretation of these facts. The "shows how" level involves using this ability to interpret facts and demonstrating clinical competence in a controlled environment designed to closely match the real world. Finally, the "does" level is what actually happens in real life when healthcare professionals are working.

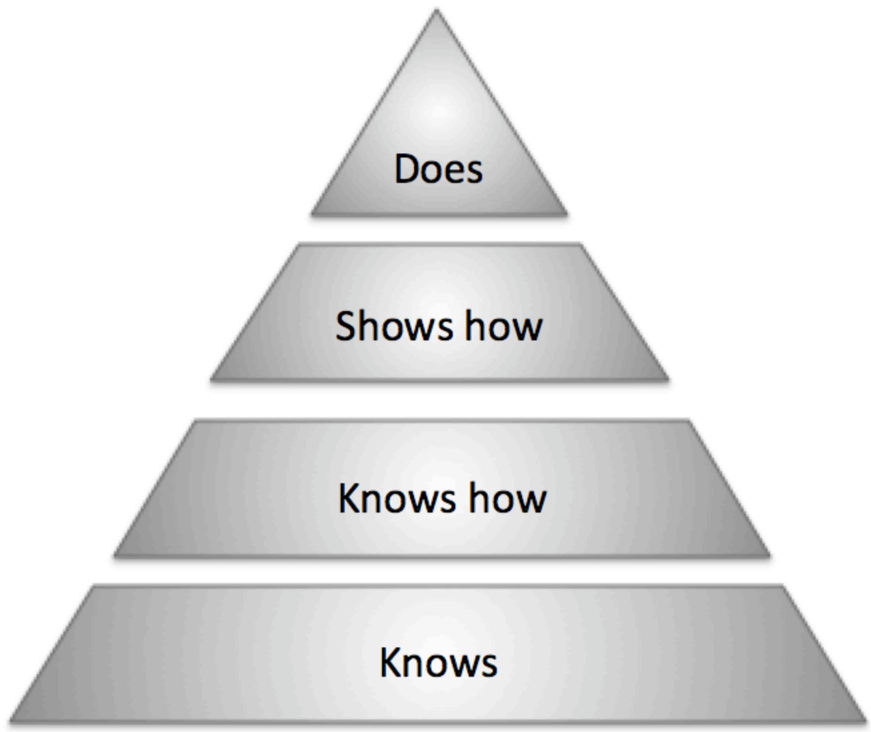

Figure 2.1 Miller's model of clinical competence

Any clinical course designed to develop CDM skills must focus on the first three levels of Miller's pyramid. It must ensure that all clinical facts are taught in a suitable manner along with the ability to interpret this knowledge. This should then be followed by the teaching of how to demonstrate these clinical skills in a safe controlled environment.

In Optometry, CDM requires a sound understanding of the following aspects:

- The basic anatomy and physiology of the ocular structures to help identify ocular abnormalities and their significance;
- The clinical characteristics including symptoms and signs of various ocular diseases;
- How to complete a case history and clinical tests during a routine eye examination in order to detect these symptoms and signs;
- How to perform the supplementary clinical tests that allow the investigation of any ophthalmic disorders detected during a routine eye examination.

Traditionally, Optometry undergraduates at Aston University were taught these four aspects, in separate modules, by different instructors. Therefore, the only opportunity they had to integrate knowledge in these four areas to develop their CDM skills was in the final year where they had the opportunity to examine real patients under supervision. In the opinion of the author, CDM skills should have been taught in a structured manner prior to students examining real patients. Therefore, the Clinical Practice Development (CPD) module was developed for second year optometry undergraduates in preparation for seeing real patients in their final year.

The CPD module was designed to break away from the rote learning of lists of symptoms, signs, or sequences of tests to be performed in an eye examination. Instead, students needed to learn how presenting symptoms and signs lead to differential diagnoses that indicate which clinical tests were required to identify the most likely tentative diagnosis and management plan, possibly involving referral with an appropriate level of urgency. This would effectively encourage the integration and application of knowledge gained from other modules. Students were traditionally taught a database style examination which, although less efficient, lends itself well to ensuring that all clinical skills were practised in students that lack the knowledge and experience to carry out a problem-orientated eye examination (more is said about the two types of examinations in chapter 4). The CPD module introduced the concept of a problem-orientated eye examination.

Recall from Chapter 1 that there had been a considerable growth in the use of CAL to aid the teaching of CDM skills in many healthcare disciplines (Rogers et al., 1998; Teichman & Richards, 1999; Maleck et al., 2001; Mattheos et al., 2001; Parker & Seifter, 2001; Kneebone, 2005; Jenkins et al., 2008). The present study was designed to determine whether CDM skills were learnt more efficiently and with greater levels of student satisfaction using (a) traditional didactic lectures or (b) VP simulation. The following sections describe the study design in more detail. See chapter 4 for more detailed information regarding development of the teaching material for the didactic lectures and the VP.

2.3 The two cohorts

Two cohorts of students took part in this study. The first cohort (the didactic lecture cohort) received a CPD module composed of didactic lectures during their second year, which took place in the academic year of 2010-11. The second cohort (the virtual patient cohort) received a second year CPD module delivered using web-based VP software in the academic year of 2011-12.

Great care was taken to ensure that both teaching formats covered the same clinical material (see chapter 4 for a more detailed description of the development of the teaching material for both cohorts).

2.4 Learning objectives

The learning objectives evaluated by this study focused upon five key areas of CDM: question selection, critical symptom recognition, test selection, critical sign recognition and referral urgency selection. The following learning objectives were assigned to each area:

1. The ability to ask the right questions in history and symptoms based upon a patient's presenting complaint (question selection);
2. The ability to recognise symptoms of serious eye disease (critical symptom recognition);
3. The ability to select the most appropriate tests needed to look for signs that aid differential diagnosis (test selection);
4. The ability to recognise signs of serious eye disease (critical sign recognition);
5. The ability to decide upon the most appropriate referral urgency for any eye disease detected (referral urgency selection).

Perceived and actual mastery of these 5 learning objectives was measured. Perceived mastery was a rating based on the student's own perception of their confidence in the 5 learning objectives derived from a questionnaire completed at the end of their second year and at the beginning of their final year (see chapter 6). Measurement of actual mastery was based on a percentage score for each of the 5 objectives derived from a MCQ examination that took place at the end of the academic year (see chapter 5). Identical MCQs and questionnaires were given to both cohorts. The relationship between perceived and actual mastery, termed self-assessment accuracy was then evaluated (see chapter 7).

2.5 Student satisfaction

Data for student satisfaction of the teaching material was obtained by a questionnaire, which was completed by students at the end of their second year. Identical questionnaires were given to both cohorts (see chapter 8 for more details).

2.6 Influencing factors

The influence of teaching method, sex, academic ability and learning style on actual mastery, perceived mastery and self-assessment accuracy of the five learning objectives along with student satisfaction was investigated in both cohorts.

2.6.1 Sex

Comparisons were made between male and female students.

2.6.2 Academic ability

Academic ability was based on the overall average mark achieved by each student across all second year modules in sessional examinations carried out at the end of the academic year. At Aston University, 5 degree classifications are assigned in the following way: first class corresponds to 70-100%, upper second class to 60-69%, lower second class to 50-59%, third class to 40-49% and less than 40% is classed as a fail. The same banding was applied to each student's overall average mark to create 5 classes of academic ability.

2.6.3 Learning styles

Every individual has a personal preference in how they learn (Felder & Silverman, 1988). For example, some learn best by seeing whereas others do so by hearing; some learn best by reflecting on ideas whereas others prefer active participation. These differences in learning preferences are referred to as learning styles.

In order to determine whether learning styles influenced the actual or perceived mastery of the 5 learning objectives, all students were required to complete and submit an Index of Learning Styles (ILS) (Felder & Soloman, 2004) which was based on the model of learning styles developed by Felder and Silverman (1988). This inventory established every student's learning preference along 4 learning style dimensions; the active – reflective learning style, the sensing – intuitive learning style, the visual – verbal learning style and the sequential – global learning style.

The scoring system of the ILS allowed students to be classified as having either a strong, moderate or balanced learning style preference for each of the 4 learning style dimensions. (See chapter 3 for more detail regarding the theory of learning style preferences, the ILS and the spread of learning styles for the study sample).

2.7 Data collection

To ensure fair comparison of the two cohorts, great care was taken to collect data at the same point in the academic year for both cohorts. This was particularly important with respect to the collection of the learning styles data, as previous research has shown that learning styles may alter over time (Felder, 1993, Honey and Mumford, 2000 p.19, Kolb, 2000).

The ILS was made available from academic week 4 until academic week 17 via the Blackboard learning management system used at Aston University. The questionnaire to determine perceived mastery was released two weeks before the end of the academic year for both cohorts via the Blackboard learning management system. The deadline to complete this questionnaire was one week. The questionnaire to determine student satisfaction was also released 2 weeks before the end of the academic year via the Blackboard learning management system. The MCQ examination to determine actual mastery took place at the end of the academic year for both cohorts. The MCQs formed part of the sessional examination for the CPD module that took place on the 8th June 2011 for the didactic lecture cohort and the 29th May 2012 for the virtual patient cohort.

2.8 Statistical analysis

All statistical analyses were performed using SPSS 21.0 (IBM, Version 21.0 2012) and findings were tested for statistical significance at the 5% alpha level. Kolmogorov-Smirnov tests were used to determine whether the data for student satisfaction, actual mastery, perceived mastery and self-assessment accuracy were normally distributed. Data did not follow a normal distribution and so square root and logarithmic transformations were attempted to convert it to a normal distribution. These failed and so only non-parametric tests were used; Chi-square tests, Kendall's correlations and Decision Tree Analyses (DTA) (see section 2.8.1). A Bonferroni correction was applied to all Chi-square tests to compensate for multiple comparisons.

Power calculations were made using GPower 3.1.0 software (Faul et al., 2007) at the 5% alpha level. As discussed in the author's published work (Prajapati et al., 2010; Prajapati et al., 2011), the determination of statistical power is dependent upon the desired effect size to be investigated. Cohen (1988, p.24) defined effect sizes for a variety of statistical tests as small, medium and large. He explained that the choice of the effect size should always be based on the context of the research study. Any changes to teaching methods would require significant resources, both time and financial. Therefore, it would only be justifiable to make changes to teaching methods based on statistically significant findings for large effects. Consequently, a large effect size was selected for all statistical power analyses. Post hoc power is quoted for all results that were not statistically significant in order to ensure that power was always equal or above the conventional value of 80% (Araujo & Froyland, 2007; Prajapati et al., 2010).

2.8.1 Decision tree analysis

DTA is a form of multivariate analysis where each node of the flowchart represents a test on an attribute and each branch represents the outcome of the test. It benefits from accounting for all variables at once, therefore, eliminating confounding. Another advantage of this analysis over other multivariate analyses is the hierarchical output, where the strongest influence is shown at the top of the classification tree.

The chi-squared automatic interaction detection (CHAID) tree-growing method was used for DTAs carried out in the present study. This method can be used for both continuous and categorical data. The data in the present study were categorical and so at each stage the Chi-square test was used as the splitting criteria and Bonferroni adjustments were applied to p-values to account for multiple testing. At each stage CHAID chooses the strongest interaction with the dependant variable. By default, the maximum tree growth levels were set at 3; however, this was increased to 5 to ensure maximum tree growth had been achieved.

The CHAID model consists of the root node (the dependant variable), parent nodes (a node that has nodes below it) and child nodes (a node extending from another node). There are no strict criteria for selecting the sizes for the parent and child nodes, but Collins (2010, p291) suggested that minimum sizes of parent and child nodes should be selected based on the overall sample size. It has been suggested that for small sample sizes the minimum size settings should be 10 for the parent node and 5 for the child node, but for larger sample sizes, minimum sizes may be increased to 20 for the parent node and 10 for the child node (the Measurement group 1999-2005). In the present study, parent nodes were set at 30 and child nodes were set at 15 as a sample size of 30 is generally accepted as being large enough to be defined as a population and so is large enough to analyse (Bailey, 2008, p.97).

There is no method to compute statistical power for DTA but Ridgman's (1975, p. 44) '15 degrees of freedom' rule, which states that any statistical test with greater than 15 degrees of freedom has sufficient statistical power, was applied to ensure all DTAs had sufficient statistical power.

2.9 Research ethics approval

This study was approved by Aston University's Ethical Committee and the study was performed in accordance to the tenets of the Declaration of Helsinki (see appendix 1).

2.10 Study Sample

Voluntary informed consent was obtained from all students before any data was used in this study. The following inclusion criteria were used:

- Student had given consent for their data to be analysed
- Student had completed an ILS;
- Student had completed the questionnaire for perceived mastery;
- Student had completed the MCQ examination to determine actual mastery;

All four criteria had to be achieved for inclusion in the study.

A total of 195 students met all of the inclusion criteria and gave informed consent for their data to be analysed. This consisted of 102 students from the didactic lecture cohort and 93 students from the virtual patient cohort. Females represented 61% of the didactic lecture cohort and 69% of the virtual patient cohort. There were no statistically significant differences in the sex mix for both cohorts (chi = 1.37, df = 1, p = 2.41, power = 99%). Table 2.1 shows the distribution of academic ability in both cohorts. This differed between the two cohorts (chi = 8.27, df =2, p = 0.02), however, the DTAs used to analyse the data include academic ability as an influencing variable and so the noted difference in the proportion of academic grades in both cohorts was accounted for in the analyses. The spread of learning styles in the study sample are described in chapter 3.

Table 2.1 Spread of academic ability of 102 consenting students from the didactic lecture cohort and 93 consenting students in the virtual patient cohort.

Cohort	First class (%)	Upper second class (%)	Lower second class (%)	Third (%)	Fail (%)
Didactic lecture cohort	39	41	20	0	0
Virtual patient cohort	56	37	7	0	0

2.11 Summary

The purpose, methodology and study sample has been described in this chapter. The measurement and classification of learning style preferences is discussed in chapter 3.

Chapter 3 – Determination of learning styles

3.1 Introduction

Chapter 3 introduces and explores the concept of learning styles. The literature is critically reviewed to identify the most appropriate instrument to measure learning styles. Learning styles of the study sample are then classified using this instrument.

3.2 Learning styles

Each individual student processes and learns new information in different ways (Felder & Silverman, 1988). For example, some individuals prefer to learn by active participation, whereas others prefer to sit back and reflect on ideas or theories. Some prefer to make written notes, whereas others prefer to use diagrams or pictures. These differences are referred to as learning styles (Felder & Spurlin, 2005).

It has been stated that all educators should have a thorough understanding of the spread of learning styles that exist in their teaching environment in order to provide the most efficient learning experience (Coffield et al., 2004). Whether teaching methods should be adapted to complement the learning styles of the students is a controversial issue. There is much disagreement in the literature. While the majority of researchers support the logical assumption that matched learning and teaching styles will lead to more efficient learning (Dunn & Dunn, 1978; Felder & Silverman, 1988; Hayes & Allinson, 1993; Allinson & Hayes, 1996; Sprenger, 2003), there are others who oppose this view (Grasha, 1972; Felder, 1993; Lawrence, 1997). There are two arguments against the matching approach. Firstly, every learning environment is likely to be made up of students with a variety of learning styles. Therefore, any attempt to match the learning style of one group will inevitably disadvantage learners of another style (Felder, 1993). Felder and Silverman (1988) showed that students whose learning styles are mismatched tend to have greater difficulty with the learning task. Secondly, learners will all have to encounter situations outside their comfort zone at some point in their lives. Therefore, mismatched learning and teaching styles will allow them to develop the ability to cope with such situations (Hayes & Allinson, 1996). It has also been argued that deliberate mismatching of teaching and learning styles can, in fact, lead to improved learning outcomes. Grasha (1972) claimed that optimal learning occurs when there is a certain amount of challenge in the learning situation. Grasha also claimed that learners will become bored and will not be able to tolerate learning and teaching styles that are always matched.

Desmeddt and Valke (2003) found that it is not essential for educators to match teaching methods to learning styles, as merely understanding or recognising the existence of different learning styles in a group can be beneficial to the educational outcome. Ramsden (1983) and Sadler-Smith (2001) have also highlighted that students who are aware of different learning styles find it easier to adapt to different learning situations.

Optometry schools typically enrol a mixture of student types; males and females, home (entrants from the UK or European Union), international (entrants from outside the UK or European Union), graduate (entrants with a bachelor's degree or higher) and non-graduate (entrants without a bachelor's degree or higher) students. It was likely that these student groups may also vary in their learning styles. Therefore, it was important to be able to classify learning styles for the study sample to later investigate whether the two different teaching methods, didactic lecture and use of the VP (see chapter 2) impacted one group more than another.

3.3 Literature review to select instrument to measure learning styles

A literature search was conducted using Google Scholar, ERIC Database, International Education Research Database, British Education Index, Science Direct, PubMed, EPPI-Centre and Review Groups. Keywords and their combination used in the search included 'learning styles', 'learning preferences', 'learning style models', 'thinking styles', 'cognitive styles', 'learning orientations' and 'learning strategies'. Once the key influential learning style models were identified (see Table 3.1), the names of the instrument to measure these learning styles were used in conjunction with the keywords 'validity' and 'reliability'.

In the early stages of this literature review, the sheer volume of research was overwhelming. There have been many contributions to the literature from a variety of disciplines including psychology, education and the commercial sector. This has resulted in a very fragmented field of research, with many subdivisions.

An initial search of learning style theories and models revealed in excess of 80 different models or concepts. These are summarised in chronological order in appendix 2. Some of these models were original ideas, which had a strong theoretical and empirical base. Others were developments of existing ideas with further improvements. A small number of these models proposed new labels to existing measures and claimed it to be a new model. A number of models designed to analyse the learning styles of students with English as a second language were also identified. These have not been included in this review as most were very similar to models that had previously been identified (i.e. based on the same theory), but had minor alterations to compensate for language.

3.3.1 Influential learning style models

It was impractical to fully investigate every model identified with regard to its theory, instrument of measure, validity and reliability. Instead, the most influential models in this discipline were identified using the criteria outlined below:

- It was based upon a widely accepted theory that logically identified the learning styles described in the model;
- It was widely quoted in the research literature;
- It had an instrument (questionnaire or test) that could classify or measure the learning styles proposed by the model;
- This instrument was still available;
- Research had been conducted on the validity and reliability of the instrument.

These models are summarised in Table 3.1

To assess whether the model was widely quoted in the research literature, a citation search was conducted in the Google Scholar database. This database was selected as it allowed articles to be picked up across a number of subject areas without it being restricted to certain journals or disciplines. For example, the Web of Science database is restricted to articles in the field of science and it does not cite any articles that are published online in open-access journals. Citation analysis is now a well-recognised method for analysing the importance and popularity of an article or researcher (Harter, 1996). However, a high citation rate did not always mean that the model was accepted by all those who cited it.

Some of the cited articles received negative reviews that disputed the validity or reliability of the model. Nevertheless, the citation number still gave a greater indication of the impact the model had on the discipline as a whole and whether it had led to further research by others. The citation search was conducted on the first article or book that introduced the model. The number of citations were ranked and then the top 5 models that met all the other inclusion criteria were identified as being the most influential.

There is a considerable amount of research that has been published evaluating the reliability and validity of these influential learning style instruments by both the authors of the instruments and independent researchers. Each study has investigated different types of reliability and validity and the methods used have varied greatly.

Table 3.1 Literature search results identifying most influential learning style models

Author (Date model first introduced)	Theory it was based on	Number of citations	Instrument	Instrument still available	Research on validity and reliability of instrument available
Kolb (1976)	Kolb's Experiential Learning	8068	Learning Styles Inventory	✔	✔
Myers-Briggs (1962)	Jung's Theory of Psychological Types	1317	Myers-Briggs Type Indicator	✔	✔
Felder & Silverman (1988)	Cognitive Styles	854	Index of Learning Styles	✔	✔
Honey & Mumford (1982)	Kolb's Experiential Learning	839	Learning Styles Questionnaire	✔	✔
Gregorc (1977)	Kolb's Experiential Learning	125	Mind Style Delineator	✔	✔

3.3.1.1 Reliability

Two methods of assessing the reliability of learning style instruments appeared in the research literature. The first method involved the investigation of the internal consistency of the items that form the instrument. This looked at whether all the items were measuring the same or two different attributes. Internal consistency can be rated using Cronbach's alpha coefficient, which can range form 0 to 1. Ideally, any test should measure different aspects of the same trait and not different traits. There are various rules of thumb in the literature about how to interpret the alpha coefficients. George and Mallery (2003) have suggested the following rules of thumb:

- > 0.90 = Excellent
- 0.80 to 0.89 = Good
- 0.70 to 0.79 = Acceptable
- 0.60 to 0.69 = Questionable
- 0.50 to 0.59 = Poor
- < 0.50 = Unacceptable

Tuckman (1999), however, suggested that different criteria are required depending on the attributes being assessed by the instrument. He suggested that a coefficient of 0.75 or greater is acceptable for instruments that measure achievement and 0.5 or greater is acceptable for instruments that measure attitude.

The desired level identified by the majority of researchers is between 0.70 and 0.90 (Nunnally & Bernstein, 1994). If the coefficient is too low (less than 0.70) then it indicates that different traits are being measured. If it is too high (over 0.90) then it indicates that there are redundant items.

The second method of assessing reliability was the test-retest reliability. This is a measure of the consistency of the outputs of the instrument when they are repeated. It is rated using Pearson's correlation coefficient, which can range from -1 to +1. A correlation coefficient of -1 represents a perfect negative relationship, a correlation coefficient of 0 represents no relationship and a correlation coefficient of +1 represents a perfect positive relationship. It is generally accepted that a coefficient greater than 0.70 is acceptable and 0.80 is good.

3.3.1.2 Validity

Validity of the instruments used to measure learning styles was assessed in a number of different ways including construct validity, face validity and criterion-related validity. Construct validity refers to the ability of an instrument to measure the theory and concept of the construct it claims to measure. Put more simply it is a measure of how far test scores can be interpreted as measuring only what they are supposed to measure. Face validity refers to whether it looks like the instrument is measuring what it is supposed to. It is based on common sense and a consensus of opinion amongst professionals. Criterion-related validity is the relationship between the instrument and another well-established instrument. It is comprised of two types of validity – predictive validity and concurrent validity. Predictive validity refers to an instrument's ability to correctly predict an expected outcome or criterion that will occur in the future. Concurrent validity refers to an instrument's ability to correctly identify an outcome or criterion that is already known or is tested at the same time.

Various studies have found conflicting results regarding the reliability and validity of the instruments used to measure learning styles. These are discussed in more detail for each of the five influential learning style models.

3.3.1.3 Kolb's Learning Style Inventory (LSI)

3.3.1.3.1 Brief overview of theory and construct of LSI

David Kolb developed his model of learning styles based on his own experiential learning theory (ELT), which was first identified in 1976 and published in 1984. His work was inspired by Kurt Lewin (1951). Kolb described learning as "the process whereby knowledge is created through the transformation of experience" (Kolb, 1984). Figure 3.1 summarises his model of learning.

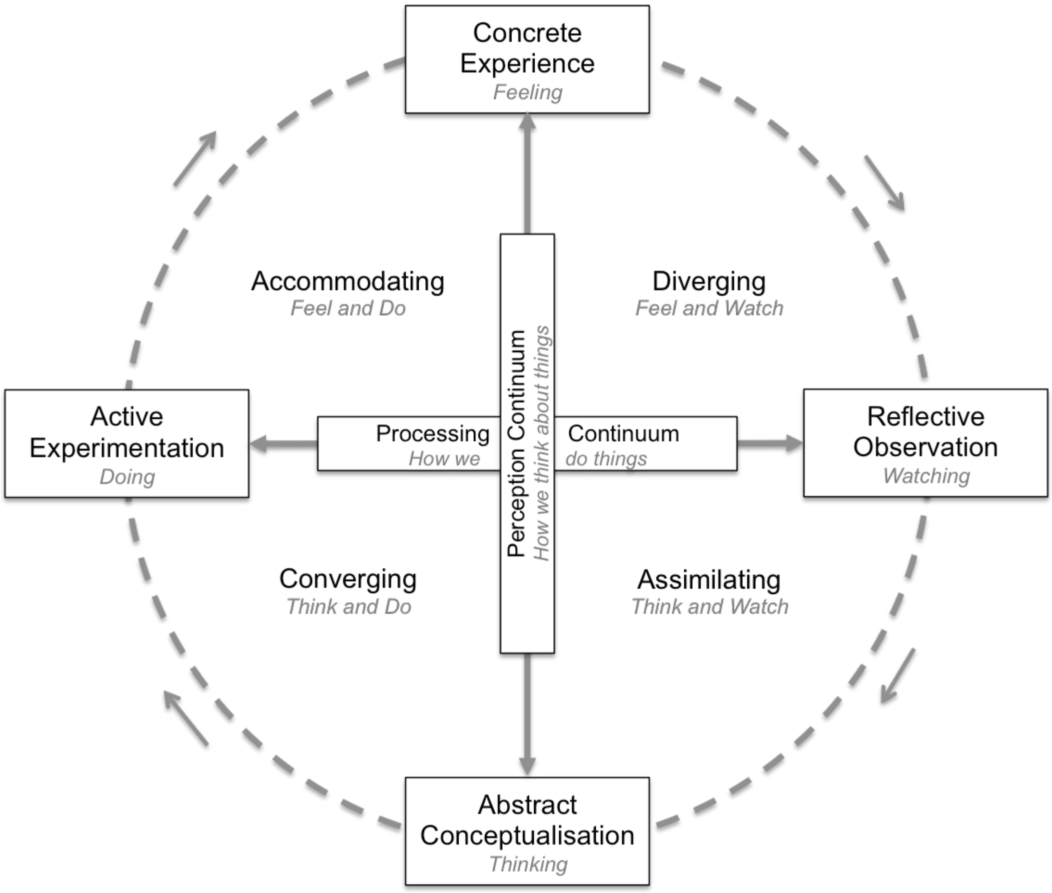

Figure 3.1 Kolb's learning style model

This model represents learning as a continuous cycle with four distinct stages:

- **Concrete experience** – in this stage the learner relies more on feelings and instincts rather than a systematic approach to problem solving. The learner is open minded and adaptable to change.
- **Reflective observation** – in this stage the learner reflects upon and understands ideas from different perspectives.
- **Abstract conceptualisation** – in this stage the learner uses a more logical approach to problem solving by referring to theories, logic and ideas. The learner approaches the task in a systematic manner.
- **Active experimentation** – in this stage the learner takes a practical approach. The learner gains more from active participation over thinking, reading and studying.

Kolb states that learning can begin at any of the four stages, but all four stages must be incorporated for efficient learning to occur. These four stages are shown to be on two continuums that Kolb calls 'grasping experience' (processing – how we prefer to approach a task) and 'transforming experience' (perception – how we think or feel about a task). Both ends of the continuum cannot occur at the same time. Kolb's model is based on the theory that we have an inherent preference for one end of each of the two continuums. Pairing these two gives rise to the four learning styles:

- **Accommodating** – Accommodators prefer concrete experience (feeling) and active experimentation (doing). They tend to be more hands-on people who prefer to learn by active participation rather than reading or studying. They rely on intuition and act on 'gut instinct' rather than logical analysis;
- **Diverging** – Divergers prefer concrete experience (feeling) and reflective observation (watching). They are able to look at things from different perspectives and come up with new ideas;
- **Assimilating** – Assimilators prefer abstract conceptualisation (thinking) and reflective observation (watching). They prefer a logical approach to problem solving and tend to be capable of understanding a wide range of information in a clear and logical format;
- **Converging** – Convergers prefer abstract conceptualisation (thinking) and active experimentation (doing). They tend to be good at problem solving and finding practical solutions.

Kolb explains that these learning styles are not fixed personality traits but are, nevertheless, relatively stable patterns of behaviour. These learning styles can be measured using the Learning Styles Inventory (LSI). There are three versions of this available. The latest version was developed in 2005. The LSI was designed specifically for management students.

3.3.1.3.2 Reliability of LSI

Dispute over the reliability of all three versions of this instrument still continues. The initial version of the LSI was heavily criticised by many researchers for its poor internal consistency and test-retest validity (Geller, 1979; Wilson, 1986). Kolb went on to develop a revised version of the LSI but there is no consensus among researchers about whether there was any improvement in reliability. While some researchers have confirmed improved reliability (Veres et al., 1991; Metallidou & Platsidou, 2008), others claim there was no improvement (Wilson, 1986; Veres et al., 1987).

Henson and Hwang (2002) looked at a number of previous studies that had evaluated both the reliability and validity of all three versions of Kolb's LSI. They concluded that although the scores were higher for the second version of the LSI with regard to its internal consistency, the test-retest scores were lower. Overall, the test-retest and internal consistency scores fluctuated considerably for all three versions of the instrument. Kolb (2000) explained that poor test-retest reliability arose because the learning styles he had described were flexible and could vary greatly depending on the learning situation. He explained that people naturally adopt different learning styles in different learning environments.

