The Design, Development and Evaluation of an Online, Interactive, Formative Assessment Tool for Medical Education

Edward Palmer

School of Medicine

PhD Thesis

December 2010

The University of Adelaide

Table of Contents

Abstract

Rapid changes in medical practices used across the world over the last twenty years have lead to significant issues in allowing medical students' access to patients with commonly occurring conditions. Equally significant changes in medical education systems have stressed the need for more authentic, problem solving curricula. One of the methods that can be used to address this apparent disconcordance between education and clinical practice is the use of technology to give students access to pedagogically sound learning activities.

Online formative assessment using case-based scenarios provides students with the ability to assess their own progress in their studies in their own time and at their own pace. Supplemented by sufficient relevant images and video, students can experience an authentic means of assessment that can be used to complement their ward-based education.

This thesis discusses the need for such online formative assessment systems and details the design, development and evaluation of such a system. The conclusion from this research is that with appropriate assessment strategies, students can be motivated to use online formative assessment for learning activities. It will be shown that there is sufficient evidence to suggest that such systems can be used to improve student outcomes in their field of interest, in this case Surgery.

Declaration

Candidate:	Edward Palmer
Student Number:	1065043
Award:	PhD
University:	University of Adelaide
Faculty:	Health Sciences
School:	Medicine
Principal Supervisor:	Associate Professor Peter Devitt
Associate Supervisor:	Professor Geoffrey Crisp
Date:	18/12/2010

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution to Edward Palmer and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. The author acknowledges that copyright of published works contained within this thesis (as listed below*) resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue, the Australasian Digital Theses Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

* (Palmer & Devitt, 2007a, 2007c, 2008)

Acknowledgements

Thanks to my supervisors Professor Geoffrey Crisp and Associate Professor Peter Devitt for encouraging me in this endeavour and providing friendship as well as their experience and knowledge.

Thanks to my mother, who showed me that it is possible to study while raising a family and to my father, who would have done much better than me had he had the opportunities provided to me.

A big thank you to Donna Velliaris, study buddy extraordinaire for her feedback and friendship and thanks to Sheila Scutter and Michelle Harvey for their feedback and encouragement.

Kind thanks are due to Anthony Dunstan who worked on the original Medici 20 years ago and John Barnard and Tom Sullivan who provided statistical assistance.

Finally thanks to Peter Murdoch for making this look as good as it does.

I dedicate this to my children Daniel, James and Michael.

List of Figures

List of Tables

Acronyms

AAC	advanced audio coding
ADSL	asymmetric digital subscriber line
ASP	active server pages
BBC	British Broadcasting Corporation
CAL	computer-assisted learning
CASE	Computer Aided Simulation of the Clinical Environment
CD	compact disc
CDC	Control Data Corporation
CD-ROM	compact disc read only memory
CEX	clinical evaluation exercise
COMET	Computerized Modular Exercise Templates
CPR	cardiopulmonary resuscitation
CTT	classical test theory
DEC	Digital Equipment Corporation
DVD	digital versatile disc
ELI	Educause Learning Initiative
FDA	Food and Drug Administration
GMC	General Medical Council
IBM	International Business Machines
ICU	intensive care unit
IRT	item response theory
JPEG	Joint Photographic Experts Group
kB	kilobyte
LCD	liquid crystal display
MB	megabyte
MBBS	Bachelor of Medicine and Bachelor of Surgery
MCQ	multiple-choice question
MEQ	modified essay question
MP3	Moving Picture Experts Group Audio Layer 3

NZ	New Zealand
OOP	object oriented programming
OSCE	objective structured clinical examination
PBL	problem-based learning
PDA	personal digital assistant
PHP	personal home page (hypertext preprocessor)
RECAP	recall, comprehension and application
SAT	Scholastic Aptitude Test
SC	script concordance
SOLO	structure of observed learning outcomes
UK	United Kingdom
US	United States
VHS	video home system

Chapter One: Introduction

The Evolution of Medical Education

Medicine and medical education are continually growing and changing entities. Subject to social, academic, political and corporate pressures, the ability to provide an education system capable of developing medical practitioners who can meet society's high expectations of the medical profession is essential. The methods used in the teaching and assessment of medical students and their learning strategies are often challenged as a result of these pressures. Even in cultures that are similar, there have been major differences in the methods used to educate future medical practitioners. In the United Kingdom (UK), the United States (US) and Australia, numerous changes in medical education have occurred over the past century, although the nature and timing of these developments has varied, reflecting the diversity that exists in the respective medical and education systems.

A century ago, the Flexner Report (1910) became the catalyst to economise and standardise medical education in the US (Moy, 2000). Flexner recommended that the special interest teaching hospitals existing at the time be abandoned in favour of highly standardised medical schools. He favoured a hybrid of scientific theory and practice as the essential components of medical education. This blended approach to instruction was an amalgam of methods from across Europe, but mainly influenced by Britain and Germany. The teaching method favoured in Britain and France was based on an apprenticeship model where students worked in hospitals learning their trade. The German approach to medical education, admired by Flexner, was based on the development of scientific approaches to Medicine, and involved substantial laboratory work focused on such disciplines as anatomy and physiology. German medical education was based in the university environment, rather than in a clinical setting. Accordingly, the model developed for medical education in the US consisted of two years of basic science instruction involving substantial laboratory work, followed by two years of hospital-based teaching. Students therefore experienced a course in two parts; one component where didactic approaches to teaching were used, where memorisation and rote learning were vital for success and then a more experiential program, requiring problem solving skills. Theoretically, this produced medical practitioners with a sound scientific knowledge that had been put to practice on the wards with real patients. This style of medical education is still part of all accredited US medical schools (Beck, 2004; Bloom, 1988).

As with many innovations and developments requiring social change, the revitalising of medical education in the US was traumatic. Hospitals that were unable to improve their admissions standards and adhere to strict curriculum requirements were forced to close. The standardisation process, which still persists today had unintended consequences for disadvantaged and rural sectors of the population, leading to a shortage of practitioners in those areas (Beck, 2004).

The US approach of encouraging university-based teaching hospitals was typified by Johns Hopkins University, where the university and Johns Hopkins Hospital combined to develop students' medical learning. A key figure in the integration of the scientific method and clinical practice in the US was Sir William Osler, who also attempted to encourage change in the UK. Osler developed ward-based teaching and created medical residencies and, with Flexner, made efforts to encourage British institutions to adopt teaching methods such as those used in the US. However, British and colonial countries, such as Australia and New Zealand (NZ), did not rapidly respond to the challenges of the Flexner report (Engel, 2000). The Royal Colleges had substantial influence in Britain and were not at all willing to allow universities to have a major influence in clinical education. Their reluctance to move to a more university-based clinical education was supported by evidence suggesting that German doctors, trained using the 'scientific' method, were deemed to have lacked sufficient clinical skills on the battlefield during World War I (Bonner, 1995). Factors such as these resulted in significant opposition to change in Britain (Weatherall, 2006) and up until 1945, only the Hammersmith Hospital had adopted a teaching model similar to Johns Hopkins Hospital.

After World War II and the fragmentation of the British Empire, attitudes in Britain and its former colonies changed and this led to the development of new methodologies for teaching medical students (Engel, 2000). The British system changed to mostly university-based medical schools and the training now consisted of three stages (Weatherall, 2006); two basic science courses were followed by a final course that included clinical practice. Although slightly different from the US model, which endeavoured to include clinical material at earlier stages in the students' education, the medical education systems on both sides of the Atlantic were now quite similar.

Problem-based Learning

Medical education in the former British colonies had begun to diverge from that developed in Britain after World War II. Canada adopted a similar approach to that of the US in the early years, but was responsible for one of the true innovations in education in the form of Problem-based Learning (PBL). PBL originated in the 1960s at McMaster University in Canada and was applied to medical education by Howard Barrows in 1969. This style of education reflects the ethos of self-directed learning, a concept encouraged by Osler before the end of the 1900s (Khoo, 2003). Rather than memorising facts, students were encouraged to seek information, assimilate new knowledge, address problems, and synthesise solutions to problems (Barrows & Tamblyn, 1980). Whereas the traditional medical education model separated the study of the basic sciences and clinical experience, PBL often sought to integrate the learning of the two areas (Nandi, Chan, Chan, Chan, & Chan, 2000). Traditional lectures were often sacrificed in a curriculum based on PBL, as they were deemed not to be consistent with the basic tenets of independent learning (Engel & Clarke, 1979).

Australia was one of the early adoptees of the PBL explosion, when the University of Newcastle chose to use this approach to teach its medical students in the 1970s (Engel & Clarke, 1979). The success of this initial foray acted as a catalyst for the implementation of PBL in many universities across the nation by 2000. Reasons for this success were based in part on the international credibility of PBL, the ease of trialling this approach in some courses, and the obvious levels of interaction observed in a PBL group as compared with a lecture group (Sanson-Fisher & Lynagh, 2005). Also influential in its adoption was the belief that it encourages self-directed learning (Sanson-Fisher & Lynagh, 2005) thus providing a sound base for life-long learning; a concept crucial to the medical profession which has a strong emphasis on continuing medical education (Jennett & Swanson, 1994). Expectations of the general public have changed as patients have become more informed of the options for treatment and are no longer prepared to take unthinking acceptance of the practitioner's word. Eminence-based Medicine has been steadily replaced by evidence-based Medicine (Keckley, 2004; Sinclair, 2004).

In 1993, the UK General Medical Council (GMC) made recommendations that dramatically changed the approach to medical teaching in that country (Towle, 1998). Although not fully endorsing PBL specifically, the GMC stated that, 'Medical schools are well aware of the merits of the learner-centred and problem-oriented approaches and are striving for their adoption, moves which are strongly encouraged' (General Medical Council, 1993, p. 11). Many of the recommendations from the GMC echoed traits associated with PBL, specifically a focus on self-directed learning, less focus on knowledge acquisition, a greater emphasis on approaches to learning based on curiosity and exploration of knowledge and a recognition that basic science and clinical teaching needed to be part of an integrated approach to learning (General Medical Council, 1993).

In the 1960s in the US, McMaster University developed PBL programs partially based on educational developments at Case Western Reserve University in the 1950s. The goal was to provide a more holistic approach to teaching medicine outside traditional disciplinary boundaries by encouraging students to solve realistic problems (Baptiste, 2003; Boud & Feletti, 1997). Educational institutions such as Harvard University adapted some of the features from this model, coupled with the pedagogy from McMaster to develop a hybrid program where lectures and laboratory sessions were integrated with problem solving tutorials (Peters, Greenberger-Rosovsky, Crowder, Block, & Moore, 2000). This method is believed to have been highly influential in the adoption of a PBL curriculum in more traditional institutions around the world (Boud & Feletti, 1997).

The greatest benefits attributed to the PBL approach relate to increased attendance at teaching sessions (Vernon & Blake, 1993), improved knowledge retention (Dochy, Segers, Van den Bossche, & Gijbels, 2003), greater satisfaction with the learning program by students (Sanson-Fisher & Lynagh, 2005) as well as staff (Nandi, et al., 2000), and social benefits such as the development of communication skills, cultural understanding and appreciation of ethics in medical care (Koh, Khoo, Wong, & Koh, 2008). However, the introduction of PBL as a method of teaching and learning has often been tinged with controversy, due to its non-traditional nature and its impact on the cost of teaching (Finucane, Johnson, & Prideaux, 1998; Norman & Schmidt, 2000). Also of concern have been change management issues, such as the effective use of human resources (Epstein, 2004), and the development of a clear understanding by staff of the role they will play in a new curriculum (Yusof, Tasir, Harun, & Helmi, 2005). The popular hybrid models of PBL may increase the workload of teaching staff (Kwan, 2002) and there is still a lack of compelling evidence to support its implementation. The latter has been of particular concern to those considering the adoption of PBL. Studies undertaken have often lacked the full rigour of a randomised-controlled trial, but meta-analysis of research meeting strict criteria has shown that PBL does not provide substantially different cognitive outcomes to traditional curricula (Albanese & Mitchell, 1993; Colliver, 2000; Nandi, et al., 2000; Newman, 2004; Norman & Schmidt, 1992).

The assessment of students involved in PBL has proven to be controversial (Colliver, 2001). Methods by which students in PBL-based courses have been assessed include modified essay questions (MEQs), essays, standardised patient simulations, oral presentations, multiple-choice questions (MCQs), case-based examinations (Gijbels, Dochy, Van den Bossche, & Segers, 2005; Nendaz & Tekian, 1999) and objective structured clinical examinations (OSCEs) (Hodges, Regehr, McNaughton, Tiberius, & Hanson, 1999). Some of these assessments have tested higher order cognitive skills, while others have tested basic knowledge recall. There are indications that students in a PBL course perform better when higher order skills are tested (Dochy, et al., 2003; Gijbels, et al., 2005). When tested on knowledge recall, PBL students perform at the same or lower levels (Dochy, et al., 2003), but they can have gaps in their knowledge base, which could be potentially problematic in clinical practice.

Regardless of the opinion on its efficacy or proven worth, PBL has not only become a popular educational methodology, but has dominated educational approaches to teaching medicine for the last two decades. PBL now forms a key component of the medical curricula in some of the oldest and most prestigious institutions in the world.

Changes in the Teaching and Learning Environment

While this educational revolution was taking place, healthcare systems across the world were changing. Managed care was introduced in the US. Rapid technological advances in many branches of medicine resulted in patients spending less time in hospital. These and other reasons have led to reduced access to patients as 'teaching aids' for medical student education. In many countries, increasing numbers of patients are cared for in the private sector, thereby further reducing access for teaching. In Australia this number is almost 40 percent, with 56 percent of all surgical procedures taking place within the private sector (Australian Private Hospitals Association, 2007). All of this has occurred at a time when the volume of medical knowledge is rapidly expanding. For nearly a century it has been acknowledged that it is not a viable educational philosophy to have students attempt to learn everything.

> 'Undoubtedly the student tries to learn too much, and we teachers try to teach him too much—neither, perhaps, with great success.'
>
> Sir William Osler. After Twenty-Five Years, *Montreal Medical Journal*, 28:823, 1899 in (Bowman, 2002).

The demographics of the student population have changed over time. The percentage of students in part-time work, while enrolled in full-time study has increased throughout Australia (Krause, Hartley, James, & McInnes, 2005; McInnis, James, & Hartley, 2000), the UK (Bonner, 1995; Curtis & Lucas, 2001; Ford, Bosworth, & Wilson, 1995) and the US (National Center for Education Statistics, 2005). Many of these students are part of the workforce in excess of 15 hours a week, particularly in the US, and these students report detrimental effects of employment commitments on their academic work (U.S. Department of Education, 2002). This results in lower attendance at classes, reduced access to teaching resources and lower grades. Students are therefore more likely to seek learning opportunities at unusual times - times when they may not have any access to patients or clinicians.

Whereas traditional medical education was dominated by lectures, ward rounds and patient-focused teaching, today's medical students are confronted by a world where there is more self-directed learning, shorter patient time in hospitals (Towle, 1998), and a correspondingly smaller opportunity to observe and learn from patients (Kirk, 2007) and clinicians (Bryant, et al., 2003). With a significant workload thrust upon clinicians, there is a belief that the teaching environment in hospitals needs to be improved by additional resources and increased staff instruction (Gibson & Campbell, 2000). Alarmingly, some medical practitioners are becoming disillusioned with their role as clinician-educators, believing that teaching is undervalued and poorly acknowledged. Many are considering leaving the field (Denniston, 2008; Lowenstein, Fernandez, & Crane, 2007). If a general perspective is taken on Lowenstein's (2007) study, it is likely that clinicians will not be inclined to participate in as many educational activities as may have been expected in the past.

Providing Students with an Authentic Learning Environment

One solution adopted by universities is to provide an authentic learning environment for students by placing them with practitioners in the community. The major issues with such placements are finding sufficient practitioners to participate, and ensuring that they have appropriate levels of training (Thistlethwaite, Jacobs, & Rudolphy, 2005). Financial considerations may also play a role in some practitioners' reluctance to be involved in student teaching, as they often have to reduce their working pace, or otherwise alter their normal practice to accommodate the students. Without the resources to address these issues, these types of placements are logistically difficult. However, concerns about the learning environment that can be provided for students by busy community-based practitioners may be unfounded, as there is evidence to suggest that rural-based students in community settings have improved outcomes in standard examinations when compared with hospital-based students (Worley, Esterman, & Prideaux, 2004).

An alternative to placing students in an environment where there are real patients is to facilitate learning by a case study or clinical scenario. This method forms the cornerstone of many PBL programmes. Students are provided with a problem, based on an authentic situation and, either alone or in groups, they attempt to resolve it. In most cases this is carried out with the guidance of a facilitator. Some PBL courses adopt this approach over many hours or several days, with information incrementally released to students as they progress through the case. This method of teaching is based on constructivist theories where the underlying premise is that learners are actively engaged in building on their own experiences and knowledge (Savery & Duffy, 2001; Sokal, 1990). A PBL curriculum with strong guidance, such as a scaffolded structure involving lectures, texts and computer-based learning materials to support group-based collaborative inquiry, is often employed (Hmelo-Silver, Duncan, & Chinn, 2007) to guide students without interfering in the constructivist learning approach. Scaffolding can play an important role in cognitive learning strategies (Reiser, 2004) as this type of support can provide feedback to learners and help to address misconceptions. Indeed there is evidence to suggest that this guidance is vital, as an unguided or 'pure discovery' approach to learning may be flawed, partially due to a lack of ability by students to direct their own learning towards productive pathways (Mayer, 2004).

The Role of Assessment

Whilst the use of guided clinical scenarios and other resources may be useful to students' learning, it is likely that they will only engage in this type of material if it forms a component of their assessment, particularly 'summative' assessment. Summative assessments award grades, marks or descriptors of levels of competence that can be used to determine whether students have met course or program requirements and are often used to judge whether students are able to proceed in their chosen study. As such, they can be powerful motivators for students' efforts. Indeed, Biggs bemoans the negative attitudes towards summative assessment, seeing the possibilities in utilising the influence of these types of tasks. '... one might have a powerful enhancement to learning, using such a synthesis to engineer backwash from [summative assessment] so that the effects were *positive,* the backwash from [summative assessment] *supporting* the feedback from [formative assessment]' (Biggs, 1998, p. 106)

Clinical scenarios as a part of a PBL curriculum may be summatively assessed, even if it is just via attendance and participation at PBL sessions. In situations where a scenario is provided as a component of a scaffolded approach to learning, the assessment may be 'formative'. Formative assessment is used to inform students and teachers about the progress of the learner. The objective is to help students attain the knowledge, skill and understanding required to be successful in their course, often in a self-directed environment. Feedback forms an important component of this type of assessment, as it allows students to make some measure of their own abilities and develop strategies for self-improvement. Another type of assessment labelled as 'diagnostic' serves similar purposes to formative assessment, but is typically carried out before learning begins so that students and teachers share a clear understanding of the work required to provide the desired learning outcomes (Leighton & Gierl, 2007).

It is important that medical graduates have the ability to rise to the expectations of their colleagues and society when they enter the workforce as medical professionals. Factors such as growing intellectually and being in a career to serve society as a whole in order are important features of the characteristics of medical students. In 1999, a study showed that these aspects were in the top 5 of medical students' determinants of career satisfaction over the period of their education (Reed, Jernstedt, & McCormick, 2004). The ability to apply their knowledge and skills to new situations will be regularly tested, so it is essential that students be assessed effectively in their learning environments, whether these assessments involve inpatients, outpatients, external attachments or scenario-based learning.

In a clinical environment, assessment can be based on observation by clinical staff who assess clinical reasoning skills, knowledge and the interpretation of clinical problems (Fowell & Bligh, 1998; Kreiter & Bergus, 2007), or more indirectly through case reports presented orally or via written work (Vu & Barrows, 1994). More rigorous and reliable methods include the use of simulated patients (Adamo, 2003; O'Connor, 2008), where actors are trained to behave as patients and the interaction with the patient is observed in a controlled environment, or through clinical evaluation exercises (CEX) or the mini-CEX, where students are observed interacting with real patients (Norcini, Blank, Duffy, & Fortna, 2003). The mini-CEX consists of a series of short observations (10–15 minutes) over an extended period, thus allowing a wide range of skills to be tested. The mini-CEX shares this trait with a common assessment, the objective structured clinical examination (OSCE) (Newble, 2004), where a series of standardised 'stations' are established, each one testing a component of clinical skill such as the ability of students to examine a patient or take a history. Collectively, these tests are used to assess the ability of students to manage a real case, which is the goal of any medical curriculum, whether it is based on PBL or more traditional methods. The main disadvantage of these types of assessments is that they are extremely labour intensive, requiring significant administrative and clinical input from a large number of staff.

In many situations, the assessment of students in clinical situations forms a portion of the overall summative assessment, but is often intended to serve a formative purpose as well, where feedback is provided to assist students in improving their skills. This formative assessment component is considered integral to the clinical environment (van der Issenberg & McGaghie, 2002; van der Vleuten, et al., 2000), but in PBL-based curricula (Nendaz & Tekian, 1999) it can be challenging to effectively implement. Students need to be advised when they are not performing satisfactorily and strategies put in place to address these issues. This tends to be costly in terms of human resources and time and if the criteria for assessment are not clearly described; supplying all students with a consistent level of feedback may be challenging. Where human resources or the nature of the learning environment does not allow this to be done effectively, technology may provide appropriate mechanisms for meaningful feedback to be available to students in a wide variety of situations. This can be done in a timely, non-threatening and objective manner. When students are struggling with a facet of a course, it may not be helpful to support them with loose guidance, such as instructing them to read a text or see additional patients. There may be more fundamental issues of understanding and misconceptions, despite the student possessing a sound knowledge base. Even in situations where students can meaningfully respond to feedback, there is always the issue of resources to consider. The type of patients that students need to consult in order to improve their abilities may not be available, or the clinician they were scheduled to see may become unavailable. Resource issues are of genuine concern in all teaching environments, but in the clinical situation they are particularly acute, leading to staff burnout (Harden, 1999) and the cessation of teaching programs (Lane & Gottlieb, 2000). A more convenient, positive environment is required, where students can practice clinical management. Ideally, such an environment should allow them to practice at any time without concern over poor results, loss of face or inflicting harm on patients. This may be of particular importance in disciplines where mistakes can be fatal, such as surgery.