More alarmingly, Ruble and Stout's (1992) study revealed that 16% of their sample were classified as having the opposite learning style when the test was repeated. A separate study by Loo (1997) also found similar results where 13% of his sample changed to the opposite style when the test was repeated approximately ten weeks later. However, in this study the students were all given a 30 minute briefing about Kolb's ELT and how the LSI could be useful in improving their learning after the first test was completed. They were also given an eight page handout so they could refer to the learning styles material over the next ten weeks. It is possible that this intervention may have been responsible for the poor test-retest reliability and so casts serious doubts on the results of this study.

3.3.1.3.3 Validity of LSI

There is just as much disagreement in the literature with regards to the validity of these three instruments. The construct validity has been disputed by a number of researchers. A study by DeCantis and Kirton (1996) cast serious doubt on the existence of the two bipolar continuums proposed by Kolb. Other studies go further to claim that they have found opposite bipolar continuums (Cornwell et al., 1991).

A number of studies have looked at the comparison of the LSI to Honey and Mumford's LSQ (see section 3.3.1.6). The results are summarised below in Table 3.2.

Table 3.2 Correlations between Kolb's LSI and Honey and Mumford's LSQ

Study *(Version of LSI)*	**Sample Size**	**Activist (LSQ) - Concrete experience (LSI)**	**Reflector (LSQ) - Reflective observation (LSI)**	**Theorist (LSQ) - Abstract conceptualisation (LSI)**	**Pragmatist (LSQ) - Active experimentation (LSI)**
Sims et al. (1989) *(LSI v2)*	279	0.22	0.28	0.11	0.01
Bokoros et al. (1992) *(LSI v1)*	44	0.23	0.09	0.36	0.38
Bokoros et al. (1992) *(LSI v2)*	44	0.43	0.14	0.23	0.38
Honey and Mumford (1982) *(LSI v1)*	29	0.23	0.73	0.54	0.68

The only study that found good correlations between the two instruments was that by Honey and Mumford (1982). Even so, the correlation between Kolb's concrete experience and Honey and Mumford's activist dimension was poor. The larger scale study by Sims et al. (1989) revealed very weak correlations across all four learning style dimensions.

The face validity of the LSI is generally accepted in literature (Henson & Hwang, 2002) but this alone was not enough to justify the use of this instrument.

3.3.1.4 Myers-Briggs Type Indicator (MBTI)

3.3.1.4.1 Brief overview of theory and construct of MBTI

Katherine Cook Briggs and Isabel Briggs Myers developed the MBTI based directly on Carl Jung's theory of psychological types (Jung, 1921). Jung's theory was written in such a style that it was difficult for most people to understand and utilise in every day life. Therefore, Myers and Briggs went on to develop this theory further with the aim of developing an instrument to easily classify the 16 psychological types proposed by Jung. The MBTI was developed for general use by the public.

The MBTI consists of 88 forced choice questions. It classifies behavioural preferences across four dichotomous dimensions and then groups these four dimensions to identify an overall personality type. The four dimensions are:

- **Extroversion (E) or Introversion (I)** – this is based on how an individual's energy is focused. Extroverts are talkative, outgoing individuals who prefer being in a fast paced environment. They tend to work out ideas with others, think out aloud and enjoy being the centre of attention. Introverts, on the other hand, are reserved and private individuals who prefer a slower pace with time for contemplation. They tend to think through ideas in their head and prefer to stand back and observe;

- **Sensing (S) or Intuition (I)** – this is based on how an individual prefers to take in new information. Sensing individuals tend to focus on reality and pay attention to concrete facts. They prefer ideas that have practical applications and to describe things in a simple way. Intuitive individuals, on the other hand, tend to imagine the possibilities of how things could be and look at the bigger picture. They prefer to describe things in a poetic or figurative way;
- **Thinking (T) or Feeling (F)** – this is based on how individuals prefer to make decisions. Thinkers tend to use logical reasoning based on value, justice and fairness. Feelers, on the other hand, base decisions on personal values and by considering how the actions affect others;
- **Judging (J) or Perceiving (P)** – this is based on how individuals prefer to live their outer life. Judging individuals prefer to have matters settled and think rules and deadlines should be adhered to. They prefer to have detailed step-by-step instructions and make detailed plans. Perceiving individuals, on the other hand, prefer to leave their options open and see rules and deadlines as being flexible. They prefer to improvise and tend to be spontaneous.

These four behavioural preferences are grouped to create a personality type. There are a total of 16 possible personality types, which are summarised in Figure 3.2.

ISTJ Quietly systematic, factual, organised, logical, detailed, conscientious, analytical, responsible, pragmatic, critical, conservative, decisive, stable, concrete, efficient	**ISFJ** Quietly warm, factual, sympathetic, detailed, dependable, organised, thorough, conscientious, systematic, conservative, realistic, caring practical, stable, helpful	**INFJ** Vision and meaning oriented, quietly intense, insightful, creative, sensitive, seeks harmony and growth, serious, loves language and symbols, preserving, inspiring	**INTJ** Vision oriented, quietly innovative, insightful, conceptual, logical, sees understanding, critical, decisive, independent, determined, pursues competence and improvement
ISTP Logical, quietly analytical, practical, adaptable, curious, cool, observer, problem-solver, exact, realistic, troubleshooter, hands-on, variety, adventurous, independent	**ISFP** Gentle, quietly caring, compassionate, adaptable, modest, aesthetic, idealistic, observant, loyal, helpful, realistic, patient with details, spontaneous, joy in action	**INFP** Deep-felt valuing, quietly caring, compassionate, pursues meaning and harmony, creative, idealistic, empathic helpers, inquisitive, enjoys ideas and language, independent, adaptable	**INTP** Logical, conceptual, analytical, objective, detached, critical, ingenious, complex, intellectually curious, loves ideas, pursues understanding, questioning, adaptable, independent
ESTP Excitement seeking, active, pragmatic, direct, easy-going, observant, concrete, realistic, adaptable, efficient, analytical, trouble-shooter, spontaneous, adventurous, experimental	**ESFP** Energetic, sociable, practical, friendly, caring, expressive open, enthusiastic, excitement seeking, spontaneous, resourceful, adaptable, observant, hands-on, generous, fun-loving	**ENFP** Enthusiastic, imaginative, energetic, creative, warm, future-oriented, individualistic, insightful, caring, optimistic, possibility focussed, open, novelty seeking, spontaneous, playful	**ENTP** Energetic, inventive, enthusiastic, abstract, logical, theoretical, analytical, complex, ingenious, verbal, novelty seeking, change oriented, global, independent, adaptable
ESTJ Active organiser, logical, assertive, fact minded, decisive, practical, results oriented, analytical, systematic, concrete, critical, responsible, take charge, common sense	**ESFJ** Actively sociable, warm, harmoniser, caring, enthusiastic, empathetic, people-oriented, practical, responsible, concrete, orderly, conscientious, co-operative, appreciative, loyal	**ENFJ** Actively sociable, enthusiastic, harmoniser, expressive, warm, idealistic, empathic, possibility-oriented, insightful, co-operative, imaginative, conscientious, tactful, appreciative	**ENTJ** Driving organiser, planner, vision focussed, decisive, initiating, conceptual, strategic, systematic, assertive, critical, logical, organised, pursues improvement and achievement

Figure 3.2 The 16 MBTI personality types and their descriptors.

3.3.1.4.2 Reliability of MBTI

Both the authors and independent studies have demonstrated high levels of internal consistency. Myers and McCaulley (1989) had a sample of over 32,000 and reported alpha coefficients of 0.79 for extrovert-introvert, 0.74 for thinking-feeling, 0.84 for sensing-intuition and 0.82 for judging-perceiving. Fleenor (2001) found alpha coefficients of mostly greater than 0.90.

High levels of test-retest reliability have also been established (McCaulley, 1981; Myers & McCaulley, 1985; Fleenor, 2001). However, these studies have looked at the four dimensions individually. The authors state that the instrument is designed to identify an individual's whole type and not individual scales. The stability of the overall personality type is less impressive; only 65% (Myers & McCaulley, 1985).

3.3.1.4.3 Validity of MBTI

The face validity of the MBTI is generally accepted by researchers (Coffield et al., 2004). Similarities have been found between this and Gregorc's MSD (Drummond & Stoddard, 1992; Harasym et al., 1996) and Felder and Silverman's ILS (Rosati & Felder, 1995). Due to the complexity of this model, it has only been possible to make comparisons of some of the four dimensions individually, rather than the whole type. This means that these studies only provide limited support for the criterion-related validity of this instrument.

The construct of the MBTI has also been questioned. Many researchers have claimed that the four dimensions are actually not bi-polar (Hicks, 1984; McCrae & Costa, 1989). A study by Girelli and Stake (1993) demonstrated this with a sample of 165 of which 25% scored highly on both ends of the scale for three of the four dimensions (introvert-extrovert, sensing-intuition and thinking-feeling).

3.3.1.5 Felder and Soloman's index of learning styles (ILS)

3.3.1.5.1 Brief overview of theory and construct of ILS

The ILS was developed by Felder and Soloman (2003) and was based on the model of learning proposed by Felder and Silverman (1988). This model was developed to understand the learning style differences of engineering students in a higher education setting and to allow for the development of a teaching approach that would address the learning needs of all students.

This model describes differences in learning style preferences across four continuous dimensions, which are:

- Active-Reflective - active learners prefer immediate participation in learning tasks and learn best in groups. Reflective learners, on the other hand, prefer to stand back and think ideas through alone;
- Sensing-Intuitive - sensing learners are practical and prefer to work with facts, whereas intuitive learners tend to be more innovative and enjoy understanding theories;
- Visual-Verbal - visual learners learn best through pictures, diagrams or charts, while verbal learners prefer words; either written or verbal;
- Sequential-Global - sequential learners absorb information in small logical steps, whereas global learners try to tackle the whole problem at once taking large, somewhat random steps, before putting everything together.

Preferences across these four dimensions are identified and then pooled together to form 16 possible overall learning style profiles. Felder (1993) explained that each of these learning style dimensions was a continuum and not designed to group individuals into discrete categories. Therefore, an individual may have a strong, moderate or no preference along the four dimensions and these preferences may change with time or in different learning environments.

This model proposed that teaching styles that are adapted to include both poles of the four learning style dimensions would be close to providing the optimal learning environment for most students in any given class (Felder and Silverman, 1998, p.675).

These learning style preferences can be identified using the ILS, which is a 44-item questionnaire.

3.3.1.5.2 Reliability of ILS

Both the authors and several independent studies have investigated the internal consistency of the ILS (Zwanenberg et al., 2000; Livesay et al., 2002; Zywno, 2003; Felder & Spurlin, 2005; Litzinger et al., 2005). The results of these studies are summarised in Table 3.3.

Table 3.3 Cronbach alpha coefficients showing internal consistency of the ILS

Study	Sample Size	Active-Reflective	Sensing-Intuitive	Visual-Verbal	Sequential-Global
Litzinger et al. (2005)	572	0.60	0.77	0.74	0.56
Livesay et al. (2002)	242	0.56	0.72	0.60	0.54
Felder and Spurlin (2005)	584	0.62	0.76	0.69	0.55
Zwanenberg et al. (2000)	284	0.51	0.65	0.56	0.41
Zywno (2003)	577	0.60	0.70	0.63	0.53

While Zwanenberg et al. (2000) found poor to moderate internal consistency across the four learning style dimensions, the other studies had more consistent findings showing moderate to strong internal consistency.

Two independent studies have shown moderate to strong levels of test-retest reliability. These results are summarised in Table 3.4

Table 3.4 Pearson's correlation coefficients showing test-retest reliability of the ILS

Study	Sample size	Active	Sensing	Visual	Sequential
Zywno (2003)	124	0.68	0.68	0.51	0.60
Livesay et al. (2002)	24	0.73	0.75	0.68	0.60

Both studies had a similar time interval; eight months for the study of Zywno (2003) and seven months for that of Livesay et al. (2002). However, the smaller sample size for the second study makes it more difficult to interpret the results due to reduced statistical power.

3.3.1.5.3 Validity of ILS

Factor analysis by Litzinger et al. (2005) established very good construct validity. Rosati and Felder (1995) established criterion related validity by comparing it to the Myer-Briggs MBTI. However, this should not be taken at face value as the construct of the MBTI itself has been questioned (see section 3.3.1.4.3).

3.3.1.6 Honey and Mumford's learning styles questionnaire (LSQ)

3.3.1.6.1 Brief overview of theory and construct of LSQ

Honey and Mumford's model of learning styles was based on Kolb's original model and his ELT. It was developed in the 1970s and published in 1982. The four learning styles proposed by Honey and Mumford are a link between the four learning stages identified by Kolb (see Figure 3.1). Honey (2002, p.116) explained that the learning cycle is "flexible and helps people to see how they can enter the cycle at any stage with information to ponder, with a hypothesis to test, with a plan in search of an opportunity to implement it, with a technique to experiment with and see how well it works out in practice".

The four learning styles identified by this model are:

- **Reflector** – Reflectors are imaginative learners who prefer to take the their time, observe, stand back, think about the problem and gather data before taking any action;
- **Theorist** – Theorists are analytic learners who prefer to think through a problem in a logical manner first. They tend to be independent learners who like to understand the theory behind actions;
- **Pragmatist** – Pragmatists tend to be common sense learners. They prefer to take the time to think about how to apply learning into practice in the real world;
- **Activist** – Activists are dynamic learners who tend to be open-minded and prefer immediate participation or experience of the learning task. They tend to be spontaneous, impatient and prefer to use the trial and error method with a very fast paced learning approach.

These learning styles can be measured using the LSQ, which is an 80 item self-report inventory. This was designed and developed specifically for the industry and management sectors.

3.3.1.6.2 Reliability of LSQ

Many independent studies have verified that the LSQ has moderate levels of internal consistency (Sims et al., 1989; Tepper et al., 1993; DeCantis & Kirton, 1996; Duff & Duffy, 2002), (see Table 3.5). All studies have used a large sample size, which gives the results greater statistical power. The only scale that was found to have relatively poor internal consistency was for pragmatists.

Table 3.5 Chronbach's alpha coefficients showing the internal consistency of Honey and Mumford's LSQ.

Study	Sample Size	Activist	Reflector	Theorist	Pragmatist
Sims et al. (1989)	279	0.68	0.68	0.78	0.75
Duff and Duffy (2002)	224	0.68	0.73	0.58	0.52
Tepper et al. (1993)	277	0.75	0.76	0.67	0.52
DeCantis and Kirton (1996)	185	0.76	0.76	0.67	0.64

The test-retest reliability was shown to be very good with a Pearson's correlation coefficient of 0.89 on a sample of 50 with a two week time interval (Honey & Mumford, 2000). It could be argued that this time interval was too short. However, as the instrument is modelled on Kolb's ELT (1984), which states that learning styles vary in different situations, therefore, it could be argued that a shorter time interval would be preferred to assess such a characteristic.

3.3.1.6.3 Validity of LSQ

Construct validity has not been evaluated by the authors, but they do state that the LSQ has good face validity (Honey & Mumford, 2000). Factor analysis by Allinson and Hayes (1990) confirmed construct validity.

Criterion-related validity for this instrument has been demonstrated by comparisons with Kolb's LSI. As already discussed the only study that found good correlations between the two was by Honey and Mumford (1982). These results are flawed in two key areas. Firstly, the sample size was too small to draw these conclusions and, secondly, other studies have already shown that Kolb's LSI was flawed in its own construction. Therefore, it was not appropriate to claim the LSQ was valid by comparing it to Kolb's LSI.

3.3.1.7 Gregorc's Mind Style Delineator (MSD)

3.3.1.7.1 Brief overview of theory and construct of MSD

Anthony Gregorc developed the MSD in 1984. His model was based on the existence of perceptions and focuses on differences in thinking and processing information. It is a modified version of Kolb's model. Gregorc proposed that the processing of information is governed by one of two perceptual qualities (concrete and abstract) and one of two ordering abilities (sequential and random). Combining these leads to four learning styles:

- **Concrete Sequential** – this type of learner prefers logical order, gathering facts, predictability, following orders and working in a step-by-step manner. They tend to work best in a structured environment;
- **Abstract Sequential** – this type of learner tends to prefer verbal communication. They prefer a systematic and logical approach to problem solving;
- **Concrete Random** – this type of learner prefers to experiment and take risks. They use their intuition to find answers and prefer to work independently;
- **Abstract Random** – this type of learner tends to prefer an unstructured learning environment. They tend to have a strong visual preference in learning.

The MSD is a 10-item inventory that was designed for use by adults.

3.3.1.7.2 Reliability of MSD

The test-retest reliability of the MSD has been established as being very good with correlations of 0.85 for concrete sequential, 0.87 for abstract sequential, 0.88 for abstract random and 0.87 for concrete random for a sample of 110 people and with time intervals ranging from 6 hours to 8 weeks (Gregorc, 1982).

Gregorc (1982) also reported very high alpha coefficients for internal consistency ranging from 0.89 to 0.92. However, when measured by independent researchers the alpha coefficients were less impressive. Joniak and Isaksen (1988) reported alpha coefficients ranging from 0.23 to 0.66 and the alpha coefficients found by

O'Brien (1990) were 0.64 for concrete sequential, 0.51 for abstract sequential, 0.61 for abstract random and 0.63 for concrete random.

3.3.1.7.3 Validity of MSD

Gregorc (1982) himself has established the validity of the MSD by a number of different methods. Firstly, to establish construct validity he asked 123 people to rate the descriptions of themselves on a scale of 1 (strongly disagree) to 5 (strongly agree). He found 29% strongly agreed with their descriptions, 57% agreed and 14% were unsure. This method has relied strongly upon self-perception and research has shown that the accuracy of self perception is poor and can be subject to bias (John & Robins, 1994). People tend to overestimate positive self-attributes (Hill et al., 1988; Cummins & Nistico, 2002).

Another method used by Gregorc to establish construct validity was by asking 110 people to respond to 40 other words that were supposedly characteristic of each style. He found moderate correlations with this method but, again, the methodology can be criticised, as there is no concrete evidence that these additional 40 words are 100% characteristic of the learning styles.

Finally, Gregorc established face validity by interviewing 475 people. 89% claimed the descriptive words were accurate but, again, this method relies on self-perception.

The MSD has been compared to other instruments such as the Myers-Briggs Type Indicator. Some similarities between the two have been reported (Drummond & Stoddard, 1992; Harasym et al., 1996) but again these comparisons are of limited value as research has already cast doubts on the construct validity of the MBTI.

3.3.2 Final selection of instrument used to measure learning styles in the present study

This detailed literature review of the five most influential learning style models identified flaws and weaknesses in all five models. The methods used to investigate the validity and reliability of each of these models have also varied significantly. There is presently a need for a large-scale study to evaluate the validity and reliability of all five instruments using the same methodology and the same study sample. However, this was beyond the scope of the current study.

It has been said that the perfect learning style model and instrument is a fantasy (Coffield et al., 2004). Nevertheless, Bostrom et al. (1993, p.119) stated that 'important research cannot always wait for the perfect measure'. Given that the perfect model and instrument did not exist, Felder and Soloman's ILS was selected. This decision was based on the fact that this instrument was designed for engineering students in a higher education setting. This was the closest match to the study sample in the present study. The ILS had been evaluated by authors and independent researchers. Parallels for all four learning style dimensions exist in other learning style models. For example, the Active-Reflective dimension is a component of Kolb's theory, the Sensing-Intuitive dimension is one of the four dimensions used in Jung's theory of psychological types, the Visual-Verbal dimension is analogous to the VARK (Visual, Aural, Read, Kinesthetic) models based on cognitive studies of information processing (Fleming & Mills, 1992) and the Sequential-Global dimension is a component of Gregorc's model. This indicated that the model had a strong theoretical basis. Furthermore, this model was developed with the intention of improving student learning in the higher education setting, which was also the purpose of the present study.

3.3.2.1 Index of Learning Styles (ILS)

The Index of Learning Styles consisted of 44 forced choice questions; 11 per learning style dimension. Each question gave two options ('a' or 'b'), which corresponded to one or the other end of a learning style dimension. Counting the number of 'a' responses for each dimension generated an integer score that ranged from 0 to 11. Using the first dimension of active – reflective as an example, Felder and Spurlin (2005) have defined a score of 0–1 to represent a strong preference for reflective learning, 2–3 a moderate preference for reflective learning, 4–5 a mild preference for reflective learning, 6–7 a mild preference for active learning, 8–9 a moderate preference for active learning and 10–11 a strong preference for active learning. Felder and Soloman (2003) have described a mild preference as being essentially well balanced, therefore this scoring system was simplified so 0–3 was classed as having a preference for one end of the dimension, 4–7 was classed as a balanced learning style preference and 8–

11 was classed as a preference for the other end of the learning style dimension. This scoring system generated three groups for each of the four learning style dimensions. For example, for the active – reflective dimension, the three groups were: active learners, balanced learners and reflective learners. This allowed analysis of each of the four learning style dimensions individually. In order to achieve a broader picture of the distribution of learning style preferences, the four dimensions were grouped to generate a learning style profile. These were generated by using scores of 0–5 to represent a preference for one learning style and 6–11 as a preference for the opposing learning style in the dimension. This approach generated 16 possible learning style profiles.

Although the ILS is a valuable tool in education, it is worth noting that it can only identify learning style preferences. These preferences only suggest behavioural tendencies rather than being concrete predictors of behaviour (Felder & Spurlin, 2005). For example, individuals with an active learning style preference may still, in certain situations, use a reflective learning style. Only a questionnaire investigating each student's self perception of the learning styles adopted in the course could address this issue, but research has shown that the accuracy of self perception is poor and can be subject to bias (John & Robins, 1994).

3.4 Study Sample

Students from both the didactic lecture and virtual patient cohorts completed the ILS. The ILS was introduced at the same point in both academic years. The ILS was made available from academic week 4 until academic week 17 via the Blackboard learning management system used by Aston University. This ensured that the learning styles were assessed at the same time across both cohorts as some studies suggest that learning styles may change with time and experience (Honey and Mumford, 2000 p19, Kolb 2000, Felder, 1993).

In the didactic lecture cohort, 112 students completed the ILS of which 102 (91%) gave informed consent for their data to be used. In the virtual patient cohort 111 students completed the ILS of which 93 (84%) gave informed consent for their data to be used.

Figure 3.3 shows the spread of the integer scores across each of the four learning style dimensions for both cohorts.

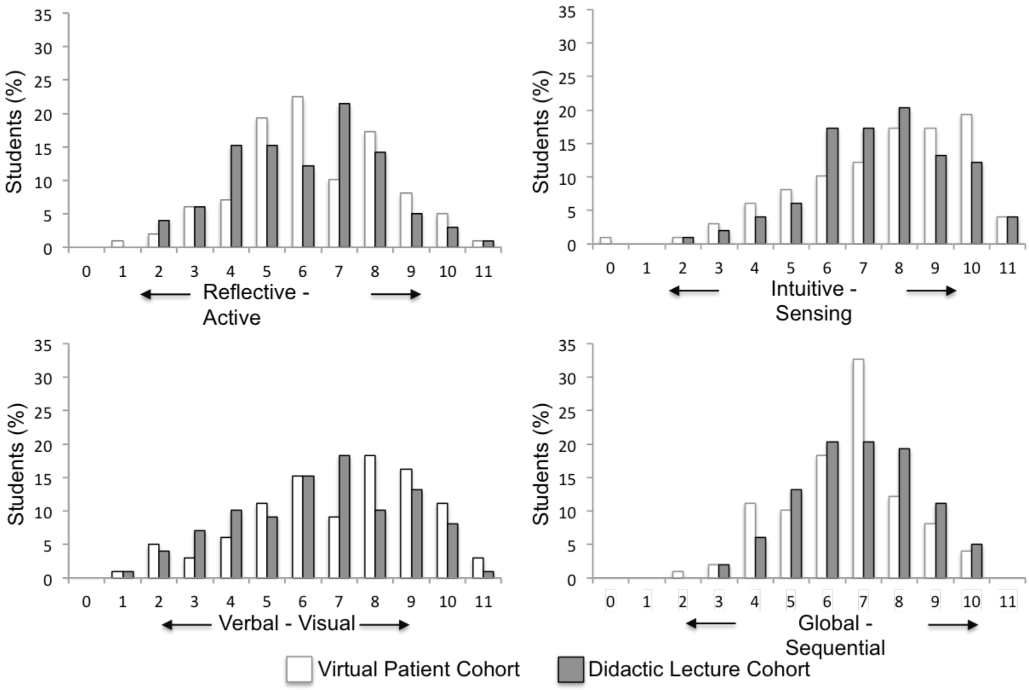

Figure 3.3 Distribution of integer scores for four learning style dimensions measured using the Index of Learning Styles questionnaire in 102 consenting optometry undergraduates from the didactic lecture cohort compared to 93 consenting optometry students from the virtual patient cohort.

Mann Whitney U tests revealed that there was no statistically significant variation in the spread of the four learning styles across the two cohorts (active – reflective, $Z = -0.92$, $p = 0.35$, power = 92%, sensing-intuitive, $Z = -0.82$, $p = 0.41$, power = 92%, visual verbal, $Z = -1.06$, $p = 0.29$, power = 92%, sequential – global, $Z = -0.93$, $p = 0.36$, power = 92%). Therefore, the data for both cohorts were pooled to identify learning style preferences in optometry undergraduates. Pooled data (Table 3.6) showed that the majority of students (41-67% across the four learning styles) had a balanced learning style. However, there was a statistically significant preference for the active (sign test, $p < 0.01$), sensing (sign test, $p < 0.01$), visual (sign test, $p < 0.01$) and sequential (sign test, $p < 0.01$) learning styles.

Table 3.6 Learning style preferences measured using the Index of Learning Styles questionnaire in 195 consenting optometry undergraduates (pooled from the didactic lecture and virtual patient cohorts).

Learning style	**Strong preference for (a) (%)**	**Balanced (%)**	**Strong preference for (b) (%)**
Active (a) – Reflective (b)	28	62	10
Sensing (a) – Intuitive (b)	55	41	4
Visual (a) – Verbal (b)	41	48	11
Sequential (a) – Global (b)	30	67	3

Table 3.7 shows the spread of learning style profiles in optometry undergraduates at Aston University. Again, this shows that a large proportion of students (27.3%) had an active-sensing-visual-sequential learning style profile. This is consistent with the findings of other studies examining the learning style profiles of students studying health science programmes (Brown et al., 2009). Brown et al. (2009) investigated the learning style preferences of students studying a variety of health science programmes including occupational therapy, physiotherapy, paramedicine, social work, nutrition & dietetics, pharmacy, radiation therapy, radiography, nursing and midwifery. Although Brown et al. (2009) have not commented on this, it seemed intuitively obvious to the author that whether a student chooses to be a medical doctor, dentist, nurse or optometrist, the core skills are the same; they all involve the application of clinical knowledge in the field of health and social sciences in order to make tentative diagnoses and relevant management plans. All are health care professionals regardless of their speciality. Similarities in the learning style profiles of students studying health sciences support Kolb's (1984, p.84) theory which states that people already have an understanding of their learning style preferences and so tend to choose to enter disciplines and careers that are consistent with these.

Table 3.7 Percentage of students with each learning style profile measured using the Index of Learning Styles questionnaire in 195 consenting optometry undergraduates (pooled from the didactic lecture and virtual patient cohorts).

REF, INT, VIS, GLO	REF, INT, VIS, SEQ	REF, INT, VRB, GLO	REF, INT, VRB, SEQ
0.5	3.6	1.0	0.5
REF, SEN, VIS, GLO	REF, SEN, VIS, SEQ	REF, SEN, VRB, GLO	REF, SEN, VRB, SEQ
5.2	19.1	2.1	6.7
ACT, INT, VIS, GLO	ACT, INT, VIS, SEQ	ACT, INT, VRB, GLO	ACT, INT, VRB, SEQ
2.6	4.1	1.5	2.6
ACT, SEN, VIS, GLO	ACT, SEN, VIS, SEQ	ACT, SEN, VRB, GLO	ACT, SEN, VRB, SEQ
8.2	27.3	2.1	12.9

ACT: active; REF: reflective; SEN: sensing; INT: intuitive; VIS: visual; VRB: verbal; SEQ: sequential; GLO: global.