Educational Resources

Educational support systems include electronic texts, such as the free books available at Student Bookworld (Student Book World, 2009), and those available by subscription such as the *Oxford Handbook of Anaesthesia* (Oxford University Press, 2006). Resources such as online question banks exist and can make useful formative assessment tools such as that available to purchasers of the book *Essentials of General Surgery* by Lawrence (2005). More sophisticated resources include the *Visible Human Project*, which provides '…complete, anatomically detailed, three-dimensional representations of the normal male and female human bodies' (U.S. National Library of Medicine, 2003). In laboratories, expensive devices such as surgical simulators may be available to some students. The CAthLabVR system allows learners to '…practice endovascular procedures such as coronary and peripheral vascular interventions, cardiac pacing, and cardiac valve replacements, using a combination of computer and haptic simulation (Immersion Corporation, 2003). Other haptic devices are used for medical applications (Laycock & Day, 2003) including the simulation of laparoscopic surgery (Basdogan, Ho, & Srinivasan, 2001) and knee surgery (Gibson, et al., 1997). Full-bodied mannequins with realistic heart, lung and bowel sounds as well as the ability to speak and breathe are in use (Clark, 2008; Trent, 2008). The downside of these teaching aids is their cost. One mannequin device can be valued up to US$250,000 (Trent, 2008) and whilst they may have educational value the cost places severe limitations on their use. Despite this expense, the use of educational resources is widespread, and mannequins in particular have been used since the late 1960s for medical learning purposes (Cooper & Taqueti, 2004). Non-technological resources often utilised include real-life simulated patients who can be used in OSCE examinations and in the formative environment, but the recurring cost of actors for formative assessment is high.

The area of computer-based and online teaching as part of the supporting scaffold for learning appears attractive from a number of viewpoints. The penetration of computers and the Internet in developed countries is now high (Chinn & Fairlie, 2010) and it is likely that students in these countries will have access to these resources. Hence, from an equity point of view, there are unlikely to be the significant issues of previous years such as the lack of sufficient access to the required technology and software. Given the need for students to be able to access educational resources at any hour of the day, often due to work as well as study commitments, a computer-based formative assessment system is a logical proposition. Of course, there is more required from a learning resource than students having access to the appropriate technology. The resources provided need to be based on sound pedagogical design, supported by a reliable technological infrastructure and regular maintenance to ensure currency and thus relevance to both students and the course. Although set-up costs can be high (Rumble, 2001), they can also be economical in comparative terms. When compared against a US$250,000 mannequin that can be used by only a few people at a time, utilises valuable teaching space and may not be available all year round, a computer-based version may be more viable.

Issues to Consider in Providing Online Formative Assessments

If educational materials are to be developed for formative purposes using online technologies, there are many factors that need to be considered. Should such a resource be text-based or are images and video an essential component of a medical education system? This question is not simple to answer and may involve trading off educational effectiveness against development costs and time. Once the developer has made a decision, questions around the nature of the learning resource need to be answered. Should the resource be didactic or interactive? Would a scenario-based approach to the system be of benefit and if so how should it be implemented? A sequence of essay or multiple-choice type questions may be appropriate, but there are important considerations about the validity and reliability of any type of educational resource that should be examined before making such decisions. Feedback in an interactive system seems essential, but how much is enough, and is scoring students' work a good idea? Providing scores for students' formative work has been suggested as potentially detrimental (Black & Wiliam, 1998; Sadler, 1989; Taras, 2002) yet many medical students are competitive, if not with their peers at least with themselves, and perhaps scoring is appropriate for this group. However, it is not obvious what type of scoring mechanism would be appropriate. Nor is it obvious whether students should use the resource independently or collaboratively. The most pertinent issue might be how students can be motivated to use a formative assessment tool. With direct links to assessable content, it is likely that students would engage with the material, but issues about how to provide this link remain unanswered. The following research questions arise:

- Can the use of an online formative assessment tool improve student learning outcomes for medical students as measured by traditional examinations? and
- If the use of an online formative tool did improve student learning outcomes, what strategies will encourage its use?

Thesis Structure

This thesis will examine the role of formative assessment as a component of self-directed learning and the role of technology in implementing this type of assessment. Chapter One has discussed the current educational situation for students and teachers and has posed questions relating to the role of online formative assessment.

Chapter Two provides an examination of assessment in medical education. The historical origins of the types of assessment are explored, drawing the conclusion that most assessment has been driven by factors other than student learning. This chapter highlights formative assessment as an essential component of a students' medical education. It concludes that online formative assessment is likely to have a significant positive impact on students' learning provided appropriate strategies can be developed to encourage their participation. Essential characteristics of online formative assessment are determined.

Chapter Three examines the use of technology in education from a historical perspective and concludes that the conditions for the use of computers to support learning are now optimal in terms of cost, social conditions and technological development.

Chapter Four illustrates the design and development of an electronic formative assessment tool based on the results of the recommendations on the use of technology in education and in formative assessment. The software used, and examples of its use, is discussed in the context of the results from Chapters Two and Three.

Chapters Five and Six detail the results of a two-year evaluation of electronic assessment tool usage in the teaching of students in their surgical ward placements at the University of Adelaide. This research work concludes with recommendations on how online formative assessment can be used to influence student learning patterns and improve learning outcomes.

Chapter Two: Assessment

Assessment tasks can be designed to serve different purposes, such as improving learning or demonstrating competency, although there can be substantial overlap between their use in practice. Educators often divide assessments into three distinct types based on their purpose: diagnostic, summative and formative. All three forms of assessment have a role to play in student learning as illustrated in Figure 2.1. Diagnostic assessment usually takes place before learning and may be used to assist the teaching, while summative assessment usually occurs after learning, and may be used for accreditation or certification purposes. Formative assessment is typically used during learning and is intended to guide students as they study, providing meaningful feedback for students to reflect upon. Students may then adjust their learning approaches based on the outcomes of the formative assessment.

NOTE:
This figure is included on page 37 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.1 Assessment types (Crisp, 2009)

Diagnostic Assessment

Diagnostic assessment in education works along similar lines to a diagnostic assessment in clinical practice. Whereas clinicians carry out a diagnostic assessment on patients by gathering important information on their condition in order to guide their management of the patient (Ashwal, et al., 2004; Beun, van Emde Boas, & Dekker, 1998), educators can use diagnostic assessments to examine the prior understanding and skill level of students in order to direct their learning (Leighton & Gierl, 2007). The goal of this type of assessment in both clinical and educational environments is to determine the most appropriate course of action based on an accurate diagnosis. For teaching purposes, a diagnostic assessment is usually carried out at the beginning of a course and can be used to determine if students require additional preparatory work or, more often, to tailor the contents of a course to suit a particular group of students. Diagnostic assessment tasks offer the opportunity to engage students with the core concepts upon which learning will be built. A shared knowledge of student ability, either at the class or individual level, established at the beginning of a course is a valuable commodity and should lead to an improved learning experience for students, as well as a more productive teaching experience for teachers (Earl, 2003).

Summative Assessment

Summative assessment is a type of assessment often linked to certification (Boud, 2000; Knight, 2002) and can be described as an 'assessment *of* learning' (Harlen, 2006). The skills, knowledge and understanding of a student may be tested with the goal of ensuring that any student that passes the assessment can be considered to have attained an appropriate level of competency in that field. The definition of what is 'appropriate' is a complex construct requiring input from professional or accrediting bodies, discipline peers and the wider community. This is a standard type of assessment and is one that every medical practitioner must pass before being permitted to practice. It is common to all disciplines and programs. The format for summative assessments is varied and may include essays, multiple-choice questions, practical demonstrations of skill, and oral presentations. They are typically held at the end of a course or program as a barrier that students must overcome, but may also be implemented during a course. Grades, marks or other descriptors of the nature of students' achievement are used as indicators of the 'level' of the achievement of students measured against grading criteria

(criteria referenced assessment) or against their peers via a preferred distribution of outcomes (norm-referenced assessment) (Cohen, Manion, & Morrison, 1996).

Formative Assessment

Formative assessment is designed to inform both student and teacher about the progress of the student and may be described as 'assessment *for* learning' (Harlen, 2006). It may manifest itself in many of the same formats as summative assessments, but the timing and intent is vastly different. Whereas summative assessment seeks to ensure that students cannot proceed with their study or profession without passing the assessment, formative assessments are designed to help students attain the knowledge, understanding and skills crucial to their subject and course. Student use of formative assessment exercises is intended to create awareness of their own weaknesses in order to formulate plans to address them. When well designed, a formative assessment process should reduce students' dependence on the teacher but, through good feedback, the teacher will still play an important role. Appropriate formative assessment provides timely and relevant feedback to the teacher, who can then make decisions on how to address learning issues that may have arisen from student responses. Indeed this step, which at the very least will allow modification of future formative assessment tasks, can be regarded as vital (Nicol & Macfarlane-Dick, 2006). Formative assessment should allow students to reflect on their own learning outcomes, ideally measured against defined standards and criteria, and for that reason should be criterion-referenced rather than measured against other learners (Boud, 2000).

The concept of formative assessment, as used today, can be attributed to Bloom (Bloom, 1971). He defined the difference between summative and formative assessment by distinguishing between the grading purposes associated with summative assessment and the process of improvement linked to formative assessment. Although studies in the early 1970s suggested that students' learning was determined by the assessment tasks they had been required to complete, rather than the teaching itself (Gibbs & Simpson, 2004), much of the emphasis on assessment over the next 20 years was on summative forms, to the extent that most educational texts did not devote significant effort to discussing feedback or formative assessment (Sadler, 1989; Wiliam & Black, 1996). In England, formative assessment was proposed as part of the national curriculum for primary and secondary students but, over a decade from the early 1990s, the formative components of the proposal disappeared, partially due to the government requirement for national standards testing (Black & Wiliam, 2003). Formative assessment has proven to be difficult to implement for educators, and it is only recently that they have endeavoured to construct a theoretical framework to support it.

According to Sadler (1989) there are three core features of a formative assessment organisational framework. Learners have to:

- possess a concept of the standard being aimed for;
- be able to compare actual performance against the standard; and
- engage in action to close the gap between the standard and their actual performance.

The first criterion can be achieved by clear communication of the required standards perhaps illustrated by exemplars, but the latter two criteria require additional processes to be in place. Figure 2.2 illustrates a visual model of a formative process, which starts with a student submitting some formative work, perhaps guided by exemplars or standards set by the teacher. Once the teacher provides feedback, students are able to compare the submitted work with the defined standard and develop strategies, perhaps with the assistance of their lecturer or tutor, to address the gap. The figure suggests that this process could be repeated many times until students are satisfied with their learning outcomes. In practice, because teacher feedback is likely to only occur once per assessment, this repetition may take place over a series of formative assessments, or as part of an internal reflective process by students.

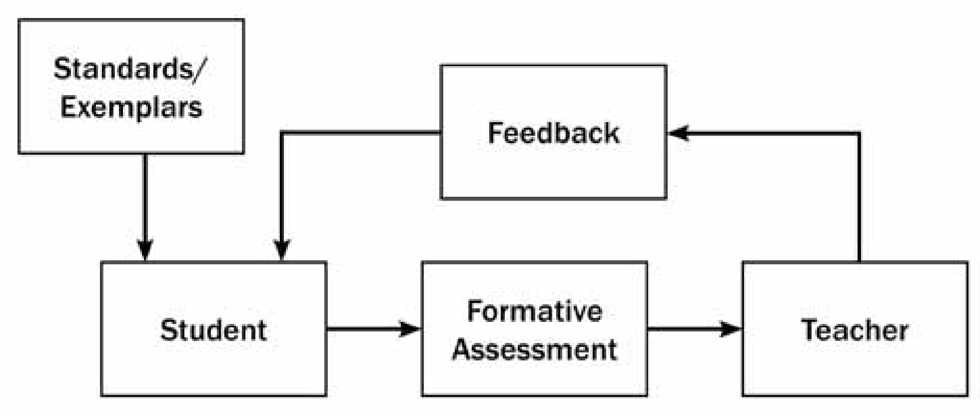


Figure 2.2 Feedback model for formative assessment

Gibbs, Simpson and Macdonald (2003) have further developed the formative assessment framework (Table 2.1). Items 4, 8 and 11 share common themes with Sadler and some of these concepts (1 and 6) echo the thoughts of Chickering and Gamson (1987) who wrote a widely cited article on good principles for teaching, emphasising the allocation of reasonable amounts of time for student tasks and providing prompt feedback. The extension that Gibbs, Simpson and Macdonald (2003) contribute to the framework primarily focuses on the nature and timing of both the tasks and the feedback obtained from them.

Table 2.1 Eleven conditions under which assessment supports student learning (Gibbs, et al., 2003)

NOTE:
This table is included on page 41 of the print copy of the thesis held in the University of Adelaide Library.

The second item in the framework ('These tasks distribute student effort evenly across topics and weeks') acknowledges issues students may have in balancing workload. In order for students to engage effectively with formative assessment, it may be more advantageous to allow the task to be spread over a period of time rather than providing large tasks that may be neglected due to higher priority summative tasks. The nature of the task (item 3) needs to be such that it appropriately directs learning, such as by encouraging additional reading or practical work, as well as being engaging (Gibbs & Simpson, 2004). Feedback from tasks needs to focus on the goal of the learning exercise rather than on grades and be prompt, clear and in sufficient detail to be of assistance to learners (items 5 to 9). Finally, methods such as requiring students to respond explicitly to feedback, or providing feedback only (not grades) for assessments may be helpful in engaging student attention to the feedback and thus improving their learning (items 10 and 11).

The issue of feedback in Gibbs' and Simpson's (2003) work dominates their recommended conditions for assessment to support student learning. Nicol, Macfarlane-Dick and Milligan (Nicol & Macfarlane-Dick, 2006; Nicol & Milligan, 2006) focused on feedback in formative assessment and isolated seven principles of good feedback (Table 2.2). Items 1 and 6 reflect Sadler's influence and item 3 has been suggested by Gibbs and Simpson (2004), but the other items reflect the importance the authors allocate to reflection (item 2), a friendly learning environment (items 4 and 5), and the possibility that feedback can not only improve learning but also the quality of teaching (item 7). The focus on students' internal thinking and motivational processes addresses themes touched on by earlier researchers, and is of great importance in encouraging students to make the most of formative opportunities.

Table 2.2 Good feedback practice (Nicol & Milligan, 2006)

NOTE:
This table is included on page 42
of the print copy of the thesis held in
the University of Adelaide Library.

The Use of Taxonomies in Assessment

The one unifying feature of diagnostic, summative and formative assessments is that they seek to enhance student learning by measuring, in some way, the ability of students to accomplish a task, demonstrate their understanding or knowledge of a topic or to solve problems. Students attempting these tasks may demonstrate different abilities in their approach. While some may attempt to complete an assessment by relying only on knowledge, others may treat the assessment as a series of unrelated tasks that need to be completed or may demonstrate the ability to link different facets of the task into a whole. This progression in how to solve a problem can be described by the 'structure of the observed learning outcome (SOLO)' taxonomy, devised by Biggs and Collis (1982).

This taxonomy considers five levels of student understanding as illustrated in Figure 2.3: pre-structural, where the student does not approach the task effectively or with any real understanding; uni-structural, where single components of the task are addressed but only additively; multi-structural where multiple components of the task may be addressed simultaneously yet still without a grasp of the understanding of the task as a whole; relational where the understanding of the task as a whole becomes apparent and finally extended abstract where students are able to apply the concepts involved in the task to new and varied situations.

NOTE:
This figure is included on page 43 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.3 The SOLO taxonomy

Another tool often associated with assessment is Bloom's taxonomy, which was devised to classify the levels of cognitive behaviour essential to learning. The SOLO and Bloom's taxonomies share many similarities. Bloom (1956) conceived of three taxonomies of domains: Cognitive (Figure 2.4), Affective (Figure 2.5) and Psychomotor (Figure 2.6). These are often used to map learning tasks and objectives to levels of affective, psychomotor and cognitive achievement. The affective domain has been used in examining behavioural objectives such as musical appreciation (Lewy, 1971), measuring the effect of patients' attitudes towards practitioners (Syrjälä, Knuuttila, & Syrjälä, 1990) and motivation (Denton, Doran, & McKinney, 2002). The psychomotor domain has been used to measure practical abilities in tasks such as training clinical researchers (Taekman, et al., 2004), running trauma life support courses (George & Doto, 2001) and training nurses (Alavi, Loh, & Reilly, 1991). The most commonly discussed domain is the cognitive domain where skills are described as ranging from the lowest level (knowledge recall) to the highest level (evaluation). Additionally, Bloom's domain of cognitive skill has been applied to perioperative nursing (Sigsby, 2004), examination design (Usova, 1997) and as a testing tool for various examination formats (Hancock, 1994; Palmer & Devitt, 2007b) amongst other uses. Although the SOLO taxonomy is more clearly linked to the measurement of learning outcomes, Bloom's taxonomy produces similar repeatable results when applied to the same task (Chan, Tsui, Chan, & Hong, 2002).

NOTE:
This figure is included on page 44 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.4 Bloom's taxonomy: Domain of cognitive ability

Figure 2.5 Bloom's taxonomy: Domain of affective ability

NOTE:
These figures are included on page 45 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.6 Bloom's taxonomy: Domain of psychomotor ability

Although the diagram for the cognitive domain implies a hierarchical structure for cognitive development, this is not always regarded as valid (Hancock, 1994; Isaacs, 1996). Other interpretations (Anderson, Krathwohl, Airasian, & Samuel, 2001) replace the nouns with verbs, to indicate active processes, and add an additional level (creating), as well as making modifications to the definitions of the lower levels. There also exist modified versions of Bloom's taxonomy; the REcall, Comprehension and Application (RECAP) model of learning suggested by Imrie (1995) and Entwistle and Brennan's (1971) classifications (Surface Passive, Surface Active, Deep Passive and Deep Active). Despite the existence of these different models, the classical version of the cognitive domain of Bloom's taxonomy is still widely used and will be the version referred to in the following discussion.

Testing Different Cognitive Levels

An assessment of any kind may aim to address any part of the cognitive domain, but some techniques are more appropriate to the cognitive skill they wish to test. A task such as 'Name the bones that form the skull', would be useful if it was desirable to test if a student had memorised the names of these bones (testing knowledge in Bloom's cognitive domain, multi-structural in the SOLO taxonomy, recall according to Imrie and Surface Active according to Entwistle and Brennan), but would offer little information about their ability to manage a patient admitted to an emergency department with a fractured skull, which would require an assessment task at a higher cognitive level.

Motivation

The issue of motivation for students is at the heart of the success or otherwise of any educational initiative, particularly one focusing on a formative approach to learning. Motivation can be broadly regarded as being either intrinsic or extrinsic. According to theory (Deci, Vallerand, Pelletier, & Ryan, 1991), an intrinsically motivated behaviour may be based on a person's interest and enjoyment of a task, with any reward being satisfaction at successfully completing the task or enhancing one's knowledge. An extrinsically motivated behaviour may be based on factors outside the person's control, such as a reward for satisfactorily finishing the task or it may also be based on internal factors, such as recognition of the value of the task. For example, consider a fictitious researcher, who enjoys designing and carrying out experimental research, but is a reluctant speaker and documenter of research, yet craves peer recognition for the work carried out. This researcher is intrinsically motivated to carry out the research. The researcher understands that as part of a research grant there is an obligation to document all results and report them at a conference. The motivation for carrying out this component of the work is extrinsically motivated by an external source since it is required as a condition of the grant, but it is also extrinsically motivated by an understanding that the work needs to be documented, as well as by the desire to reinforce the researcher's own ego by demonstrating the research to peers. The researcher understands that in order to carry out the work, strict occupational health and safety standards need to be met in the laboratory. This is another extrinsic motivation, but according to the model advocated by Ryan and Deci (U.S. Department of Education, 2002), is of a different nature to other motivators. The model and the example have been integrated in Figure 2.7.

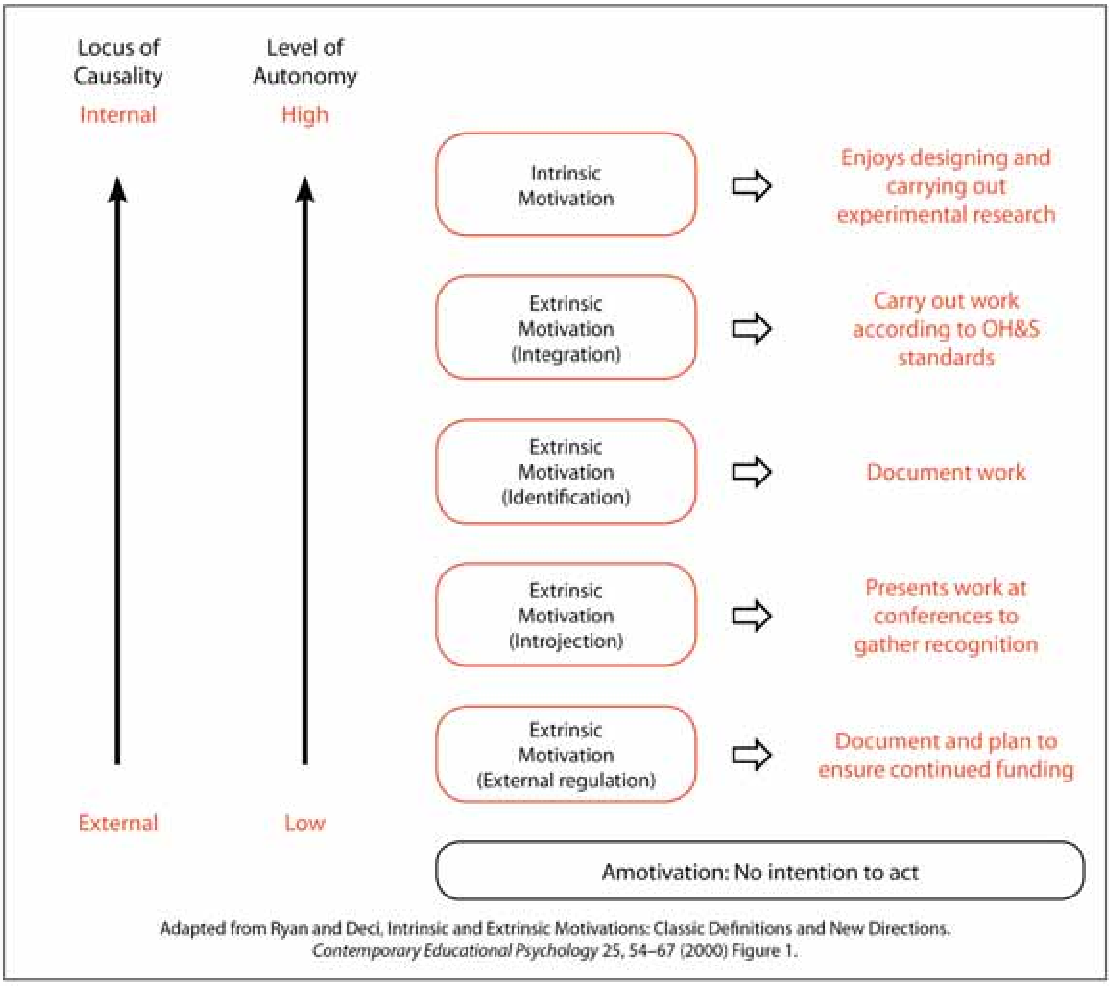


Figure 2.7 Levels of motivation

Key motivators for students using self-directed learning have been identified as attending a good lecture, having an interesting subject, doing work relevant to the chosen profession, curiosity, achieving the end result of being a professional, having a good mentor in professional life, receiving clear guidance and feedback (Bengtsson & Ohlsson, 2010; Regan, 2003) and wanting to pass an assignment (Mann, 1999). It is unlikely that any student would be intrinsically motivated by every learning situation they will encounter in their education and thus it is important for formative assessment tasks to provide additional motivation. An authentic assessment approach has been suggested as a useful approach to assist in providing this extra stimulus for students (Anfara, Anfara, Anfara, Andrews, & Mertens, 2005; Dochy & McDowell, 1997; Martin, Kulinna, & Cothran, 2002) and is intuitively appealing, although the empirical evidence supporting the motivational impact of undertaking authentic assessments is limited or contradictory at this time (Erwin & Wise, 2002; Peacock, 1997).