Unlike previous research, however, the present study flags up two additional significant learning style profiles; reflective-sensing-visual-sequential (19.1% of students) and active-sensing-verbal-sequential (12.9% of students). Examination of these two profiles shows that both only differ from the most frequent profile in one aspect of their learning style; so it appears that they are subgroups of the main profile. One of these subgroups is reflective rather than active; that is they prefer to sit back and think ideas through over immediate participation. The other subgroup is verbal rather than visual; that is they prefer words and text over pictures and diagrams. All three profiles together, 59.3% of the students, have in common a preference for sensing and sequential learning; that is, they prefer to work with facts delivered in a logical order. In the author's opinion, this is not at all surprising given the way optometrists carry out their daily tasks. They are focused on arriving at diagnoses from presenting symptoms and signs. This, often confusing, clinical information is most easily handled if it is approached in a logical sequence using factual knowledge about the way symptoms and signs are connected to specific diagnoses. That there should be subgroups which differ subtly in the way they prefer to be taught these clinical facts is also, in the author's opinion, not at all surprising. Having settled on some form of explanation as to why these three profiles might exist, this information should be taken with caution as these profiles only emerged when students with mild preferences were included; mild preferences meaning balanced learners. No groups emerged when students with mild preferences were removed from the analyses. Therefore, it may be more accurate to say that the majority of students are balanced learners and no strong profiles really exist.

3.5 Summary

The research presented in this chapter has shown that the majority of optometry students have balanced learning style preferences with a tendency towards the active-sensing-visual-sequential profile. Given the disagreement in literature regarding matching of teaching to learning styles, both the didactic lecture and the VP (see Chapter 4) were developed to cater for both poles of each of the four learning style dimensions. This adhered to the proposal of Felder and Silverman (1988) that this would lead to optimal learning. Chapter 4 provides a description of how the teaching material for both the didactic lecture and the virtual patient cohort was developed.

Chapter 4 – Development of the teaching material

4.1 Introduction

Recall from chapter 2 that the present study was designed to investigate differences in learning that occurred when the CPD module, which was designed to teach CDM skills, was taught using two different methods; didactic lectures or a VP. It was imperative that both teaching methods imparted the same knowledge to students. The development of this teaching material is described in this chapter. The PBL approach and the problem-orientated eye examination system that formed the basis of both teaching methods are discussed.

4.2 Literature review

The CPD module was designed to develop the CDM skills of second year optometry undergraduate students prior to them examining real patients in their final year. It was designed to pool the clinical knowledge that was already taught across several different modules in the undergraduate program with the intention of encouraging critical thinking. Previous research had already demonstrated that the complex task of CDM is taught more efficiently when the PBL approach is utilised (Morales-Mann & Kaitell, 2001; Wells et al., 2009; Nango & Tanaka, 2010; Carrió et al., 2011; Oja, 2011).

PBL is a student centred approach to teaching and learning (Tang & Sung, 2012) in which students learn by attempting to solve an open ended problem. It is based on the assumption that better learning occurs when students are asked to solve the sorts of problems encountered every day in their chosen careers (Barrows and Tamblyn, 1980, p. 1). It is different from traditional teaching methods in that it relies on encouraging students to use self-directed learning skills (i.e. to source the answer themselves from literature or discussions with colleagues).

Research has shown that PBL provides a highly motivational environment for the acquisition of knowledge (Pedersen, 2003; Hwang & Kim, 2006; Martin et al., 2008) and it is well received by those who take part in it (Chang et al., 2004). Furthermore, PBL can increase students' self-efficacy, in respect that is raises their confidence in their ability to perform future study tasks (Liu et al., 2006).

Given these findings it was important that the teaching material for both cohorts should incorporate the PBL approach to yield the best learning outcomes.

In Optometry, there has been an increasing focus on the differences between a database style eye examination and a problem-orientated eye examination along with the benefits of each style (Eperjesi et al., 2007 p.3, Elliott, 2013, p.14). A database style eye examination involves the completion of a list of questions and tests in a specific order. While this approach lends itself well to ensuring that all clinical skills are practiced, it is not an efficient way to develop CDM skills. A problem-orientated eye examination is better suited to this, as it involves identifying the reason for visit (RFV) first and then tailoring the rest of the examination to efficiently investigate, diagnose and manage the presenting concerns. The use of the SOAP system is encouraged when performing a problem-orientated eye examination (Macalister and Wickham, 2008, Rosenfield et al., 2009, p.423) as it improves CDM time and quality (Elson & Connelly, 1997). The mnemonic SOAP represents:

- SUBJECTIVE – the process of obtaining information from a patient verbally. It includes identifying the patient's chief complaint or RFV and completing a detailed case history including information regarding other symptoms, ocular history, medical history, family ocular history, family medical history and social history.
- OBJECTIVE – these are the results of the optometrist's observations and clinical tests. It is based on selecting relevant tests based on the case history findings. These can include visual acuity, cover test, motility, pupils, external eye examination, internal eye examination, refraction, near add determination, convergence, accommodation, tonometry, visual field assessment and supplementary tests such as pupil dilation, slit lamp examination, fluorescein instillation, tear break up time assessment, Van Herrick test, lid eversion, keratometry, indirect ophthalmoscopy, colour desaturation test and the Amsler test.
- ANALYSIS – this is the process of differential diagnosis based on a combination of the subjective data collated in the case history and the results of the objective tests performed by the optometrist.

- PLAN – this involves deciding on the final management plan and its urgency. This decision is based on clinical knowledge, experience and judgement.

A detailed case history (the subjective element of the SOAP system) is an essential part of the problem-orientated eye examination. Asking pertinent questions allows the optometrist to efficiently tailor the rest of the examination; they can select the most appropriate tests and eliminate any irrelevant ones. The process of differential diagnosis begins in the case history and in some situations a definitive diagnosis can be reached by a thorough case history alone.

A common approach used to obtain the relevant information for each presenting symptom is the LOFTSEA system (Eperjesi et al. 2007, p.3-4). This mnemonic represents:

- LOCATION/LATERALITY – where does the presenting symptom occur? Does it affect one side/eye or both sides/eyes?
- ONSET – when did the presenting symptom begin? Did it start suddenly or gradually?
- FREQUENCY AND OCCURRENCE – how often does the presenting symptom occur? How long does it last?
- SELF-TREATMENT AND ITS EFFICACY – does anything make the presenting symptom go away or reduce in intensity? How well does this work?
- TYPE AND SEVERITY – for example is pain throbbing, sharp or dull?
- EFFECT ON PATIENT – does the presenting symptom affect the patient's ability to do certain things? Has the patient consulted a doctor?
- ASSOCIATED OR SECONDARY SYMPTOMS – are there any other complaints associated with the presenting symptom?

Recall from chapter 2 that the 5 learning outcomes of this study were:

1. The ability to ask the right questions in history and symptoms based upon a patient's presenting complaint;
2. The ability to recognise symptoms of serious eye disease;
3. The ability to select the most appropriate tests needed to look for signs that aid differential diagnosis;
4. The ability to recognise signs of serious eye disease;

5. The ability to decide upon the most appropriate referral urgency for any eye disease detected.

These learning outcomes were centred on performing a problem-orientated eye examination and therefore reflect aspects of the SOAP system. The first objective was based on the subjective element of the SOAP system, the second and fourth objectives were based on the analysis element of the SOAP system, the third objective was based on the objective element of the SOAP system and finally, the fifth objective was based on the plan element of the SOAP system. Consequently, the learning material for both teaching methods was based on the SOAP system and encouraged students to adopt the LOFTSEA mnemonic.

The following sections provide a detailed description of how the learning material for both teaching methods was developed.

4.3 Construction of a master database

As discussed earlier, it was imperative that both teaching methods covered the same clinical material. With this in mind, a master database was constructed which linked demographic risk factors (e.g. age and sex), symptoms and signs to a list of eye conditions and their referral urgencies (see appendix 3). This database was then used as the basis to create the teaching material for both cohorts.

The list of eye conditions fell within 22 specific presentations. These were:

1. Vision loss in a white eye
2. Non-traumatic red eye
3. Diplopia
4. Watery, itchy and gritty eye
5. Floaters
6. Flashes
7. Visual field loss
8. Distorted vision
9. Pain in or around eye
10. Headache
11. Rash around the eye
12. Abnormal eyelid position

13. Red swollen eye
14. Eyelid lump
15. Eyelid spasm
16. Patient cannot close eyelids
17. Bulging eye
18. Spot on the eye
19. Unequal pupils
20. Abnormal optic disc
21. Hypertensive retinopathy
22. Diabetic retinopathy

This database followed the SOAP system, in that it was segmented into a list of subjective items that should be asked, a list of objective signs that should be looked for, a list of eye conditions to be considered during analysis (i.e. the differential diagnosis) and a list of management plans.

A total of 255 cases with various eye conditions and associated plans were included in the master database. This was not a fully comprehensive list of every eye condition but included those most likely to be encountered and those that should be recognised by optometrists. Some eye conditions were duplicated as varying patterns of presentation may warrant different management plans. For example, consider a patient who presented complaining of reduced vision; examination revealed that the reduction in vision was due to a cataract but the visual acuity improved with a change in spectacle prescription. Here the patient would be advised of the presence of a cataract and new spectacles would be prescribed. However, if visual acuity could not be improved with new spectacles, then a routine referral for the cataract removal would be considered.

The content of the master database was initially derived from a number of popular optometry textbooks (Bezan et al., 1999; Singh et al., 2002; Loewenstein & Lee, 2004; Pane & Simcock, 2005; Jackson, 2007). This was followed by an extensive search using Google Scholar, Science Direct and PubMed databases. Keywords and their combination used in the search included the name of each eye condition with the terms “signs”, “symptoms”, “characteristics” and “management”. This allowed the identification of the classic presentations for each eye condition and its associated management plan. The author was aware that atypical disease presentations were possible and that comorbidity could occur. Variations in the presentation of eye conditions were included (e.g. some conditions could be unilateral or bilateral) but the existence of comorbidity was considered beyond the scope of the teaching material, which was designed for inexperienced optometry undergraduates.

4.4 The nine key presentations

Due to the number of academic teaching weeks available each year, some of the 22 specific disease presentations covered in the master database were grouped in 9 key presentations:

1. Vision loss in a white eye,
2. Non-traumatic red eye,
3. Diplopia,
4. Watery, itchy and gritty eyes,
5. Other symptoms (floaters, flashes, visual field loss, distorted vision, pain in and around eye, headaches)
6. External eye signs I (rash around the eye, abnormal eyelid position, red swollen eye, eyelid lump, eyelid spasm, patient cannot close eyelids)
7. External eye signs II (bulging eye, spot on the eye)
8. Anisocoria (unequal pupils),
9. Posterior eye signs (abnormal optic disc, hypertensive retinopathy and diabetic retinopathy).

4.5 Development and delivery of the didactic lectures

Nine traditional didactic lectures, which were based on the information from the master database, were written using Microsoft PowerPoint (2010) (see appendix 4). Each of these lectures highlighted the importance of the SOAP system and the LOFTSEA mnemonic. They were designed to teach the basic rules of CDM and listed the key questions an optometrist should ask, the key signs they should look out for, critical indicators of serious disease and the appropriate management plans for each eye condition.

The development of the didactic lectures involved the following steps:

1. Mark Dunne created, for each of the 9 key presentations, databases linking demographics, symptoms, history, relevant tests and signs with diagnoses and their associated referral urgencies (see appendix 3);

2. Bhavna Pancholi converted these databases into diagnostic flow charts based upon her clinical experience (see appendix 5);
3. Mark Dunne then wrote PowerPoint lectures and accompanying MCQs based upon the diagnostic flow charts (see appendix 4);
4. Bhavna Pancholi double checked the PowerPoint lectures and MCQs, drawing upon her experience as a clinical trainer, prior to incorporating the same steps into the VP.

This 4-step process ensured compatibility between the didactic lectures, the VP and the MCQs.

The 9 lectures for the didactic lecture cohort were delivered in a traditional didactic format. At the end of each lecture, student-directed learning was encouraged in preparation for a tutorial that would take place the following week. These tutorials (see appendix 6) were designed to support the PBL approach and so were student-centred. Each tutorial included three case studies relevant to the key presentation covered in the previous week. Prior to attending each tutorial, students used the lecture notes to diagnose and select the most appropriate referral urgency for the case studies. They submitted their answers via a BlackBoard quiz. The same case studies were then reviewed in the tutorial. This pedagogical approach allowed students to demonstrate the application of the basic rules of CDM and to reflect on what they had learnt.

4.6 Development of the virtual patient

4.6.1 Programming the virtual patient

The VP was a web-based simulator developed using the Visual Basic language in Microsoft Visual Studio (2010). Software development was carried out by the author and took 14 months to complete. The master database was first converted to a .DAT file. The VP software was programmed to read data from this .DAT file to ensure that responses of the VP were identical to the teaching material for the didactic lecture cohort.

4.6.2 Testing and refinement of the of the virtual patient

The first version of the VP underwent vigorous testing by both the author and a focus group of 9 final year optometry undergraduate students as part of their elective study projects. Each focus group member was allocated to 1 of the 9 key presentations and was asked to complete all VP episodes included in that presentation and to look for any inconsistencies in the questions and tests prompted for by the VP and covered in the corresponding didactic lecture. The focus group was also asked to provide recommendations to improve the learning experience. A feedback form ensured consistent feedback for each of the 9 key presentations (see appendix 7). Feedback from the focus group (see appendix 8) was reviewed by the author and led to the development of the final version of the VP. The author also ran all 255 VP scenarios to check for any final inconsistencies or errors.

The most common issue identified during this testing process was inconsistency in differentiating between some VP episodes that could present in different ways, which altered the management plan. For example the management of a retinal detachment with the macula on is different to that of a retinal detachment with the macula off. These inconsistencies were caused by minor programming errors rather than incorrect clinical information. Therefore, additional coding was required to resolve these issues.

4.6.3 Functions of the virtual patient

The VP was programmed to operate in 4 modes; teaching, intermediate, advanced and assessment. The welcome screen allowed the student to select the simulator mode and also gave the option to select the presentation group or the specific eye condition that they wished to examine (Figure 4.1). Each VP episode had randomly generated elements to cover presentation variations and any two episodes were rarely the same.

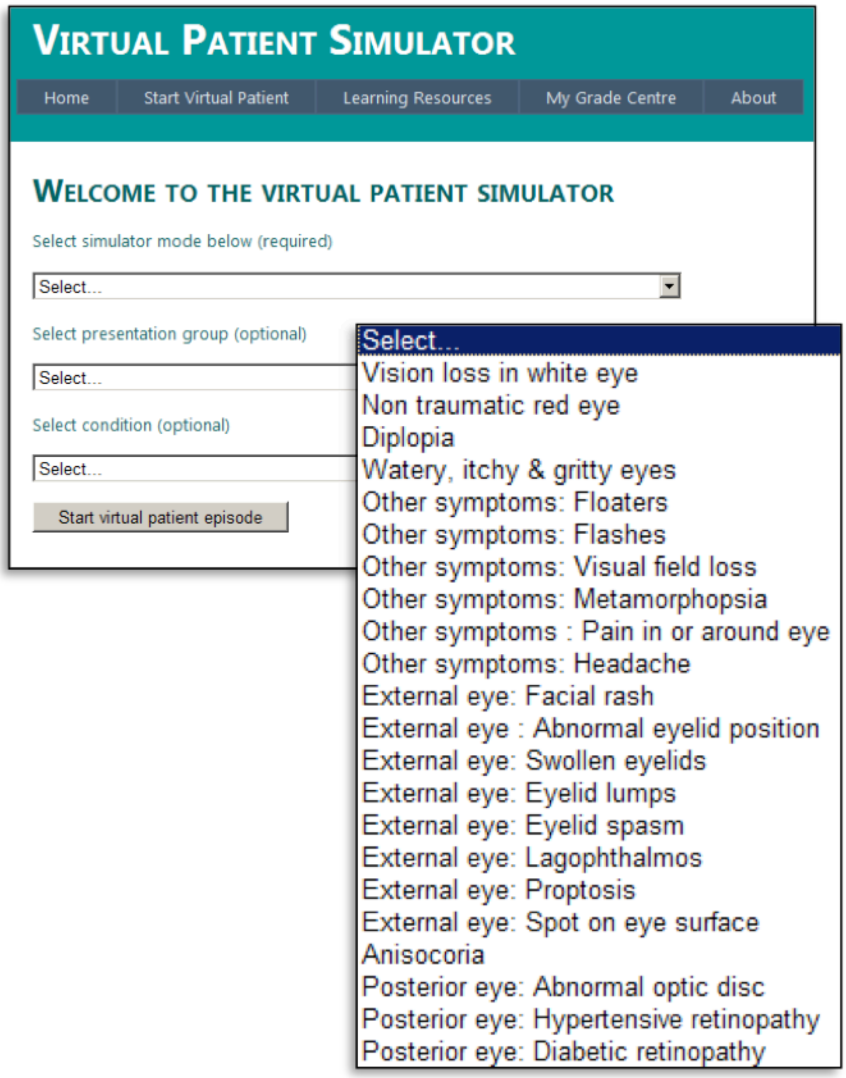

Figure 4.1 Virtual patient welcome screen which allowed selection of the simulator mode, presentation group or eye condition

4.6.3.1 Teaching mode

In teaching mode, a virtual tutor was designed to guide the student through the eye examination, explaining what the next steps should be. It flagged up any critical points that were relevant to that specific disease case or presentation at appropriate points of the eye examination. These critical points were identical to those raised in the didactic lectures (see appendix 4). Live differential diagnosis was shown at each stage. The differential diagnosis list changed with each patient response demonstrating how asking pertinent questions and performing relevant clinical tests could eliminate certain disease cases. In all four modes, the virtual tutor created a tuition log (Figure 4.2) that recorded the whole VP episode showing good and bad clinical decisions made during the eye examination. It also created an eye examination record card (Figure 4.3) demonstrating how the clinical findings should be recorded. Students were able to save these for revision purposes.

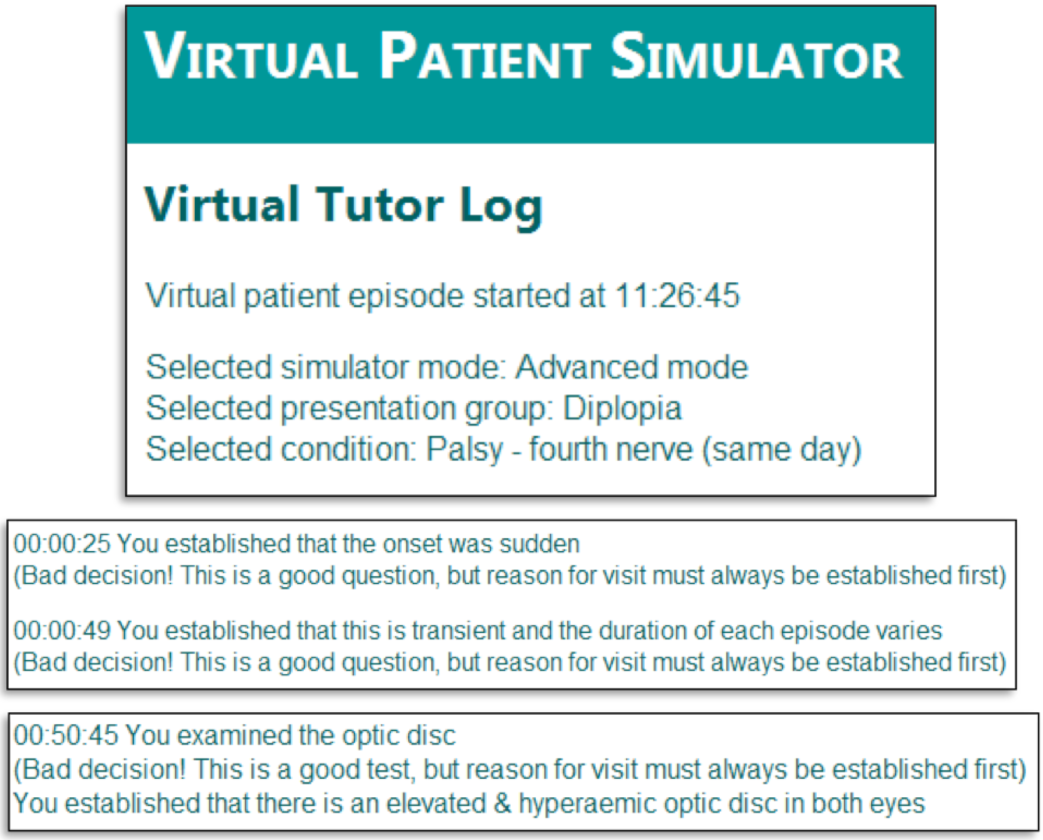

Figure 4.2 Virtual tutor log

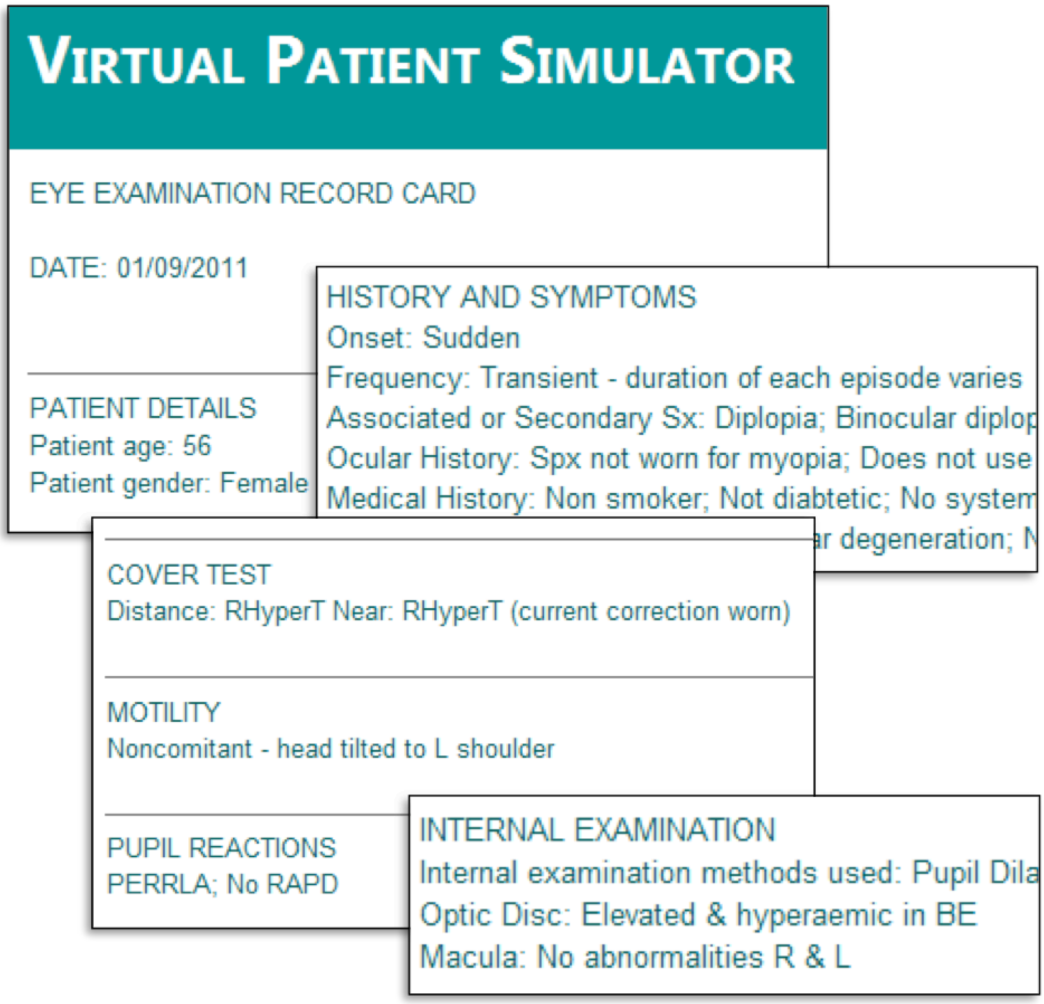

Figure 4.3 Extracts of the virtual patient eye examination record card

4.6.3.2 Intermediate mode

In intermediate mode, the virtual tutor no longer told the student what the next step of the eye examination should be. Instead, it provided immediate feedback for all actions taken by the student indicating whether this was a good or bad clinical decision. The virtual tutor log was available throughout the eye examination for the student to refer to.

4.6.3.3 Advanced mode

In advanced mode, the virtual tutor was programmed to remain silent and the virtual tutor log was hidden. It monitored all actions taken by the student and provided instant detailed feedback at the end of the eye examination (Figure 4.4). This feedback included a percentage mark for their performance which was based on the number of correct questions asked, the number of correct tests performed, the final diagnosis reached, the management plan and procedure penalties (see later). The feedback highlighted the test duration, any relevant items (questions or clinical tests) that were missed, any irrelevant items (questions or clinical tests) that were included and any procedure penalties that were applied. It also highlighted which clinical competencies associated with CDM were achieved.

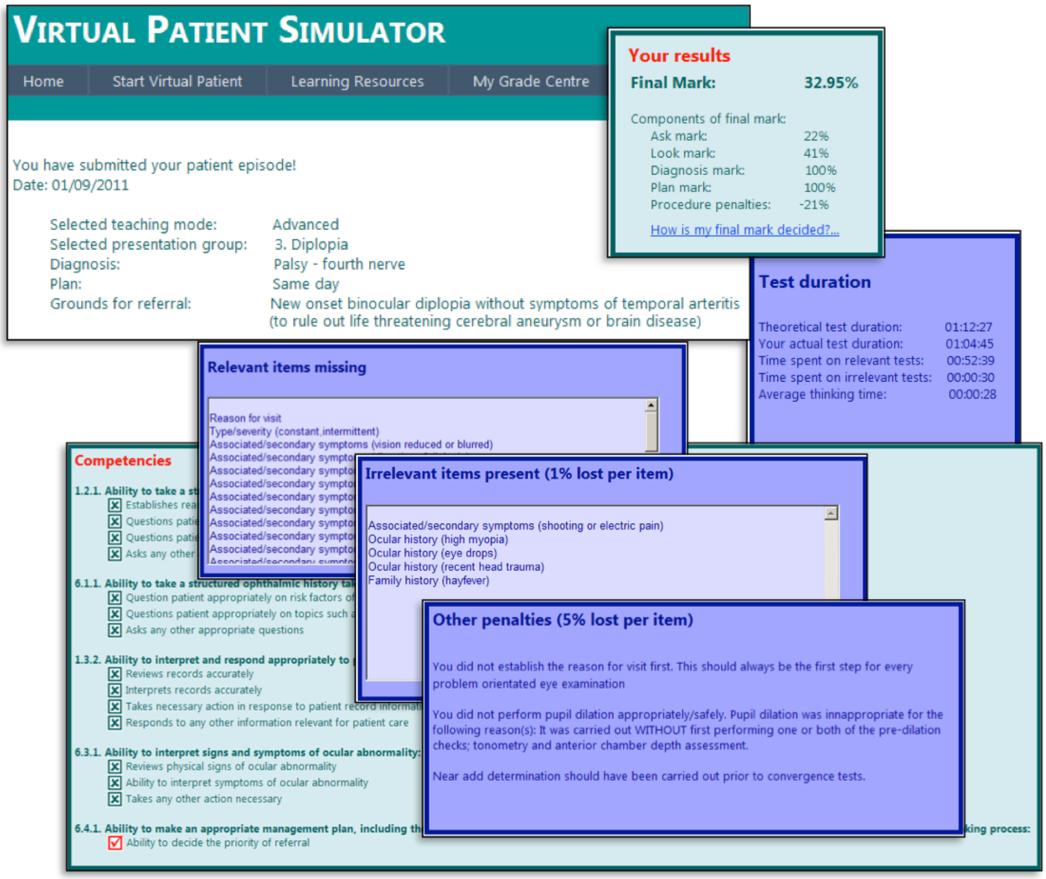

Figure 4.4 Extracts from the virtual patient feedback

Procedure penalties were applied for inappropriate test sequences or for inappropriate examination techniques. These included:

- Not asking RFV first
- Not performing visual acuity before any other clinical test
- Not performing pupil dilation safely which included the following errors:
    - Performing pupil dilation before refraction
    - Not performing pre-dilation checks such as tonometry or anterior chamber depth assessment

    - Performing pupil dilation when known contraindications exist such as intra-ocular pressure over 40mmHg, narrow or closed angles or iritis
    - Not performing post dilation tonometry
- Determining near addition prior to refraction
- Testing convergence prior to refraction
- Testing accommodation prior to refraction or in a presbyope
- Using both indirect and direct ophthalmoscopy
- Missing clinical signs due to the use of inappropriate internal or external examination methods

4.6.3.4 Assessment mode

Finally, assessment mode was similar to advanced mode, but was used for assessments only. Students only had one attempt at this mode for each of the 22 disease presentations. Once completed the grades achieved were automatically stored in a central results database, which could only be accessed by the module instructors.

It was impractical to expect students to learn the key characteristics of over 250 VP episodes by simply repeating them. Therefore, a learning resources section was created. This contained ‘at a glance guides’ for the 22 presentations. These guides highlighted what to ask, what to look for, referral urgencies and critical points (Figure 4.5). The extra information in these guides helped to ensure consistency with the learning material that was delivered in the didactic lectures.