Authentic Assessment

At the heart of authentic assessment lies the themes of worthy and engaging exercises and a relevance to real world situations. One of the earliest definitions of authentic assessment comes from Archbald and Newman (1988, p. 1) who stated that assessments based on 'worthwhile, significant and meaningful' tasks could be regarded as authentic. Whilst their examples of what they considered to be authentic were based on high school situations, they clearly identified analysis and problem solving based on real-world problems as meeting their criteria. They also suggested that tasks would have greater authenticity if they had value in addition to instructional evaluation (Archbald & Newmann, 1988, p. 4); that is if the tasks were not fully summative. Newman and Wehlage (1993) defined a framework for authentic instruction, which has direct relevance to authentic assessment methodologies (Gulikers, Bastiaens, & Kirschner, 2004). They defined five standards as shown in Table 2.3.

Table 2.3 Standards for authentic instruction (Newmann & Wehlage, 1993)

NOTE:
This table is included on page 48
of the print copy of the thesis held in
the University of Adelaide Library.

An authentic assessment framework proposed by Gulikers (2004) suggested that authentic assessments should have five dimensions (Table 2.4)

Table 2.4 Dimensions of authentic assessments Gulikers (2004)

1	the assessment task,
2	the physical context,
3	the social context,
4	the assessment result or form, and
5	the assessment criteria.

According to this framework, the task should resemble real life problems as experienced by professionals in the field. This may entail complex tasks but high levels of complexity are not essential (Cronin, 1993; Van Merrienboer, Kirschner, & Kester, 2003). Importantly, the framework stresses that the task must be meaningful and relevant to the student. The criteria for authentic instruction mirror this approach. Ideally, authentic instruction encourages higher order thinking (Biggs & Collis, 1982; Bloom, 1956) and deeper learning (Marton & Saljo, 1976) and should have relevance to students outside the learning environment.

The physical environment used for the assessment needs to be faithful to the real-life environment encountered by professionals. This is an important motivator in medical education where medical students spend substantial amounts of time engaged in activities that involve patient contact. This relates to part of the authentic instruction framework that suggests learning needs to take place in the world beyond the classroom. This does not mean that every assessment task must be carried out in a surgery or laboratory, but the assessment should acknowledge the physical realities of the professional environment.

The social context for an authentic assessment suggests that social requirements of real-life problems be mirrored in the assessment. Where group work is required or social interactions may affect the successful addressing of an issue, it should be acknowledged in the assessment. This recognition of social importance is addressed in Newman's and Wehlage's (1993) framework, but is not a direct reflection of the social context dimension in the framework for authentic assessment.

The authentic instruction framework that Gulikers (2004) proposes suggests that there should be interaction between students and the instructive process either via a teacher or some other mechanism and social support needs to be provided in terms of a clear establishment of high standards and clear expectations. The social aspects of learning, in particular the opportunity to compare answers and rationales with others, appears as a strong theme in authentic assessment (Gulikers, et al., 2004) and is one of the core components of a problem-based learning approach using a constructivist framework (Savery & Duffy, 2001). Another part of the social side to instruction is that students with varying capacities for completing tasks should be equally encouraged to put significant effort into them and not be subject to behaviour that might damage their self-esteem.

The authentic assessment framework suggests that there should be sufficient tasks to be able to determine competency in the work. The assessment should be of sufficient quality and justifiable in order to demonstrate competence. Presenting the results of an assessment to fellow students or staff is one way this could be achieved. Clear criteria and standards, related to real-world situations also need to be attached to the assessment to ensure a transparent and defensible process.

To be effective, authentic instruction and assessment both need to be implemented (Gulikers, et al., 2004). They form integral parts of the feedback process illustrated in Figure 2.3 where students learn from the assessment and apply that learning to other modes of instruction. In particular, authentic assessment is a logical method of assessment for students in a PBL program. Combined with a formative assessment approach delivering appropriate levels of feedback, authentic assessment supports the constructivist approach of PBL. This clearly defines the nature of assessment that would be beneficial to medical students. The types of individual assessment tasks best suited for this needs to be determined. The remainder of this chapter will describe the types of assessment tasks that are best suited to this constructivist approach to PBL.

Important Concepts in Assessment

An assessment is composed of test items that may be of a wide variety of types, some of which include essays, multiple-choice questions, demonstrations of practical skills, and oral presentations. Irrespective of the type or style of item, there are important characteristics that help define how useful that item will be in measuring student ability, namely the (i) fidelity, (ii) validity and (iii) reliability of the assessment item (Mulholland, 1997; Palmer, Duggan, Devitt, & Russell, 2010).

Fidelity is a term used to show how relevant a task is to real-life situations. An assessment which focused on the use of old techniques to treat the acute abdomen would have a low fidelity if the goal was to help students learn how to handle the situation in modern hospitals. It may have high fidelity if students were studying medical history or were going to be working in conditions where modern techniques are not viable for resource reasons. An authentic assessment is by definition a high fidelity methodological process.

Validity is the term commonly used to indicate that an assessment item will measure what it was designed to measure. Consider three questions set to determine if students understand how to manage the acute abdomen:

> Question 1: ‘Explain briefly how you would manage a patient who presents with the symptoms of an acute abdomen?’
>
> Question 2: ‘What are the symptoms of the acute abdomen?’
>
> Question 3: ‘A patient presents at the Emergency Department complaining of stomach pains. How would you manage this patient?’

Questions 1 and 2 are clearly asking about the acute abdomen, but if the goal was to design a test question to see if a student understood how to the manage the condition, then Question 2 is not valid. The question is clear and worded well, yet does not test what was desired. This question fails the basic test of *face validity*. At close inspection it does not look as if it would provide the information required. Question 3 fails the test of *content validity* (Carmines & Zeller, 1979). Superficially, it may appear that it will lead to information on the students’ management skills of the acute abdomen, but in reality it may only be testing how they handle a patient who arrives in The Emergency Department. If it were possible to show that the questions did indeed test the ability of a student to manage a patient with the acute abdomen, it would have *predictive validity* and if it produced outcomes consistent with other tests designed to measure the management of the acute abdomen, it could be said to have *concurrent validity*. Both of these terms are components of *criterion validity* (Carmines & Zeller, 1979). By comparing these questions with other measured items and any underlying theories or rationale relevant to the acute abdomen it would be possible to determine if the test items had *construct validity* (Carmines & Zeller, 1979; Cronbach & Meehl, 1955). If there were a hypothesis that students who had not seen patients with the acute abdomen would do poorly answering this question, then it would be possible to test this type of validity.

Reliability. Even if the questions were not valid, they may be *reliable*. If that were the case and they were applied to groups of students of equivalent ability repeatedly, then they would yield results that were the same. If the management of the acute abdomen changed from one year to the next, then Question 1 may not be reliable over an extended time period (poor *test-retest* or *inter-rater* (*stability) reliability*), but may be reliable within it. Question 1 may also suffer from one of the most important concepts in reliability, poor inter-rater reliability (Downing, 2004), where one marker would grade the response in a completely different way to another. It is possible that Questions 1 and 3 are effectively testing the same thing and at the same level. In that case, they could be said to have *equivalency* or *parallel reliability*.

Figure 2.8 shows a 'Robin Hood' description of validity and reliability (Trochim, 2006) which is analogous to arrows being fired at a target. It is a simple visual aid that assists in understanding these tightly linked concepts.

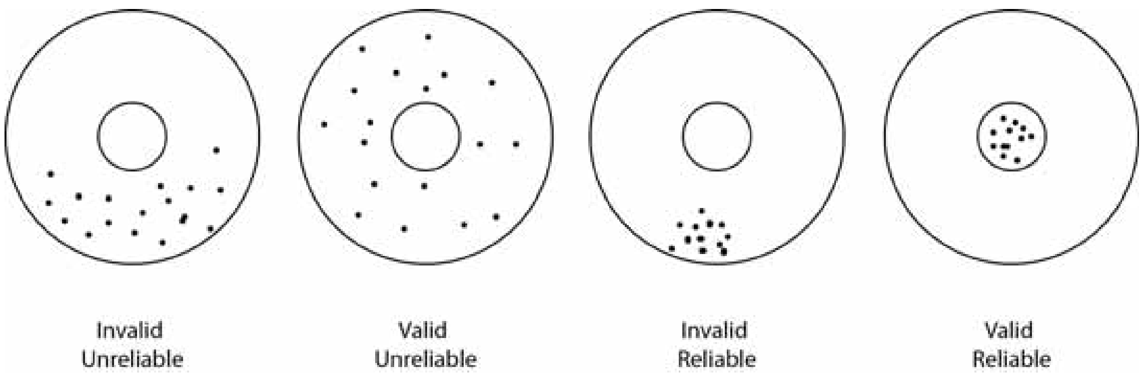


Figure 2.8 'Robin Hood' definition of reliability and validity

If the concepts illustrated by Figure 2.8 are applied to an assessment task, where a dot represents the outcome of a student response then the accurate completion of the task would be indicated by a dot in the centre of the 'target'. The first part of the figure shows the assessment to be unreliable, as students who attempted the task were inconsistent in their outcomes (widespread dots on the diagram) and invalid because on average students fell below the ideal centre point. Conversely, the last circle in Figure 2.8 shows a reliable assessment task (dots closely bound together) that was also valid (dots on average are in the centre).

In a summative assessment, especially in high stakes situations such as exit examinations, the assessment must have high validity and reliability (Downing, 2004), and should also have a high fidelity. Formative assessments may be able to have lower reliability, if they provide sufficient feedback to address misconceptions that may have been introduced due to the poor wording of tasks, or the lack of sufficient tasks to produce high reliability.

In addition to the commonly used characteristics of reliability and validity, an individual test item has other characteristics that can be measured. This measurement can be carried out using two distinct psychometric theories; classical test theory (CTT) or item response theory (IRT). The former theory has been used since the early twentieth century (Traub, 1997) and the characteristics of many types of assessments have been measured using it. Characteristics include the item difficulty, (a difficult item is one in which the majority of the class failed to provide a correct answer), and the item discrimination (a discriminating question is one where the correct response was provided by the majority of students who achieved high marks overall, but not by those students with low marks overall (Fan, 1998)). Reliability is often determined by a measure of the internal consistency of a series of test items and is often measured by Cronbach's alpha (Cronbach, 1951). CTT has the advantage of providing constructs that are easy to understand and calculate, but it lacks the rigorous theoretical background of IRT (Bock, 1997; Crisp & Palmer, 2009; Fan, 1998). Furthermore CTT makes no distinction between the characteristics of the person taking the test and the test items themselves. IRT makes this distinction and, at the cost of far greater complexity of calculation, provides a sound theoretical basis for item analysis. Item response theory provides mathematical constructs such as the item characteristic curve, which shows the probability that a person of a certain ability will be able to answer an item correctly. This curve provides an equivalent result to the CTT derived difficulty. Point bi-serial correlation is another important output from IRT and in some circumstances can provide equivalence with the CTT parameter discrimination. Researchers have examined the correlation between CTT and IRT and have found significant overlap in the results from the two methods (Fan, 1998; Macdonald & Paunonen, 2002), particularly when comparing difficulty and person statistics. Discrimination was comparable only when the difficulty of test items was confined to a narrow range.

A high quality assessment need not always have items of high discrimination or difficulty. For example, in a diagnostic assessment where one may be testing how well prepared a group of students is for a new topic, it may be desirable to have low levels of difficulty, testing only if students have a basic knowledge of core background material. The use of test items should always be considered in the context of the full assessment regime. Diagnostic, summative and formative assessments may utilise similar test items to achieve their goals, and may aim to test different educational domains from Bloom's taxonomy, but the aim of each type is not the same, although they may be used in conjunction with each other. It could be argued that all three types should be used throughout any type of tuition. Historically, this has often been the case in areas other than higher education and the following historical review of assessment provides insight into the use of different kinds of test items to help determine the most appropriate assessment items for formative assessment.

A Historical Perspective of Assessment

The words 'test' and 'examination' in the context of measuring or evaluating have been in existence for close to 400 years (Harper, 2001), yet the concept of 'testing' has been in existence for far longer. Formal assessment processes have been effectively used since Roman times. Roman soldiers were arguably some of the finest and well-disciplined soldiers in history (Campbell, 1994; Le Bohec & Bate, 2000) and the Roman Empire was built upon its conquests (Le Bohec & Bate, 2000). This was partly due to the assessment structure built into the profession of soldiering.

As a diagnostic assessment, prospective Roman soldiers were chosen as those who were deemed to have the characteristics necessary to be soldiers, such as intellect, physical attributes and social background (Le Bohec & Bate, 2000). Those with skills in horse riding may have progressed to the cavalry as a result of diagnostic assessment. Training included marching, swimming and weapon-handling (Campbell, 1994; Cavazzi), and it was only when soldiers were deemed competent with their weapon work against inanimate objects that they were permitted to advance to training against human combatants. This formative assessment ensured that recruits, in conjunction with their trainers, understood their capabilities and improved their abilities (Le Bohec & Bate, 2000). This authentic assessment regime provided a strong framework for soldiers to perform well in their duties after having completed training.

Apprentices in trades such as blacksmithing and candle making underwent similar processes. Guilds in England ensured the quality of work produced by guild members through a monitoring process (Rorabaugh, 1986), which could be deemed to be an ongoing summative assessment. Guild membership was only available to those who had produced work of a masterful quality, reflecting not just the skill of the creator of such work, but also their trainer or master. Circumstances in the English colonies were substantially different due to a small and widely dispersed population, but middle class children such as a young Benjamin Franklin were expected to become craftsmen. Benjamin Franklin did not meet the physical requirements to be a blacksmith, however his love of books provided his father with a rationale to apprentice him as a printer (Rorabaugh, 1986). In this case a diagnostic process proved fruitful for the prospective apprentice, master and to society as a whole.

Modern soldiers and apprentices undergo stringent levels of formative assessment, as do modern professionals such as airline pilots and surgeons. In some apprentice situations, works crafted by the apprentice are kept from them until they are deemed to be of a high enough standard by the master; the whole process highlighting student strengths and areas needing improvement for both student and master (Lancy, 1980). Summative assessment for the soldier is still as potentially disastrous as ever, but improved simulations and training techniques provide a solid formative ground. Research into combat outcomes provides greater insight into some of the diagnostic approaches that may be necessary to ensure combat group cohesion (Griffith, 1988).

The examples above show how the diagnostic, summative and formative assessment approaches carried out in an authentic instruction and assessment environment may lead to improved outcomes for individuals and society. Although the situations encountered by soldiers and apprentices are quite different from those usually encountered in higher education, it seems that the above approaches and types of assessment could translate naturally to higher education. However, students at higher education institutions often receive a broader education. Without the tight focus provided by a trade or military service, they may focus their learning efforts on assignments, essays, examinations or some other form of assessment, rather than the long-term outcomes. Despite the efforts of teachers to provide interesting, practical and relevant courses that link to desirable graduate attributes or professional requirements, student focus will usually be on 'How do I do well in this course?' For some students 'doing well' may be attaining a pass grade, for others it may be achieving high standards of excellence, but in the majority of cases it is likely that students will measure their achievement based on successful results of graded assessments rather than their enjoyment or perceived mastery of topics. In other words, the driving force behind student learning is summative assessment and although this has been the case in higher education for many years, it was not so when universities were first established. At that time the assessment more closely mirrored the apprentice model.

Early Assessment Methods in Higher Education

Assessment in the universities in medieval times (11th to 15th centuries) relied typically on the oral disputation, where the student would use their knowledge of texts and their intellectual skill to make an academic argument (Kilpatrick, 1993). Although a series of academic arguments is potentially beneficial for developing critical analysis skills, the didactic method of learning was based around memorisation and recitation (Wilbrink, 1997) and as a result the depth of learning outcomes was potentially limited. Nonetheless, recitations in class and amongst peers were a form of formative assessment that was frequently carried out as essential practice for the summative examination. Encouragement for students was usually based around punishment, often by humiliation (Wilbrink, 1997) and there were clear signs that this methodology had deteriorated from a method of demonstrating understanding to an opportunity to build and demonstrate power. The great Humanist, Juan Luis Vives, was scathing of the environment at the University of Paris, decrying long hours of lectures, the nature of the disputations, and the conceit of the richer students (Niss, 1993; Noreña, 1970). On the other hand, Powicke (1934) commended the thoroughness of the disputation system, particularly the mental capacity of lecturers and students who often relied on memory to deliver and receive lectures.

Final examinations in medieval universities revolved around the disputation. When students were deemed ready to sit an examination, they were provided with questions they could be asked and a topic on which they would argue. The examination was a public event and this summative exam must have been a major motivator for students. Evidence supporting this lies in the existence of exam compendia in the fourteenth century, providing examples of questions and answers (Wilbrink, 1997). Despite the public nature of the examination, the pressure on students sitting examinations was likely to be less than that of their modern counterparts due to the fact that students could only have the examination when the masters decided that the student was ready. As a consequence few students failed, although there may be less savoury reasons for this fact. According to Osipian (2004), the Paris Medical Faculty has no records of failing a student over a one hundred year period and this may be in part due to corruption as much as the ease of passing examinations.

Class recitation as an assessment tool would likely meet many of the conditions for assessment supporting learning as shown in Table 2.1. To prepare, students would have had to devote significant time and effort over a number of weeks. The expectations were high and both lecturer and peers provided feedback, even if it was not uniformly positive. From the descriptions provided above, the environment for learning could be hostile, unsupportive and even inequitable, meaning the criteria for good feedback practice as shown in Table 2.2 were only partially met. The cognitive level of the task appears to have been at the lower levels with memorisation and recitation being valued skills, but the task could be considered authentic because this was the way lecturers taught and the skills obtained from class recitations and disputations were directly transferable to the professional world. These assessments could be considered to be valid and of high fidelity for this reason as well. Reliability would have been a major issue with this form of assessment, due to the limited number of questions that could be asked of the student in a given time period and the now well known issues with inter-marker reliability, but as preparation for the summative disputation, the formative recitations in class could be considered appropriate. Strangely, the formative component of medieval universities, as cruel as it might seem by current standards, can therefore be perceived as a positive approach to assessment. It was tightly linked to the summative method of examination, as well as the goals of the universities, which was to produce people who could lecture and teach in a similar style to the way in which they themselves had been taught.

Medieval universities ranked their students in final examinations, usually according to their place in society (Wilbrink, 1997). Within classes, students could be given the honour of winning a dispute and the loser obtained the symbolic, but no less damaging title of *asinus* (meaning jackass), which was retained until the next dispute. It is easy to envisage students retaining such a title for many months and struggling for any positive recognition for their efforts. Lack of recognition for lesser performers was acknowledged as a problem with the giving of prizes only to top students and a ranking system based on merit eventuated, although the lowest ranked individual still gained the dubious gift of a man-sized wooden spoon (Wilbrink, 1997). In the late nineteenth and early twentieth centuries the ranking system evolved into providing marks to students, and this convention has been retained at most educational institutions since that time.

The medieval style of learning and assessment continued at universities for hundreds of years and it was not until the eighteenth century that written examinations were introduced. Larger numbers of students enrolling at universities began to make the oral type of examination unwieldy and on occasion pointless. One candidate in the 1700s passed their examination by answering only two short answer questions, one of which was contentious (Ballard, 1944). Written examinations became more popular, particularly in the sciences and mathematics (Kilpatrick, 1993), and by the mid-nineteenth century were becoming a feature of education at all levels, from elementary school to university. Although there was a massive backlash at the over-examination of young children towards the end of the nineteenth century (Ballard, 1944) that resulted in the use of written examinations being reduced markedly, the written examination eventually became a standard method of assessment at universities.

Up to the twentieth century, university assessment evolved from a form similar to apprenticeships and military training, with a significant formative basis, to something more similar to modern methods. The latter stages of this evolution were in part dictated by increasing numbers of students attending universities. This was caused by many factors, including the growth of society, marketing and the greater opportunities for education (Thelin, 1994). American higher education had a thirty-fold increase in enrolments over the period 1870-1940 (Hofstadter & Hardy, 1952). Responding to the growth in enrolments by altering the assessment regime would have made sense from a purely logistical point of view, but the changes in assessment were bound to have an effect on how students would study. Commenting on grading examinations an inquirer stated

> ‘We should prefer not to use the marking system, but we do not think it possible to do away with it altogether at present. We have reduced it to its lowest limits. The objections to it are obvious. Working for marks is not an elevating occupation’ (n.a. cited in (Hawkins, 1960, p. 249).

The student focus would have changed from preparing for regular disputations to ensuring that the knowledge required to pass a written exam was memorised and understood. This is typified by the words of a young Woodrow Wilson at Johns Hopkins University in 1884, where he discussed the wasted effort of cramming for an exam,

> ‘The examination (in colonial history) took place on Saturday last, and a very fair, sensible examination it was. Adams gave out five topics and told us to choose, each man for himself, one from the number and spend the two hours in writing an informal essay upon it. That's the sort of examination I like. But it wasn't the sort I had expected; and I went in crammed with one or two hundred dates and one or two thousand minute particulars about the quarrels of nobody knows who with an obscure Governor, for nobody knows what. Just think of all that energy wasted! The only comfort is that this mass of information won't long burden me. I shall forget it with great ease’ Wilson cited in (Hawkins, 1960, p. 283)

Modern assessment arose, at least partially, through a number of completely unrelated factors including increased student numbers at universities, rather than being a natural assessment technique arising from traditional teaching methods. Didactic methods of teaching based on an era where memorisation was an essential component of learning, continued to exist, and became more focused on grades as the formative method of learning, up to that time the disputation, faded.

Written assessment was only the first of many different assessment types developed, implemented and evaluated over the period of the twentieth century. All of these types of assessment have application in diagnostic, summative and formative assessments and are discussed below.

Types of Assessment

Common types of assessments are often divided into the categories of constructed and selected response formats (Osterlind, 1997). Constructed response items for assessment rely on students developing answers to questions using their own words and without cueing (Cronbach, 1984). Assessment items in this category include essays, short essays, short answer questions, modified essay questions (MEQs) and clinical evaluation exercises (CEX). Disputations, oral presentations and objective structured clinical examinations (OSCEs) could be regarded as existing in this category as well, due to the requirement for students to provide information and demonstrate understanding and problem solving skills by completing tasks that are essentially open ended. Selected response items, in contrast, provide students with boundaries that cannot be exceeded by suggesting possible solutions to the test item (Popham, 2003). Items in this category include multiple-choice questions (MCQs), true-false questions, multiple response questions, matching items questions, ranking questions, extended matching questions and script concordance questions.

Constructed Response Items

Essays

Essay questions were the first kind of written questions used for assessment. They are a logical extension of the disputation method of assessment and usually require a student to argue a position and support it with evidence. Essays are usually substantial pieces of work, often more than 1000 words in length. In an exam situation, several pages may be required, but in an assignment an essay could range from 1000–10,000 words depending on the requirements of the assessment designer. A possible essay question in clinical medicine might be:

> Explain the common diseases that affect the human liver. Describe the mechanisms causing these diseases, how patients may present with them and explain the appropriate management, including investigations and treatment.

There are numerous advantages to this type of assessment. Students are able to utilise their knowledge and understanding of a topic, develop an argument, apply critical analysis to the issue, and use their skills in writing to communicate this to others. As a method of testing higher order cognitive abilities, the essay is highly regarded (Brown, Bull, & Pendlebury, 1997; Popham, 2003; Terry, 1933) and is commonly used (Biggs & Tang, 1999). The major disadvantage with essays is the time required to mark them (Fowell & Bligh, 1998; Norton, 1990) and the ability to apply a consistent marking scheme to a large number of essays (Popham, 2003). Moderation of essays is important, but also resource intensive, and where multiple markers are employed, inter-marker reliability is an issue. A piece of work regarded as deserving a credit standard from one marker may be regarded as being worth a distinction or a pass by other markers. These difficulties have resulted in the essay recently being dropped from the MRCPsych Examinations held by the Royal College of Psychiatrists in the UK (Palaniyappan, 2008), although they still persist in other postgraduate institutions such as the Royal Australasian College of Surgeons (Royal Australasian College of Surgeons, 2007).

Henderson (1980) proposed three strengths of the use of essays for learning, one of which was their use for summative assessment. He suggested that the essay is also intended to aid student learning, and to provide a diagnostic dialogue between students and teachers. Although the words ‘formative assessment’ are not used, these two features are clear strengths of the essay in a non-summative situation. A clear communication channel can be established between students and teachers, and if sufficient time can be devoted to both the writing and reading of the essay, a clear understanding of the capabilities of a student can be established. The difficulties of doing this in a class of several hundred cannot be underestimated. Some markers would not see the students for whom they are marking papers in a face-to-face situation and there can be little doubt that the effects of fatigue and tight deadlines would have an effect on the markers ability to look insightfully at student capabilities. Lackey and

Lackey (2002) highlight other issues with essay marking, such as the place in the marking stack where the essay is, and whether it follows an excellent essay or not as being key issues. Nonetheless, as a formative tool, the essay can be powerful for learning.

Short essays

Short essay questions require a coherent, reasoned argument as an answer to a question much as a standard essay would, but are unlikely to extend beyond 500 words. This is a challenging form of assessment as it can be difficult for the student to distil an argument down to such a short form. In a summative exam, there could be ten or more short essay questions set for a three-hour exam, thus allowing for greater breadth of content to be assessed. A possible short essay question in clinical Medicine might be

> Describe the mechanism causing each of the three major forms of alcoholic liver disease.

Institutions such as the Royal College of Obstetricians and Gynaecologists in the UK use short answer questions requiring answers of 250–450 words to assess prospective members (Khaled, Ellis, & Buck, 1999; Royal College of Obstetricians and Gynaecologists, 2009). The Irish College of General Practitioners asks four questions over one and three quarter hours and restricts the examinee to two sides of A4 paper per answer (Irish College of General Practitioners, 2009). The advantages and disadvantages of short essay questions are similar to those of long essays. In many cases, the former have greater need for moderation, as there is less opportunity for a marker to assess whether a student thoroughly understands what they are writing.

Modified essay questions

Modified essay questions (MEQ) are a variation on short essay questions and are often used in medical examinations. They were initially developed for the Royal College of General Practitioners in the UK (The Board of Censors of the Royal College of General Practitioners, 1971). They pose a problem and stage-by-stage, additional features of the problem are revealed, which require analysis by the student. A possible MEQ in clinical Medicine might be:

Stage 1: A 44-year-old man presents with his mother at your clinic with a three-month history of dyspepsia and an infected wound on his elbow which has failed to heal over the last 2 weeks. He is dishevelled and has a number of bruises on his body. His mother states he isn't taking care of himself and drinks too much. What are likely differential diagnoses at this stage and what are relevant questions you would ask based on the patient's history?

Stage 2: On closer examination, the patient is undernourished. He responds in a hostile manner to questions about alcohol. What is your next immediate step and what are your plans for long term management of this patient?

In theory, these questions simulate the management of a clinical case and test higher order problem solving skills (Stratford & Pierce-Fenn, 1985), thus making the assessment process more valid. MEQ papers can be easier to develop, but in reality MEQ questions are challenging to write effectively, require substantial marking time, pose inter-marker reliability issues and may not test higher level cognitive skills any more than other formats of test items which are substantially easier to mark (Feletti & Smith, 1986; Palmer & Devitt, 2007b). MEQs have been criticised for their inability to consistently and rigorously test clinical material to the high standards required of a medical graduate where the emphasis is expected to be on assessment of the higher cognitive skills (Epstein, 2007; Ferguson, 2006), but they are still used in assessing problem-based learning in some universities as they are deemed to be valid and reliable (The University of Sydney, 2006)

Short answer questions

Short answer questions typically require one word or one sentence responses. They can be viewed as a compromise between essay questions and selected response questions such as the MCQ (Fowell & Bligh, 1998), as they are relatively easy to set and mark. A possible short answer question in clinical Medicine might be:

'What are the three main types of alcoholic liver disease?'

The main perceived benefit of these types of questions is that they require students to provide simple answers, without any form of cueing (Rademakers, Cate, & Bär, 2005). In terms of reliability, short answer questions have been shown to have reliabilities greater than essay questions (Wakeford & Roberts, 1984), but are deemed no more reliable than MCQs (Bacon, 2003; Wakeford & Roberts, 1984; Webber, 1992). The nature of this type of question makes it likely that they test recall only and not higher order skills.

Objective structured clinical examination

The objective structured clinical examination (OSCE) (Newble, 2004) was first introduced in the late 1970s as a potentially reliable and authentic method of assessing the clinical skills of students (Harden, Stevenson, Downie, & Wilson, 1975). It is an assessment form that uses a series of short (5–10 minute) standardised 'stations', each one testing a component of clinical skills such as the ability of a student to examine a patient or take a history. Actors often portray patients and this ensures the consistency of information and responses provided to students as they progress from one station to the next (Adamo, 2003; Harden, et al., 1975). A possible OSCE in clinical medicine might be:

> 'Take the history of this patient who presented to your clinic having been found unconscious on the street 2 hours ago.'

Collectively, the results from each station are used to assess the ability of students to manage real cases. The main drawback with OSCEs is that they require significant amounts of time and resources to run, including actors and assessors. An OSCE requires a large number of stations to form a reliable assessment. A ninety minute long OSCE may only have a reliability of around 0.6 (Brown, Manogue, & Martin, 1999; Newble, 2004; Schwartz, et al., 1998) and, in summative assessments it is recommended to combine an OSCE with other forms of assessment to ensure higher reliability (Newble, 2004) or alternatively, run them over longer periods with larger numbers of stations in order to obtain higher reliabilities (Cohen, Reznick, Taylor, Provan, & Rothman, 1990). An OSCE would, however, be a formidable formative assessment tool. It is based completely in the clinical realm and thus satisfies authenticity, tests higher order cognitive ability and could theoretically provide immediate feedback from an assessor if sufficient resources were available. The resource issues dominate the implementation of a formative OSCE. Developing an OSCE purely for this purpose may be beyond the capabilities of many areas.

Clinical evaluation exercises

Clinical evaluation exercises (CEX) and the mini-CEX are assessment tasks where students are observed interacting with real patients (Norcini, Blank, Duffy, & Fortna, 2003). The original CEX proposed in 1995 by the American Board of Internal Medicine (Norcini, Blank, Arnold, & Kimball, 1995) involved a practitioner evaluating the performance of a student with one patient and providing feedback over a period of two hours. Significant issues with this method revolved around the marker reliability and the fact that only one clinical scenario was involved (Norcini, et al., 1995). The mini-CEX addressed these shortcomings by introducing a series of short observations (10–15 minutes) over an extended period with different assessors, thus allowing a wider range of skills to be tested, and overcoming the marker reliability issues, but at the expense of each patient encounter being a more superficial history-taking or examination process. The main disadvantage of these types of assessments is that they are extremely labour intensive, requiring significant administrative and clinical input from a large number of staff, although it has been shown that a high reliability can be obtained with ten or fewer assessments (Nair, et al., 2008). A mini-CEX would be quite similar to a ward round, and is clearly an authentic task, but lacks the consistency that a formal assessment procedure would provide. If ward specialists could provide sufficient constructive feedback, the mini-CEX would be a useful formative tool.

Summary of constructed response assessments

Constructed response items as assessment tools meet many of the conditions for assessment supporting learning as shown in Table 2.1. To develop an essay or prepare for an OSCE or CEX, students need to devote significant time and effort over a number of weeks. The expectations from these types of assessments are high but the ability to supply meaningful feedback in a formative situation can be difficult. The use of rubrics to assist in both marking and feedback provision can be helpful and assist in alleviating some of the resource issues associated with marking essays. Peer or self-assessment is another strategy that can be used to reduce academic marking time by replacing it with students acting as assessors for both themselves and their peers as a learning exercise (Dochy, Segers, & Sluijsmans, 1999; Nilson, 2003).

Most of the conditions for good feedback practice as shown in Table 2.2 can be met when using constructed response assessments. The cognitive level of the task is aimed at assessing higher order skills and the task may be considered authentic because in many cases the essay or other constructed response item could be written to reflect real world problems. These assessments could be considered valid and of high fidelity for this reason. Reliability is a major issue with this form of assessment, due to the limited number of questions that could be asked of students in a given time period and the problem with assuring inter-marker consistency.

The time taken to grade or provide feedback on these types of assessment is significant and there have been many efforts to try and automate some of these processes. Until recently it had been assumed that such marking could only be done manually and it has been estimated that teachers may spend up to 30 percent of their professional time on marking (Mason & Grove-Stephensen, 2002). In addition, there are concerns about the quality of marking essay type questions as illustrated by studies on the reliability of double marking (Brooks, 2004). Software programs have been constructed to mark essay-type material and some are available commercially (Educational Testing Service, 2009; Page, 2005; Pearson Education, 2009; Valenti, Neri, & Cucchiarelli, 2003; Vantage Learning, 2006).

These programs are complex tools that analyse essays using varied methods such as natural language processing (Association for the Advancement of Artificial Intelligence, 2009), ‘proxes’ in Project Essay Grade (Page, 2005), singular value decomposition in Intelligent Essay Assessor (Pearson Education, 2009) or Bayesian analysis in the program BETSY (Rudner & Liang, 2002). They primarily focus on the provision of summative marking for the benefit of those who have set the questions.

Unfortunately many of these systems have to be ‘trained’ to mark. In other words, a percentage of papers need to be hand-marked in order for key material to be identified and fed into the program. Most of these systems require 100-200 essays to be marked to complete the training. After this point, marking of additional essays is extremely quick and has been reported as extremely reliable (Attali & Burstein, 2006; Rudner, Garcia, & Welch, 2006; Valenti, et al., 2003). For formative assessment, where there may be 5 or 6 short essay questions per student per month, or where there are fewer than 100 students, these tools are of little benefit due to the training requirement of the software and the lack of feedback mechanisms. The focus of these programs is to provide a marking process for teachers rather than to provide immediate feedback on performance to students (Mitchell, Aldridge, & Broomhead, 2003). Thus many of these automated marking tools have relatively limited value to students who seek immediate feedback, particularly on formative material. For formats of assessment that are easy to mark and have high reliability, there is a need to look at selected response items.

Selected Response Items

At the beginning of the twentieth century, concerns about the essay grew rapidly. In part, these concerns were due to the growing belief that it was possible to scientifically measure many factors to do with education, including intelligence. The subjective nature of the essay was at odds with this new mindset. Edward L. Thorndike, a psychologist and an early educational researcher, was one of the most significant figures of the early 1900s in this area, encouraging this scientific approach towards education and the measurement of outcomes (Januszewski, 2001; Thorndike, 1923). His influence played a role in the development of the first MCQ by Kelly in 1914-1915 (Sokal, 1990), who used it as part of a reading test administered to children (Hall, 2003). The ability to mark this new test item quickly and without ambiguity was a powerful lure for evaluators, and in a few short years was put to practice on a large scale. An intelligence test, based on the early work of Binet and Simon in France (Minton, 1905; Wasserman & Tulsky, 2005), and Otis, Goddard and Terman in the US (Hall, 2003), consisting of MCQs, was used by Robert Yerkes on over one million servicemen in the US during World War I (Encarta; Feldhusen, 2003; Minton, 1990). The successful development, use and analysis (Thorndike, 1921) of this measurement tool was to influence educational assessment across the world, starting with entrance examinations to colleges in the US.

At the beginning of the twentieth century, US colleges across the country had non-uniform entrance requirements, causing significant difficulties for high school officials trying to prepare students for university (Linn, 2001). The newly formed College Entrance Examination Board devised standards and over a period of years, an increasing number of students sat examinations in an attempt to gain entrance to a university based upon those standards. Primarily consisting of short essay questions (Frontline, 2008), the examinations were subject to criticism as the measurement mentality swept the world, which drew attention to the deficiencies in essay questions, namely their subjective nature, as well as the significant amount of time required to mark them.

A MCQ exam, the Scholastic Aptitude Test (SAT), was devised but it took some time for the College Board to implement the SAT at the most highly regarded universities in the US (Ravitch, 2006). The loss of many people during World War II meant that it became logistically difficult to have essay-based entrance examinations marked, whereas the SAT could be marked rapidly with limited resources. Machine reading of the examination papers was a development that occurred rapidly after their introduction (Paris & Stahl, 2005). The MCQ would hold a primary position in educational assessment from that time and influence the development of other selected response assessment items over the coming decades.

Multiple-choice questions

MCQs, or 'objective tests' as they were initially termed, began to be used in many areas of society. These included aptitude tests (Gleason, 1926; Hull, 1928), tests of reasoning on children (Dashiell, 1928), and interpretation of Rorschach ink blots (Harrower-Erickson, 1943). Today, they are a common method of assessment. The MCQ consists of a stem which outlines the problem, a question and a correct answer situated amongst other potential answers or distracters. The number of available choices is typically four or five, but occasionally two has been used (Hull, 1928), in which case they are often referred to as alternate-choice items. Evidence suggests that there is little difference in the discrimination and reliability of test items with four or five choices (Ramos & Stern, 1973), but the effect of guessing would be heightened by fewer distracters. Students are expected to select one of these choices as the correct answer to the question. These questions are typically associated with testing knowledge but, correctly constructed, are capable of testing deeper forms of learning (Palmer & Devitt, 2007b). Seemingly easy to construct, in reality these types of questions require careful design (Haladyna &

Downing, 1989a, 1989b; Haladyna, Downing, & Rodriguez, 2002) and testing to ensure students are reliably assessed. A MCQ to assess medical students may look like this:

> A patient presents with a three day history of right upper quadrant pain and dark urine. He is 44 years old and his skin and the sclera have a yellowish tinge. He has been homeless for several years. What is the most likely diagnosis?
>
> A. Alcoholic liver disease
>
> B. Common duct stones
>
> C. Viral hepatitis
>
> D. Cholecystitis
>
> E. Carcinoma of the head of the pancreas

There are a number of advantages associated with these types of questions, typically the ease of marking, the apparent ease of construction and the ability to test many aspects of a topic in a short period of time. Disadvantages include that poorly constructed questions are often used and that many educators feel that they test knowledge only. Students can guess the correct answer to a question. Although the odds against passing an examination by blind-guessing are high if sufficient test items are included in the assessment (less than three chances in 10,000 in scoring better than 15 in a 30 item assessment, assuming four distracters of equal quality and one correct answer), if a student is capable enough to score close to half marks, it is quite possible to guess sufficient correct answers to pass. In the earlier example of 30 MCQs, if students know ten of the answers, they only need five correct guesses from 20 questions and the odds for this are substantially better (nearly 4 in 10), and can be further improved if the student can reject some choices before guessing.

An additional problem is the effect of cueing (Schuwirth, van der Vleuten, & Donkers, 1996), supplying students with potential answers that they may not have considered if they were not placed before them. MCQs have also come under a recent legal challenge. A dyslexic student has challenged the General Medical Council in the UK, claiming that the assessment format discriminates against her, and does not allow her to demonstrate her true skills (Lipsett, 2008). A successful legal challenge could leave institutions, not just educational ones, with significant problems, as many tests worldwide involve the use of MCQs. According to Collins (2006, p. 544), MCQs are the most commonly used test items in continuing medical education exams.

True-false questions

True-false questions are usually a series of statements that are either correct or incorrect. Often considered a poor cousin to multiple-choice questions, they have nevertheless been used for many years, at least since World War I, when they formed part of the army tests run by Yerkes (Hull, 1928). Students answering these questions need to make simple choices, but a series of these types of questions can provide a sound method for ensuring that basic knowledge or understanding is present and as such are useful questions for barrier type examinations. A true-false question may be:

> A patient presents at clinic with a history of alcohol abuse. He is 44 years old and his skin and the sclera have a yellowish tinge. The most likely diagnosis is Jaundice caused by gallstones (True or False?)

True-false questions have the advantage that they are the easiest to construct and the fastest of all types of objective tests to answer (Hull, 1928), thus allowing many more to be included in an assessment and sampling the relevant material more broadly. Although figures vary somewhat, an average of 1.5 true-false questions can be answered in the time it takes to complete one multiple-choice question (Frisbie & Becker, 1991), although this will depend on content (Frisbie, 1973). There is rarely any ambiguity in marking this type of question, unless the question has been poorly worded (a concern that applies to all types of assessments). From that perspective, true-false questions are easier to create than MCQs due to the additional work required to design effective distracters in the latter type of test item (Frisbie & Becker, 1991).

Concerns that the true-false question is not useful for testing were allayed in the early days of their use in higher education by measurement studies showing an acceptable reliability and correlation with written assessments (Gates, 1921; Kellogg & Payne, 1938; Knight, 1922). Concerns about the triviality of such questions have more to do with question design than item format (Ebel, 1970). Informed guessing in this format can be influential, although the literature offers opposing views regarding this tactic (Burton, 2002; Downing, 2003). Strategies for dealing with guessing include reducing marks if questions are not answered, or if they are answered incorrectly. Consider the example above where there are 30 test items with a pass mark of 15. The chances of blind-guessing 15 answers is nearly 60 percent, but if the student knows 10 answers and needs to guess 5 of the remaining 20, the odds are significantly improved (nearly 99 percent) that they will pass.

Extended matching questions

Extended matching questions address the cueing problems inherent in multiple-choice questions by providing a much longer list of distracters (Case & Swanson, 1996). By providing an exhaustive list of possible answers, the extended matching question places greater onus on a student's ability to provide their own answer, but at the same time leveraging many of the benefits of MCQs, such as reduced marking time and reproducible marking (Wood, 2003). These questions take longer to construct and consume significantly more space in a written examination and take longer for students to answer. They come with the benefit of being more discriminating than true-false or multiple-choice questions (Fenderson, Damjanov, Robeson, Veloski, & Rubin, 1997). An example of an extended matching question is:

A. Alcoholic liver disease

B. Common duct stones

C. Viral hepatitis

D. Cholecystitis

E. Acute Alcohol Poisoning

F. Myocardial Infarction

G. Iron deficiency

H. Gilbert Syndrome

I. Zellweger Syndrome

J. Fatty Liver

K. Enlarged Liver

L. Liver Cysts

M. Cirrhosis of the liver

> A patient presents with a three day history of right upper quadrant pain and dark urine. What is the most likely diagnosis? (Select your answer A-M from the list above)

More than one question would usually be asked relating to the extended matching list, thus providing benefits of economy in developing the list. The list may be of any length, but is often limited by practical matters, such as the maximum size that can fit on paper or the number of options that can be read by scanning devices (Fowell & Bligh, 1998). An extended matching assessment item does reduce the chance of a student guessing their way to success. Using the 30 test items example from above, with a pass mark of 15 and an extended matching list of 13 items, the chance of blind-guessing 15 answers is one in a billion and even if students know ten answers the chance of them guessing the remaining five required to pass is less than 2 percent.