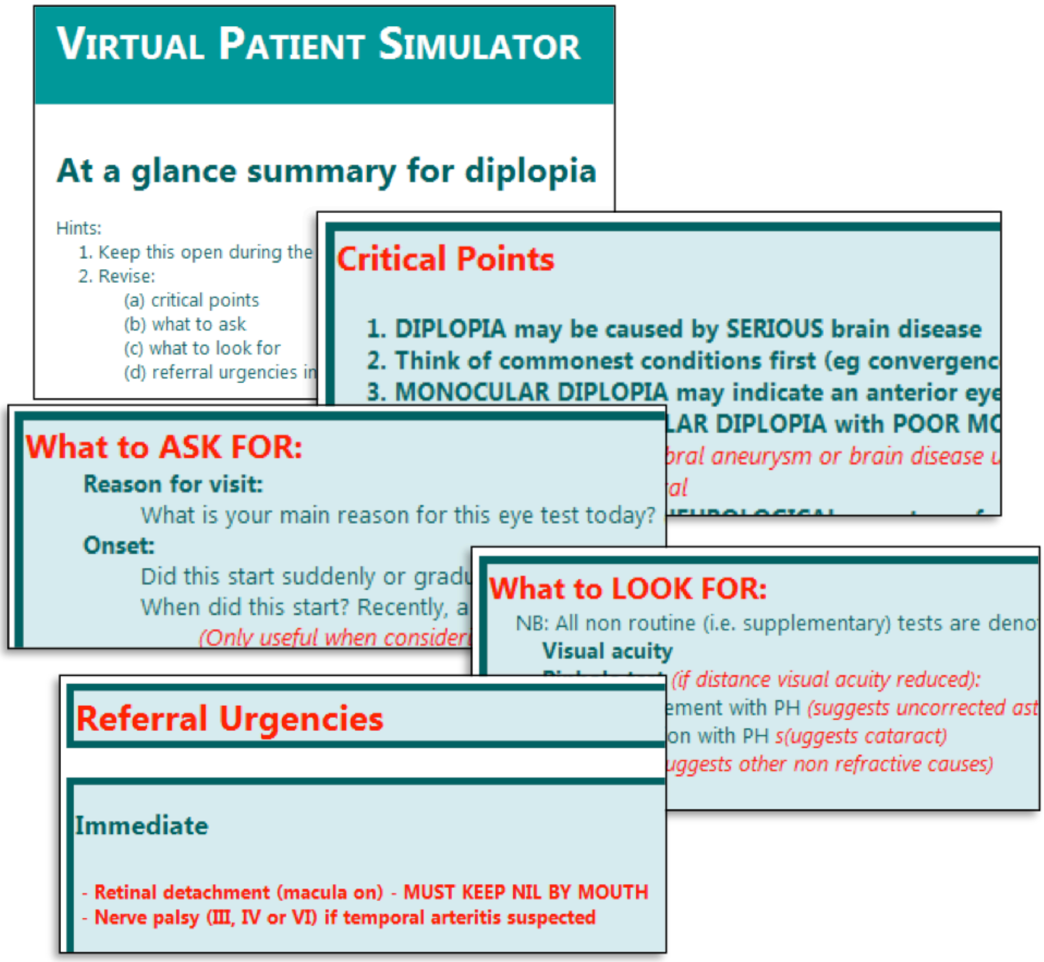

Figure 4.5 Extracts from at a glance guides showing the critical points, what to ask for, what to look for and referral urgencies

4.7 Delivery of the CPD module using the virtual patient

This module was taught via 18 tutorials during which the procedures for each presentation were discussed with students. These tutorials supported the PBL approach. Students were allocated individual login details for the online VP to enable unlimited access in their own time and student-directed learning was encouraged. These individual logins also allowed module instructors to easily access and link student profiles to assessment results, which were stored in a central results database.

4.8 Measures taken to ensure both teaching methods favoured all learning styles equally

Recall from chapter 3 that Felder and Silverman's (1998, p.675) model of learning style preferences proposed that teaching styles that are adapted to include both poles of the four learning style dimensions (active-reflective, sensing-intuitive, visual-verbal and sequential-global) would be close to providing the optimal learning environment for most students in any given class. Therefore, great care was taken to ensure that the teaching material for both cohorts catered for all 8 elements of the learning styles identified by Felder and Soloman's (2003) ILS.

The first learning style dimension was 'active-reflective' where active learners prefer immediate participation in learning tasks and learn best in groups, while reflective learners prefer to stand back and think ideas through alone. Both poles of this learning style dimension were accommodated in the didactic lecture cohort. Active learners had the opportunity to practice case studies in a group as part of the tutorials and reflective learners had the opportunity to work alone through the didactic lectures at their own pace prior to the tutorials. In the virtual patient cohort, active learners had the opportunity to practice various clinical scenarios in the interactive group tutorials and in their own time. Reflective learners had the opportunity to use the VP in teaching mode at their own pace and use the 'at a glance guides' (figure 4.5) in the learning resources to think ideas and concepts through.

The second learning style dimension was 'sensing-intuitive' where sensing learners are practical and prefer to work with facts, while intuitive learners tend to be more innovative and enjoy understanding theories. The didactic lectures favoured sensing learners by providing clear lists of signs, symptoms, tests and referral urgencies. Intuitive learners had the opportunity to understand theories by working through the group tutorials. The VP favoured sensing learners by again providing clear lists of what to ask, what to look for and referral urgencies in the 'at a glance' guides in the learning resources. Intuitive learners benefitted from critical points that were flagged up for specific disease cases or presentations at appropriate points of the eye examination. The live differential diagnosis shown at each stage of the examination in teaching mode also aided intuitive learners understanding theories.

The third learning style dimension was ‘visual-verbal’ where visual learners learn best through pictures, diagrams or charts, while verbal learners prefer words; either written or verbal. Both the didactic lectures and the VP consisted of relevant images to benefit visual learners. Both teaching methods provided written text in the lectures or in the ‘at a glance guides’ to benefit verbal learners.

The final learning style dimension was ‘sequential-global’ where sequential learners absorb information in small logical steps, whereas global learners try to tackle the whole problem at once taking large, somewhat random steps, before putting everything together. The tutorials in the didactic lecture cohort were delivered in a clear logical manner, guiding students through each stage of the patient episodes. Similarly, teaching mode in the VP guided students through each stage. These methods benefitted sequential learners. The advanced mode of the VP allowed global learners to jump in and tackle patient episodes by asking questions and performing tests in whichever order they preferred before making diagnoses and determining referral urgency. Similarly, the didactic lectures allowed global learners to learn signs, symptoms, tests and referral urgencies in any order before putting it all together and attempting the tutorials.

4.9 Assessment of the CPD module

The same assessment methods were used for both the didactic lecture cohort and the virtual patient cohort to ensure that this did not influence the learning experience. Forty percent of the module mark was derived from (a) completion of case studies for the didactic lecture cohort or (b) completion of VP episodes for the virtual patient cohort. The remaining 60% of the module mark was derived from a MCQ examination, consisting of 80 MCQs, which took place at the end of the academic year. This will be described in more detail in chapter 5.

4.10 Summary

The development of the teaching material for both cohorts and the measures taken to ensure they both covered the same clinical material have been described in this chapter. Methods used to evaluate the teaching material will be discussed in chapter 5.

Chapter 5 Evaluation of teaching material: Actual mastery

5.1 Introduction

The first measure used to evaluate the teaching materials in both cohorts was actual mastery. The concept of actual mastery is introduced and factors that influenced it are explored in this chapter.

5.2 Learning objectives

Recall from chapter 2 that the 5 learning objectives were:

- The ability to ask the right questions in history and symptoms based upon a patient's presenting complaint (question selection);
- The ability to recognise symptoms of serious eye disease (critical symptom recognition);
- The ability to select the most appropriate tests needed to look for signs that aid differential diagnosis (test selection);
- The ability to recognise signs of serious eye disease (critical sign recognition);
- The ability to decide upon the most appropriate referral urgency for any eye disease detected (referral urgency selection).

These learning objectives were evaluated in two ways; actual mastery and perceived mastery (see chapter 6).

5.3 Actual mastery of the learning objectives

Actual mastery was determined based on a MCQ examination that took place at the end of the second academic year. It has been argued that MCQs tend to measure factual recall and recognition of isolated facts. However, research has shown that carefully constructed MCQs are capable of assessing deeper levels of knowledge and understanding (Peitzman et al., 1990). MCQs also have the advantage of being capable of assessing broad domains of knowledge efficiently and reliably (Vleuten et al., 1996; Abdel-Hameed et al., 2005).

A total of 25 MCQs were created which included 5 MCQs per learning objective (see tables 5.1 to 5.5).

Table 5.1 MCQs used to determine actual mastery for question selection

MCQ	The correct answer is shown in bold text
(1)	Which question is LEAST LIKELY to be useful when investigating reduced vision in a white eye? a) do you have high blood pressure or diabetes? b) does anyone in your family have migraine? c) did this start after your cataract operation? d) **do you ever have discharge from your eye?** e) did this come on suddenly or gradually?
(2)	Which question is LEAST LIKELY to be useful when investigating diplopia? a) **do you wear contact lenses?** b) do you have jaw ache when chewing? c) does it get worse when you are tired? d) does this still happen if one eye is closed? e) do you have diabetes or high blood pressure?
(3)	Which question is LEAST LIKELY to be useful when investigating a complaint of watery, itchy or gritty eye(s)? a) does tend to happen at certain times of the year? b) have you had eye surgery? c) do you suffer from rheumatoid arthritis? d) **do you have high blood pressure?** e) do you use eye drops?
(4)	Which question is LEAST LIKELY to be useful when investigating floaters or flashes? a) is this in one or both eyes? b) **do you experience pain with eye movements?** c) do you have diabetes? d) have you had cataract surgery? e) have you ever suffered from migraine?
(5)	Which question is LEAST LIKELY to be useful when investigating facial rashes around the eye? a) do you use eye drops? b) do you have pain around the eye? c) do you suffer from light sensitivity? d) do you have blurred vision? e) **do you have watery eyes?**

Table 5.2 MCQs used to determine actual mastery for critical symptom recognition

MCQ	The correct answer is shown in bold text
(6)	Which symptom is LEAST LIKELY to indicate serious eye or brain disease in a patient with reduced vision in a white eye? a) light flashes b) loss of peripheral vision c) **floaters** d) distortion e) pain
(7)	Which symptom is LEAST LIKELY to indicate serious eye disease in a patient with non traumatic red eye? a) reduced vision b) haloes seen around lights c) **subconjunctival haemorrhage** d) light sensitivity e) pain
(8)	Which symptom is LEAST LIKELY to indicate serious brain disease in a patient with diplopia? a) transient diplopia b) horizontal diplopia c) new onset diplopia d) **monocular diplopia** e) binocular diplopia
(9)	Which is LEAST LIKELY to indicate a sinister headache? a) **chronic daily headache** b) onset within last 6 months c) very young or elderly patient d) atypical headache e) high blood pressure
(10)	Which symptom or sign is MOST LIKELY to occur with Horner's syndrome? a) diplopia b) poor motility c) blurred vision d) **mild ptosis** e) glare

Table 5.3 MCQs used to determine actual mastery for test selection

MCQ	The correct answer is shown in bold text
(11)	Which test is LEAST LIKELY to be useful when investigating reduced vision in a white eye? a) motility b) **near point of convergence** c) tonometry d) refraction e) ophthalmoscopy
(12)	Which test is LEAST LIKELY to be useful when investigating non traumatic red eye? a) visual acuity b) **visual fields** c) pupil responses d) ophthalmoscopy e) tonometry
(13)	Which test is LEAST LIKELY to be useful when investigating diplopia? a) convergence b) cover test c) **tonometry** d) pupil responses e) motility
(14)	Which is MOST LIKELY to be useful when investigating a complaint of watery, itchy or gritty eye(s)? a) pupil responses b) **external examination** c) cover test d) internal examination e) motility
(15)	Examination of which is LEAST LIKELY to be useful when investigating a complaint of pain in or around the eyes? a) eyelid b) motility c) **retinal blood vessels** d) optic nerve e) pupils

Table 5.4 MCQs used to determine actual mastery for critical sign recognition

MCQ	The correct answer is shown in bold text
(16)	Which sign is LEAST LIKELY to indicate serious eye disease in a patient with non traumatic red eye? a) **injected blood vessels that move over sclera** b) corneal oedema c) shallow anterior chamber d) keratic precipitates on corneal endothelium e) choroiditis
(17)	Which sign is LEAST consistent with a foreign body sensation? a) **punctual stenosis** b) trichiasis c) corneal staining d) poor TBUT e) corneal stromal infiltrate
(18)	Which form of visual field loss is LEAST consistent with brain disease affecting the optic radiation? a) respects vertical midline b) **unilateral** c) homonymous d) hemianopia e) quadrantanopia
(19)	Which is LEAST LIKELY to be an eyelid tumour? a) a rounded, pearly lump with abnormal blood vessels b) a warty growth c) a lesion with rolled edges and central ulceration d) **a tender swelling on the eyelid margin** e) a black round lesion
(20)	Which sign is MOST LIKELY to indicate raised intracranial pressure? a) bilateral crowded optic disc b) unilateral optic disc cupping c) **bilateral optic disc swelling** d) unilateral optic atrophy e) bilateral optic disc drusen

Table 5.5 MCQs used to determine actual mastery for referral urgency selection

MCQ	The correct answer is shown in bold text
(21)	Which cause of reduced vision in white eye(s) requires the HIGHEST referral urgency? a) advanced chronic glaucoma b) dry macular degeneration c) central macular oedema after cataract surgery d) cataract e) **central retinal artery occlusion**
(22)	Which cause of diplopia requires the HIGHEST referral urgency? a) keratoconus b) **nerve palsy** c) persistent migraine d) myasthenia gravis e) convergence insufficiency
(23)	Which requires the HIGHEST referral urgency? a) flashes associated with severe migraine b) visual field loss and optic disc damage in chronic glaucoma c) **headache with transient reduced vision and optic nerve swelling** d) scotoma caused by macular hole e) distortion with central serous retinopathy
(24)	Which requires the HIGHEST referral urgency? a) orbital tumour b) conjunctival malignant melanoma c) pingueculum d) conjunctival squamous cell carcinoma e) **orbital cellulitis**
(25)	Which requires the HIGHEST referral urgency? a) anisocoria in which one eye has keratic precipitates b) anisocoria that is likely to be due to Adie's pupil c) anisocoria with head and neck pain d) **anisocoria in which one eye is red with a hazy cornea and an oval pupil** e) anisocoria with heterochromia

5.3.1 Item analysis of the MCQs

Item analysis was used to check the quality indices for each MCQ. Item analysis involves the examination of student responses to individual items of an MCQ test to assess the quality of those items and the test as a whole (Singh et al., 2009, p. 70-77). There are two main components to item analysis; difficulty factor and discrimination index. Difficulty factor is simply the proportion of students who answered a question correctly; the higher the difficulty factor, the easier the question is. The recommended range is between 0.30 and 0.70 (Hingorjo & Jaleel, 2012). Discrimination index can be described as almost being a correlation that shows the relationship of a student's response to a single item in relation to their overall performance in the test. It measures the difference in the percentage of students who answered a question correctly in the overall top and bottom performing 27%. Discrimination index can range from -1.00 to +1.00. The higher the discrimination index, the better the test item is at discriminating between students with high and low test scores. A high positive discrimination index suggests that students who answer that question correctly will also have a relatively high score in the overall test. A negative discrimination index would suggest that the question was not capable of discriminating the student's overall test performance based on the response to that question. This would occur where the most knowledgeable students would answer the question incorrectly while the least knowledgeable students would answer it correctly. A discrimination index above +0.30 is generally considered good (Abdel-Hameed et al., 2005).

The results of item analysis for the 25 MCQs are shown in tables 5.6 to 5.10. The quality indices for some MCQs fell outside the desired range, but this was considered acceptable, as having perfect quality indices for MCQs in both the VP and didactic lecture cohort would potentially mask the effect that was being investigated; whether actual mastery differed with the two teaching methods. MCQ 6 highlights this where the difficulty factor of 0.86 for this MCQ in the didactic lecture cohort shows that the question was too easy, but the same question had a difficulty factor of 0.23 in the virtual patient cohort showing that this question was too difficult. It is likely that the different teaching methods might have caused this difference. It could be argued that MCQ 6 was flawed in it's design as the symptoms 'flashes' and 'floaters' are usually used together, however, these two symptoms were taught separately in the CPD module.

Table 5.6 Difficulty factor and discrimination index for the 5 MCQs used in both the didactic lecture cohort and the virtual patient cohort to determine actual mastery of question selection

MCQ	Difficulty factor (didactic lectures cohort)	Discrimination index (didactic lectures cohort)	Difficulty factor (virtual patient cohort)	Discrimination index (virtual patient cohort)
1	0.80	0.36	0.88	0.33
2	0.72	0.41	0.80	0.23
3	0.86	0.37	0.81	0.32
4	0.44	0.33	0.53	0.47
5	0.55	0.22	0.71	0.48

Table 5.7 Difficulty factor and discrimination index for the 5 MCQs used in both the didactic lecture cohort and the virtual patient cohort to determine actual mastery of critical symptom recognition

MCQ	Difficulty factor (didactic lectures cohort)	Discrimination index (didactic lectures cohort)	Difficulty factor (virtual patient cohort)	Discrimination index (virtual patient cohort)
6	0.86	0.39	0.23	-0.14
7	0.59	0.48	0.76	0.57
8	0.74	0.33	0.75	0.42
9	0.53	0.48	0.50	0.47
10	0.84	0.35	0.64	0.54

Table 5.8 Difficulty factor and discrimination index for the 5 MCQs used in both the didactic lecture cohort and the virtual patient cohort to determine actual mastery of test selection

MCQ	Difficulty factor (didactic lectures cohort)	Discrimination index (didactic lectures cohort)	Difficulty factor (virtual patient cohort)	Discrimination index (virtual patient cohort)
11	0.64	0.32	0.79	0.06
12	0.37	0.23	0.54	0.19
13	0.84	0.15	0.74	0.01
14	0.92	0.13	0.95	0.19
15	0.47	0.44	0.50	0.25

Table 5.9 Difficulty factor and discrimination index for the 5 MCQs used in both the didactic lecture cohort and the virtual patient cohort to determine actual mastery of critical sign recognition

MCQ	Difficulty factor (didactic lectures cohort)	Discrimination index (didactic lectures cohort)	Difficulty factor (virtual patient cohort)	Discrimination index (virtual patient cohort)
16	0.62	0.38	0.75	0.22
17	0.82	0.39	0.93	0.13
18	0.75	0.28	0.76	0.55
19	0.63	0.22	0.60	0.26
20	0.82	0.38	0.60	0.36

Table 5.10 Difficulty factor and discrimination index for the 5 MCQs used in both the didactic lecture cohort and the virtual patient cohort to determine actual mastery of referral urgency selection

MCQ	Difficulty factor (didactic lectures cohort)	Discrimination index (didactic lectures cohort)	Difficulty factor (virtual patient cohort)	Discrimination index (virtual patient cohort)
21	0.91	0.39	0.93	0.22
22	0.85	0.21	0.96	0.11
23	0.56	0.42	0.77	0.38
24	0.92	0.41	0.94	0.36
25	0.28	0.38	0.40	0.38

5.4 Study sample findings

Appendix 9 shows the complete findings for all 102 students in the didactic lecture cohort and all 93 students in the virtual patient cohort. The distributions of actual mastery scores for all 5 learning objectives are shown below (figure 5.1). One-sample Kolmogorov-Smirnov tests revealed that the scores did not follow a normal distribution. Various transformations were attempted but the data could not be transformed to follow a normal distribution. Therefore, only non-parametric tests were used to analyse the data.

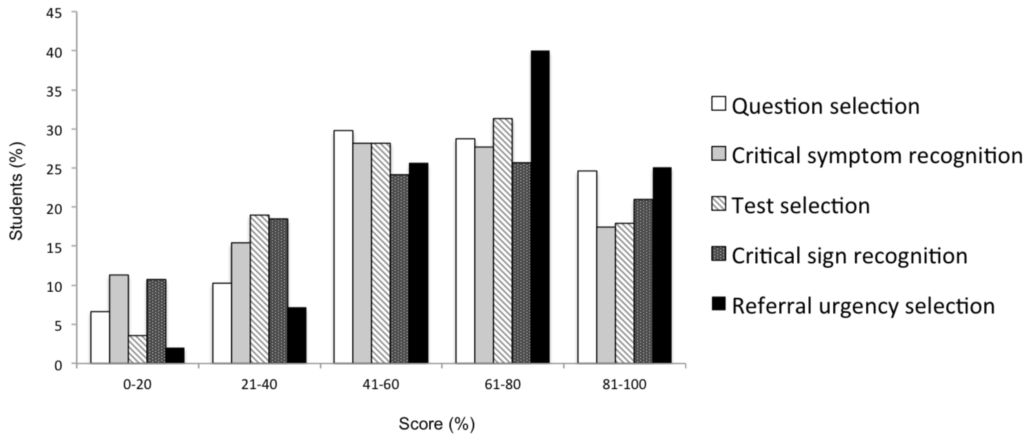

Figure 5.1 Distribution of actual mastery scores for the 5 learning objectives of 195 students, including 102 students in the didactic lecture cohort and 93 in the virtual patient cohort.

Actual mastery scores were classified as high or low by splitting the data at the 60% mark. A midpoint of 60% was used to split the data as the scores for actual mastery ranged from 20-100%.

DTA was used to investigate the influence of teaching method, academic ability, sex and learning style profile upon actual mastery scores for each of the 5 learning objectives. The only statistically significant findings were that the academic ability altered actual mastery for all five learning objectives (see figures 5.2 to 5.6: higher for higher academic achievers) and that the teaching method altered actual mastery of critical symptom recognition (figure 5.3: higher for didactic lectures) and referral urgency selection (figure 5.6: higher for VP).

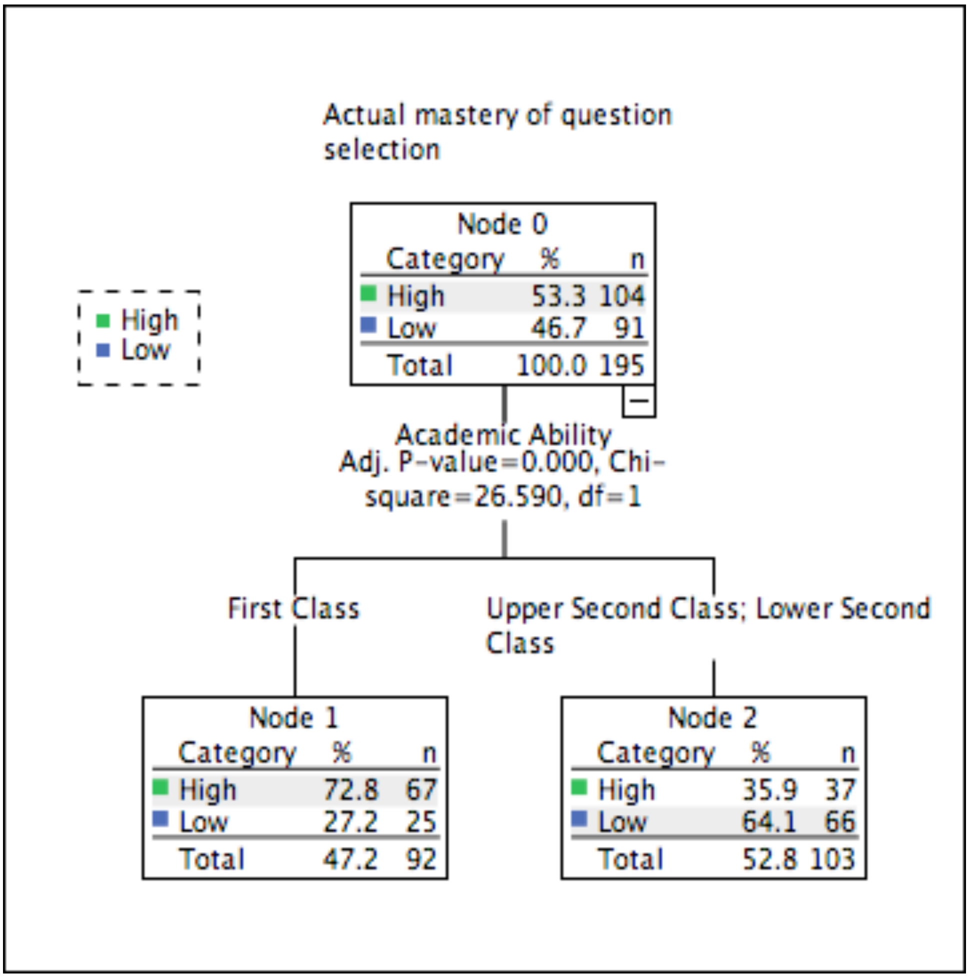

Figure 5.2 Decision tree analysis investigating the influences of teaching method, sex and learning style profile on actual mastery scores of all 195 students pooled from the didactic lecture and the virtual patient cohorts for question selection.

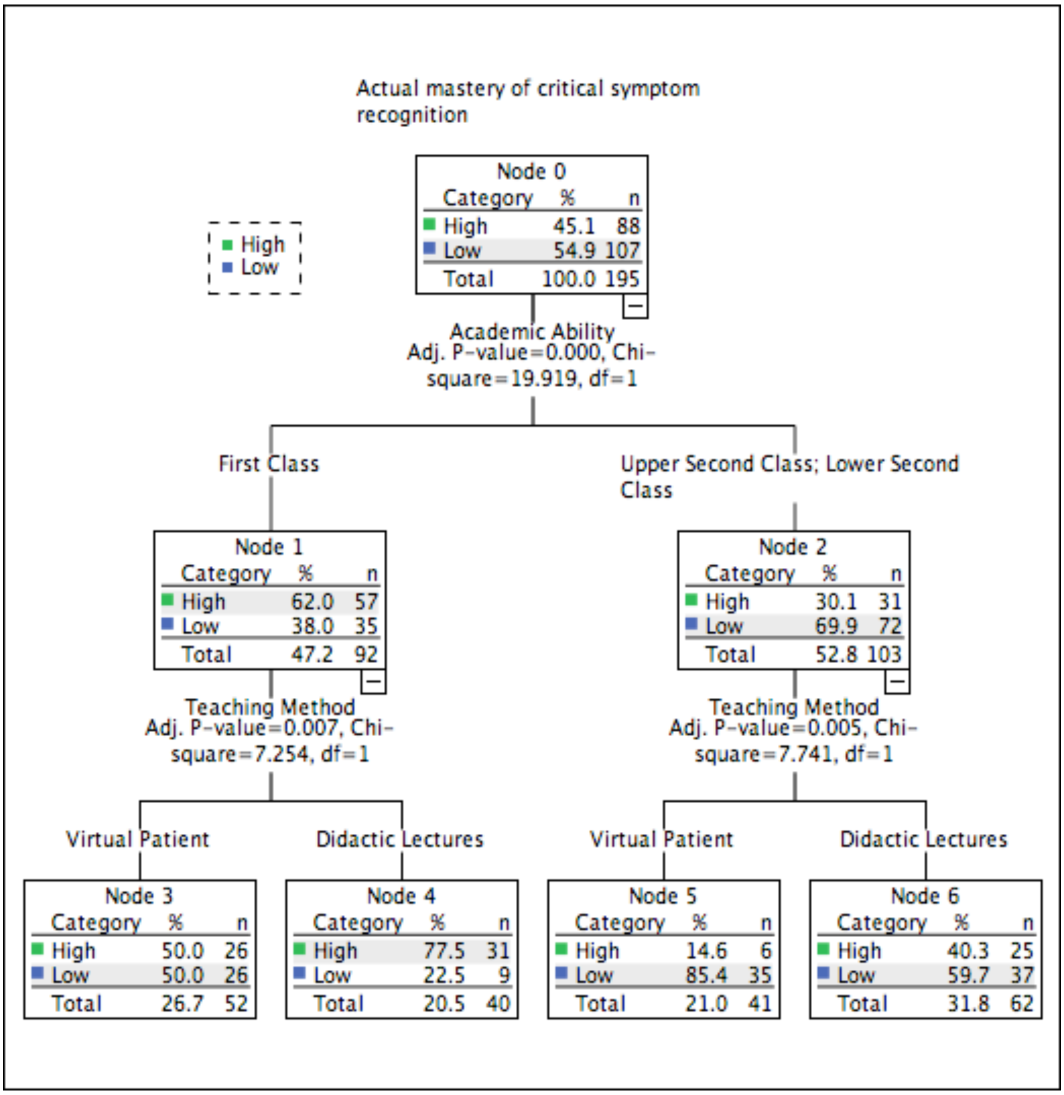

Figure 5.3 Decision tree analysis investigating the influences of teaching method, sex and learning style profile on actual mastery scores of all 195 students pooled from the didactic lecture and the virtual patient cohorts for critical symptom recognition.