Multiple response questions

Multiple response questions can be regarded as questions that ask the student to pick all appropriate options ('pick N' questions) (Case & Swanson, 1996), or a collection of multiple true-false questions linked thematically by a stem as shown below:

> A patient presents with a three day history of right upper quadrant pain and dark frothy urine. He is 44 years old and his mother states that he is normally in a confused state. Which of the following are likely to be present in this patient?
>
> A. Jaundice
>
> B. Gallstones
>
> C. Malnutrition
>
> D. Cirrhosis of the liver
>
> E. Spider naevi

In this case, more than one option is correct. These types of items are considered useful for measuring concept mastery (Richardson, 1992) and are comparable with MCQs when it comes to measuring achievement (Frisbie & Sweeney, 1982), and offer higher reliability (Downing, Baranowski, Grosso, & Norcini, 1995). In summative assessments, it has been observed that multiple true-false questions may be a better method for testing higher order cognitive skills (Downing, et al., 1995), but are nonetheless regarded as ‘a very efficient method of testing examinee knowledge of characteristics of complex phenomena’ (Albanese & Sabers, 1988, p. 111).

Script concordance tests

In a script concordance (SC) test, students are presented with a clinical vignette, incompletely describing a clinical problem (Sibert, Darmoni, Dahamna, Weber, & Charlin, 2005). Students are then asked to determine what effect different pieces of information, combined with a clinical hypothesis, would have on their decision making. Students identify the impact of the information on a scale running from –2 to +2 as shown in the example below:

> A patient presents at clinic with a long history of alcohol abuse. He is 44 years old and his mother states that he is normally in a confused state. In your management if you were thinking of:
>
> a diagnosis of cirrhosis of the liver and found gallstones then your diagnosis is +2 to –2 where;
>
> -2 indicates the diagnosis is absolutely incorrect
>
> -1 indicates it needs to be reconsidered
>
> 0 indicates your diagnosis is unaffected
>
> +1 indicates your diagnosis is supported and
>
> +2 indicates that your diagnosis is completely correct.

This form of assessment was first described in 2000 (Charlin, Roy, Brailovsky, Goulet, & van der Vleuten, 2000) and its goal is to compare whether the items (or scripts) of knowledge students possess are linked appropriately and match those of experienced clinicians. The questions require 10 to 15 qualified clinicians to answer them in order to define the scale (Meterissian, 2006) and may appear at first glance to be resource intensive to develop. The advantage provided by this type of question, however, is that it deals with everyday clinical situations and this can make the process of completing these questions quite rapid (Fournier, Demeester, & Charlin, 2008). SC tests require up to 100 items to obtain sufficient reliability in a summative assessment and have been shown to distinguish effectively between experts and non-experts (Meterissian, Zabolotny, Gagnon, & Charlin, 2007). It is a potentially useful formative assessment item if it provides specialised feedback on choices selected by students.

Summary of selected response assessments

Selected response items as assessment tools meet many of the conditions for assessment supporting learning as shown in Table 2.1. Although much of the discussion of the selected response assessment items in the literature clearly relates to summative assessment, where higher reliabilities and the effects of guessing are important, these types of assessments can be easily applied to formative assessment. Feedback to students on these types of items has not typically been a feature of the literature. However, if feedback was provided to students on these types of questions, clearly articulating why a specific response is the best answer or not, there can be obvious benefit for both student and educator. A student will benefit from feedback through improved learning, and their educator will benefit by setting a formative assessment that is easy to grade.

Selected response items, as shown in Table 2.2, meet most of the conditions for good feedback practice. The cognitive level of these kinds of items can be aimed at assessing higher order skills, and they can be written as clinical scenarios, thus becoming authentic assessments. These assessments could be considered valid and of high fidelity for this reason. Reliability is rarely a problem with this form of assessment, due to the large number of questions that could be asked of the student in a given time period. Issues with inter-marker reliability are non-existent as a simple template is all that is required to ensure consistent and accurate grading.

Scoring

Whilst it is recognised that scoring is essential in summative assessments, there is no such agreement around the scoring of formative assessments. Studies of primary school children have shown that the provision of grades only is not as effective as the provision of meaningful feedback (Butler, 1987). Although this may not directly translate into issues with scoring accompanying feedback in higher education, the results have been considered in the context of motivation and the effect on the psychology of the student and cannot be lightly dismissed. For medical trainees, it has been proposed that feedback should not contain any of the qualities typically associated with a grade, such as judgement or comparisons with others (Wood, 2000). This is to ensure that it does not become a measure of worth. There are additional concerns that scoring may reduce student engagement with feedback and authors have suggested a variety of techniques that can be used in lieu of scoring (Black, Harrison, Lee, Marshall, & Wiliam, 2004).

Is scoring automatically a poor option? As a guide to performance and highlighting areas which can be improved by the learner (the gap as defined by Sadler (1989)) it can be useful in subjects such as maths (Foster & Poppers, 2009). Removing scoring may not be a good option because little is known of the effects this type of extrinsic motivator can have on student performance. On the contrary, since a score does contain some information about competence it can have positive effects (Stipek, 1996) It is unlikely students will fully understand the concepts behind a problem the first time they study it. If a scoring system can motivate them to tackle the problem again, as demonstrated by Devitt and Palmer (1998), then such a strategy may have a positive outcome in terms of student learning. A scoring system may form part of more comprehensive feedback by encouraging student reflection and to allow them to comprehend the relevant weightings given to assessment results. It is likely that the use of scoring is dependent on the context in which it is used and any additional feedback provided. As such, it may be appropriate to allow options for scoring in formative assessments bearing in mind that scoring can have unwanted negative effects, such as taking the students attention from the rationale behind the score (Sadler, 1989).

Discussion

Accuracy and objectivity is of crucial importance in many assessment processes. The so-called 'high stakes' examinations in Medicine are examples where variations of small percentage points can mean the difference between success and failure. Achieving this accuracy in marking is relatively simple when the process depends on multiple-choice questions and, provided the questions have appropriate statistical power, degrees of difficulty and validity, the results can be highly reliable (Norcini, Swanson, Grosso, & Webster, 1985; Wass, Van der Vleuten, Shatzer, & Jones, 2001). These examinations are used extensively to test candidates' basic knowledge of a subject. Testing higher order cognitive skills (e.g. deduction, synthesis) (Bloom, 1956) can be more difficult using MCQs only (Martinez, 1999; van der Vleuten, 1996).

Apart from the need to provide an accurate summative assessment process, it is also important to give students the opportunity to test themselves formatively in preparation for the high stakes assessment.

One of the key goals of a medical curriculum is to provide motivation and direction for student-centered learning. In the absence of appropriate direction, learning can be an inefficient and time-consuming process, and without suitable goals and guidelines, learners can easily drift away from areas in which they should be focused. An important role of the teacher is to assist and guide students in their learning; to develop and define appropriate strategies for students and to help them make the most effective use of the time they have available to study. For better or worse, a strong stimulus to encourage 'learning' is some form of assessment. Traditionally, this has been in the form of summative assessment such as an end-of-course barrier examination. This method focuses students' minds towards a single goal, but tends to foster rote learning with the inevitable 'is this going to be in the exam?' approach to the choice of material studied. This barrier assessment method governs student decisions on what they will attempt to learn and how they will learn it (Terry, 1933), but is 'essentially passive and does not normally have immediate impact on learning' (Sadler, 1989).

The impact of summative assessment on the learning process for students should not be underestimated and may have a negative impact for some students on the motivation to learn (Black & Wiliam, 1998). The effect summative assessments can have on some students is shown by the extraordinary efforts some may go to in order to obtain information about the content of examinations. Apart from quizzing staff and previous students, and looking at old examination papers, methods of subverting a fair and equitable system include plagiarism (Roig & Caso, 2005), using mobile phones and other technology to assist in exams (Hrabak, et al., 2004), selling and buying examination questions from the 'black market' (Chaput De Saintonge & Pavlovic, 2004) and bribery (Hrabak, et al., 2004). The proportion of students cheating in examinations is substantial (Diekhoff, Labeff, Shinohara, & Yasukawa, 1999; Genereux

& McLeod, 1995). To prevent students from feeling that such extreme measures are necessary, they should be made to feel they are capable of completing assessment tasks.

In the early 1900s, Thorndike (1911) developed a number of laws based on animal studies, which he applied to humans. One, the ‘law of effect’, proposed that if an action was taken that had a positive result it was more likely to be repeated. This law formed the heart of Skinner’s theory of reinforcement (Januszewski, 2001). Similarly Vegetius, an ancient scholar, wrote ‘No one is afraid to perform what he is confident he has learned well’ (in Campbell, 1994, p. 15). Along these lines of thought, a preferred stimulus for learning should be some sort of formative assessment process coupled with reinforcement of the positive aspects of the performance of the task.

The concept of formative assessment has been promoted as a means of raising the standards of achievement within the classroom, particularly in primary and secondary education (Black & Wiliam, 1998). Formative assessment can be defined as some form of self-assessment by the student, which will provide feedback to both the student and teacher. This feedback is then used to modify teaching and learning to meet students’ needs. Designers of medical curricula have grasped this strategy with enthusiasm as an apparent means of ensuring deeper learning and understanding.

Within the clinical context formative assessment might be used to encourage appropriate professional behaviour, to foster clinical competence and to stimulate acquisition of knowledge and reasoning. Formative assessment may take many forms and, as has been made apparent, can use a variety of selected and constructed assessment items. Feedback can vary from informal comments made at the end of a case presentation on a ward round to that using highly complex and formally structured computer-based learning tools (Greenhalgh, 2001; Letterie, 2003).

With regard to the latter category, provision of learning materials and study guides are frequently considered suitable tools for formative assessment. However, this is often done without any evidence of their efficacy or an understanding of what an online assessment requires in terms of educational methodology. In reality, what is required is that the educational tool be built to an informed template and evidence is gathered that the material is used to learn, to stimulate further enquiry and with advantage being taken of feedback.

Summary

This chapter has shown that an authentic formative assessment method, providing feedback on an ongoing basis as often and as comprehensive as possible, is a useful learning tool. The nature of assessment items can be of constructed or selected response types and may include scoring as long as it is accompanied by immediate feedback.

Characteristics of good formative assessment methodologies:

- provide material in a case-based environment
- allow scoring if the author desires. Scoring must be accompanied with immediate feedback
- support constructed and selected response items
- provide substantial feedback to the user on both correct and incorrect decisions
- allow students to revisit cases as often as desired
- allow students to investigate alternative choices without penalty
- allow lecturers to monitor student progress and isolate areas of weakness in both student performance and in the learning material itself

Chapter Three: Technology in Education

> 'If, by a miracle of mechanical ingenuity, a book could be so arranged that only to him who had done what was directed on page one would page two become visible, and so on, much that now requires personal instruction could be managed by print.' (Thorndike, E.L. (1912). Education: A First Book. New York: Macmillan Co. in Howard, et al., p. 160)

Introduction

This chapter discusses the uses of technology in education, with a final focus on computer-based educational initiatives.

If educators are to make practical and worthwhile use of new and often expensive teaching tools, it is worth their while to understand the potential impact of such technology and to anticipate the benefits, or disadvantages, if any, of its use. If a technology is new, the only thing an educator can rely on is predicted outcomes based on previous evidence from similar technologies and experience. Harry Truman wrote 'There is nothing new in the world except the history you do not know' in (Miller, 1974, p. 26), and in order to better understand the motivations of, and restrictions on, educators of the past, it is necessary to look at how technology has been used historically. The perspective gained provides an insight into the methodologies and technologies likely to be useful in the present and near future. A historical examination of the paths taken by educational scholars provides an appropriate perspective for the use and potential of educational technology today.

The Impact of Gutenberg's Printing Press

Universities have been in existence for nearly 1000 years. The first university was founded in Bologna in 1088 (Alma Mater Studiorum - Università di Bologna, 2009) with many others, including Cambridge and Oxford, being created in the following 200 years. The existence of universities created a need for books to study with, but the lack of appropriate technology to mass-produce them meant it was unlikely the demand for books could ever have been met. In order to reproduce a book a scribe was required, but the ability to reproduce a work uniformly was often absent (Eisenstein, 1979). Textbooks themselves, as books a student or school might possess, did not broadly influence education until after the invention of the moveable type printing press by Gutenberg in 1452. This invention is still considered one of the most influential devices of all time, especially in terms of educational advancement. Whereas students would previously have had access to a limited library of books (less than 130 books were available at Cambridge University in 1424 (Meggs, 1998)), they now began to have access to a wider selection at a vastly cheaper cost. Prior to Guttenberg, books were beyond the reach of most people due to price: a Bavarian elector once offered to trade a town for one book...and was refused! (Shane, 1940)

Undergraduates now had the ability to access large bodies of knowledge potentially greater than that of their teachers (Eisenstein, 2005), and so the opportunity for true self-directed learning had arrived. The benefits of the printing press extended beyond universities. One of the earliest books printed was a version of the Bible, and within a few years of the creation of the printing press, pamphlets, grammar books and certificates used to assist the bearer in absolving sins were printed. By 1500, even though most of the world was illiterate, over 20 million printed works may have been produced (Williams, 2003). An examination of the library catalogue of Cambridge University from 1600 to 1960 shows an exponential increase in the number of books (Figure 3.1). This trend, reflected in the growth of the British Library catalogue (Pollock, 2009), is likely to represent a world-wide pattern of growth and it is estimated that in 2000 there could have been 65 million titles in the world (Lyman, Varian, & Swearingen, 2003). With the advent of the printing press, technology began to have a massive influence on education, as it would continue do so for the next 500 years.

NOTE:
This figure is included on page 81 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.1 Growth in library catalogue for Cambridge University from 1600-1960
Adapted from (Pollock, 2009)

The first book designed for students was a spelling book called *The English Schoolmaster* by Cootes, first published in Europe in 1596 (Department of English, 1997). One of the most influential texts, first published in 1658, was an illustrated encyclopaedia, *Orbis Pictus* (The World in Pictures). Its popularity was at least partially due to its lavish illustrations. The concept of using pictures to illustrate concepts was deemed by the author, Comenius, to be motivational for children (Anderson, 1962). The book remained on sale for over 150 years and is often regarded as the first instance of a tool providing audio-visual approaches for the classroom (Nodelman, 1988). Comenius devised a series of instructional methods that encouraged education from an early age. The methods had a practical relation to life and were based on visual instruction and instructional examples. He felt the educational environment should be cheerful and not punish failure (Saettler, 1968). In many ways, his ideas mirror modern approaches to instructional design, particularly authentic instruction and the constructivist approach to learning. Comenius himself was one of the great educators of the 1600s although, perhaps unsurprisingly, he was not acknowledged as such during his lifetime.

Although the moveable type printing press did not create books or manuscripts, it did make them more readily available for general use. This ability to provide a large number of consumers with cheap versions of identical material created a large global community of educated people, who had the opportunity to direct their own learning. This became a recurrent theme in the successful educational technologies that were to follow, but until the late 19th century books were the main instruments of learning, accessible eventually to students from early school through university and continuing professional education, especially through scholarly journals.

Even though material in a book is often presented in didactic fashion, this is more often due to the author than the medium. The book has the potential to be very useful in formative assessment. From an assessment perspective, a book can present material in case-based format and can support the use of selected response items by 'branching' from one page to another based on decisions a student makes, thus providing a measure of authenticity. This 'choose your own adventure' style of book has been used successfully in medical education (Devitt, Barker, Mitchell, & Hamilton-Craig, 2003).

Books can also provide constructed responses by posing questions and providing ideal answers at another location in the book. This requires the student to write the answer in another location, perhaps in an exercise book or loose piece of paper, but is still a viable option. Feedback on decisions made by the reader can be given as answers to questions, and this feedback can be as substantial as desired by the author. Books can obviously be used as often as desired by students, and it is possible for lecturers to collect answer sheets in order to provide feedback to students.

The book can therefore meet most of the criteria for formative assessment outlined in Chapter Two, but nonetheless suffers from some significant disadvantages. There is very little content written in a style conducive to formative assessment. Most books are presented in a non-interactive style, full of information but providing little support for developing an understanding of the content therein. Books with questions and substantial feedback may be extremely useful to students, but lecturers are unlikely to have the time to monitor students' answers and provide additional feedback. Feedback and information is usually limited to textual information. So although the book has potential for providing some levels of authenticity and feedback, other devices may be useful for additional stimulus and interaction. Nowadays there are many technologies that can be used away from the classroom, but in the beginning those having the biggest impact were located within it.

Early Technological Aids in the Classroom

Whilst early universities may have developed large libraries, the educational situation away from Western Europe was notably different. Anderson (1962) provides an interesting perspective of life in the early years of North America and the USA. He describes how in the period from the 1600s to the 1700s, the style of education in North America was based mainly on recitation and memorisation, and was often provided on an individual basis due to low numbers of students attending schools. The only instructional devices likely to be available in a classroom at that time were the pen, made from goose quills, and ink, made from maple bark (Anderson, 1962). Paper was expensive and poorer students would write on birch bark. Paper teaching materials themselves were valuable to the extent that primers used to teach the alphabet needed to be protected. Hornbooks (Figure 3.2), invented circa 1450 (Horner, 1973), were named because of the thin, transparent layer of horn used to provide such protection. As paper became cheaper, educational pamphlets called Battledores (Figure 3.3) appeared and proved extremely popular with learners.

Figure 3.2 Hornbook used with permission (Northern Illinois University, 2007)

NOTE:
These figures are included on page 84 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.3 Battledore used with permission (Northern Illinois University, 2007)

In the 1800s proponents for teaching by using examples began to voice their desire to utilise technologies such as blackboards, slate tablets and world globes. School Boards were concerned about the cost of such items and the use of blackboards was considered to be experimental even up to the middle of the nineteenth century (Skrabec, 2009), but blackboards and slates in particular revolutionised education across the world. The thought of using visual aids did cause some consternation, with concerns aired that teachers would be unable to adapt to new teaching methods and would become little more than manipulators of new educational technologies (Anderson, 1962). On the other hand, some educators took the new technologies to such extremes that, in one case, blackboards covered every wall of a classroom even though only a fraction of that space was ever required. Inexorably though, blackboards, slates and other tools, including the abacus, progressed from being novelties or luxuries to essential teaching tools. Teaching aids such as slates and blackboards allowed large classes to carry out the same work at the same time, providing greater teaching efficiency. By the 1840s, blackboards in particular were in use all over the USA and texts on how they could be effectively used for teaching were being printed, including a five issue run on ‘Slate and Black Board Exercises for Common Schools’ in the *Connecticut Common School Journal* (Anderson, 1962).

The realisation that teachers worked best using some teaching aids was at least partially due to a wave of new educational publications and exhibitions of educational products. This mixture of scholarly work and marketing assisted in the change in perception that took place regarding the use of these tools for teaching. This perception was further enhanced by the large number of museums that came into existence from the mid 1800s (Ellenbogen, 2002). Unlike Universities, which were perceived as insular entities, museums gave the general public an opportunity to see objects, thus indirectly raising awareness and support for new ways of educating. Some American museums opted for an instructional role in the late 1800s where conducted tours and lectures began to be organised following a lead begun by the Crystal Palace Exposition in London in 1851 (Saettler, 1968). This lead to the creation of School Museums, designed to supplement the school system. Although actively supported by various cities such as St Louis and Cleveland in the early 1900s, school museums are an unrealised adjunct to education, with very few coming into existence since this time.

As paper production became cheaper and transportation infrastructure improved, the ability to produce textbooks cheaply and efficiently increased. New technologies such as printing plates replaced movable type, and techniques to reproduce photographs in books became available. The use of visual media in books progressed from the somewhat crude illustrations in the *Orbis Pictus* to real life images and reached its culmination in the 1920s when colour printing in texts was first used. Outside of books, visual media was beginning a new revolution. The magic lantern (Anderson, 1962; Northern Illinois University, 2007) was an expensive apparatus, which could project slides onto a screen using a kerosene lamp. Another invention, the stereoscope, allowed the projection of 3D images. In 1856, the latter was hailed by Sir David Brewster, the famous scientist, inventor of the kaleidoscope and critic of non-visual teaching methods, as a tool for teaching a variety of disciplines.

> 'The teacher…may pour it into the ear, or extract it from the printed page, or exhibit it in a caricature in the miserable embellishments of the school book, but unless he teaches through the eye...no satisfactory instruction can be conveyed' Brewster in (Anderson, 1962, p. 48)

By the 1900s, technology had already had a massive impact on education. The most significant invention had been the printing press, which was welcomed by many around the world who could now access the printed word. Many innovations used in education became superseded but each replacement nearly always provided additional benefits. As paper became cheaper, slates disappeared. The arrival of duplicating devices in schools lead to a significant reported decrease in the use of blackboards (Fildes, 1935). The fountain pen, introduced in the 1830s, became popular in the 1930s, replacing quills and ink. It too, was fated to be replaced: ball point pens made their mark in the 1950s.

The innovations and technical changes occurring during this period are of importance more for the transformation in thinking they stimulated rather than their ability to project a picture or write words down on paper. Teachers adjusted their approach to teaching. It became important to teach by example, providing practical applications of theory rather than expecting students to recite facts. With many countries stipulating compulsory education for children by 1900, the ability to engage and motivate students in their learning increased in importance.

The evidence indicates that education was in part guided by the technology of the time. New technologies demanded new approaches. Blackboards encouraged whole class exercises where every student could benefit from a single example in many different subjects (Wells, 1862), while slates allowed students to practice their learning independently with a variety of exercises (Beard, 1860).

Many of these newer technologies proved to be useful in providing authenticity and feedback. Much of the learning provided in the classroom was still rote style, but the opportunity for feedback on a group or individual level was possible. Growing class sizes would have made this more challenging and the formative work was likely to have been controlled by the teacher or lecturer and not the student, but the potential for true formative assessment was growing with both teacher and student sharing some common understanding of what was required to learn. Working at the blackboard or writing on a slate allowed teachers to observe student learning and provide immediate feedback. From an assessment perspective, these methods could potentially present material in an authentic instruction setting and could support the use of selected response items by 'branching' from one part of a blackboard to another based on the decisions a student would make. Blackboards could also provide constructed responses by posing questions and the teacher writing ideal answers at a later time, after students had responded on slates or via discussion. And, although visual media were not commonplace, teachers were looking to access them, so the written word could be supported by physical examples, such as a world globe, a human skull or other artefacts, thus providing additional authenticity.