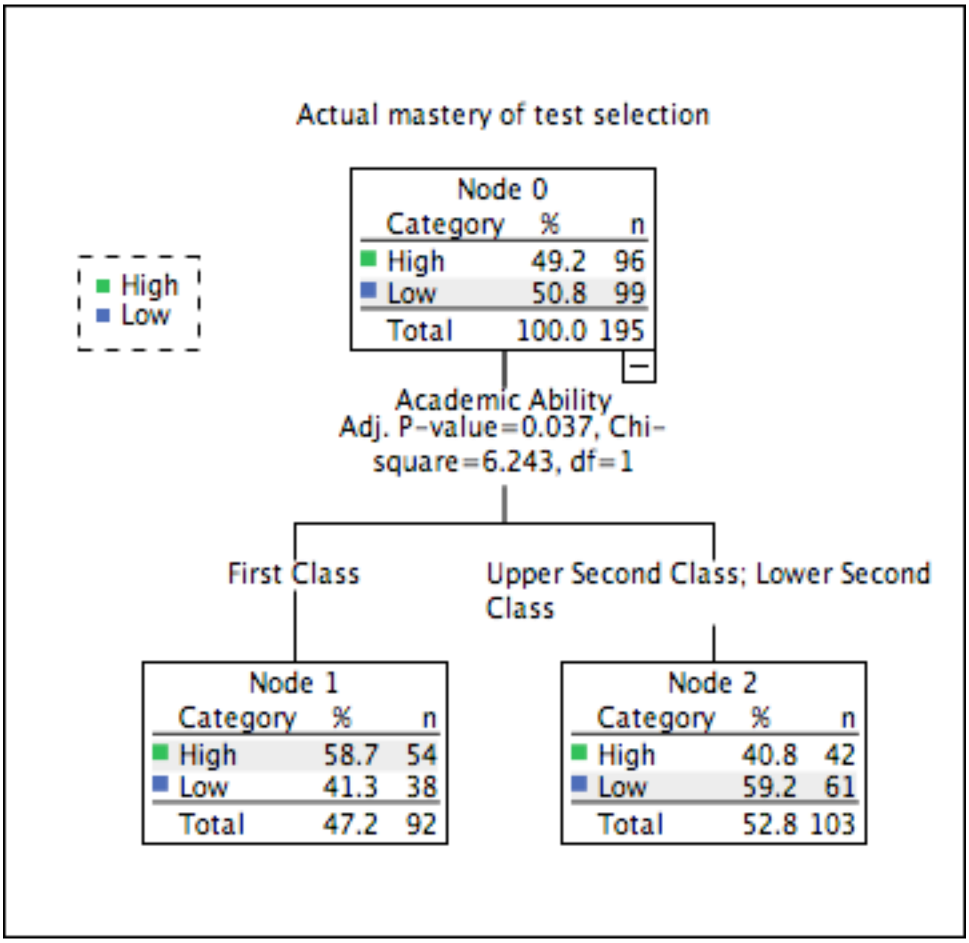

Figure 5.4 Decision tree analysis investigating the influences of teaching method, sex and learning style profile on actual mastery scores of all 195 students pooled from the didactic lecture and the virtual patient cohorts for test selection.

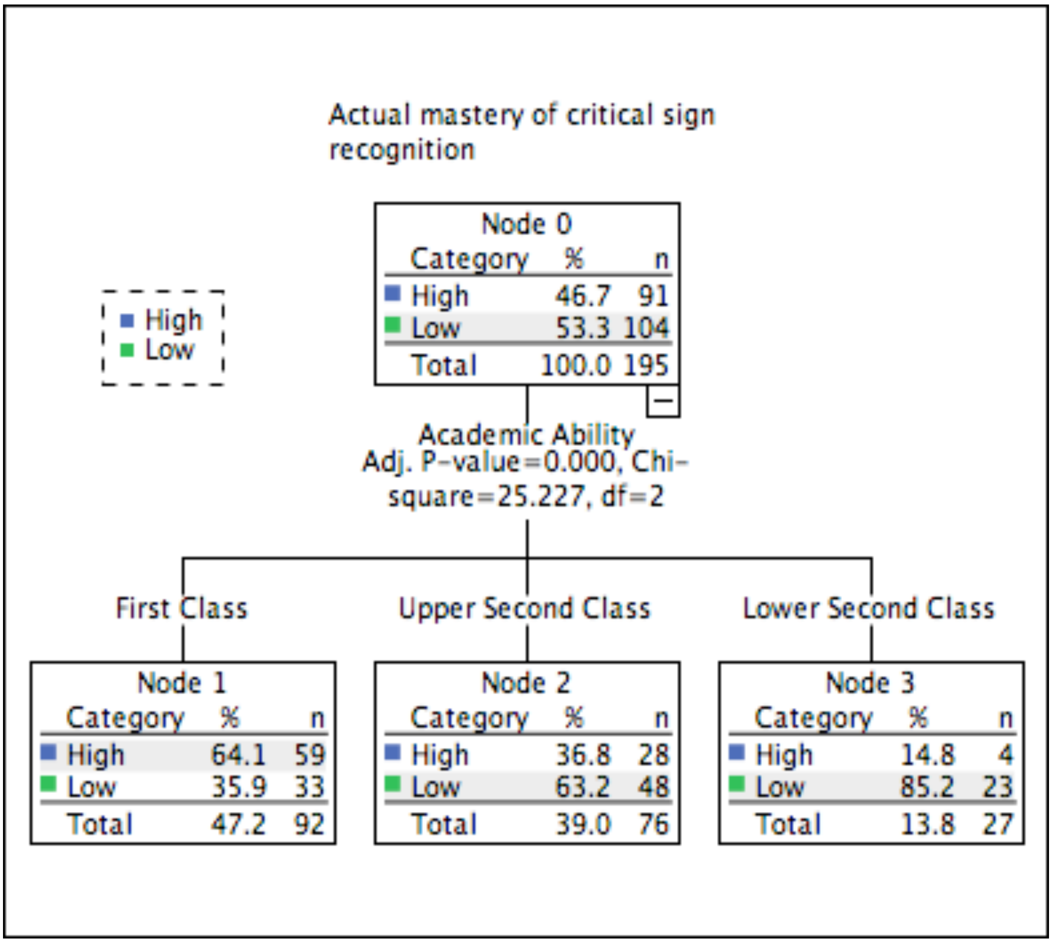

Figure 5.5 Decision tree analysis investigating the influences of teaching method, sex and learning style profile on actual mastery scores of all 195 students pooled from the didactic lecture and the virtual patient cohorts for critical sign recognition.

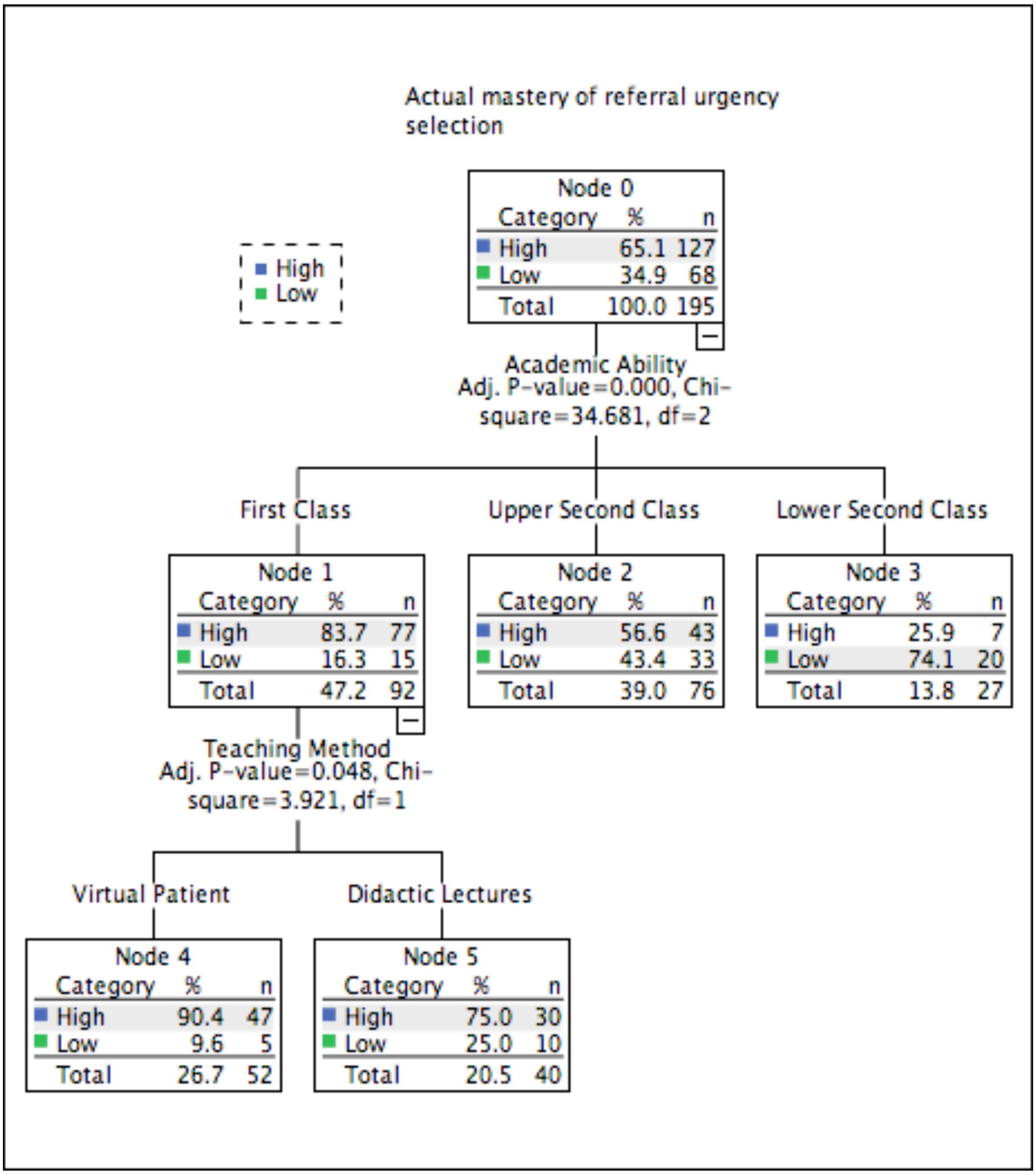

Figure 5.6 Decision tree analysis investigating the influences of teaching method, sex and learning style profile on actual mastery scores of all 195 students pooled from the didactic lecture and the virtual patient cohorts for referral urgency selection.

Table 5.11 shows the number of students from the didactic lectures cohort with high and low actual mastery scores across all 5 learning objectives. Application of Chi-square for R x C tables to the data shown in table 5.11 showed that levels of actual mastery did not differ between the 5 learning objectives (Chi = 2.70, df =4, p = 0.61, power = 99.9%).

Table 5.11 Table showing the number of students with high (60%+) and low (<60%) actual mastery scores across all 5 learning objectives for 102 students in the didactic lecture cohort.

	Question selection	Critical symptom recognition	Test selection	Critical sign recognition	Referral urgency selection
High	52	56	51	46	56
Low	50	46	51	56	46

Table 5.12 shows the number of students from the virtual patient cohort with high and low actual mastery scores across all 5 learning objectives. Application of Chi-square for R x C tables to the data shown in table 5.12 showed that levels of actual mastery varied significantly between the 5 learning objectives (Chi = 35.11, df =4, p < 0.01). Post hoc testing using Chi-square for 2 x 2 tables showed that actual mastery was highest for referral urgency selection compared to question selection (Chi = 8.67, df =1, p < 0.01), critical symptom recognition (Chi = 33.09, df =1, p < 0.01), test selection (Chi = 15.48, df =1, p < 0.01) and critical sign recognition (Chi = 15.48, df =1, p < 0.01). Actual mastery was also lower for critical symptom recognition compared to question selection (Chi = 8.68, df =1, p < 0.01).

Table 5.12 Table showing the number of students with high (60%+) and low (<60%) actual mastery scores across all 5 learning objectives for 93 students in the virtual patient cohort.

	Question selection	Critical symptom recognition	Test selection	Critical sign recognition	Referral urgency selection
High	52	32	45	45	71
Low	41	61	48	48	22

5.5 Discussion

Academic ability influenced actual mastery of all 5 learning objectives showing that a greater proportion of high academic achievers had high actual mastery. This is not surprising as recall from chapter 2 that academic ability was determined based on overall performance across 6 different modules. Therefore, it seems logical that the same trend would continue for this form of assessment.

Referral urgency selection was influenced by teaching method where more students in the virtual patient cohort had high actual mastery scores than in the didactic lecture cohort. However, the opposite influence of teaching method was found for actual mastery scores of critical symptom recognition, where more students in the didactic lectures cohort had high actual mastery scores than in the virtual patient cohort. Actual mastery did not differ between the five learning objectives in the didactic lecture cohort but did differ in the virtual patient cohort. Therefore, only teaching using the VP appeared to raise actual mastery of referral urgency selection and lower actual mastery of critical symptom recognition.

This disparity in the influence of teaching methods for the different learning objectives may be explained by taking a closer look at the learning objectives and the assessment method used in this study. It could be argued that critical symptom recognition relies upon simple memorisation of lists of symptoms and that didactic lectures allowed easier rote learning of lists of tests, signs and symptoms than the VP. This could be the reason why the VP led to lower actual mastery scores for this learning objective. Referral urgency selection, on the other hand, relies on the ability to evaluate a multitude of presenting signs and symptoms along with the ability to correctly select the appropriate tests to aid diagnosis and then finally select the most appropriate management plan. Subtle differences in the presenting signs or symptoms could easily change the management plan; for example, the management of a “macula on” retinal detachment would be different to a “macula off” retinal detachment. It could be that the VP encouraged a deeper level of interaction while completing patient episodes, which might have led to the higher actual mastery scores for this learning objective.

This disparity in the outcome of the influence of teaching method for critical symptom recognition versus referral urgency selection has highlighted the importance of setting appropriate assessment criteria and using appropriate assessment methods to evaluate the efficiency of new teaching methods. This was highlighted by a study by Rosenberg et al. (2010) who evaluated a new form of CAL in orthodontics. They developed an electronic tutorial and then assessed students using a traditional paper based test and a computer based test, which were both matched for level of difficulty and topics covered. They found that females performed significantly better than males when examined with the traditional paper based test, but no gender differences occurred when examined with the electronic tutorial. This revealed how choice of assessment method can influence the learning outcomes.

The efficiency of new teaching methods is commonly determined by examining learning outcomes, which can be assessed in a number of different ways. Current literature has established that learning outcomes are influenced significantly and maybe to a greater extent by the assessment method chosen (Rust, 2002; Wood, 2003) as opposed to the teaching methods used. A study by Scouller (1998) concluded that by the time students have reached higher education, they have already developed the skills and knowledge of how best to adapt their learning approaches to the assessment methods that will be used. Students find it difficult to cope with the sheer volume of learning materials that most curricula contain and so resort to learning only what they believe they will be assessed on (Wass et al., 2001).

Therefore, although it could be concluded that use of the VP led to a reduction in actual mastery scores for critical symptom recognition, it can be argued that this finding may have resulted from inappropriate assessment methods. These learning objectives encouraged the memorisation of lists of symptoms, while good clinical decision-making is dependent upon the ability to interpret and process combinations of presenting signs and symptoms. Further investigation would be required to determine the impact of changing the assessment methods upon actual mastery scores.

Although there has been a large amount of research examining gender differences on the academic performance of healthcare students, very little research has focused specifically on the differences in CDM skills of males and females. A recent study by Klemenc-Ketis and Kersnik (2014) examined the decision making process of undergraduate medical students using virtual clinical cases. They found that females performed significantly better than males. However, a study by Noohi et al. (2012) found no gender differences in CDM skills of nursing students. This present study supports the findings of this latter study.

The present study found that learning style profiles did not influence actual mastery. Recall from chapter 3 that Felder and Silverman, (1998, p.675) suggested that teaching styles that are adapted to include both poles of the four learning style dimensions would be close to providing the optimal learning environment for most students in any given class. Therefore, the lack of influence of learning style on actual mastery, reported in the present study, is a positive finding as it suggests that the measures taken to ensure that both teaching methods were adapted to include both poles of the four learning style dimensions (refer back to chapter 4.8) were successful.

5.6 Summary

The findings of this chapter indicate that academic ability and teaching method influenced actual mastery scores. It has highlighted the importance of selecting appropriate assessment methods when investigating the efficacy of new teaching methods. The concept of perceived mastery and the factors that influenced it will be explored in the next chapter.

Chapter 6 Evaluation of teaching material: Perceived mastery

6.1 Introduction

The second measure used to evaluate the teaching materials in both cohorts was perceived mastery. The concept of perceived mastery is introduced and factors that influenced it are explored in this chapter.

6.2 Learning objectives

Recall from chapter 2 that the 5 learning objectives were:

- The ability to ask the right questions in history and symptoms based upon a patient's presenting complaint (question selection);
- The ability to recognise symptoms of serious eye disease (critical symptom recognition);
- The ability to select the most appropriate tests needed to look for signs that aid differential diagnosis (test selection);
- The ability to recognise signs of serious eye disease (critical sign recognition);
- The ability to decide upon the most appropriate referral urgency for any eye disease detected (referral urgency selection).

These learning objectives were evaluated in two ways; actual mastery (see chapter 5) and perceived mastery.

6.3 Perceived mastery of the learning objectives

Perceived mastery was a self-assessed measure of the student's own perception of their confidence in the 5 learning objectives. Previous literature has referred to this as self-efficacy, which was defined by the psychologist Albert Bandura (1977) as an individual's belief in their ability to perform or accomplish a task. Perceived mastery was determined via a 5-item Likert scale questionnaire. This questionnaire was comprised of the following 5 items:

1. I am very confident in my ability to ask the right questions in history and symptoms based upon a patient's presenting complaint.
2. I am very confident in my ability to recognise symptoms of serious eye disease.

3. I am very confident in my ability to select the most appropriate tests needed to look for signs that aid differential diagnosis.
4. I am very confident in my ability to recognise signs of serious eye disease.
5. I am very confident in my ability to decide upon the most appropriate referral urgency for any eye disease detected.

These items corresponded directly to the 5 learning objectives. The following 5 Likert levels were used for each item:

1. Strongly disagree
2. Disagree
3. Neither agree or disagree
4. Agree
5. Strongly agree

These levels were converted to a score of 1 to 5 where a score of 1 corresponded to strongly disagree and a score of 5 corresponded to strongly agree. These scores were then converted to a percentage.

This questionnaire was released via the Blackboard learning management system used by Aston University and was completed by students in both cohorts at the end of their second academic year. This same questionnaire was then repeated at the beginning of their third academic year to see if this perceived confidence was retained after the long summer break.

6.4 Study sample findings

Appendix 9 shows the complete findings for all 102 students in the didactic lecture cohort and all 93 students in the virtual patient cohort. The distribution of perceived mastery scores for all 5 learning objectives are shown below (figure 6.1). One-sample Kolmogorov-Smirnov tests revealed that the scores did not follow a normal distribution. Various transformations were attempted but the data could not be transformed to follow a normal distribution. Therefore, only non-parametric tests were used to analyse the data.

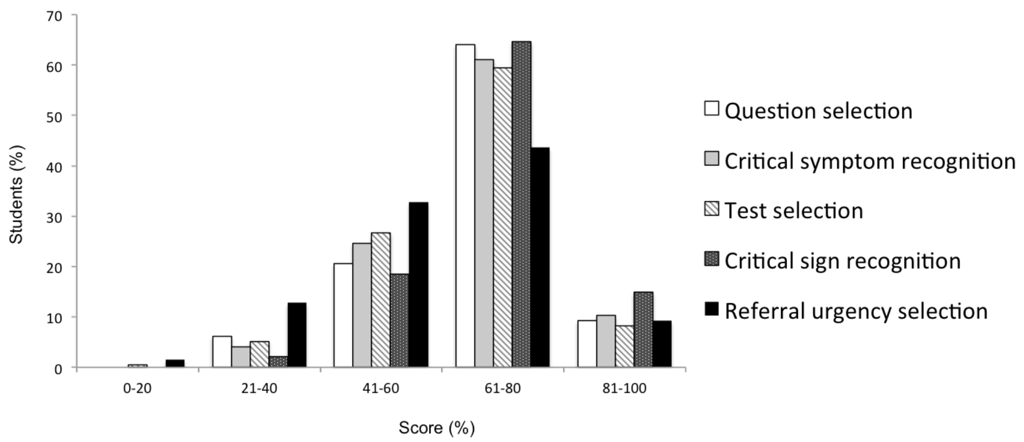

Figure 6.1 Distribution of perceived mastery scores for the 5 learning objectives of 195 students, including 102 in the didactic lecture cohort and 93 in the virtual patient cohort.

Perceived mastery scores were classified as high or low by splitting the data at the 60% mark. A midpoint of 60% was used to split the data as the scores for perceived mastery ranged from 20-100%.

DTA was used to investigate the influence of teaching method, academic ability, sex and learning style profile upon perceived mastery scores for each of the 5 learning objectives (see figures 6.2 and 6.3).

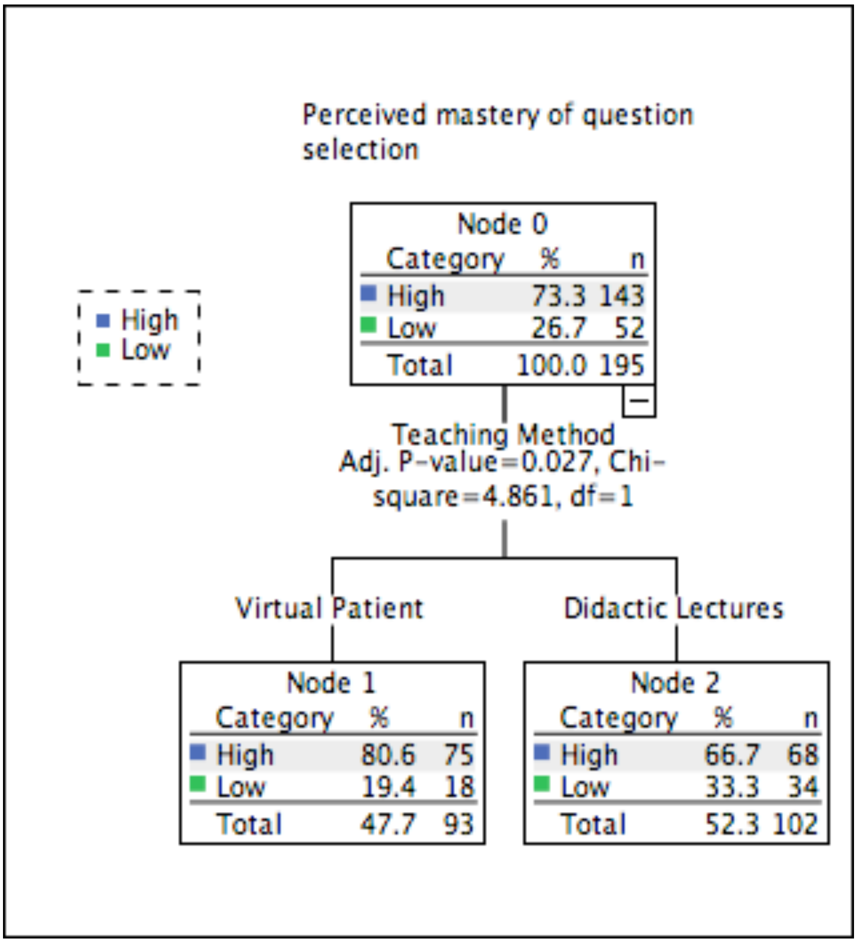

Figure 6.2 Decision tree analysis investigating the influences of teaching method, academic ability, sex and learning style profile on perceived mastery scores of all 195 students pooled from the didactic lecture and the virtual patient cohorts for question selection.

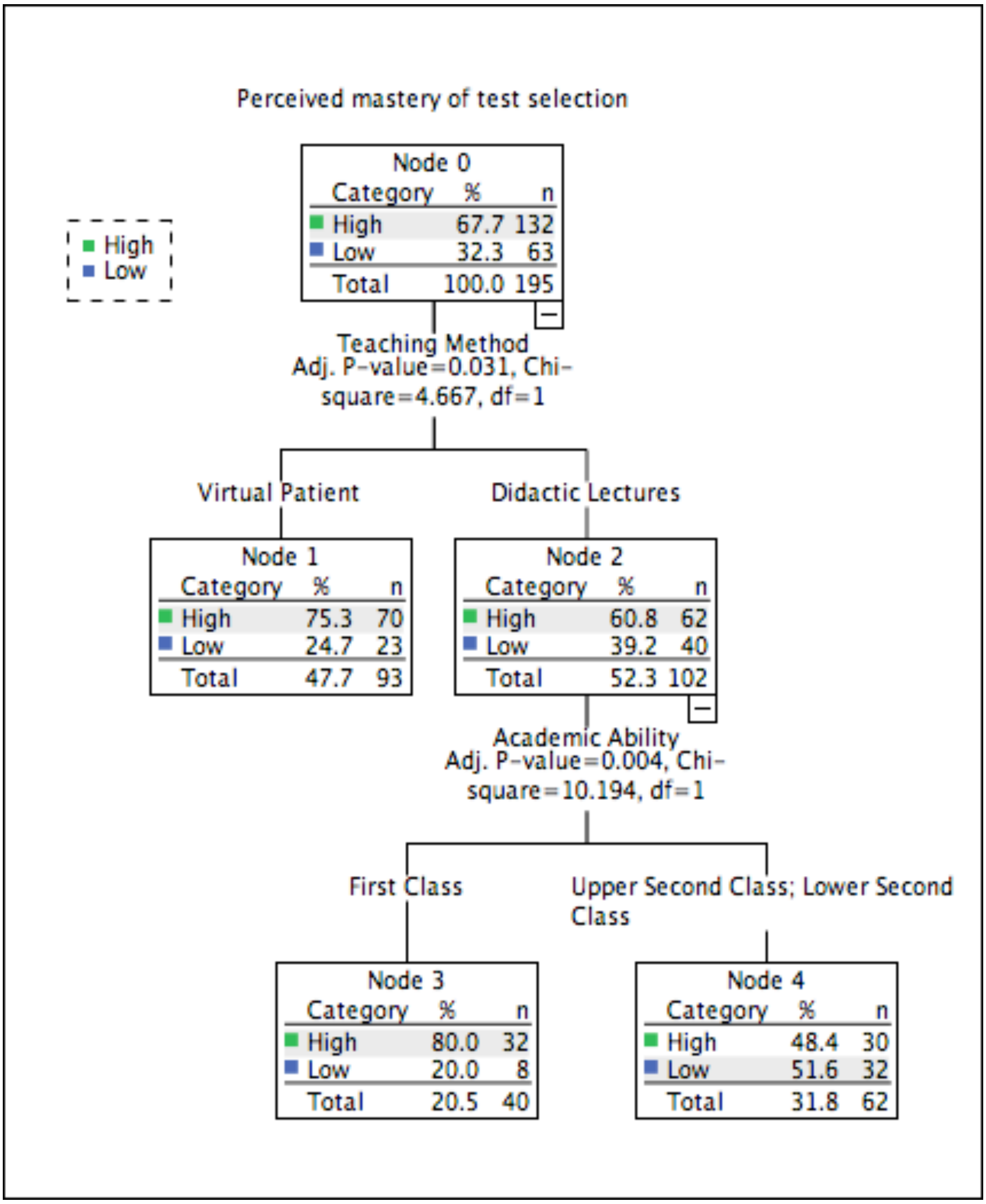

Figure 6.3 Decision tree analysis investigating the influences of teaching method, academic ability, sex and learning style profile on perceived mastery scores of all 195 students pooled from the didactic lecture and the virtual patient cohorts for test selection.

Only teaching method and academic ability had statistically significant influences on perceived mastery. The VP raised perceived mastery of question selection (Chi = 4.86, df = 1, p = 0.03) and test selection (Chi = 4.67, df = 1. P = 0.03). High academic ability raised perceived mastery of test selection for students in the didactic lecture cohort (Chi = 10.19, df = 1, p < 0.01).

Table 6.1 shows the number of students from the didactic lecture cohort with high and low perceived mastery scores across all 5 learning objectives. Application of Chi-square for R x C tables to the data shown in table 6.1 showed that levels of actual mastery differed significantly between the 5 learning objectives (Chi = 14.23, df =4, $p = 0.01$). Post hoc testing using Chi-square for 2 x 2 tables showed that perceived mastery was lower for referral urgency selection compared to question selection (Chi = 4.57, df =1, $p = 0.03$), critical symptom recognition (Chi = 3.97, df =1, $p = 0.05$) and critical sign recognition (Chi = 13.33, df =1, $p < 0.01$). Perceived mastery was higher for critical sign recognition compared to test selection (Chi = 5.83, df =1, $p = 0.02$).