With change inevitably came some resistance, sometimes on financial grounds, such as was the case with the blackboard (Skrabec, 2009), but also from those challenged or threatened by the introduction of new technologies. For example, the introduction of the printing press was met with opposition from scribes, whose livelihood was threatened and spiritual leaders who felt challenged by the thought that they would not be able to guide people in the reading of religious texts (Briggs & Burke, 2009). Other objections came on philosophical grounds. Echoing Socrates, John Henry Newman attacked cheap printed books because he did not believe that reading without educated guidance could truly allow a person to learn effectively (Landow, 1996; Wright, 2007). Finally, there were also educational issues, such as those raised by Hall (1947), who was concerned that writing answers to arithmetic as opposed to verbal calculations was detrimental to students developing the ability to solve problems.

It should also be noted that each technology had its problems and solutions could be years in arriving. Finding large enough boards and appropriate writing materials for blackboards was an issue (Alcott, 1843), as were breakages that occurred with slate tablets (Chambers, 1840). Writing implements such as the ballpoint pen suffered from poor design both structurally and in ink development, causing significant reliability issues which directly affected uptake (Souder, 1955). Every step forward created opportunities, but also provided issues that needed to be addressed. Audio technologies, which emerged in the late nineteenth century, were no exception.

Audio Instruction

In 1878, one year after its invention, Thomas Edison suggested that the phonograph could be used as an educational tool. Reports from the early 1900s show it was utilised in classrooms for language instruction and, later, within the Armed Forces during World War II (Agard, et al., 1945; Cipriani, 1912; Clarke, 1918; Furness, 1952; Kunze, 1929; Oberhelman, 1960). It was also used to assist with music education (Hands, Scholes, & Whittaker, 1934; Sunderman, 1943; unknown, 1942), though not always to the extent desired (Dent, 1947; Mathews, 1947). The reason for its lack of use in what may be considered the most logical area was that the reproduction was not of sufficient fidelity and quality to adequately instruct students (Kuhn, 1954). Phonographs were also used to present a series of science lectures from some of the leading scientists at the time including Millikan (unknown, 1932) and were suggested as a means of teaching American history (Landman, 1927). The benefits of self-paced and individual learning were deemed positive outcomes from this style of learning (Cipriani, 1912), as well as being a motivational tool (Gosch, 1940; Landman, 1927; Pattee, 1923). Evidence supporting the use of these devices as learning aids was mostly anecdotal, but was uniformly positive.

One of the most useful outcomes reported in one study (Cipriani, 1912), was the ability for students to listen to French language on the phonograph and record one's own phonograph to test their pronunciation of words. This sentiment was echoed by others (Kunze, 1929; Whyte, 1931), but the ability to record would be denied to students some years later. Cylindrical records, which supported recording, were replaced by disc records, which were easier to produce, easier to handle and marketed more effectively (Thompson, 1995). Recorders for disc records were more expensive and were not successful as consumer products. This is notable because it was one of the few instances when technology had moved in a direction that resulted in the loss of a methodological teaching practice valued by teachers.

The Phonograph in Medical Education

In medical education, a phonograph was used to record a history from a patient (unknown, 1890a) and to record dog heart sounds (Detweiler, 1964). There is evidence that medical resources were available in this format (Allahbadia, 1959; James, 1959; Morrissey, 1960; Wintersgill, 1959) and the use of phonographs to deliver information passively to students was seriously considered (Williamson, 1936). In 1890 recordings of various vocalisations characteristic of certain diseases such as whooping cough were played for medical practitioners. It was suggested that they may be useful for teaching (unknown, 1890b) and in 1891, an announcement of phonograph recordings of two patient histories was made (unknown, 1891). Measuring heart sounds for teaching also occurred (Phillips, 1949). There were those specifically interested in the technical side of the phonograph (McKendrick, 1895) or as a means of taking dictation (unknown, 1889) which was how Edison marketed his invention, but the most interesting outcome may have come from a physician’s suggestion to record heart sounds using the phonograph in order to investigate the timing of heart sounds and murmurs (Bramwell & Murray, 1888). Much of the interest in the technology used for phonographs was diverted to assisting patients rather than education, such as testing the hearing of school children (Rodin, 1927) or as a relaxation tool for asthma (Morwood, 1953).

NOTE:
This figure is included on page 90 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.4 Phonograph (Dja65, nd)

Recommendations from the US Army's experience using phonographs for training (Agard, et al., 1945) led to educational phonographs being sold to the general public after World War II (Furness, 1952). Phonographs had become cheap and a feature of modern households, but there was a new medium for transmitting audio, which had captured the imagination of consumers, educators and advertisers alike. Radio had arrived and soon became the dominant medium for audio transmission (Thompson, 1995).

Radio

Radio was the first form of communication that could rapidly and simultaneously reach a vast population. The first short radio broadcast was made in 1906 (Grant, 1907) and such was the enthusiasm with which the medium was adopted that it was only three decades later that Orson Welles was able to paralyse and panic his listeners with his broadcast of HG Wells' War of the Worlds and its threat of imminent Martian invasion (Corliss, 2001).

Classified as radio education, instructional radio, or the fourth 'R', after 'Reading, Riting and Rithmatic' by some supporters (Schinnerer, 1945), radio in the armed forces played a part in boosting morale and continued to be used after World War II (Culbert, 1979) in preference to film programs and other social science innovations trialled during World War II. Even though the cost of phonographs was relatively low (Kunze, 1929) some teachers found a combination of phonograph and radio to be a more cost effective solution, particularly for non-core teaching (Lemieux, 1949) or for the sake of convenience (Miller, 1944).

The use of radio in medical education was instigated in 1921 when the first broadcast, focused towards public health, was aired (Rickards, 1923). In 1929, processes were set in place that eventuated in the creation of the National Advisory Council on Radio in Education, Incorporated (Taylor & Mason, 1934). This council was a major advocate for providing fixed time on air for educational material and was instrumental in broadcasting medical information for the public (Taylor & Mason, 1934) such as a series of broadcasts for mouth health education (Wisan, 1940).

Problems for education radio were immediately obvious as untrained speakers attempted to engage a remote audience. There was a clear recognition by broadcasters that for the public to pay attention to these messages, speakers needed to train and develop new skills (Blanchard, 1935) and that different tactics needed to be attempted to engage the public (Bauer, 1935; Routzahn, 1924). Broadcasting was carried out by medical societies and voluntary organisations as well as state departments (Turner, Drenckhahn, & Bates, 1935), but much of the health information was sponsored by corporate interests such as the Metropolitan Life Insurance Company, Horlicks and

Sharp and Dohme (Turner, et al.). State sponsored messages seemed to be either unmemorable or ignored altogether (Turner, et al.). This may not be surprising when alternative programs with music and entertainment were likely to be far more appealing (Wisan, 1940).

In countries such as Australia and the US, with populations spread over vast areas, the 'tyranny of distance' often made formal school impractical, and for many people radio offered advantages. In the 1920s, 'Schools of the Air' (Clifton, 1930; Harshbarger, 1941) became popular and were particularly useful in educating students in widespread locations. However, radio was a commercial enterprise with all the support and infrastructure that entails, and although educational radio became widespread, it did not attract large audiences (Reed, 1941). Radio continued to be used in distance education (Lewis, 1946) and was occasionally used in other formats, such as allowing distant physicians to participate in question and answer sessions at conferences (Farber, 1966).

Reports on Australia's 'School of the Air' suggested that the use of technology in education in Australia at the time, and for the next 20 years, was focused more on 'Action Research'; trying to find general applications and establishing services rather than evaluating applications to defined teaching tasks (Connor, 1970). Worldwide, research into radio faltered due to the advent of World War II (Carle, 1959) but up to that time 205 studies and other literary articles had been reported in the use of broadcasting in schools (Carle, 1959). Much of the focus of initial research was in the impact of the educational medium on primary and secondary age students and teachers. An evaluation run by the Evaluation of School Broadcasts Project during the 1940-1941 series of educational radio broadcasts from CBS came to the conclusion that although many of the broadcasts had educational value, there was a lack of a defined purpose for the broadcast and a mismatch in the level of the content and the educational background and listening conditions of the intended audience (Saettler, 1968). Saettler discusses two other significant evaluations of radio in education. The Wisconsin project ran from 1937-39 and ran a comparative study with primary and secondary aged students. The results showed no significant difference in outcomes for the groups that used radio to enhance their learning compared with traditional methods. The Princeton Project, which ran from 1937-1941, assessed the attitudes of society to educational broadcasts. It found that 'non-serious' broadcasts were not likely to be listened to by low-income earners who didn't read much, higher income groups who found most radio programs 'rather boring' (Saettler, 1968, p. 315) or those who did not read serious content.

When examining the effectiveness of radio as compared with other methods, the evidence did not consistently favour radio learning over face-to-face teaching (Garfinkel, 1972) except, perhaps unsurprisingly, in the area of music (Ewbank, 1930). Although the educational effectiveness of radio may have been in doubt, the social impact of radio in general was not (Siepmann, 1941). Radio's influence over society was far greater than any previous form of communication.

Recording Audio

Radio had the ability to reach the masses, but unless the educator had access to a radio station, the impact an individual staff member could have on students using the technology was limited. It was beyond the capabilities of staff or the technology to tailor courses to individual groups of students. Phonographs had provided educators with this ability temporarily but tape recording devices were to have a longer life. The ability to record broadcast material had been desired for many years (Dent, 1947) and even though magnetic tape recording devices were invented in 1934, it was the invention of the compact audiocassette which became available in 1962 that brought a dramatic change to its use and perceived value by consumers and educators. Audiotapes have been used as part of a process to assess students in paediatrics by providing clinical vignettes (Margolis, Cook, Barak, Adler, & Geertsma, 1989), in student study groups (McKee, 1965) and in psychiatry (Pinkerton, 1961) amongst other areas, such as libraries of tapes to educate general practitioners (Graves, 1965), physiology (Asher, 1962) and recording lectures (Donsky, 1971). Although occasionally used in isolation with reported successes such as improved outcomes for law students (Ackers & Oosthoek, 1972) and increasing the cognitive skills of doctors and teachers in the area of the teaching resources (Scott, Brascho, & Brown, 1978), evaluations on the use of audiocassettes in education often used educational methodologies dependent on using the cassette with complementary material, such as learning booklets or slides (Blank, Kirk, & Weinswig, 1975; Gale & Clarke, 1977; Orr, 1968; Tomlinson, 1979). Tomlinson's non-comparative study provided some interesting feedback suggesting that learners were keen on using their inactive time to engage with audio material and made the best use of the mobile technology. Comments from participants included ‘enjoyed tape while commuting’ and ‘surprised how much time I used to waste in my car’.

Audiocassettes complemented with additional material appeared to have educational benefits. When used with an active learning strategy, such as a booklet in which students were asked to draw diagrams or answer problems, significant positive short-term benefit was obtained (Gale & Clarke, 1977; Tomlinson, 1979) but this did not translate into any long-term gain (Ackers & Oosthoek, 1972; Tomlinson, 1979). A study in physiology (Davies, Gale, & Clarke, 1977) showed that students could learn using audiocassettes and a booklet, but interestingly showed that students learned as well using written notes and a booklet studying haemostasis, yet less well when studying Indicator Dilution. This strongly suggests that there were weaknesses in some of the resources supplied to the students and indicates that the technology is only as useful as the content developed for it. Educators rarely appeared interested in using audiotapes in isolation, and in areas such as radiology, where there was an obvious requirement to provide visual support to students (Cockshott, 1973) such attempts would have been counterproductive. Students found the use of audio recording of lectures useful, although the paper copies of lecture notes were deemed more so in a study of 275 law and education students (Roberts, 1994). In particular students found them a useful aid to catching up if they had missed a lecture.

One novel method of using audiocassettes was for lecturers to provide feedback to students. One study found little difference in the amount of time it took a lecturer to provide feedback in audio form compared to written form, but found the feedback was richer although outcomes for students in the different groups as measured by final grades were the same (Kirschner, Brink, & Meester, 1991). One English class prepared oral presentations on cassette and then received their feedback in the same format with reported student satisfaction and improvement in language skills (Allan, 1991). As an aid for patients audio feedback was also considered by recording specialist consultations and allowing the patient to play them back at a time of their choosing (Tattersall & Butow, 2002). Audio had a strong place in education and for a time, the audiocassette was the medium most often used.

As with many other technologies, the time where audiocassettes commanded a significant market share ended and in 2010 Sony finally ceased production of their portable audiocassette player, the Walkman (Indvik, 2010). Audiocassettes are still used in education, and even recently they had been listed amongst the top methods for delivering CME instruction to medical practitioners, ranking well above video methods, the Internet and CD-ROM (Mamary & Charles, 2003), but their time has now passed.

Audio Compact Discs (CDs)

The invention of the CD continued an impressive market presence for audio media starting from 1920 and extending to 1980. As shown in Figure 3.5, there was a linear growth in sales over 60 years, resulting in an exponential growth in profits for the recording industry over the same period. Audiotapes had been popular as recording devices over a long period of time, but were eventually replaced by audio compact discs (CDs) as shown by sales figures over a 20- year period in the US (Figure 3.6). The CD was initially a play only device but quickly became an inexpensive consumer product able to record sound. The CD-ROM technology also had a significant impact on computer applications and this will be discussed later in the chapter.

NOTE:
This figure is included on page 95 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.5 Audio sales and profits 1920-1980 (Gronow, 1983)

NOTE:
This figure is included on page 96 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.6 Audiotape sales compared with CD sales 1990-2007 (Weisenthal, 2008)

In essence, the use of audio CDs has followed the same path as other audio technologies, particularly in how they are used with other types of media. Like audiocassettes, audio CDs were supplemented by textual information, although it was more often PowerPoint slides rather than paper handouts (Bond, et al., 2004; De Muth & Bruskiewitz, 2006). Audio CDs have been used in medical education and have been shown to be effective in learning heart sounds via repetition (Barrett, et al., 2006). They are also still published in formats to allow students to practice their language skills as preparation for entry to University (Cambridge Esol, 2007).

MP3 Players

Figure 3.6 shows a drop in CD sales around the year 2000. This can be attributed to the advent of MP3 music devices, which play encoded audio files. These files are easily recorded and shared on the Internet and occupy very small amounts of disk space, thus making them cheap and easy to download. Unlike audiocassettes and CD players, the devices for playing these files are very small and modern phones are capable of playing them with no additional costs. This enhanced portability and convenience at a low cost has lead to these devices dominating the market for the latter half of the 2000s. Apples iPod is the name synonymous with MP3 players and Apple has made significant efforts to position itself as a supporter of education by allowing educational audio files to be published and accessed via its popular iTunes software.

Portable music devices began becoming popular and affordable around 2005, as shown in Figures 3.7-3.8. Over a few short years CD sales plummeted by 50% as a result of consumers embracing newer technologies. Audio purchases are now heading away from physical media with greater emphasis on downloadable audio files (Figure 3.8). Books too are being downloaded as evidenced by the existence of popular electronic book readers such as Amazon's Kindle (Amazon.com, 2010) or more versatile devices such as Apple's iPad (Apple Inc, 2010b) and associated book selling software (Apple Inc, 2010a).

NOTE:
This figure is included on page 97 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.7 iPod sales 2002-2010 (MySchizoBuddy, 2010)

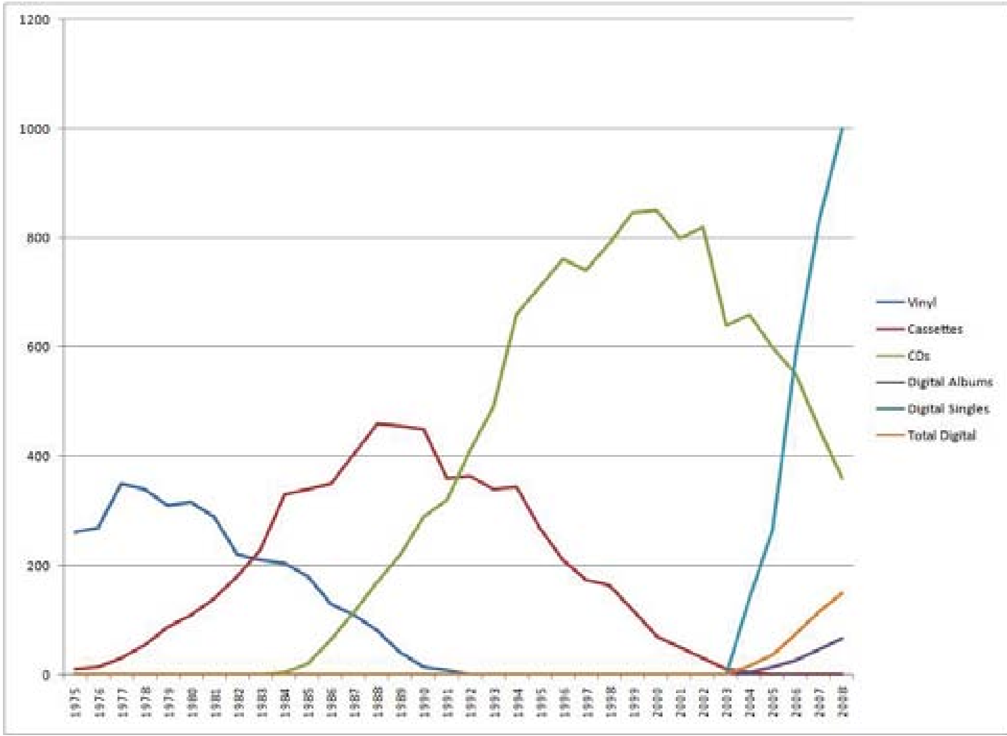


Figure 3.8 Music sales by format 1975-2008
(RIAA New Shipments Statistics Database, 2008; Stensrud, 2008)

The iPod is one of the icons of the current decade and is viewed as a symbol of youth and vibrancy. Although now being replaced by smartphones such as the iPhone, the iPod has been a hugely successful technological device. This is largely due to the development of an attractive product, combined with the effective marketing of what is now seen as a highly desirable, if not essential, piece of technology in terms its versatility in its delivery of music to the consumer. Portability, innovative tactile controls, ease of use and design appeal are all key features in the iPod's popular appeal. With the iPod, Apple made a deep impact into modern culture even to the point of generating new words such as podcast and podict. The New Oxford American Dictionary proclaimed the word 'podcasting' as word of the year in 2005, ahead of other words such as 'bird-flu' and 'sudoku' (BBC News, 2005).

In essence, a podcast is an audio file that can be purchased or downloaded for free via iTunes (Apple Inc, 2007b) or other media source (Jham, Duraes, Strassler, & Sensi, 2008). Podcasts only differ from an audiocassette or audio CD in the method chosen to deliver the content to the user. Potentially, anyone with *iTunes* on their computer can listen to a selection from over half a million hours of audio content and carry it with them on their iPod. The medium has been adopted widely, with users varying from institutions such as the British Broadcasting Corporation (BBC) to individuals marketing their talents such as the comedian Ricky Gervais (Gervais, 2010) or just giving their opinions on everyday matters. Educational bodies are also leveraging the technology. Harvard's medical school has made its entire series of lectures available as downloadable audiofiles (Rainsbury & McDonnell, 2006). The Educause Learning Initiative (ELI) states clearly 'Whenever possible, ELI podcasts reports, participant interviews, and entire sessions to expand member access to its events.' (Educause, 2007). Some students have indicated that although they are satisfied with paper and audiocassettes, they prefer online resources (Weinberg & Knoerr, 2003), so the time may now be right for podcasts to have an impact in education.

The number of podcasts available grew rapidly from the early days (Nesbitt, 2005). From the 2nd to 3rd quarters of 2005 alone, the number of podcasts available on *iTunes* quadrupled from 5000 to 20,000. This growth continued and by Sept 2006, there were 82,000 podcasts available on *iTunes* (Cajiao, 2006), with a projection of between 20 to 80 million podcasts by 2010 (Chapman, 2006). There are large numbers of podcasts available for current use and some are clearly aimed at the educational market. For example, the US Food and Drug Administration (FDA) has a website devoted to podcasts on drug safety (U.S. Food and Drug Administration, 2007).

Podcasts in education are used in a variety of ways and are used with a variety of names including narrowcast, mobcast and skypecast depending on the intended audience and target devices (McLoughlin & Lee, 2007). One typical way is to record a lecture and make it available to the students via a podcast (Scutter, Stupans, Sawyer, & King, 2010; Williams & Bearman, 2008). This is an asynchronous method unlike audio conferencing where the audio of a lecture is transmitted 'live' or synchronously (Burge & Howard, 1990). The recording is often called a podcast but is often referred to as an MP3 file or other terminology relating to the file format. Students find podcasts useful and in a dental setting felt that they enhanced their learning (Whitney & Pessina, 2008). Part of this may have come from the fact that some students felt they didn't need to take notes but could participate more in class. This leads to fears that some academics have; that students will not attend lectures if they are available online as podcasts or other equivalent format (Maag, 2006; Scutter, et al., 2010). Although there is evidence to suggest that this may occur (Fietze, 2009), it may be that this is because students feel that they learn as much from podcasts as lectures. From the previous discussion on audiotapes and audiocassettes it is reasonable to surmise that if additional learning resources were provided to supplement podcasts, students could learn as effectively as in traditional lectures.