Table 6.1 Table showing the number of students with high (60%+) and low (<60%) perceived mastery scores across all 5 learning objectives for 102 students in the didactic lecture cohort.

	Question selection	**Critical symptom recognition**	**Test selection**	**Critical sign recognition**	**Referral urgency selection**
High	68	67	62	78	53
Low	34	35	40	24	49

Table 6.2 shows the number of students from the virtual patient cohort with high and low perceived mastery scores across all 5 learning objectives. Application of Chi-square for R x C tables to the data shown in table 6.2 showed that levels of perceived mastery varied significantly between the 5 learning objectives (Chi = 26.30, df =4, $p < 0.01$). Post hoc testing using Chi-square for 2 x 2 tables showed that perceived mastery was lower for referral urgency selection compared to question selection (Chi = 15.25, df =1, $p < 0.01$), critical symptom recognition (Chi = 11.53, df =1, $p < 0.01$), test selection (Chi = 9.39, df =1, $p < 0.01$) and critical sign recognition (Chi = 18.10, df =1, $p < 0.01$).

Table 6.2 Table showing the number of students with high (60%+) and low (<60%) perceived mastery scores across all 5 learning objectives for 93 students in the virtual patient cohort.

	Question selection	Critical symptom recognition	Test selection	Critical sign recognition	Referral urgency selection
High	75	72	70	77	50
Low	18	21	23	16	43

6.5 Discussion

Use of the VP appeared to raise students' confidence in question and test selection. This is consistent with previous literature that has shown that use of CAL and simulators can significantly increase student self-confidence (Engel et al., 1997; Valizadeh et al., 2013). A study by Valizadeh et al. (2013) established that, compared to conventional teaching methods, simulation based teaching increased self-confidence in nursing students learning peripheral venous catheterization.

Academic ability only appeared to increase confidence in test selection for students in the didactic lecture cohort. While this is in keeping with previous research which has shown that that high academic achievers tend to have a more realistic image of their own performance (Saavedra & Kwun, 1993; Furnham & Stringfield, 1994), this finding was not consistent across all 5 learning objectives. Studies by Freund and Kasten (2012) and Holling and Preckel (2005) have shown that alongside academic ability, past experiences and social comparison processes (how students evaluate their own opinions and abilities by comparing themselves to others in order to reduce uncertainty) can also influence students' self-perception of their performance and ability. Therefore, further investigation into the influences of these factors would be beneficial.

Previous research has found that females report lower levels of self-assessed performance compared to males (Lind et al., 2002; Hargittai & Shafer, 2006). A study by Colbert-Getz et al. (2013) found that sex alone did not influence self-assessed performance; anxiety also played a role. They found that females with high anxiety had lower levels of self-assessed performance in comparison to males with high anxiety, but no gender differences emerged for students with moderate or low anxiety. The present study found that sex did not influence perceived mastery. However, further investigation including measurement of anxiety would be beneficial.

Previous studies examining the influence of learning style upon self-assessed performance have focused on Marton and Saljo's (1976) deep and surface learning approaches. Deep learning involves seeking to understand and connect concepts, being critical and relating ideas to previous knowledge and experience. Surface learning, on the other hand, involves learning specifically to meet course requirements and memorising facts and figures to repeat. Marton and Sajlo have shown that students who use a surface learning approach reported lower levels of self-assessed performance compared to those that used the deep learning approach (Sadler-Smith et al., 2006; Cassidy, 2007). The learning approaches described by Marton and Saljo cannot be compared directly to Felder & Silverman's (1988) learning style profiles investigated in the present study. The present study found that learning style profiles did not influence self-assessed performance. Recall from chapter 3 that Felder and Silverman, (1998, p.675) suggested that teaching styles that are adapted to include both poles of the four learning style dimensions would be close to providing the optimal learning environment for most students in any given class. Therefore, the lack of an influence of learning profile could be, once again, regarded as a positive finding as it suggests that the measures taken to ensure that both teaching methods were adapted to include both poles of the four learning style dimensions (refer back to chapter 4.8) were successful.

Perceived mastery was lower for referral urgency selection compared to all other learning objectives in both the didactic lecture and virtual patient cohorts. Recall from chapter 5 that actual mastery was highest for referral urgency selection in the virtual patient cohort. Previous research has shown that students tend to under-estimate their confidence for tasks with a lower level of difficulty (Lichtenstein & Fischhoff, 1977; Baranski & Petrusic, 1994; Suantak et al., 1996; Yang et al., 2012). The findings of the present study appear to support this.

6.6 Summary

The VP raised perceived mastery for some of the learning objectives. That perceived mastery was lowest for referral urgency selection, for which actual mastery was highest (chapter 5), hinted that there might be an inverse relationship between both forms of mastery. Indeed, this had been reported before in the literature (Lichtenstein & Fischhoff, 1977; Baranski & Petrusic, 1994; Suantak et al., 1996; Yang et al., 2012). For this reason, the relationship between perceived and actual mastery is explored further in the next chapter.

Chapter 7 Evaluation of the teaching material: Relationship between perceived and actual mastery

7.1 Introduction

The relationship between perceived and actual mastery scores and factors that influenced this relationship are explored in this chapter.

7.2 Literature review

Self-assessment is considered an important aspect of medical education (Blanch-Hartigan, 2011) as it allows health professionals to understand their own strengths and weaknesses and helps to highlight areas they need to focus on to aid their development (Eva & Regehr, 2005). There have been many studies that have investigated the accuracy of self-assessment (Paradise et al., 1997; Young et al., 2002; Biernat et al., 2003; Barnsley et al., 2004; Woods et al., 2004; Leopold et al., 2005).

A meta-analysis by Blanch-Hartigan (2011) showed that previous studies had used different methods to evaluate the differences between self-assessed and actual performance. Three common methods were identified. The first method involved correlations between self-assessed scores and actual performance scores. This was the most commonly used method (Ward et al., 2002). Although this method easily indicates how well actual performance could be predicted from self-assessed scores, it provides no indication of the direction of the inaccuracy (i.e. was actual performance over or under-estimated by self-assessment). The second method, termed independent means comparisons, involved comparison of the mean of the self-assessed score and the mean of the actual performance score. Although this method provides an indication of the direction of the inaccuracy of the group as a whole, it does not consider students individually and so does not allow investigation of factors that could affect self-assessment accuracy such as gender or academic ability. The third method, termed paired comparisons, involved calculating the difference between self-assessed and actual performance scores for each student. This method provides an indication of the direction of any differences and also allows the investigation of factors that could affect self-assessment accuracy. Blanch-Hartigan (2011) suggested that using a combination of both correlations and paired comparisons was the ideal way to investigate differences between self-assessed and actual performance. Therefore, this methodology was adopted in the present study.

7.3 Learning objectives

Recall from chapter 2 that the 5 learning objectives were:

- The ability to ask the right questions in history and symptoms based upon a patient's presenting complaint (question selection);
- The ability to recognise symptoms of serious eye disease (critical symptom recognition);
- The ability to select the most appropriate tests needed to look for signs that aid differential diagnosis (test selection);
- The ability to recognise signs of serious eye disease (critical sign recognition);
- The ability to decide upon the most appropriate referral urgency for any eye disease detected (referral urgency selection).

These learning objectives were evaluated in two ways; actual mastery (see chapter 5) and perceived mastery (see chapter 6). Self-assessment accuracy was determined by subtracting the actual mastery score from the perceived mastery score. Therefore, a positive score indicated that perceived mastery was greater in value than actual mastery meaning that self-assessed accuracy is an over-estimate of the truth. Self-assessment accuracy was then classified into three groups; 'over-estimation' for scores greater than zero, 'under-estimation' for scores less than zero and 'accurate' for scores equal to zero.

7.4 Study sample findings

Appendix 9 shows the complete findings for all 102 students in the didactic lecture cohort and all 93 students in the virtual patient cohort. The distribution of the perceived and actual mastery scores for all 5 learning objectives is shown below (figure 7.1). One-sample Kolmogorov-Smirnov tests revealed that the scores for perceived and actual mastery did not follow a normal distribution. Various transformations were attempted but the data could not be transformed to follow a normal distribution. Therefore, only non-parametric tests were used to analyse the data.

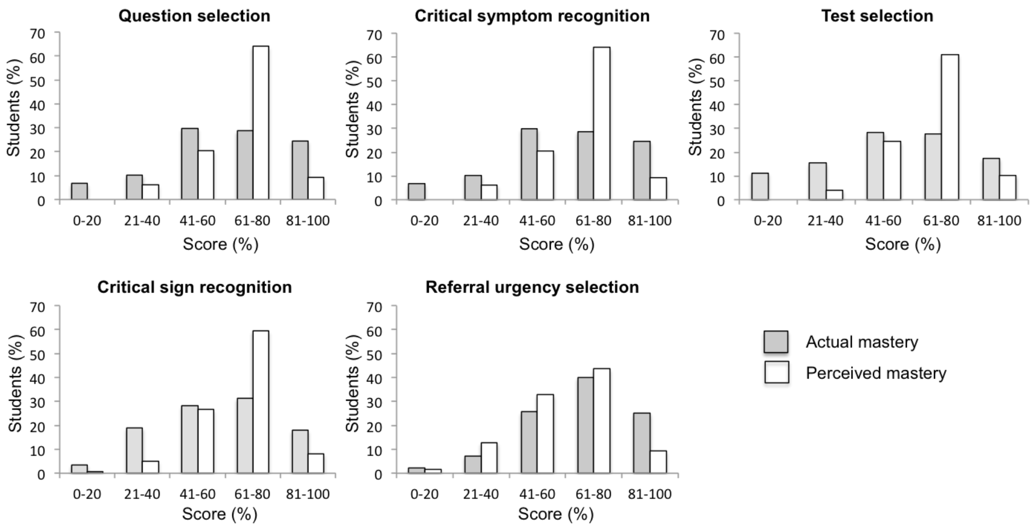

Figure 7.1 Distribution of perceived and actual mastery scores for the 5 learning objectives of 195 students pooled from the didactic lecture and virtual patient cohorts

Kendal's tau tests revealed that there were no statistically significant correlations between perceived and actual mastery of the 5 learning objectives in students from both the didactic lecture and virtual patient cohorts (table 7.1).

Table 7.1 Kendall's tau (r_t) correlations for perceived and actual mastery of all 195 students pooled from the didactic lecture and the virtual patient cohorts for each of the 5 learning objectives. Statistical power = 91.0%.

Learning objective	Kendall's tau
1	$r_t < 0.01$, $p = 0.97$
2	$r_t = 0.04$, $p = 0.54$
3	$r_t = 0.06$, $p = 0.35$
4	$r_t = 0.10$, $p = 0.11$
5	$r_t = -0.03$, $p = 0.71$

In order to investigate factors influencing self-assessment accuracy the paired comparisons method recommended by Blanch-Hartigan (2011) was used. Multivariate analysis was conducted in the form of DTA as factorial ANOVAs were unsuitable for the non-parametric data (see figures 7.2 to 7.5).

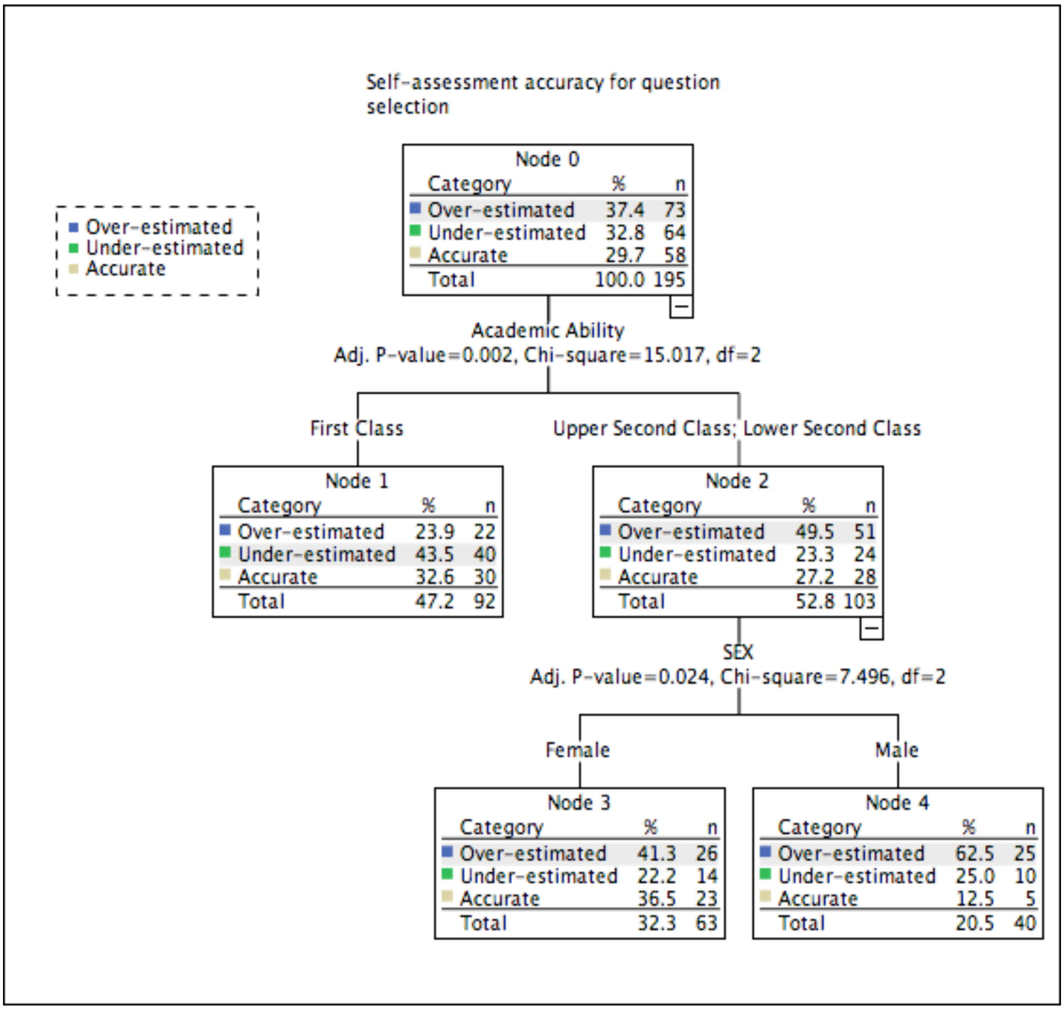

Figure 7.2 Decision tree analysis investigating the influences of teaching method, academic ability, sex and learning style profile on self-assessment accuracy of all 195 students pooled from the didactic lecture and the virtual patient cohorts for question selection.

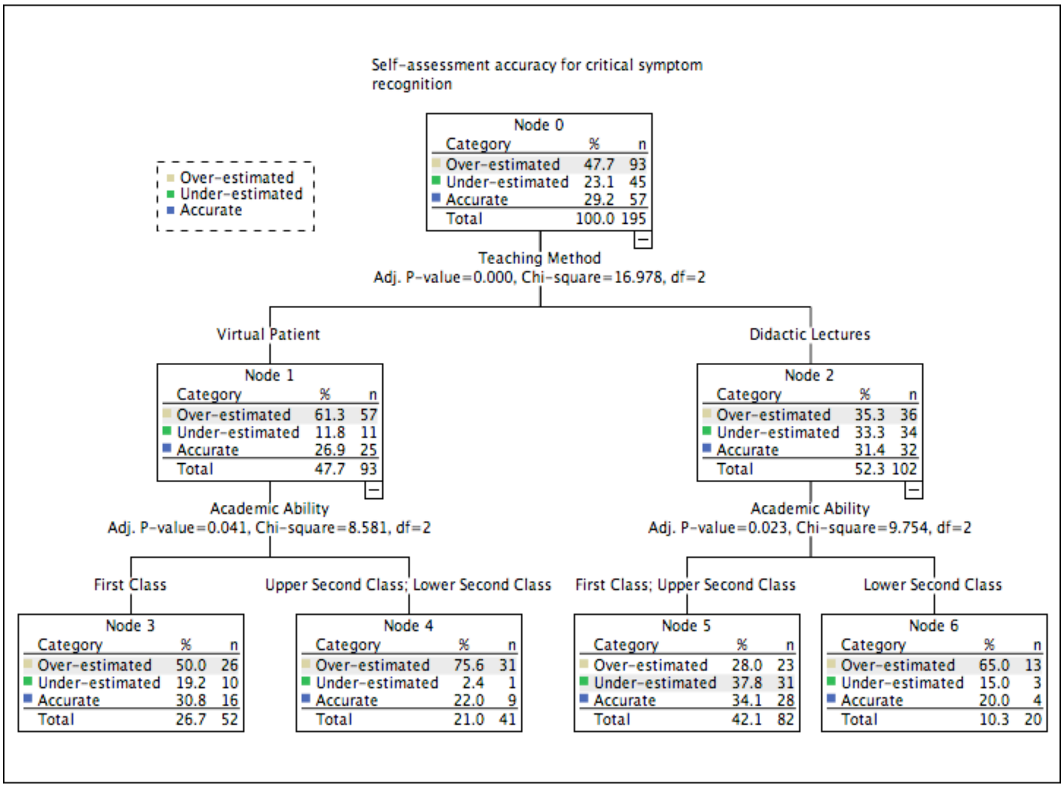

Figure 7.3 Decision tree analysis investigating the influences of teaching method, academic ability, sex and learning style profile on self-assessment accuracy of all 195 students pooled from the didactic lecture and the virtual patient cohorts for critical symptom recognition.

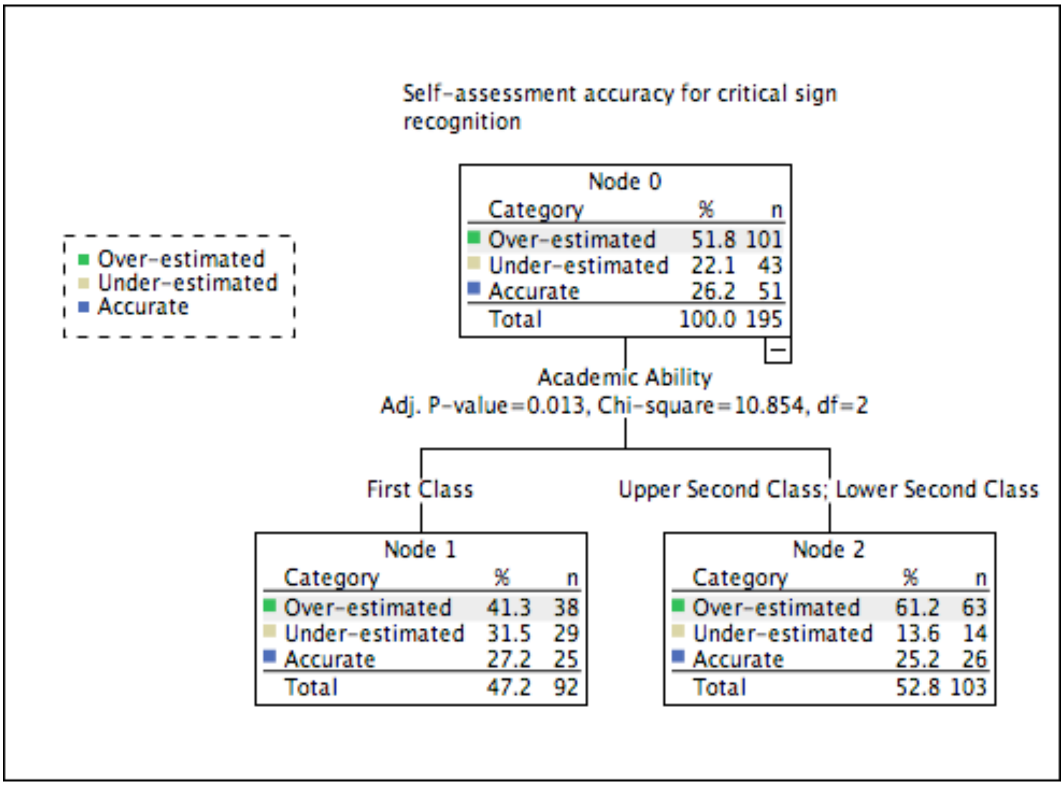

Figure 7.4 Decision tree analysis investigating the influences of teaching method, academic ability, sex and learning style profile on self-assessment accuracy of all 195 students pooled from the didactic lecture and the virtual patient cohorts for critical symptom recognition.

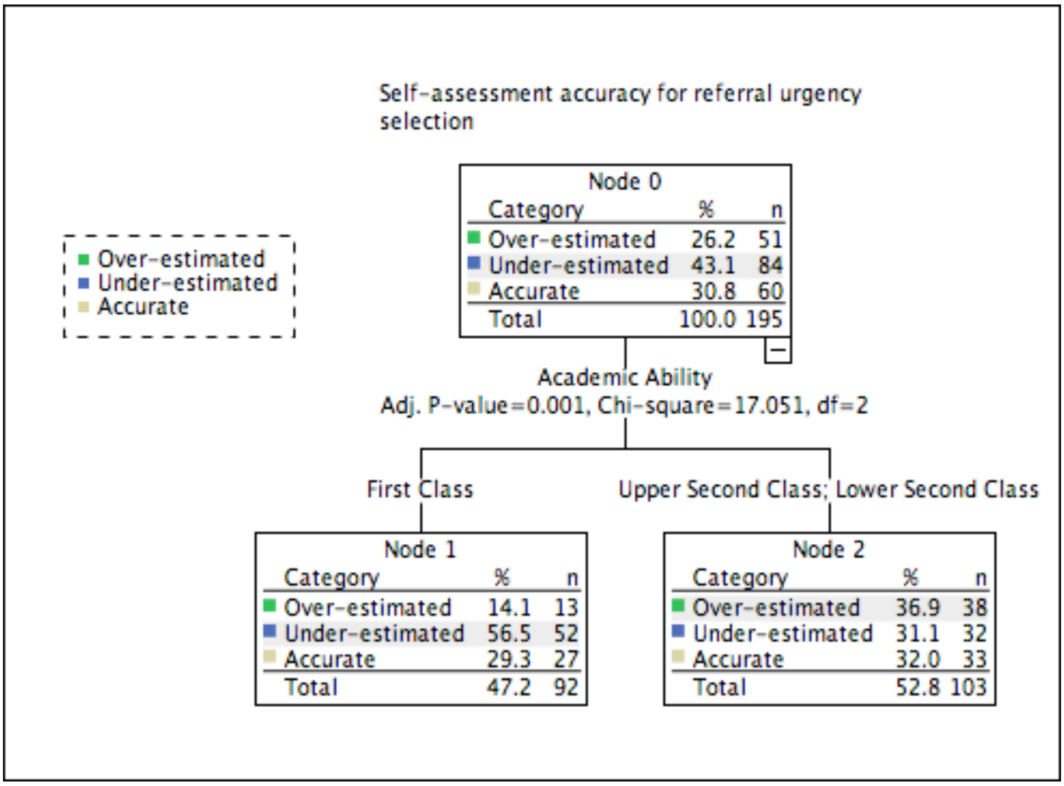

Figure 7.5 Decision tree analysis investigating the influences of teaching method, academic ability, sex and learning style profile on self-assessment accuracy of all 195 students pooled from the didactic lecture and the virtual patient cohorts for referral urgency selection.

Teaching method, academic ability, sex and learning style profile did not appear to have any influence on self-assessment accuracy of test selection (which is why a decision tree is not shown).

Academic ability influenced all other learning objectives (figures 7.2-7.5) where over-estimated self-assessment accuracy was lowest in the highest academic achievers.

Sex only appeared to influence self-assessment accuracy for question selection (figure 7.2) of mid to low academic achievers (Chi = 7.50, df = 2, p = 0.02), where over-estimated self-assessment accuracy occurred more often in males than females

Teaching method only appeared to influence self-assessment accuracy for critical symptom recognition (Chi = 16.98, df = 2, $p < 0.01$) where over-estimation was more common in the virtual patient cohort than in the didactic lecture cohort.

Table 7.2 shows the number of students from the didactic lecture cohort with over-estimated, accurate and under-estimated self-assessed accuracy across all 5 learning objectives. Application of Chi-square for R x C tables to the data shown in table 7.2 showed that levels of self-assessment accuracy did not differ between the 5 learning objectives (Chi = 5.98, df =8, p = 0.65, power = 97.1%).

Table 7.2 Table showing the number of students with over-estimated, accurate and under-estimated self-assessment accuracy across all 5 learning objectives for 102 students in the didactic lecture cohort.

	Question selection	Critical symptom recognition	Test selection	Critical sign recognition	Referral urgency selection
Over-estimated	40	36	41	48	34
Accurate	28	32	31	29	31
Under-estimated	34	34	30	25	37

Table 7.3 shows the number of students from the virtual patient cohort with over-estimated, accurate and under-estimated self-assessed accuracy across all 5 learning objectives. Application of Chi-square for R x C tables to the data shown in table 7.3 showed that levels of self-assessment accuracy varied significantly between the 5 learning objectives (Chi = 55.69, df =8, $p < 0.01$). Post hoc testing using Chi-square for 2 x 2 tables showed that referral urgency selection had a greater proportion of under-estimated self-assessment accuracy compared to question selection (Chi = 8.89, df =1, p = 0.01), critical symptom recognition (Chi = 44.26, df =1, $p < 0.01$), test selection (Chi = 17.78, df =1, $p < 0.01$) and critical sign recognition (Chi = 32.41, df =1, $p < 0.01$). Critical symptom recognition also had a greater proportion of over-estimated self-assessment accuracy compared to question selection (Chi = 15.66, df =1, p = 0.01) and test selection (Chi = 8.06, df =1, p = 0.02) while critical sign recognition had a greater proportion of over-

estimated self-assessment accuracy compared to question selection (Chi = 8.88, df =1, p = 0.01).

Table 7.3 Table showing the number of students with over-estimated, accurate and under-estimated self-assessment accuracy across all 5 learning objectives for 93 students in the virtual patient cohort.

	Question selection	Critical symptom recognition	Test selection	Critical sign recognition	Referral urgency selection
Over-estimated	33	57	43	53	17
Accurate	30	25	24	22	29
Under-estimated	30	11	26	18	47

7.5 Discussion

This study revealed that there were no correlations between the perceived mastery and the actual mastery scores for any of the 5 learning objectives. The findings of previous studies have been varied. While some have found no relationship between self-assessed and actual performance (Young et al., 2002; Barnsley et al., 2004), others have reported either a positive correlation (Biernat et al., 2003; Woods et al., 2004) or negative correlation (Paradise et al., 1997; Leopold et al., 2005). Blanch-Hartigan (2011) suggested that these inconsistencies might have been due to variations in individual student characteristics. Therefore, the present study explored the influence of teaching method, sex, learning style profile and academic ability on self-assessment accuracy.

Academic ability influenced self-assessment accuracy of almost all learning objectives, where high academic achievers were more accurate with self-assessment and low academic achievers had a greater tendency to over-estimate their perceived performance. This finding was consistent with previous research (Lichtenstein & Fischhoff, 1977; Robbins et al., 1994; Davis et al., 2006). This is an interesting finding, as it demonstrates that certain students may simply be under-achieving due to their inability to accurately self-assess their performance. It seems logical to assume that those students, who consistently over-estimate their self-assessed performance, would focus less on revision to improve these skills as they feel there is no need to do so. Therefore, actual performance would be reduced, as they have already wrongly perceived themselves to be better at a certain skill than they actually were. Many medical schools have already realised the importance of self-assessment skills and its link to actual performance, and so they have conducted a vast amount of research focused on teaching students to improve self-assessment skills (Lane & Gottlieb, 2004; Zick et al., 2007; Hawkins et al., 2012). This study has highlighted that Optometry schools would also benefit from educating their students how to improve their self-assessment skills.

Teaching method influenced self-assessment accuracy of critical symptom recognition. The VP had a greater proportion of students over-estimating their self-assessed performance compared to the didactic lectures cohort, while the didactic lectures cohort had a greater proportion of students who were accurate with their self-assessment.

There are two possible explanations for these differences. Firstly, recall from chapter 6 that teaching method had no influence on perceived mastery scores for this learning objective. Therefore, use of the VP did not falsely inflate perceived mastery. However, findings of chapter 5 revealed that students in the virtual patient cohort had reduced actual mastery scores for this learning objective. Although it seems that use of the VP led to a reduction in performance, and supports Greenhalgh's (2001) suggestion that CAL may just be an 'expensive disaster', this may not be the case, as it may have been due to inappropriate assessment criteria (see chapter 5). Further investigation would be required to determine the impact of the assessment criteria and methods on these findings.

Secondly, research has shown that by the time students have reached higher education, they have already developed the skills and knowledge regarding how best to adapt their learning approaches to the assessment methods that will be used (Scouller, 1998). Students were already familiar with the didactic teaching method and so had already adapted their learning styles to cope with transferring knowledge from this type of teaching material to assessments. The VP was a totally new teaching method and so did not have this benefit.