Indeed, it is found that podcasts are not always used in isolation. Enhanced podcasts, which contain images have been used successfully for surgical education (Shantikumar, 2009) and vodcasts, containing video, have also been used (Crampton, Vanniasinkam, & Ragusa, 2008). Other formats include podcasts linked to blogs (McLoughlin & Lee, 2007) and web sites containing support materials for the audio such as the Instant Anatomy site (Whitaker, 2010)

Also like audiocassettes, podcasts or online audio have been used to provide feedback to students (Ice, Curtis, Phillips, & Wells, 2007; McLoughlin & Lee, 2007; Middleton & Nortcliffe, 2009; Rotheram, 2007) (Ice, Curtis, Phillips, & Wells, 2007; McLoughlin & Lee, 2007; Middleton & Nortcliffe; Rotheram, 2007). One study in English found that students were divided over whether audio or written feedback was preferable (Kim, 2004). Interestingly, students failed to recognise the teacher when they swapped from written to audio feedback, a finding that may be very important in what can be a depersonalised environment. Another study, in pharmacology, reported positive outcomes for students using podcasts but only in specific areas (Meade, Dianne, & Joanne, 2009). Students reported that some lecturer's podcasts were superior to others and student use depended on the subject material. Another study reported significant time benefits in providing audio feedback online (Ice, et al., 2007), in contrast to that reported for audiocassettes. This study also echoed the benefit of increased richness of feedback.

For authentic assessment with feedback, audio provided new opportunities for teachers. Allowing an entire class to hear heart sounds played on a phonograph or iPod provides an opportunity for all involved to share a clear understanding of what is being listened to. Similarly, hearing a piece of music or a language spoken correctly allows students to connect to the world in which they will work when they graduate. Students can work at their own pace in their own time to close the gap between their current capabilities and those demonstrated by the audio recording. Some students will learn better from audio stimuli than written ones and the use of audio will make the learning environment more inclusive. Feedback provided by audio means is likely to be a useful addition to lecturers' options. It seems logical then for any learning tool to provide support for audio materials.

Of course, for most people there is more to the world than just text and audio. The visual medium potentially provides a richness of experience, and there have been significant levels of research around this area.

Visual Instruction

As discussed earlier, visual illustrations in books were important aids to learning from 1658 when the illustrated encyclopaedia, *Orbis Pictus* (The World in Pictures) was published. Its popularity was at least partially due to its lavish illustrations. Visual aids such as world globes, maps and photographs also became important in the 1800s in the United States. The magic lantern, a method for projecting images from glass plates was highly regarded (Highley, 1863) and used to enhance lectures such as ones in dermatology (Plumpers, 1870), natural history (Verney, 1895) and fatigue (Galloway, Cantlie, Graham, & Sambon, 1905) and recommended for use in displaying microorganisms to encourage learning about hygiene (De Garay, 1910). The magic lantern evolved, as did film technology, into more familiar forms such as the carousel slide projector used widely in the 1960s and 1970s.

Colour slides and projectors were used as educational aids for groups of two with paper assignments to teach microscopic anatomy (Stinson & Smith, 1968), with reported successes such as students asking more in depth questions, thus enhancing understanding rather than just gathering knowledge. The partially self-paced approach allowed students to collaborate and discuss material on slides. Basically, a process of repetition and discussion was used as the learning process. Slides for a physics course, used in conjunction with printed materials and film were shown to save 40% of the normal teaching time once the materials were developed (Kenshole, 1969). The author also drew attention to the fact that once developed, the resources could be reused which is a significant benefit. A major disadvantage is that if this type of resource becomes outdated it may need to be completely replaced at a potentially high cost. A dermatology course in London used slides and audiocassettes to teach the subject (Phillips, 1968). The author extolled the virtues of students being able to repeat the viewing of the resources until they were satisfied with their learning. He also left a warning about tailoring content specifically to meet the slides already available in a library. This is a useful comment as it illustrates that slides were either purchased directly from agencies or photographed by staff and added to a library and thus certain images would not have always been readily available. Creating libraries was not a trivial exercise and as those who have attempted to build them may attest, it is an open-ended task; one is always looking for better or more revealing images to add to the library (Hood, 1968).

Warnings came from other authors, suggesting that projected slides lacked excitement as compared with drawing pictures on the blackboard and that blackboards still remained the best choice for visual aids (Barabas, 1965). Certainly some would still agree with this assertion, which may be why overhead projectors were staple items in classes for many years and why document cameras (electronic overhead projectors) are still present in the most sophisticated of today's lecture theatres. The lecturer should not be an automaton; simply showing slides to an audience, but rather a dynamic instructor, able to respond to student queries using the tools available. Tools that allow the lecturer to write and draw something unanticipated during a lecture are essential items and encourage discussion and provide immediate feedback to students in a way that a slide cannot. The overhead projector allowed this, much as the blackboard had over 100 years earlier, but one major advantage with this technology, was that the lecturer was able to face the audience as they worked (Essex-Lopresti, 1979).

Overhead projectors were developed in the US Army during World War II and only became mainstream devices some 20 years later (Olsen & Bass, 1982). They became almost ubiquitous in classrooms with the benefits illustrated many times in the literature, particularly in applications in chemistry (Slabaugh, 1951; Walker, 1964) but also in teaching patient management problems in groups by laying out the clinical case on separate projection sheets and revealing extra information as the teaching session proceeded (Biran, Biran, Dunn, & Harden, 1985). Physicians were also encouraged to use overhead projector transparencies as well as slides to present lectures as part of their continuing professional education (Gelula, 1997). It should be noted that lecturers needed to learn a new set of rules or guidelines on how to use the projector most effectively. Articles such as one by Laidlaw (1987) provided such information, such as what colours projected well and what size fonts were appropriate. These were the kinds of guidelines which served lecturers in good stead when slides and overhead projectors were to a large extent superseded briefly by computerised Liquid Crystal Displays (LCDs) (Robertson, 2000; Wong, 1992) then by Microsoft's PowerPoint application (Microsoft, 2010) which was first introduced in 1987 (Gaskins, 2010).

PowerPoint rapidly became an essential tool for lecturing; and in November 2010, there were nearly 5 million hits obtained just searching for the filetype 'ppt' using Google. Able to combine visual and audio information as well as text, PowerPoint opened up a new world for lecturers in terms of being able to engage their students more actively in classes (DenBeste, 2003). Naturally, there were those who opposed this view, emphasising that these types of presentations should not be used solely to demonstrate the use of technology but to serve an educational purpose (McDonald, 2004). To emphasise this, articles appeared with guidelines on how to use PowerPoint effectively (Holzl, 1997). An example supporting the use of PowerPoint includes a study in a business school (James, Burke, & Hutchins, 2006). The results of this showed that although students were less positive than staff regarding the cognitive benefits of using PowerPoint, there was agreement that its use encouraged engagement, better note taking and helped to emphasise key concepts. Although there is little doubt that there are bad PowerPoint-based lectures, PowerPoint or other forms of projected information are not the sole cause of bad lectures (Lanius, 2004). In the end, the lecturer is responsible for ensuring that students are able to learn effectively regardless of medium.

As briefly mentioned, PowerPoint can include visual information; one component of which is video clips. As indicated by the sales of cinema tickets and DVDs, film is a powerful medium to entertain and had been used extensively in the past for the purpose of education.

Educational Film

It is difficult to appreciate the impact moving pictures had in the world of the late nineteenth century, but it was immense. By way of illustration, one of the earliest films was by Antoine Lumiere in 1895, depicting a train pulling into a station. The perception of the audience at the time was that the train was coming out of the screen and it caused panic, with people apparently fainting and screaming because of the perceived reality of what they were seeing (Brownlow, 1968, p. 2).

Film was primarily used for entertainment, but instructional films did exist alongside these. The educational library of films was quite small, initially being reworked government or welfare films. Cohen in 1918 bemoaned the fact that little ‘systematic effort to supply a definite series of pictures for a definite purpose’ existed and Gregory in 1922 was scathing about the use and selection of educational film (Saettler, 1968, p. 97). Judd (1923), suggested that movies for entertainment were changing students’ mental habits and that teachers needed to respond to that and incorporate film into teaching.

Instruction films did exist, and the one of the earliest catalogues of instructional films was produced in 1910, listing 1065 titles (Berg, 2002). The enthusiasm, support and drive of Thomas Edison was a key motivator for others to enter this new area of education, and indeed The Edison Company produced the Edison Film Library, which were still being distributed in 1929. Claims that ‘Books will soon be obsolete in the schools’ made by Edison in 1913 (Ryan, 2006), were not to prove prophetic, despite the presence of successful businessmen such as Henry Ford and George Eastman entering the field. Technical difficulties, particularly the flammable nature of celluloid and the high levels of heat coming from the projector equipment made film storage extremely dangerous (Adamson, 1915), made adoption of the technology difficult, as well as criticism that the films were not ‘systematically built to supplement existing courses of study’ (Saettler, 1968, p. 103). This lack of immediate relevance to existing curricula was emphasized by Judd (1923) and is a theme that has recurred time and time again with educational resources. In this case, the lack of an ability of companies to produce films exactly as required by their educational customers may have been the catalyst for enterprising universities to become involved in the production of educational film and for providing free or subsidised films. For professional companies, the cost of making films and the inability to make significant financial returns caused many companies to go bankrupt (Saettler, 1968).

The earliest instructional film was a public health film on tuberculosis, which was produced by Thomas Edison in 1910 (Pernick, 1978). Film was also was used for training in World War I as part of an effort to have soldiers learn about venereal disease (Sand, 1923; Spector, 2008) and reels of surplus film were supplied to the university system in the US, which may have been one of the catalysts for the introduction of visual instruction courses being offered during the period 1918-1924 (Saettler, 1968). Pernick (1978) argues persuasively that these films had a very real impact on public health education, not only because of the importance of the content but because the world was in awe of films and paid great attention. In addition, the films were long-lived, with health films made in the period 1917-1921 still in circulation over 15 years later (Pinney, 1936).

The importance of the field of visual education can be measured to some extent by the growth of a new field of journals in this area. The first appeared in 1918 and by 1921 there were 4 journals with similar themes being published. The most long-lived was Educational Screen, which although renamed during its production run, was published until 1971 as Educational Screen and Audiovisual Guide.

Educational films started off being silent which was considered to be a positive characteristic by those who preferred silent films and magic lanterns for both cost and pedagogical reasons; some of which had to do with being able to customise a soundtrack or narration for the context in which the students were in (Bernard, 1936). Nonetheless, the use of sound synchronised to the picture was introduced in the 1920s and clinicians felt that they might be used to reproduce clinical lectures (unknown, 1925) and became very popular from the 1930s. It is worth noting that silent films were still considered useful late into the 1960s for medical education, although there were issues in sourcing appropriate material (Hall, 1969).

The first experimental studies on films were by Sumstine in public schools in 1918 (Sumstine, 1918). Hoban and Van Ormer (1970) and Saettler (1968) detailed the results of a large series of studies into film for educational purposes before concluding that film offered an equivalent learning experience for students at reduced cost. Some of these studies included one based on the World War I film designed to instruct military personnel on venereal diseases. The study, by Lashley and Watson in 1919 indicated that the characteristics of the audience played a part in the amount learnt using the medium. A study by Freeman and McClusky in 1922 suggested that where film was used in education it would be best to break the content into small units and a study by Wood and Freeman in 1929 suggested that film needed to be added to the curriculum in an organised way, acting as a necessary step in the development of a subject. Wood and Freeman also suggested that the use of film in classes enhanced student outcomes, a claim refuted by McClusky who was scathing of the experimental procedure used. A study by Yale University in 1929 showed improvement in learning outcomes, although the authors, Knowlton and Tilton, were quick to point out that the involvement of the teacher was key to the success. Rulon discovered evidence that complementing films with textbooks lead to improved learning outcomes in 'informational and conceptual learning' (Saettler, 1968, p. 228). A four year study from 1929 financed by a private corporation, the Payne Fund, and run by WW Chaters as well as representatives of many US universities provided startling evidence as to the effects of motion pictures on children, particularly regarding attitudinal changes, and the influence of inappropriate content (sex and crime). These results significantly influenced further research in what was known as the Motion Picture Project (Saettler, 1968). This project ran from 1934-1941 in several stages and unfortunately proved to be underwhelming in its significance. Although the aim of the project was to define and facilitate the use of motion pictures in education, the key outcome appears to have been an encyclopaedia of 500 films in education, which had been judged as valuable by over five thousand staff and 120000 student survey results. Lorraine Noble, one of the key researchers, suggested to Saettler that '[the outcomes] were not very vital' (Saettler, 1968, p. 237).

Film instruction during World War II included gunnery training in the Navy and special training in the Army (Flanagan & Berger, 1948). The Air Force developed over 400 films on its own during that time (Reiser, 2001). The armed forces must have found some value in these aids as they launched one of the biggest educational studies of all time after the war. The US Navy and Army supported studies by Pennsylvania State University over 1947-1955 and the Air force organised an equivalent program that ran from 1950-57 (Saettler, 1968). Over 80 research studies and over 150 publications resulted from this alliance as well as some key conclusions, including that well made films could be used on their own to teach, students learn more if the film is complemented by a study guide, repetition of viewing increases learning and that short films viewed repetitively can be used effectively for drills or practice. Additional conclusions suggested that some form of testing after viewing a film can enhance learning and that note taking during a film distracts students and should not be permitted. Some of these issues are direct reflections of the outcomes of audio research, particularly that of requiring some sort of additional material such as slides or written material (Kenshole, 1969).

Film had its place and is still being used today in some form, but it does rely on having a film of worth at hand as well as projection facilities. That was no longer going to be an issue because in the late 1920s Logie Baird played a key role in inventing a world changing device: the television.

Educational Television

Television was slow to start, with somewhere around 7000 sets present in the USA up to the end of World War II (Genova, nd), but after the war production and purchasing increased dramatically with over 67 million sets sold by 1959. Today it is estimated over a billion sets have been sold across the world (Genova, nd). This is an impressive base to start with and educationalists in the 1950s decided they would take advantage of it by using the television for education.

The era of educational television, with content delivered either over closed circuit TV or on educational channels, and often using 'talking heads' gave rise to a large series of evaluations of both the social and educational benefits (Schramm, 1962). Many of the educational television channels were operated by universities, resulting in a wealth of critical reviews of the medium for use in education. The effect on education, perhaps paralleling that on the general public, was huge. More than 100 stations were devoted to educational TV (Magnuson, 2000). Entire curricula were run on television, a precursor to today's online curricula. Hundreds of systematic reviews comparing educational television programs with conventional methods of teaching were carried out. (Schramm, 1962; Wetzel, Radtke, & Stern, 1994). Despite much enthusiasm by its supporters, and studies suggesting positive outcomes for those using this medium, the overall results suggested that educational TV as a learning medium provided only marginal improvement over traditional methods. The opposing view to this was that educational TV was at least as good as traditional delivery of material to students and therefore could be considered to be a viable means of allowing distance students to learn effectively. This is a significant outcome as it supplied validity to the use of resources made available to students outside of universities and provides guidance as to potential outcomes and risks of the use of similar technologies.

For a decade, with substantial funding and government support (Cuban, 1986), educational television was used widely in America and consisted of closed circuit television used to deliver curricula as well as public television curricula for university students (Reiser, 2001). Television was mostly used as a teaching aid rather than a replacement for the teacher (Cuban, 1986) and as such was used in much the same way as audio and films. For medical students, the benefits of television was that a popular lecture or lecturer could be viewed in many locations at once, that medical details could be shown with great clarity and that it could be recorded for later use (Brong, 1964; Special Correspondant, 1967). In particular, procedures that students may not normally have been be able to view *en masse* and discuss were available for teaching purposes, such as endoscopies (Berci & Davids, 1962), viewing of the history taking of patients (Yonge, 1965) or the complete case management of surgical patients (Smith & Wyllie, 1965). For continuing medical education, the benefits were considered to be similar. By allowing a physician to see areas outside of their normal expertise or new innovations in their own field, television was seen to be an excellent medium (Brayton, 1967).

The advent of colour television, although very popular, did not seem to have any effect on learning outcomes as measured by an Army study ran in 1959 (Kanner & Rosenstein, 1960), although it did appear to enhance students' perceptions of peripheral, or non-essential visual detail (Katzman & Nyenhuis, 1972). Somewhat disturbingly it was found that another variable, the teacher, was perhaps not as essential as one would hope. Classes in sociology with trained communicators performed as well as those taught by experienced teachers (Blenheim, 1969)!

It should be noted that the use of closed circuit TV in particular required an elaborate amount of equipment to be used, as illustrated in Figure 3.9. Along with technical support, a director and cameraman were often used (Phillips, 1956). Educational TV was clearly not an *ad hoc* adjunct to learning. On the contrary, it was a demanding exercise requiring great organisation and support from a wide variety of specialists. It was also important for lecturers to be able to control the material seen by students, as film was often purchased or television broadcast without any consideration for the specific needs of a curriculum, being too long or containing information not deemed necessary (Jones, 1968). The advantage of being able to record video was considered to be very important in order to customise the content and get maximum value from educational material that was expensive to produce In this way a diverse library of useful, reusable content could be produced (Jones, 1968).

Figure 3.9 Closed circuit TV equipment and required personnel (Phillips, 1956)

Recording Video

The first video recorder was called the Ampex and was first successfully demonstrated in 1956 (Ginsburg, 1957). It was a massive piece of equipment (Figure 3.10) with an equally massive price of $50,000 and not easily within the reach of education budgets.

Consumer models followed throughout the 1960s, able to record in black and white and usually only items of short duration (BBC & The British Museum, 1957). The Sony-U-matic recorder was released in late 1969 and penetrated the education market (Gayeski, 1990), but it wasn't until the 1970s that mass market use of videocassettes occurred. The famous battle between the shorter duration Betamax format and the VHS format was eventually won and VHS cassettes became the norm in homes (Armes, 1988) and educational institutions. The adoption of videotapes by the general public and the drop in prices of the players and tapes in the 1970s provided a new avenue for educators to explore. Nonetheless a video recorder, camera and half a dozen tapes cost an Obstetrics and Gynaecology unit $22,000 to set up in 1976 (Lim & Ratnam, 1979), so educators needed to be mindful of the benefits that could be obtained. One of these was to record live presentations and play them back to different audiences at their convenience; another was the ability to record pre-existing material into a new and more robust format.

NOTE:
This figure is included on page 110 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.10 Ampex Mark IV videotape recorder (IEEE Global History Network, 2010a)

Videocassettes were used to record clinical demonstrations for remote audiences (Mir, Marshall, Evans, Hall, & Duthie, 1984) and for educating learners on endourology (Kennedy & Preminger, 1987). The ability to record and share such instructional content was highly regarded (Forrest, Ryan, Glavin, & Merritt, 1974) and films were recorded onto videocassettes for both preservation and convenience (Lim & Ratnam, 1979). The videocassette recording was found to be as effective as the original live demonstration in an experimental trial (Mir, et al., 1984). The use of video for continuing medical education was welcomed as a convenient alternative to guided learning with audiocassettes and supplementary material (Block, 1975).

Videotapes were also used in conjunction with other material to guide learning, much as audiotape was used (Forrest, et al., 1974) and as costs dropped they were shown in waiting rooms as educational material for patients (Koperski, 1989). Also, as was the case with audiocassettes, video was used to provide feedback to students. One such example, where students recorded their history taking skills and then reviewed their efforts occurred in a paediatrics attachment at United Arab Emirates University (Paul, Dawson, Lanphear, & Cheema, 1998). A large majority found it useful, but as mentioned in Chapter Two, exemplars are also important and students believed that if they had seen videos of instructors doing the same task, they would have benefited. In this case it is no surprise that students made some fundamental errors in interviewing patients. In this situation feedback from instructors would have been critical to the learning experience.

Unsurprisingly, there were no reported differences discovered in outcomes where videocassette was used as opposed to film. In essence, they are the same type of media, although videocassettes were more robust, cheaper to produce and to record. Educators still needed to develop materials that would benefit the learner using these tools and the format of how visual materials would be delivered was not that important. This would not change when optical discs arrived in the marketplace.

Digital Versatile Discs (DVDs)

Videocassettes were soon to be superseded by an optical disc format as audiocassettes were to CDs, but there was little to distinguish Videodiscs or Digital Versatile Discs

(DVDs) from their predecessors when used solely to deliver video. Videodiscs or laserdiscs were large optical discs (Figure 3.11) essentially the same as modern DVDs, though holding far less content and not storing the data digitally. Popular in Japan, the videodisc players were still made there until 2009, and some 2% of homes in North America had them in 1998 (Brancolini, 1998). DVDs replaced videodiscs and once they had been introduced in 1997 rapidly replaced the use of VHS cassettes (Figure 3.12). Whilst DVD players were initially expensive, consumers adopted the new technology and a similar scenario to that which had occurred in the audio market took place.

NOTE:
This figure is included on page 112 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.11 Phillips employee holding videodisc (Gizmo Highway, 2005)

NOTE:
This figure is included on page 113 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.12 US video sales by medium. VHS and DVD (Weisenthal, 2008)

Although celluloid films had been available for purchase, significant sales of devices to the general public related to visual learning only occurred with the advent of the television (Figure 3.13). Sales in most media grew linearly with time and declined as they met the end of their life cycles.

NOTE:
This figure is included on page 113 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.13 Market penetration of consumer media in the USA (Society of Motion Picture and Television Engineers, nd)

DVDs have been used as educational tools in a variety of different ways. In one, they were used to display the day-to-day running of an Intensive Care Unit (ICU) (Carroll, Iedema, & Kerridge, 2008). Nearly 200 hours of footage was recorded and after editing a ten-minute long DVD was produced. This had a great effect on the understanding of many people on issues that were relevant to the ICU, and although the final result could have been easily projected onto a screen using film in the 1950s, the ability to record on a digital camera and burn onto a DVD as part of an in-house project is something that could not reasonably have been accomplished in another era. In another example, student OSCE examinations were recorded for marking as a means of ensuring inter-marker reliability (Schwartz, Fernandez, Kouyoumjian, Jones, & Compton, 2007). In a similar vein, interviewers for graduate entry into medical schools across the world used a DVD of interviews as part of their training (Roberts, et al., 2008). DVDs have also been used to ensure students have access to high quality content for their learning (Triola, et al., 2004). The amount of information that can be stored on a DVD is 4.7GB for single layer discs, and can thus hold many video clips or animations. By providing it to students in this format there is no need for them to download this data, which could be expensive and slow for some students.