A greater proportion of females were accurate with their self-assessment while males had a tendency to over-estimate. This is in keeping with previous research (Beyer, 1990; Coutts & Rogers, 1999; Lind et al., 2002).

Learning style profile had no influence on self-assessment accuracy. Recall from chapter 3 that Felder and Silverman, (1998, p.675) suggested that teaching styles that are adapted to include both poles of the four learning style dimensions would be close to providing the optimal learning environment for most students in any given class. So, once again, the lack of any influence of learning style could be seen as a positive finding as it suggests that the measures taken to ensure that both teaching methods were adapted to include both poles of the four learning style dimensions (refer back to chapter 4.8) were successful.

Self-assessment accuracy only varied between the 5 learning objectives in the virtual patient cohort showing that referral urgency selection was the only learning objective where a greater proportion of students under-estimated their performance. Recall from chapter 6 that this difference may be explained by examining the differences between actual mastery scores for the learning objective. Findings of chapter 5 revealed that actual mastery was highest for referral urgency selection showing that this was the easiest learning objective. Previous research has shown that students tend to under-estimate their confidence for tasks with a lower level of difficulty (Lichtenstein & Fischhoff, 1977; Baranski & Petrusic, 1994; Suantak et al., 1996; Yang et al., 2012). Critical symptom and critical sign recognition differed from all other learning objectives where these two learning objectives had a greater proportion of students who over-estimated their performance. Recall from chapter 5, that it could be argued that good performance in these two learning objectives may have relied on simple memorisation of lists of signs and symptoms which was favoured more by the didactic lectures than the VP. Hence, this explains why a greater proportion of students in the virtual patient cohort wrongly over-estimated their performance for these learning objectives.

Self-assessment accuracy did not vary between the learning objectives for the didactic lectures cohort, where there was a similar proportion of students who were accurate or under or over-estimated their performance. This is an interesting finding as it suggests that choice of teaching method itself may have a significant impact on self-assessment accuracy. As discussed earlier, previous research has already demonstrated that self-assessment accuracy can influence actual performance. Therefore, these findings suggest that further investigation into how changes in teaching methods influence self-assessment accuracy would be beneficial to improving student learning and overall performance.

7.6 Summary

The findings of this chapter indicate that while there were no statistically significant correlations between perceived mastery and actual mastery, perceived mastery was higher than actual mastery for most learning objectives. It has highlighted that students' inability to accurately self-assess performance may be the cause of under-achievement and educating them on how to improve self-assessment accuracy may be the key to improving academic performance. Student satisfaction will be explored in the following chapter.

Chapter 8 Evaluation of teaching material: Student satisfaction

8.1 Introduction

The final measure used to evaluate the teaching materials in both cohorts was student satisfaction. Student satisfaction is evaluated and factors that influenced it are explored in this chapter.

8.2 Literature review

It has long been established that higher levels of student satisfaction with traditional teaching methods result in greater learning outcomes (Ramsden & Entwistle, 1981). Student satisfaction influences student motivation to continue interacting with the learning material of a course (Donohue & Wong, 1997) and motivation is considered a very important factor in academic success (Busato et al., 2000). Johnson (2006) established that this relationship also applies to online and CAL. Given this important link between student satisfaction and learning outcomes, it is essential that developers of any new teaching materials aim for high levels of student satisfaction (Chiu et al., 2007).

Student satisfaction is considered to be a complex concept which can be influenced by a number of factors (Saadé & Kira, 2006). It seems logical to assume that different factors are likely to have varying influences on different teaching methods. Therefore, the student satisfaction levels are compared between the two teaching methods used in this present study and the influences of sex, academic ability and learning style preference are investigated.

8.3 Data collection

Surveys to collect data for student satisfaction were released via the Blackboard learning management system two weeks before the end of the final academic term for both cohorts.

Student satisfaction in the didactic lecture cohort was determined via a 5-item Likert scale questionnaire, which comprised of the following 2 items:

1. I rate the CPD lectures given by Dr Dunne very highly (referred to as didactic lectures in this chapter)

2. I rate the CPD case record tutorials very highly (referred to as didactic tutorials in this chapter)

The following 5 Likert levels were used for each item:

1. Strongly agree
2. Agree
3. Neither agree or disagree
4. Disagree
5. Strongly disagree

Student satisfaction in the virtual patient cohort was determined via a 5-item Likert scale questionnaire, which comprised of the following item:

1. I rate the CPD simulator tutorials given by Dr Dunne very highly (referred to as virtual patient tutorials in this chapter)

The following 5 Likert levels were used:

1. Strongly agree
2. Agree
3. Neither agree or disagree
4. Disagree
5. Strongly disagree

These levels were converted to a score of 1 to 5 where a score of 1 corresponded to strongly agree and a score of 5 corresponded to strongly disagree.

8.4 Study sample findings

Appendix 9 shows the complete findings for all 102 students in the didactic lecture cohort and all 93 students in the virtual patient cohort. Figures 8.1 to 8.3 show the distribution of student satisfaction scores for the CPD lectures and tutorials in the didactic lecture cohort and the CPD simulator tutorials in the virtual patient cohort. One-sample Kolmogorov-Smirnov tests revealed that the scores did not follow a normal distribution. Various transformations were attempted but the data could not be transformed to follow a normal distribution. Therefore, only non-parametric tests were used to analyse the data.

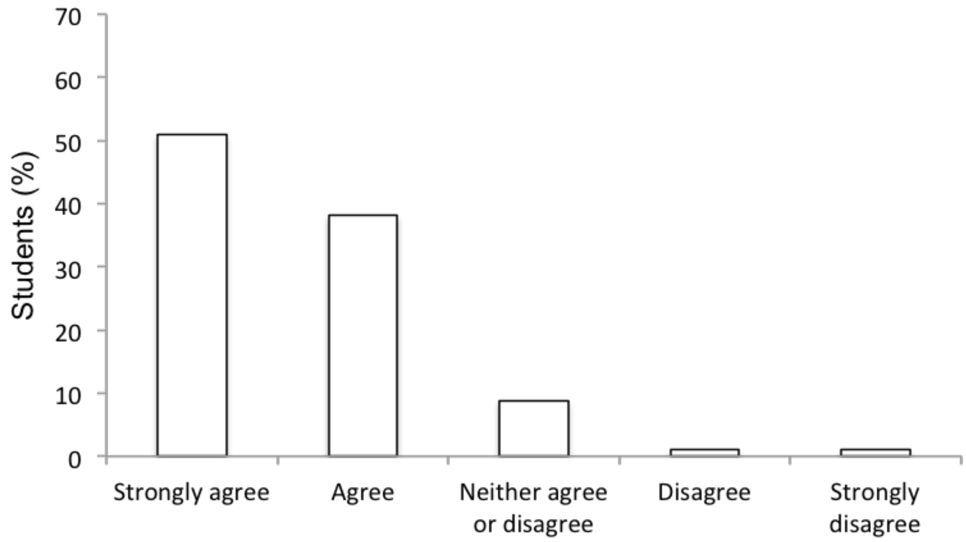

Figure 8.1 Distribution of student satisfaction scores for the CPD lectures of 102 students in the didactic lecture cohort (didactic lectures).

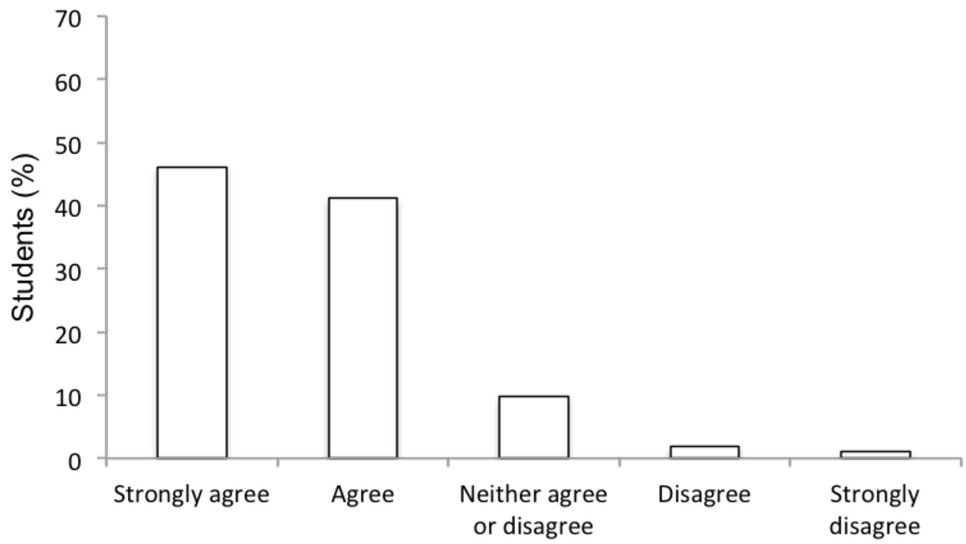

Figure 8.2 Distribution of student satisfaction scores for the CPD tutorials of 102 students in the didactic lecture cohort (didactic tutorials).

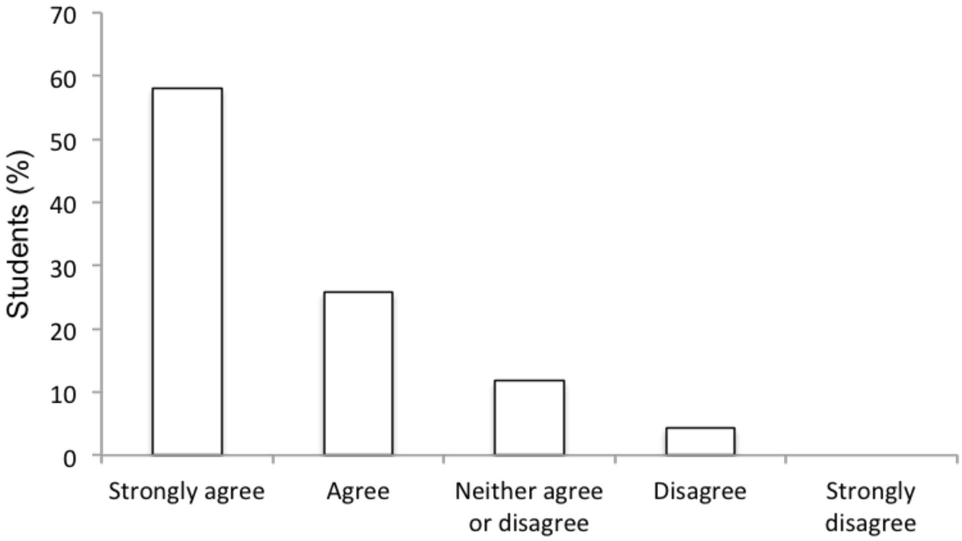

Figure 8.3 Distribution of student satisfaction scores for the CPD tutorials of 93 students in the virtual patient cohort (virtual patient tutorials).

Kruskal-Wallis test revealed that differences in median student satisfaction levels between didactic lectures, didactic tutorials and virtual patient tutorials were not statistically significant (Chi = 1.37, df = 2, p = 0.50, power = 95%).

DTA was used to investigate the influence of academic ability, sex and learning style profile upon student satisfaction scores for didactic lectures, didactic tutorials and virtual patient tutorials (see figure 8.4).

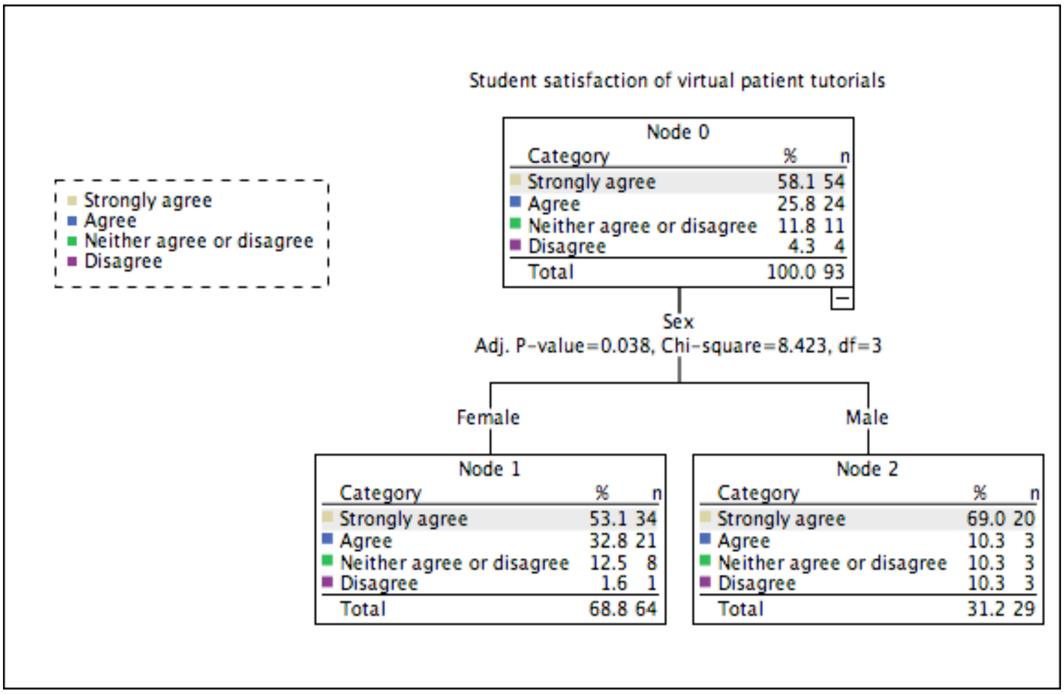

Figure 8.4 Decision tree analysis investigating the influences of academic ability, sex and learning style profile on student satisfaction scores of virtual patient tutorials of all 93 students in the virtual patient cohort.

Only sex had statistically significant influences on student satisfaction scores of virtual patient tutorials (Chi = 8.42, df = 3, p = 0.04) where a greater proportion of male students strongly agreed that they rated the teaching material highly.

Student satisfaction of didactic lectures and tutorials were not influenced by academic ability, sex and learning style profile.

8.5 Discussion

Student satisfaction levels did not differ significantly between the two teaching methods. This is consistent with a number of previous studies that compared satisfaction levels of online CAL and traditional lectures (Baumlin et al., 2000; Rose et al., 2000; Woo & Kimmick, 2000). However, a number of studies have found higher levels of student satisfaction with CAL compared to traditional lectures (Kallinowski et al., 1997; Bell et al., 2000; Fleetwood et al., 2000). A study by Sinclair and Ferguson (2009) found increased satisfaction when nursing students received traditional lectures combined with simulation. Given these varying findings, CAL can be considered a valuable addition to the teaching toolkit, but should not simply replace traditional methods (Chumley-Jones et al., 2002).

Student satisfaction levels for both teaching methods were very high with over 87% of students strongly agreeing or agreeing that they rated the teaching material very highly. This implied that the new CPD module was highly rated regardless of how it was taught.

Only sex influenced student satisfaction of the virtual patient tutorials where there was a significantly greater proportion of male students who strongly agreed that they rated the teaching material very highly. No gender differences were found for the didactic lecture cohort. Palmer and Holt (2009) found that a student's confidence and comfort level with using and interacting with technology was critical to satisfaction with online courses. Lower levels of computer self-confidence in females has still been reported by a number of recent studies (Birol et al., 2009; Mason et al., 2015) which may explain this finding.

Learning style preference did not influence student satisfaction scores for either teaching method. Recall from chapter 3 that Felder and Silverman, (1998, p.675) suggested that teaching styles that are adapted to include both poles of the four learning style dimensions would be close to providing the optimal learning environment for most students in any given class. So, once again, the lack of any influence of learning style could be seen as a positive finding as it suggests that the measures taken to ensure that both teaching methods were adapted to include both poles of the four learning style dimensions (refer back to chapter 4.8) were successful.

Academic ability also had no influence on student satisfaction scores for either teaching method. There has been very little research investigating the link between academic ability and student satisfaction of different teaching methods. Given that it has already been identified that greater student satisfaction is linked to greater learning outcomes, further research investigating whether a student's overall academic ability influences this relationship would be beneficial.

While the present study compared student satisfaction of the delivery of the two different teaching methods, further investigation into student satisfaction of the actual didactic lectures and the VP simulator itself would also have been beneficial.

8.6 Summary

Student satisfaction for both teaching methods was very high and did not differ. However, male students had higher levels of satisfaction with the virtual patient tutorials, indicating that a combination of traditional and VP teaching would be the best way to satisfy all types of students.

Limitations of the current study and future work will be discussed in the following chapter.

Chapter 9 – Summary and future work

9.1 Introduction

The key findings and the limitations of the present study are discussed and current and future research is described in this chapter.

9.2 Summary of findings, study limitations and future work

The effectiveness of the two teaching methods, didactic lectures and VP, was evaluated based upon differences in actual mastery, perceived mastery, self-assessment accuracy and student feedback.

9.2.1 Actual mastery

Teaching method influenced actual mastery of only two learning objectives; referral urgency selection and critical symptom recognition. Use of the VP appeared to raise actual mastery scores of referral urgency selection and reduce actual mastery scores of critical symptom recognition.

Recall from chapter 5 that these differences were believed to have been caused by choice of the assessment method rather than the teaching method itself. It is likely that the MCQs used to assess actual mastery (see tables 5.1 to 5.5 in chapter 5) relied on rote learning or memorisation of lists which was favoured by didactic lectures, hence, the use of the VP resulted in lower actual mastery scores for critical symptom recognition which relied heavily upon memorisation of lists of symptoms. It can be argued that referral urgency selection, on the other hand, relied more upon the ability to evaluate a multitude of presenting signs and symptoms along with the ability to correctly select the appropriate tests to aid diagnosis and then finally select the most appropriate management plan. Subtle differences in the presenting signs or symptoms could easily change the management plan; for example, the management of a “macula on” retinal detachment would be different to a “macula off” retinal detachment. It could be that the VP encouraged a deeper level of interaction while completing patient episodes, which might have led to the higher actual mastery scores for this learning objective.

Further investigation into the influence of assessment method is necessary to help establish whether the influence of teaching method upon the two learning objectives found in the present study was actually caused by the teaching method or as a result of the assessment method chosen.

There are two ways in which this theory could be further investigated. Firstly, complex MCQs aimed to assess deeper levels of CDM for each learning objective could be used. For example questions such as the following would assess deeper learning and break away from rote learning of lists.

"If a patient presented with sudden, recent onset, central vision loss with metamorphopsia, what would be the most appropriate management plan?"

a) *Immediate referral*
b) *Same day referral*
c) *Routine referral*
d) *Advise/treat*
e) *Soon referral*

Secondly, it would be beneficial to compare the learning outcomes (actual mastery scores) when two different assessment methods were used; complex MCQs and use of the VP in assessment mode. Recall from chapter 5 that a study by Rosenberg et al. (2010) who evaluated a new form of CAL in orthodontics highlighted how the choice of assessment method could influence the learning outcomes. They developed an electronic tutorial and then assessed students using a traditional paper based test and a computer based test, which were both matched for level of difficulty and topics covered. They found that females performed significantly better than males when examined with the traditional paper based test, but no gender differences occurred when examined with the electronic tutorial.

9.2.2 Perceived mastery

Use of the VP appeared to raise students' confidence for two learning objectives; question and test selection. This is consistent with previous literature that has shown that use of CAL and simulators can significantly increase student self-confidence (Engel et al., 1997; Valizadeh et al., 2013).

Interestingly, perceived mastery was lowest for referral urgency selection, for which actual mastery was highest (chapter 5). This hinted that there might be an inverse relationship between both forms of mastery. Therefore, the relationship between perceived and actual mastery was explored further (chapter 7).

9.2.3 Self-assessment accuracy

Teaching method influenced self-assessment accuracy of critical symptom recognition. The virtual patient cohort had a greater proportion of students over-estimating their self-assessed performance compared to the didactic lectures cohort, while the didactic lectures cohort had a greater proportion of students who were accurate with their self-assessment.

There are two possible explanations for these differences. Firstly, recall from chapter 6 that teaching method had no influence on perceived mastery scores for this learning objective. Therefore, use of the VP did not falsely inflate perceived mastery. However, findings of chapter 5 revealed that students in the virtual patient cohort had reduced actual mastery scores for this learning objective. Although it seems that use of the VP led to a reduction in performance, and supports Greenhalgh's (2001) suggestion that CAL may just be an 'expensive disaster', this may not be the case, as it may have been due to inappropriate assessment criteria (see chapter 5). As discussed earlier, further investigation would be required to determine the impact of the assessment criteria and methods on these findings.

Secondly, research has shown that by the time students have reached higher education, they have already developed the skills and knowledge regarding how best to adapt their learning approaches to the assessment methods that will be used (Scouller, 1998). Students were already familiar with the didactic teaching method and so had already adapted their learning styles to cope with transferring knowledge from this type of teaching material to assessments. The virtual patient cohort was exposed to a totally new teaching method and so did not have this advantage. Therefore, it would be beneficial to investigate whether these findings altered if more modules of the optometry curricula were taught using CAL over a prolonged period of time.

A very interesting finding was the influence of academic ability on self-assessment accuracy, which indicated that high academic achievers were more accurate with self-assessment and low academic achievers had a greater tendency to over-estimate their perceived performance. This highlighted that students' inability to accurately self-assess performance may be the cause of under-achievement and educating them on how to improve self-assessment accuracy may be the key to improving academic performance.

Previous research has shown that self-assessment is linked to learning outcomes (Dochy et al., 1999). Self-assessment increases student motivation (McMillan & Hearn, 2008) and this in turn results in greater commitment to the learning tasks which leads to improved learning outcomes (Longhurst & Norton, 1997). McMillan and Hearn (2008) highlighted that clear learning goals and objectives are essential in order to encourage accurate self-assessment.

It has been established that self-assessment improves with feedback (Dochy et al., 1999) and over time (Boud & Falchikov, 1989; Griffee, 1995; Birenbaum & Dochy, 2012; Hanley et al., 2014).

Many medical schools have already realised the importance of self-assessment skills and its link to performance, and so they have conducted a vast amount of research focused on teaching students to improve self-assessment skills (Lane & Gottlieb, 2004; Zick et al., 2007; Hawkins et al., 2012). This study has highlighted that Optometry schools would also benefit from educating their students on how to improve their self-assessment skills.

Fitzgerald et al. (2003) showed that with no interventions, the self-assessment skills of medical students remained stable over a two year period. However, they found that a change from the familiar classroom based examinations to different clinical examinations in the third year led to a drop in self-assessment accuracy implying that it was influenced by task familiarity.

Interventions such as videotaped clinical encounters have been used by medical and pharmacy schools and have proven to be an effective means of improving the self-assessment skills of both qualified healthcare professionals and students (Ward et al., 2003; Zick et al., 2007; Ozcakar et al., 2009; Mort & Hansen, 2010). Ward et al. (2003) videotaped surgeons performing a surgical technique and then obtained self-assessment scores before and after self-observation of their videotaped performance. There was a significant improvement in self-assessment accuracy post observation. Mort and Hansen (2010) suggested that all curricula should include opportunities for students to develop self-assessment skills early in the program, and this should be reinforced throughout the curriculum. This seems a logical suggestion as numerous studies have shown that self-assessment improves over time (Boud & Falchikov, 1989; Griffee, 1995; Birenbaum & Dochy, 2012; Hanley et al., 2014). Therefore, it seems that it is possible to train students to improve self-assessment accuracy if the right measures are used at the right stages. This in turn would improve students' academic performance.

Given these findings, it seems that Optometry undergraduates would also benefit from such interventions. Further investigation into the effects of videotaping students performing eye examinations on virtual, standardised or real patients, followed by the comparison of self-assessment scores pre and post videotape observation would be beneficial.

9.2.4 Student satisfaction

Student satisfaction for both teaching methods was very high and did not differ. However, male students had higher levels of satisfaction with the virtual patient tutorials, indicating that a combination of traditional and virtual teaching would be the best way to satisfy all types of students.

The questions to ascertain student satisfaction levels focused on the delivery of the two teaching methods. While this was an important factor and a key difference in the two teaching methods, further investigation into student satisfaction levels with the actual teaching material, the didactic lectures and the VP simulator, would have been beneficial.

9.3 General limitations of the present study

As described in chapter 2, academic ability was determined based on the overall average mark achieved by each student across all second year modules in sessional examinations carried out at the end of the academic year. This included the CPD module that included marks from the MCQs used to determine actual mastery in the present study. Therefore, in hindsight it would have been better to have based the determination of academic ability on the overall average mark across all modules from the end of the first year instead. However, the 25 MCQs used to determine actual mastery in the present study only accounted for 19% of the CPD module marks and therefore only accounted towards 3% of the overall determination of academic ability, as it was spread across the average of 6 modules. As academic ability was added as an influential variable to the DTAs used for all the analyses, the design of the DTAs would have compensated for this potential confounding.

Another limitation of the present study was the difference in the spread of academic ability between the two cohorts; didactic lectures and virtual patient. There were a greater proportion of higher academic achievers in the virtual patient cohort. But once again, the design of the DTAs would have compensated for this potential confounding as academic ability was entered as an influencing variable.

9.4 Summary

This chapter has included detail of how to enhance future research in this area along with the summary of the data and the conclusions drawn from the present study. The thesis contributes to an area of research that has previously received little attention in Optometry teaching and has highlighted how future research in improving self-assessment skills of students could be the key to improving academic performance.

References:

Abdel-Hameed, A. A., Al-Faris, E. A., Alorainy, I. A., & Al-Rukban, M. O. (2005). The criteria and analysis of good multiple choice questions in a health professional setting. *Saudi medical journal, 26*(10), 1505-1510.

Allinson, C. W., & Hayes, J. (1990). Validity of the Learning Styles Questionnaire. *Psychological Reports, 67*, 859-866.

Allinson, C. W., & Hayes, J. (1996). The Cognitive Style Index. *The Journal of Management Studies, 33*(119-135).

Araujo, P., & Froyland, L. (2007). Statistical power and analytical quantification. *Journal of Chromatography B, 847*(2), 305-308.

Bailey, K. (2008). *Methods of social research* (4th Edition ed.): Simon and Schuster.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics, 55*(4), 412-428.

Barnsley, L., Lyon, P. M., Ralston, S. J., Hibbert, E. J., Cunningham, I., Gordon, F. C., & Field, M. J. (2004). Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. *Medical Education, 38*(4), 358-367.

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach to medical education*: Springer Publishing Company.

Baumlin, K. M., Bessette, M. J., Lewis, C., & Richardson, L. D. (2000). EMCyberSchool An Evaluation of Computer-assisted Instruction on the Internet. *Academic Emergency Medicine, 7*(8), 959-962.

Bell, D. S., Fonarow, G. C., Hays, R. D., & Mangione, C. M. (2000). Self-study from web-based and printed guideline materials: a randomized, controlled trial among resident physicians. *Annals of Internal Medicine, 132*(12), 938-946.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology, 59*(5), 960.

Bezan, D., LaRussa, F. P., Nishimoto, J. H., Sendrowski, D. P., Spear, C. H., Talley, D. K., & Than, T. P. (1999). *Differential diagnosis in primary eye care*. Boston: Butterworth-Heinemann.

Biernat, K., Simpson, D., Duthie Jr, E., Bragg, D., & London, R. (2003). Primary care residents self assessment skills in dementia. *Advances in Health Sciences Education, 8*(2), 105-110.

Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology, 63*(1), 3-19.

Birenbaum, M., & Dochy, F. (2012). *Alternatives in assessment of achievements, learning processes and prior knowledge* (Vol. 42): Springer Science & Business Media.

Birol, C., Bekirogullari, Z., Etci, C., & Dagli, G. (2009). Gender and Computer Anxiety, Motivation, Self-Confidence, and Computer Use. *Eurasian Journal of Educational Research*(34), 185-198.

Blanch-Hartigan, D. (2011). Medical students' self-assessment of performance: results from three meta-analyses. *Patient Education and Counseling, 84*(1), 3-9.

Bokoros, M. A., Goldstein, M. B., & Sweeney, M. M. (1992). Common factors in five measures of cognitive style. *Current Psychology: Research & Reviews, 11*(2), 99-109.

Bostrom, R. P., Olfman, L., & Sein, M. K. (1993, p.119). Learning Styles and End-User Training: A First Step. *MIS Quarterly, 17*(1), 118-120.

Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education, 18*(5), 529-549.

Bracewell, R. M., & Anderson, I. R. C. (2003). Eye simulator. *Journal of Neurology, Neurosurgery and Psychiatry, 74*, 1573.

Brown, T., Zoghi, M., Williams, B., Jaberzadeh, S., Roller, L., Palermo, C., Holt, A. (2009). Are learning style preferences of health science students predictive of their attitudes towards e-learning? *Australasian Journal of Educational Technology, 25*(4), 524-543.

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

Carrió, M., Larramona, P., Banos, J., & Pérez, J. (2011). The effectiveness of the hybrid problem-based learning approach in the teaching of biology: a comparison with lecture-based learning. *Journal of Biological Education, 45*(4), 229-235.

Cassidy, S. (2007). Assessing 'inexperienced'students' ability to self-assess: Exploring links with learning style and academic personal control. *Assessment & Evaluation in Higher Education, 32*(3), 313-330.

Chang, C.-H., Yang, C.-Y., See, L.-C., & Lui, P.-W. (2004). High satisfaction with problem-based learning for anesthesia. *Chang Gung medical journal, 27*(9), 654-662.

Chiu, C.-M., Sun, S.-Y., Sun, P.-C., & Ju, T. L. (2007). An empirical analysis of the antecedents of web-based learning continuance. *Computers & Education, 49*(4), 1224-1245.

Chumley-Jones, H. S., Dobbie, A., & Alford, C. L. (2002). Web-based learning: Sound educational method or hype? A review of the evaluation literature. *Academic Medicine, 77*(10), 86-93.

Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning: a systematic and critical review*. London, UK: Learning and Skills Research Centre

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second ed.). New Jersey: Lawrence Erlbaum Associates Inc. Publishers.

Colbert-Getz, J. M., Fleishman, C., Jung, J., & Shilkofski, N. (2013). How Do Gender and Anxiety Affect Students' Self-Assessment and Actual Performance on a High-Stakes Clinical Skills Examination? *Academic Medicine, 88*(1), 44-48.

Collins, K. M. (2010). *Toward a broader understanding of stress and coping: Mixed methods approaches.* North Carolina: Information Age Publishing Inc..

Conrick, M. (1996). Clinical decision making: Issues in teaching. *The Australian Electronic Journal of Nursing Education, 2*(1).

Cornwell, J. M., Manfredo, P. A., & Dunlap, W. P. (1991). Factor analysis of the 1985 revision of Kolb's Learning Style Inventory. *Educational and Psychological Measurement, 51*.

Coutts, L., & Rogers, J. (1999). Predictors of student self-assessment accuracy during a clinical performance exam: comparisons between over-estimators and under-estimators of SP-evaluated performance. *Academic Medicine, 74*(10 Suppl), 128-130.

Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive cognitive bias. *Journal of Happiness studies, 3*(1), 37-69.

Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *Jama, 296*(9), 1094-1102.

DeCantis, S. M., & Kirton, M. J. (1996). A psychometric re-examination of Kolb's experiential learning cycle construct: a separation of level, style and process. *Educational and Psychological Measurement, 56*, 809-820.

Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology, 40*(2), 203-214.

Desmedt, E., & Valcke, M. (2003). *Learning style awareness: why would it work?* Paper presented at the Eighth Annual European Learning Styles Information Network Conference, Hull.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education, 24*(3), 331-350.

Donohue, T. L., & Wong, E. H. (1997). Achievement motivation and college satisfaction in traditional and nontraditional students. *Education, 118*(2), 237.

Drummond, R. J., & Stoddard, A. H. (1992). Learning style and personality type. *Perceptual and Motor Skills, 75*, 99-104.

Duff, A., & Duffy, T. (2002). Psychometric properties of Honey & Mumford's Learning Styles Questionnaire (LSQ). *Personality and Individual Differences, 33*(1), 147-163.

Dunn, R., & Dunn, K. (1978). *Teaching students through their individual learning styles: A practical approach*. Reston, VA: Reston Publishing Company.

Elson, R. B., & Connelly, D. P. (1997). *The impact of anticipatory patient data displays on physician decision making: a pilot study.* Paper presented at the Proceedings of the AMIA Annual Fall Symposium.

Engel, S. S., Crandall, J., Basch, C. E., Zybert, P., & Wylie-Rosett, J. (1997). Computer-Assisted Diabetes Nutrition Education Increases Knowledge and Self-Efficacy of Medical Students. *The Diabetes Educator, 23*(5), 545-549.

Eperjesi, F., Bartlett, H., & Dunne, M. C. (2007). *Ophthalmic Clinical Procedures: A Multimedia Guide*: Elsevier Health Sciences.

Eva, K. W., & Regehr, G. (2005). Self-assessment in the health professions: a reformulation and research agenda. *Academic Medicine, 80*(10), 46-54.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Felder, R. M. (1993). Reaching the second tier: learning and teaching styles in college science education. *College Science Teaching, 23*(5), 286-290.

Felder, R. M., & Silverman, L. K. (1988). Learning and Teaching Styles in Engineering Education. *Engineering Education, 78*(7), 674-681.

Felder, R. M., & Soloman, B. A. (2003). Index of learning styles questionnaire. Retrieved from http://www.ncsu.edu/felder-public/ILSdir/ilsweb.html 27/08/10

Felder, R. M., & Soloman, B. A. (2004). Index of learning styles. Retrieved from http://www.ncsu.edu/felderpublic/ILSpage.html Accessed 22/12/08

Felder, R. M., & Spurlin, J. (2005). Applications, reliability and validity of the Index of Learning Styles. International *Journal on Engineering Education, 21*(1), 103-112.

Findl, O. (1998). Vienna Ophthalmic Case Studies. Available at: http://www.meduniwien.ac.at/augen/englisch/steps.htm. Retrieved from http://www.meduniwien.ac.at/augen/englisch/steps.htm

Fitzgerald, J. T., White, C. B., & Gruppen, L. D. (2003). A longitudinal study of self-assessment accuracy. *Medical Education, 37*(7), 645-649.

Fleenor, J. W. (2001). Review of the Mypers-Briggs Type Indicator, Form M. In J. C. Impara & B. S. Blake (Eds.), *The fourteenth mental measurements yearbook*. Lincoln: Buros Institute of Mental Measurement.

Fleetwood, J., Vaught, W., Feldman, D., Gracely, E., Kassutto, Z., & Novack, D. (2000). MedEthEx Online: a computer-based learning program in medical ethics and communication skills. *Teaching and Learning in Medicine, 12*(2), 96-104.

Fleming, N., & Mills, C. (1992). Not another inventory, rather a catalyst for reflection. *Improve the Academy, 11*, 137.

Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological bulletin, 138*(2), 296.

Furnham, A., & Stringfield, P. (1994). Congruence of self and subordinate ratings of managerial practices as a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology, 67*(1), 57-67.

Geller, L. M. (1979). Reliability of the Learning Style Inventory. *Psychological Reports, 44*, 555-561.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Boston: Allyn & Bacon.

Girelli, S. A., & Stake, J. (1993). Bipolarity of Jungian type theory and the Myers-Briggs Type Indicator. *Journal of Personality Assessment, 60*(2), 290-301.

Glittenberg, C., & Binder, S. (2006). Using 3D computer simulations to enhance ophthalmic training. *Ophthalmic Physiological Optics, 26*(1), 40-49.

Grasha, A. F. (1972). Observations on relating teaching goals to students response styles and classroom methods. *American Psychologist, 27*, 144-147.

Greenhalgh, T. (2001). Computer assisted learning in undergraduate medical education. *BMJ (Clinical research ed.), 322*(7277), 40-44.

Gregorc, A. F. (1982). *Mind styles model: Theory, principles and applications*. Maynard, MA: Gabriel Systems.

Griffee, D. T. (1995). A Longitudinal Study of Student Feedback: Self-Assessment, Course Evaluation and Teacher Evaluation.

Hanley, K., Zabar, S., Charap, J., Nicholson, J., Disney, L., Kalet, A., & Gillespie, C. (2014). Self-assessment and goal-setting is associated with an improvement in interviewing skills. *Medical Education Online, 19*.

Harasym, P. H., Leong, E. J., Juschka, B. B., Lucier, G. E., & Lorsheider, F. L. (1996). Relationship between Myers-Briggs Type Indicator and Gregorc Style Delineator. *Perceptual and Motor Skills, 82*, 1203-1210.

Hargittai, E., & Shafer, S. (2006). Differences in actual and perceived online skills: the role of gender*. *Social Science Quarterly, 87*(2), 432-448.

Harter, S. P. (1996). The Impact of Electronic Journals on Scholarly Communication: A Citation Analysis. *The Public-Access Computer Systems Review, 7*(5).

Hawkins, S. C., Osborne, A., Schofield, S. J., Pournaras, D. J., & Chester, J. F. (2012). Improving the accuracy of self-assessment of practical clinical skills using video feedback--the importance of including benchmarks. *Medical Teacher, 34*(4), 279-284.

Hayes, J., & Allinson, C. W. (1993). Matching learning style and instructional strategy: an application of the person-environment interaction paradigm. *Perceptual and Motor Skills, 76*, 63-79.

Hayes, J., & Allinson, C. W. (1996). The implications of learning styles for training and development: a discussion of the matching hypothesis. *British Journal of Management, 7*, 63-73.

Henson, R. K., & Hwang, D. (2002). Variability and Prediction of Measurement Error in Kolb's Learning Style Inventory Scores: A Reliability Generalization Study. *Educational and Psychological Measurement, 62*(4), 712-727.

Hicks, L. E. (1984). Conceptual and empirical analysis of some assumptions of an explicit typological theory. *Journal of Personality and Social Psychology, 46*, 1118-1131.

Hill, T., Smith, N. D., & Hoffman, H. (1988). Self-image bias and the perception of other persons' skills. *European Journal of Social Psychology, 18*(3), 293-298.

Hingorjo, M. R., & Jaleel, F. (2012). Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. *Journal Of Pakistan Medical Association, 62*(2), 142-147.

Holling, H., & Preckel, F. (2005). Self-estimates of intelligence—methodological approaches and gender differences. *Personality and Individual Differences, 38*(3), 503-517.

Honey, A. C., & Mumford, A. (2000). *The learning styles helper's guide*. Maidenhead: Peter Honey Publications Ltd.

Honey, M. (2002). *Ecotourism & certification: Setting standards in practice*: Island Press.

Honey, P., & Mumford, A. (1982). *The manual of learning styles*. Maidenhead, Berkshire: Peter Honey.

Howe, K. R., Holmes, M., & Elstein, A. S. (1984). Teaching clinical decision making. *Journal of Medicine and Philosophy, 9*(2), 215-228.

Hwang, S. Y., & Kim, M. J. (2006). A comparison of problem-based learning and lecture-based learning in an adult health nursing course. *Nurse education today, 26*(4), 315-321.

IBM. (Version 21.0 2012). IBM SPSS Statistics for Macintosh (Version 21.0 Released 2012). Armonk, NY: IBM Corp.

Jackson, T. L. (2007). *Moorfields Manual of Ophthalmology*: Elsevier Health Sciences.

Jenkins, S., Goel, R., & Morrell, D. S. (2008). Computer-assisted instruction versus traditional lecture for medical student teaching of dermatology morphology: A randomized control trial. *Journal of American Academy of Dermatology, 59*(2), 255-259.

John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology, 66*, 206-219.

Johnson, G. M. (2006). College Student Psycho-Educational Functioning and Satisfaction with Online Study Groups. *Educational Psychology, 26*(5), 677-688.

Jones, R., Higgs, R., De Angelis, C., & Prideaux, D. (2001). Changing face of medical curricula. *The Lancet, 357*(9257), 699-703.

Joniak, A. J., & Isaksen, S. G. (1988). The Gregorc Style Delineator: internal consistency and its relationship to Kirton's adaptive-innovative distinction. *Educational and Psychological Measurement, 48*, 1043-1049.

Jung, C. G. (1921). Psychological Types, Vol. 6, The collected works of C. G. *Jung*.

Kallinowski, F., Mehrabi, A., Glückstein, C., Benner, A., Lindinger, M., Hashemi, B., . . . Herfarth, C. (1997). Computer-based training: a new method of surgical education. *Der Chirurg, 68*(4), 433-438.

Klemenc-Ketis, Z., & Kersnik, J. (2014). Deficiency areas in decision making in undergraduate medical students. *Advances in Medical Education and Practice, 5*, 223-227.

Kneebone, R. (2005). Evaluating Clinical Simulations for Learning Procedural Skills: A Theory-Based Approach. *Academic Medicine, 80*(6), 549-553.

Kneebone, R., Kidd, J., Nestel, D., Asvall, S., Paraskeva, P., & Darzi, A. (2002). An innovative model for teaching and learning clinical procedures. *Medical Education, 36*, 628-634.

Kolb, A. Y. (2005). The Kolb learning style inventory—version 3.1 2005 technical specifications. *Boston, MA: Hay Resource Direct, 200*.

Kolb, D. A. (1984). *Experiential learning: experience as the source of learning and development*. Englewood Cliffs, New Jersey: Prentice Hall.

Kolb, D. A. (2000). *Facilitator's Guide to Learning*. Boston: Hay/McBer.

Lane, J. L., & Gottlieb, R. P. (2004). Improving the Interviewing and Self-Assessment Skills of Medical Students: Is it Time to Readopt Videotaping as an Educational Tool? *Ambulatory Pediatrics, 4*(3), 244-248.

Lasslo, R., Keltner, J., & G., H. (2002). *Virtual patient/Eye simulator - Evolution of a futuristic teaching tool.* Paper presented at the AMIA Annual Conference, San Antonio.

Lawrence, M. V. (1997). Secondary school teachers and learning style preferences: action or watching in the classroom? *Educational Psychology, 17*(1/2), 157-170.

Leopold, S. S., Morgan, H. D., Kadel, N. J., Gardner, G. C., Schaad, D. C., & Wolf, F. M. (2005). Impact of educational intervention on confidence and competence in the performance of a simple surgical task. *The Journal of Bone & Joint Surgery, 87*(5), 1031-1037.

Levinson-Rose, J., & Menges, R. J. (1981). Improving college teaching: A critical review of research. *Review of educational research, 51*(3), 403-434.

Lewin, K. (1951). *Field theory in social sciences*. New York: Harper and Row.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational behavior and human performance, 20*(2), 159-183.

Lind, D. S., Rekkas, S., Bui, V., Lam, T., Beierle, E., & Copeland, E. M., 3rd. (2002). Competency-based student self-assessment on a surgery rotation. *Journal of Surgical Research, 105*(1), 31-34.

Litzinger, T. A., Lee, S. H., Wise, J. C., & Felder, R. M. (2005). *A study of the reliability and validity of the Felder-Soloman Index of Learning Styles.* Paper presented at the Annual Conference of American Society of Engineering Education.

Liu, M., Hsieh, P., Cho, Y., & Schallert, D. (2006). Middle school students' self-efficacy, attitudes, and achievement in a computer-enhanced problem-based learning environment. *Journal of Interactive Learning Research, 17*(3), 225-242.

Livesay, G. A., Dee, K. C., Nauman, E. A., & Hites, L. S. (2002). *Engineering student learning styles: a statistical analysis using Felder's Index of Learning Styles*. Paper presented at the Annual Conference of Americal Society of Engineering Education, Montreal, Quebec.

Loewenstein, J., & Lee, S. (2004). *Ophthalmology: Just the facts*. New York: McGraw-Hill.

Longhurst, N., & Norton, L. S. (1997). Self-assessment in coursework essays. *Studies in Educational Evaluation, 23*(4), 319-330.

Loo, R. (1997). Evaluating change and stability in learning style scores: a methodological concern. *Educational Psychology, 12*(1 and 2), 95-100.

Maleck, M., Fischer, M. R., Kammer, B., Zeiler, C., Mangel, E., Schenk, F., & Pfeifer, K. (2001). Do Computers Teach Better? A Media Comparison Study for Case-based Teaching in Radiology. *Radiographics, 21*, 1025-1032.

Martin, L., West, J., & Bill, K. (2008). Incorporating problem-based learning strategies to develop learner autonomy and employability skills in sports science undergraduates. *Journal of Hospitality, Leisure, Sport & Tourism Education., 7*(1), 18-30.

Marton, F., & Saljo, R. (1976). On qualitative differences in learning-I: Outcome and Process. *British Journal of Educational Psychology, 46*(4-11).

Mason, R., Cooper, G., Comber, T., Hellou, A., & Tucker, J. (2015). Gender differences in experiences of TAFE IT students: a work in progress.

Mattheos, N., Nattestad, A., Schittek, M., & Attstrom, R. (2001). A virtual classroom for undergraduate periodontology: a pilot study. *European Journal of Dental Education, 5*, 139-147.

McCaulley, M. H. (1981). Jung's theory of psychological types and the Myers-Briggs Type Indicator. In P. McReynolds (Ed.), *Advances in personality assessment* (Vol. 5, pp. 294-352). San Francisco: Josey-Bass.

McCrae, R. R., & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the Five Factor Model of Personality. *Journal of Personality, 57*(1), 17-37.

McGraw, R. C., & O'connor, H. M. (1999). Standardized patients in the early acquisition of clinical skills. *Medical Education, 33*(8), 572-578.

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons, 87*(1), 40-49.

Metallidou, P., & Platsidou, M. (2008). Kolb's Learning Style Inventory-1985: Validity issues and relations with metacognitive knowledge about problem-solving strategies. *Learning and Individual Differences, 18*(1), 114-119.

Microsoft. (2010). Microsoft PowerPoint (Version 14.0): Microsoft.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*(9), S63-67.

Morales-Mann, E. T., & Kaitell, C. A. (2001). Problem-based learning in a new Canadian curriculum. *Journal of Advanced Nursing, 33*(1), 13-19.

Mort, J. R., & Hansen, D. J. (2010). First-year Pharmacy Students' Self-Assessment of Communication Skills and the Impact of Video Review. *American journal of pharmaceutical education, 74*(5), 78.

Myers, I. B., & McCaulley, M. H. (1985). *Manual: a guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto: Consulting Psychologists Press.

Myers, I. B., & McCaulley, M. H. (1989). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Nango, E., & Tanaka, Y. (2010). Problem-based learning in a multidisciplinary group enhances clinical decision making by medical students: a randomized controlled trial. *Journal of Medical and Dental Science, 57*(1), 109-118.

Noohi, E., Karimi-Noghondar, M., & Haghdoost, A. (2012). Survey of critical thinking and clinical decision making in nursing student of Kerman University. *Iranian Journal of Nursing and Midwifery Research, 17*(6), 440-444.

Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (Third ed.): McGraw-Hill Humanities.

O'Brien, T. P. (1990). Construct validation of Gregorc Style Delineator: an application of LISREL 7. *Educational and Psychological Measurement, 50*, 631-636.

Oja, K. J. (2011). Using problem-based learning in the clinical setting to improve nursing students' critical thinking: an evidence review. *Journal of Nursing Education, 50*(3), 145-151.

Ozcakar, N., Mevsim, V., Guldal, D., Gunvar, T., Yildirim, E., Sisli, Z., & Semin, I. (2009). Is the use of videotape recording superior to verbal feedback alone in the teaching of clinical skills? *BMC Public Health, 9*, 474.

Palmer, S. R., & Holt, D. M. (2009). Examining student satisfaction with wholly online learning. *Journal of Computer Assisted Learning, 25*(2), 101-113.

Pane, A., & Simcock, P. (2005). *Practical ophthalmology: a survival guide for doctors and optometrists*: Elsevier Health Sciences.

Paradise, J. E., Finkel, M. A., Beiser, A. S., Berenson, A. B., Greenberg, D. B., & Winter, M. R. (1997). Assessments of girls' genital findings and the likelihood of sexual abuse: agreement among physicians self-rated as skilled. *Archives of pediatrics & adolescent medicine, 151*(9), 883-891.

Parker, M. J., & Seifter, J. L. (2001). An interactive, web-based learning environment for pathophysiology. *Academic Medicine, 76*, 550.

Pedersen, S. (2003). Motivational orientation in a problem-based learning environment. *Journal of Interactive Learning Research, 14*(1), 51-77.

Peitzman, S. J., Nieman, L., & Gracely, E. (1990). Comparison of 'fact-recall' with 'higher-order' questions in multiple-choice examinations as predictors of clinical performance of medical students. *Academic Medicine, 65*(9), 59-60.

Penn, B. K. (2008). *Mastering the teaching role: A guide for nurse educators*: FA Davis.

Prajapati, B., Dunne, M., & Armstrong, R. A. (2010). Sample size estimation and statistical power anlaysis. *Optometry Today*. Retrieved from http://www.optometry.co.uk/articles/docs/STATISTICAL ARTICLE.pdf16/07/2010

Prajapati, B., Dunne, M., Bartlett, H., & Cubbidge, R. (2011). The influence of learning styles, enrolment status and gender on academic performance of optometry undergraduates. *Ophthalmic and Physiological Optics, 31*(1), 69-78.

Ramsden, P. (1983). Context and strategy: situational influences on learning. In N. Entwistle & P. Ramsden (Eds.), *Understanding student learning*. London: Croom Helm.

Ramsden, P., & Entwistle, N. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology, 51*, 368-383.

Ridgman, W. J. (1975). *Experimentation in biology*: Blackie.

Robbins, J. M., Kirmayer, L. J., Cathebras, P., Yaffe, M. J., & Dworkind, M. (1994). Physician characteristics and the recognition of depression and anxiety in primary care. *Medical Care, 32*(8), 795-812.

Rogers, D. A., Regehr, G., Yeh, K. A., & Howdieshell, T. R. (1998). Computer-assisted Learning versus a Lecture and Feedback Seminar for Teaching a Basic Surgical Technical Skill. *American Journal of Surgery, 175*(6), 508-510.

Rosati, P. A., & Felder, R. M. (1995). *Engineering student responses to an Index of Learning Styles.* Paper presented at the ASEE Annual Conference, Washington DC.

Rose, M. A., Frisby, A. J., Hamlin, M. D., & Jones, S. S. (2000). Evaluations of the effectiveness of a web-based graduate epidemiology course. *Computers in Nursing, 18*(4), 162-167

Rosenberg, H., Posluns, J., Tenenbaum, H. C., Tompson, B., & Locker, D. (2010). Evaluation of computer-aided learning in orthodontics. *American Journal of Orthodontics and Dentofacial Orthopedics, 138*(4), 410-419.

Ruble, T. L., & Stout, D. E. (1992). Changes in learning-style preference: comments of Gieger and Pinto. *Psychological Reports, 70*, 697-698.

Rust, C. (2002). The Impact of Assessment on Student Learning. *Active Learning in Higher Education, 3*(2), 145-158.

Saadé, R. G., & Kira, D. (2006). The emotional state of technology acceptance. *Issues in informing science and information technology, 3*, 529-539.

Saavedra, R., & Kwun, S. K. (1993). Peer evaluation in self-managing work groups. *Journal of Applied Psychology, 78*(3), 450.

Sadler-Smith, E. (2001). The relationship between learning style and cognitive style. *Personality and Individual Differences, 30*, 609-616.

Sadler-Smith, E., Evans, C., & Cassidy, S. (2006). Learning style and student self-assessment skill. *Education and Training, 48*(2/3), 170-177.

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education, 35*(4), 453-472.

Sharaf, A. A., AbdelAziz, A. M., & El Meligy, O. A. S. (2007). Intra- and Inter-Examiner Variability in Evaluating Preclinical Paediatric Dentistry Operative Procedures. *Journal of Dental Education, 71*(4), 540-544.

Sims, R. R., Veres, J. G., & Shake, L. G. (1989). An exploratory examination of the convergence between the Learning Styles Questionnaire and the Learning Style Inventory II. *Educational and Psychological Measurement, 49*, 227-233.

Sinclair, B., & Ferguson, K. (2009). Integrating simulated teaching/learning strategies in undergraduate nursing education. *International Journal of Nursing Education Scholarship, 6*(1).

Singh, K., Smiddy, W. E., & Lee, A. G. (2002). *Ophthalmology review: a case-study approach* (Vol. 134). New York: Thieme.

Singh, T., Gupta, P., & Singh, D. V. (2009). Test and item analysis *Principles of Medical Education* (3rd ed., Vol. 3rd, pp. 70-77): Jaypee Brothers Medical Publishers (P) Ltd.

Smith, M., Higgs, J., & Ellis, E. (2008). Factors influencing clinical decision making. In J. Higgs (Ed.), *Clinical reasoning in the health professions* (pp. 90): Elsevier Health Sciences.

Sprenger, M. (2003). *Differentiation through learning styles and memory*. Thousand Oaks, CA: Corwin Press.

Suantak, L., Bolger, F., & Ferrell, W. R. (1996). The hard–easy effect in subjective probability calibration. *Organizational Behavior and Human Decision Processes, 67*(2), 201-221.

Tang, L.-C., & Sung, H.-C. C. (2012). The Effectiveness of problem-based learning on nursing students' critical thinking: a systematic review. *The JBI Database of Systematic Reviews and Implementation Reports, 10*(57), 3907-3916.

Tanner, C. A., Padrick, K. P., Westfall, U. E., & Putzier, D. J. (1987). Diagnostic reasoning strategies of nurses and nursing students. *Nursing Research, 36*(6), 358-365.

Teichman, J. M. H., & Richards, J. (1999). Multimedia to teach urology to medical students. *Urology, 53*(2), 267-270.

Tepper, B. J., Tetrault, L. A., Braun, C. K., & Romero, J. E. (1993). Discriminant and convergent validity of the problem solving style questionnaire. *Educational and Psychological Measurement, 53*, 437-444.

Trigwell, K., & Prosser, M. (1991). Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes. *Higher Education, 22*(3), 251-266.

Tuckman, B. W. (1999). *Conducting Educational Research* (5th ed.): Harcourt Brace College Publishers.

Valizadeh, L., Amini, A., Fathi-Azar, E., Ghiasvandian, S., & Akbarzadeh, B. (2013). The Effect of Simulation Teaching on Baccalaureate Nursing Students' Self-confidence Related to Peripheral Venous Catheterization in Children: A Randomized Trial. *Journal of Caring Sciences, 2*(2), 157-164.

Veres, J. G., Sims, R. R., & Locklear, T. S. (1991). Improving the reliability of Kolb's revised LSI. *Educational and Psychological Measurement, 51*, 143-150.

Veres, J. G., Sims, R. R., & Shake, L. G. (1987). The reliability and classification stability of the Learning Style Inventory in corporate settings. *Educational and Psychological Measurement, 47*.

Vleuten, C. V. D., Verwijnen, G., & Wijnen, W. (1996). Fifteen years of experience with progress testing in a problem-based learning curriculum. *Medical Teacher, 18*(2), 103-109.

Ward, M., Gruppen, L., & Regehr, G. (2002). Measuring self-assessment: current state of the art. *Advances in Health Sciences Education, 7*(1), 63-80.

Ward, M., MacRae, H., Schlachta, C., Mamazza, J., Poulin, E., Reznick, R., & Regehr, G. (2003). Resident self-assessment of operative performance. *The American Journal of Surgery, 185*(6), 521-524.

Wass, V., Van der Vleuten, C., Shatzer, J., & Jones, R. (2001). Assessment of clinical competence. *The Lancet, 357*(9260), 945-949.

Wells, S., Warelow, P., & Jackson, K. (2009). Problem based learning (PBL): A conundrum. *Contemporary Nurse, 33*(2), 191-201.

Wilson, D. K. (1986). An investigation of the properties of Kolb's Learning Style Inventory. *Leadership and Organisation Development Journal, 7*(3), 3-15.

Woo, M. A., & Kimmick, J. V. (2000). Comparison of Internet versus lecture instructional methods for teaching nursing research. *Journal of Professional Nursing, 16*(3), 132-139.

Wood, D. F. (2003). Problem based learning. *British Medical Journal, 326*(7384), 328-330.

Woods, R., McCarthy, T., Barry, M. A., & Mahon, B. (2004). Diagnosing smallpox: would you know it if you saw it? *Biosecurity and bioterrorism: biodefense strategy, practice, and science, 2*(3), 157-163.

Yang, H., Thompson, C., & Bland, M. (2012). The effect of clinical experience, judgment task difficulty and time pressure on nurses' confidence calibration in a high fidelity clinical simulation. *BMC Medical Informatics and Decision Making, 12*(1), 1-9.

Young, J. M., Glasziou, P., & Ward, J. E. (2002). General practitioners' self ratings of skills in evidence based medicine: validation study. *BMJ, 324*(7343), 950-951.

Zary, N., Johnson, G., Boberg, J., & Fors, U. (2006). Development, implementation and pilot evaluation of a Web-based Virtual Patient Case Simulation environment - Web-SP. *BMC Medical Education, 6*(1), 10.

Zick, A., Granieri, M., & Makoul, G. (2007). First-year medical students' assessment of their own communication skills: A video-based, open-ended approach. *Patient Education and Counseling, 68*(2), 161-166.

Zwanenberg, N. v., Wilkinson, L. J., & Anderson, A. (2000). Felder and Silverman's Index of Learning Styles and Honey and Mumford's Learning Style Questionnaire: How do they comapre and how do they predict? *Educational Psychology, 20*(3), 365-381.

Zywno, M. S. (2003). *A contribution of validation of score meaning for Felder-Soloman's Index of Learning Styles,.* Paper presented at the Annual Conference of American Society of Engineering Education.

Appendices on CD

1 – Research ethics approval

2 – Learning style models

3 – Master database

4 – Didactic cohort lectures

5 – Flow charts

6 – Didactic cohort tutorials

7 – Focus group feedback template

8 – Focus group reports

9 – Complete study findings

10 – Authors presentations and publications