Another parallel to the progression of audio is the vodcast. Vodcasts are a video subset of podcasts and like the audio podcast can be downloaded from servers such as Apple's iTunes. Vodcasts contain visual data such as video or still images with the potentially useful ability to be downloaded onto mobile devices (Brown & Green, 2007). Students appear to value the use of vodcasts (Copley, 2007; Gross & Dobozy, 2009; Schreiber, Fukuta, & Gordon, 2010), even long lecture length ones, although the study by Copley indicates that there are a significant group of students who are unable to download such large files. As with podcasts, students do not necessarily see vodcasts as a replacement for lectures, but they find it very useful as a revision tool (Copley, 2007). Schreiber's study (Schreiber, et al., 2010) suggested no significant difference to undergraduate medical student outcomes based on watching a vodcast compared with a live lecture, echoing the results of years of studies on educational television. His students did comment that they did find the content less engaging than a live lecture implying they value the interactivity present in that type of learning situation.

Summary of Audio and Visual Technologies

Both audio and visual technologies have their place in education. Many types of technologies have had some impact on student learning and they have tended to receive predominantly positive reception by students, although it is reasonable to suggest that students are likely to value new and novel educational resources or any other direct attention paid to them by lecturers. Equally important is that use of these technologies has consistently been shown to cause no harm and that they produce similar learning outcomes to traditional methodologies. From the lecturer's point of view, nearly all of these technologies require significant time to develop and are not cheap to implement on an institutional basis. Table 3.1 summarises the use of different technologies over the years based on a table developed for distance education by Taylor (1995). Flexibility refers to the students being able to choose the time and place in which they use the technology for learning. Interactivity refers to some kind of programmed interaction that has commonly been shown to occur with that technology. So although books can be written to be interactive and provide feedback, they rarely are. The measure of cost effectiveness will be explored in the next section, but here it is taken to refer to the cost of the technology at its peak, not at its introduction.

As can be seen from Table 3.1, most of the technologies became cheap to use (but not develop material for). It is apparent that any interactivity or feedback provided by using these technologies was only going to occur if a lecturer ensured that it occurred outside of the medium itself. Such an example would be to provide a question and answer booklet to accompany an audiocassette then provide feedback either by audio or other means. Nothing in the technologies themselves even suggests that feedback is a requirement or even desirable.

If we assume that educators will always strive to produce the best content for their students, and that there is now a reasonable feedback framework for educators to work in, the most important trend that has occurred for audio and visual technologies is portability and flexibility of use. The ability for students to study when and where they want provides them with far greater control over their learning which is important in a world where students choose to fit growing work commitments into their study regime (Krause, et al., 2005).

Table 3.1 Different attributes of audio and visual technologies

Type	Portable	Flexible	easy to update	Cost effective	Interactive	Feedback
Book	Y	Y	N	Y	N	N
Phonograph	N	Y	N	Y	N	N***
Radio	Y*	N	Y	Y	N	N
Audio Tape	N	Y	N	N	N	N***
Audiocassette	Y*	Y	N	Y	N	N***
Audio CD	Y*	Y	N	Y	N	N***
Podcast/MP3	Y	Y	Y	Y	N	N***
Posters	Y	Y	N	Y	N	N
Slides	N	N	N	N	N	N
Overhead Projector	N	N	Y	Y	N	N
Powerpoint	Y	Y	Y	Y	N**	N**
Film	N	N	N	N	N	N***
VideoDisc	N	N	N	N	N	N***
DVD	N	Y	N	Y	N	N***
Vodcast	Y	Y	Y	Y	N	N***

*portable radios, cassette players and CD players were in common use even though they were not pocket sized
** it is possible to program interactivity and feedback into PowerPoint
***yes, by comparing student work with a recording, particularly for language and music practice.

In terms of learning outcomes for students, a study by Mir suggested that video-based instruction is superior to audio-only instruction (Mir, Marshall, Evans, Hall, & Duthie, 1986) but results of a large meta-analysis (Cohen, Ebeling, & Kulik, 1981) showed that although visually-based instructional methods provided some benefit, the effect size was very small for most strategies when compared with face-to-face learning. The only methodology providing an effect size of real educational significance was when video was used to provide feedback as part of training or skill gaining exercises (Cohen, et al., 1981) and this lack of distinct beneficial educational outcomes led some researchers to conclude that the additional costs of video outweighed the small positive effect (Spencer, 1977).

Other concerns about the use of audio and visual media continued to occur, particularly in the early days. 'Educational racketeers' (Lemieux, 1949) were scorned for trying to influence administrators to replace lecturers with records and these concerns were also echoed elsewhere (Smith, 1934). The use of phonographs was decried for changing the social environment where live music was played at home (Sousa, 1993). Radio, phonographs and movies were cast as items damaging the educational motivation of students (Neilson, 1946) or treated with off-hand contempt as incidental and unimportant items for education (Hammer, 1945). Concerns about cost have always abounded; It was estimated that the development cost for educational television for the Air Force in 1974 was ten times greater than that for slides and over one hundred times the cost to update the content (Rose, 1976). These are important voices, opinions and facts to hear, because educators must not forget that there needs to be a strong underlying educational reason for utilising technology for learning.

Books, audio media and visual media all have the capability to provide a formative assessment environment with the nature of any feedback provided proving to be the major difference between them. The evidence supporting one medium over the other leans slightly towards the visual methods, but the costs of developing content in this way may outweigh the benefits. There is however, one type of technology that can utilise all of these types of media and it had its genesis back in the 1920s and the cost for producing content for it has dropped dramatically since that time: The computer.

Computers

The first teaching machine is credited to Pressey (1926). The machine, which had been first presented in 1924, mechanically provided a sheet of paper with a question on it and up to 4 possible responses to the question. The machine registered when a correct answer was input and then presented the next question. In its formative mode, the machine would not move onto the next question until the correct answer was input. There was a small container for a piece of candy, which would be released should the user achieve a high enough score on the machine.

Pressey proclaimed the benefits of the machine, the most significant of these being the reduced time spent marking questions and the greater accuracy afforded, thus freeing up teaching time for other educational activities. The benefit of giving instant feedback to the user was a key point of this invention. This machine was not technically a computer, rather it was a mechanical device, but clearly had great similarities to some educational uses of modern machines. Mechanical 'computers' began to be used more often for calculations using punched cards as 'codes' for the machine and although they were time savers for scientists and mathematicians the first electronic computer was still twenty years away.

The first electronic computer was the Mark 1 computer made in 1944 at Harvard (Molnar, 1997) but ENIAC, developed in 1946, is considered to be the first programmed computer (Ceruzzi, 2003). It had a very limited memory of 20 decimal numbers and its input/output device was a card reader. This computer weighed in at an impressive 30 tons and took up a significant amount of space[1] (Figure 3.14). Although slow and unwieldy by today's standards (see Figure 3.15) it is reported that a Professor from Penn University speculated 'that during the 80,223 hours ENIAC operated, it crunched more calculations than had been performed by all humanity since time began' (Kanellos, 2006, p. 1). The ENIAC had a series of connecting cables that were essentially calculating pathways that had to be set up anew for each problem that it was required to solve. This methodology did not last, and later computers used cards to give the computer instructions on how it was to calculate. This software/hardware combination was to continue to the present day although significant changes were to occur on how data was input and output from the computer.

The first American commercial computer was the UNIVAC. It was significantly cheaper, ranging up to $50,000 dollars and used punched cards, but could also use magnetic tape. The price made it unwieldy for universities to purchase, especially as complete systems could cost up to a million dollars, but some were donated and used well into the 1960s in some cases. The limited number of computers available and the cost meant that processing often had to take place separately from where cards were punched or data was stored onto tape. As the technology for computers evolved, transistors replaced the valves that were used as amplifiers and switches and by the end of the 1960s transistors were cheap and reliable enough to be used. Memory had changed from powered valves to magnetic cores that were smaller and cheaper to run. Furthermore, in 1957 International Business Machines (IBM) invented the first hard disk drive, an invention that was to define the way data would be stored for the next 50 years.

[1] As a type of tribute, a computing team reconstructed the nearly 30m^2 ENIAC using modern technology on a 35mm^2 silicon wafer.

Figure 3.14 ENIAC computer (IEEE Global History Network, 2010b)

NOTE:
These figures are included on page 119 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.15 ENIAC capabilities (Kanellos, 2006)

The size and cost of these computers and the dearth of educational software meant that use in education was very limited, but in 1959, Donald Bitzer, an engineer at the University of Illinois, helped to develop PLATO which was an authoring system that eventually allowed easier authoring of educational material for students (Rahmlow, Fratini, & Ghesquiere, 1980). It was run on a computer called ILLIAC that was designed and built by the University of Illinois and ran on a mainframe/terminal paradigm as shown in Figure 3.16, where the computer could service more than one student at a time via terminals. The student could interact with the system by using a television and a specially designed keyboard. Although PLATO could initially only support one student on a terminal by 1964 it could support 20 terminals as the technology improved. In many ways PLATO was ahead of its time, due to the technical support available where it was developed and a clear vision of the type of education it was required to deliver as well as some sophisticated control technology. Examples of this technology were registering student touch on the screen and using computer controlled compressed air systems to physically move images (slides) as part of coursework for students.

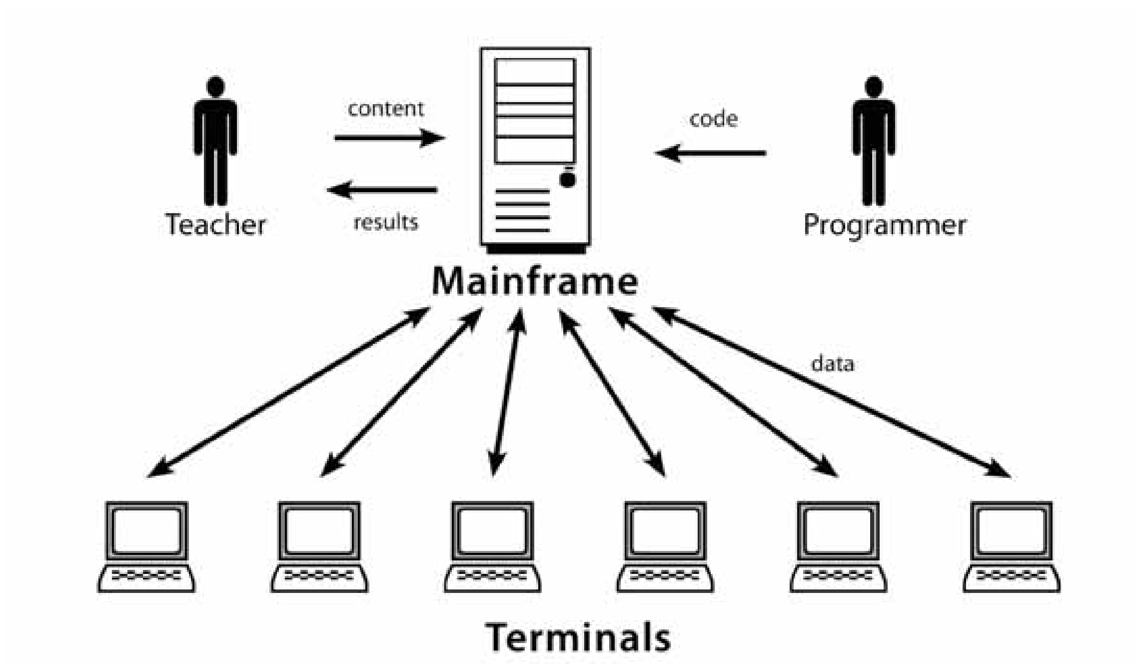


Figure 3.16 Schema illustrating learning using a mainframe/terminal paradigm

The PLATO system had an unusually long life for computer-assisted learning (CAL) software. From its early beginning in 1959 where it could only provide learning for one student at a time (Watson, 1972), it progressed rapidly. In 1972 it was an integral part of a nine million dollar project to teach first year english and mathematics (Hammond, 1972) and could provide learning to science students at an estimated fifty cents per hour (Bitzer, Sherwood, & Tenczar, 1973). Although it isn't clear if these estimates were realised, the cost effectiveness of using PLATO was highlighted by Levine (Levine, Jones, & Morgan, 1987) as a benefit of using computers in dentistry. PLATO was used for the learning in subjects such as respiratory disease and the microbiology of peptic ulcers with positive outcomes reported (Essex & Sorlie, 1979). In 1985 there were more than 100 PLATO systems in use worldwide (Woolley, 1994) and the last PLATO system wasn't turned off until 2006. There is now a new PLATO corporation using the historical name and selling online courseware (PLATO Learning Inc, 2010).

Educational programs were often written in authoring languages developed for educational purposes rather than less user-friendly computer code. In addition to PLATO (Alpert & Bitzer, 1970; Bitzer, Lyman, & Easley, 1965; Brigham & Kamp, 1974; Zinn, 1967), early examples of these were Coursewriter (Gilchrest, Deykin, & Bleich, 1972), MUMPS (Gilchrest, et al., 1972; Hoffer, Barnett, & Farquhar, 1972) and Logo (Cosmann, 1996) as well as lower level languages such as BASIC and FORTRAN. The first three allowed educators to write their educational programs without necessarily requiring a programmer. The IBM 360/367 was a computer often used and was likely the only computer in a school (Brigham & Kamp, 1974) due to cost (Seltzer, 1971) and despite researchers claims that it could be available for use by students 24 hours a day (Reed, Collart, & Ertel, 1972), the reality of having one machine for perhaps 100 students clearly equated to limited access.

By the 1960s, desk-size computers were becoming commercially available such as the IBM 1620. Some, such as the Digital Equipment Corporation (DEC) PDP-8 were available for $30000 (Ceruzzi, 2003) and these cheaper machines were the catalyst for a new era in computer sales as other manufacturers such as the Control Data Corporation (CDC) competed with DEC and IBM to capture the market. This competition allowed users at educational institutions to be able to afford computers and use them for education. Examples include Starkweather's use of the IBM 1620 for student learning of medical history taking (Starkweather, 1967), de Dombal's Elliot 903C computer for teaching the same kind of material (De Dombal, Hartley, & Sleeman, 1969a) and a method for teaching haemostasis on the PDP-15 (Gilchrest, et al., 1972). In a survey carried out in 1972, 20 percent of schools in the US and Canada were using CAL, with 30 percent anticipating future implementation.

The invention of the microprocessor in 1972 by Intel changed the way computing was to develop in the future. Rather than hard wiring functions into a computer, most of these functions could be stored as instructions supported by subroutines. Small components of memory called registers could hold information including where a computer program was up to and what it needed to do next and support a wide range of computational tasks (Rojas & Hashagen, 2002). Software written for mainframe computers such as PLATO and Coursewriter had been written for mainframe architectures and thus they did not retain their usefulness once microcomputers were developed. Indeed an attempt to introduce PLATO to microcomputers in the 1980s failed and the company eventually went out of business partly because it was built around a mainframe architecture (Timmins, 2010). Languages used to develop educational content for microcomputers in the medical field included The Computer Aided Simulation of the Clinical Environment (CASE), PILOT, MACgen (Prentice & Kenny, 1986) Teachers Toolkit, Computerized Modular Exercise Templates (COMET) (Holmes, 1983) and Hypercard (Clark & Raffin, 1992; Croninger, Tumiel, & Sowa, 1995). All of these languages were developed to do one thing, which was to make the creation of computer-based learning tools easier for academics. In addition more conventional programming languages designed for specialist programmers, such as BASIC and Pascal, were used to develop computer programs for learning in areas such as cardiology and clinical chemistry (Kordas & Lajovic, 1987; Simpson, Cookson, Percy-Robb, & Henderson, 1987). The 8080 processor released in 1974 was at the front of the wave of personal computing that has carried us to the modern era.

Interactivity with images and video came about through the use of laserdiscs integrated with computers in the late 1970s. Laserdisc players were released on the world market in 1978. Computers of this time could interface with these and play the laserdiscs, which were 12-inch optical discs containing up to half an hour of video or thousands of full colour images. Initially they were non-recordable. By the mid 1980s, these devices were being used in many areas of medical education, including paediatrics (Blackman & Huntley, 1984), nursing (van Reenen, 1990) and cardiology (Kwan, et al., 1988) as well as in general medical and science education (Miller & Willett, 1987; O'Neill, 1990; Thomson & Bryce, 1987). The types of interaction provided to students ranged from answering multiple-choice questions to interactive branching case-based scenarios (Abdulla, Watkins, & Henke, 1984; Verbeek, 1987). Whilst positive educational outcomes for laserdisc usage were also reported by others (McNeil & Nelson, 1990; Yoder, 1993), laserdiscs were being challenged by a newer technology. By 1988, the compact disc or CD was already being proclaimed as the new interactive media (Piemme, 1988). Laserdiscs became superfluous as computers grew so versatile and negated the need for peripheral devices to store images.

During the 1980s, new educational programs were developed regularly. PLATO was used to develop an interactive tutorial system with laserdisc technology for pathology (Thursh, Mabry, & Levy, 1986). Another pathology program began by being developed in PLATO but was then redeveloped into a HyperCard program also with laserdiscs to provide images (Levy & Thursh, 1989). Teaching tools developed in HyperCard were used in other fields; in pharmacology a study showed high levels of use of the developed program but did not show improved learning outcomes as measured by examinations (Moore, Waechter, & Aronow, 1991). Dentistry in the late 1980s used microcomputers connected to a videocassette player to help students how to identify structures in teeth (Watt & Watt, 1987) and pulmonary auscultation was successfully taught to first year students using a computer program developed for the MacIntosh computer (Mangione, Nieman, & Gracely, 1992). Radiology teaching material was developed using PILOT and evaluated. Results showed that student performance was as good as that obtained from traditional teaching methods. Students valued the computer-based resource because of its interactive nature (Jacoby, Smith, & Albanese, 1984). The fact that computer-based resources appeared to be as useful as traditional methods of teaching had significant benefits in some situations; one study showed no significant difference between animal experiments done live or via a laserdisc system, thus sparing the existence of experimental animals (Fawver, Branch, Trentham, Robertson, & Beckett, 1990). Overall, this was a time when there was a significant amount of experimentation of how microcomputers and education would fit together. By the time the 1990s arrived, most of the tools that would be recognisable in today’s computers, such as CD-ROM, colour graphics on monitors, significant storage space and fast processing speeds were combined and becoming affordable.

Figure 3.17 shows the uptake of computers from 1975 to 2005. It shows that uptake has increased dramatically in a power relationship over the years. Processing speed in computers has doubled every 18 months for the last 4 decades uncannily mirroring Moore's purely empirical law, which stated that the number of transistors per integrated circuit would double ever 18 months. Processor speed and other parameters of computing growth have followed this trend as well (Intel, 2005). This has allowed the calculations required behind learning exercises to be executed hundreds of times faster than even a decade earlier.

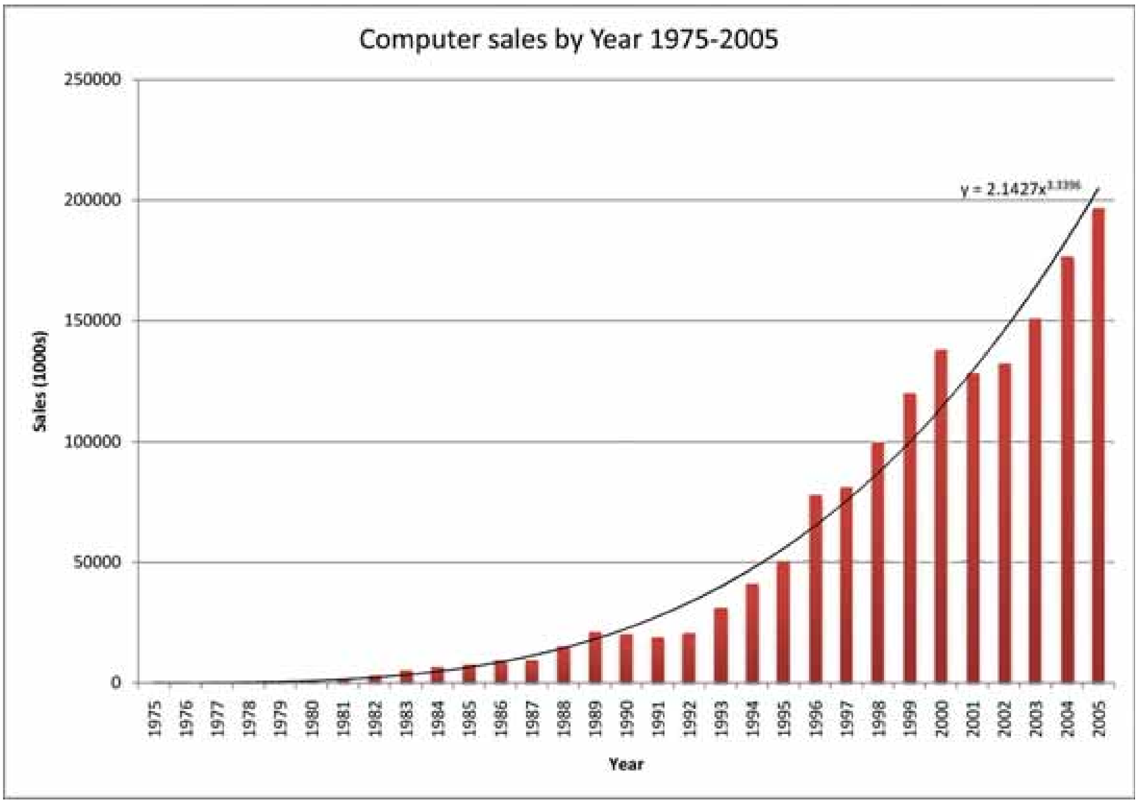


Figure 3.17 Computer sales by year 1975-2005 (Reimer, 2005, 2009)

Part of the growth in the computer market has been due to the invention of a series of very useful peripheral devices. Floppy disks allowed up to 1.4 MB of data to be transferred from one computer to another and save data in a location remote from the computer. Over time, these were replaced by Zip drives, which could carry up to 250MB of data and CD-ROMS (700 MB) and DVD-ROMS (4.6 GB). Currently USB drives and portable hard drives are used for this purpose carrying up to a terabyte of data. Data input still remains via a keyboard, but it was once possible from a light pen, which interacted with the screen, then the now ubiquitous mouse, and more recently via trackpads, tablets and touch screens. To see what the computer is outputting no longer requires special cards, or a typewriter style device like a Teletext. Laser printers are common and computer displays have progressed from televisions and plasma displays to LCD screens. Hard drive storage capacities, the memory onboard computers and processing power have grown dramatically allowing users to collect, collate, analyse and display large datasets in swift time.

Displaying graphical information was not straightforward even up to the 1990s; most machines were not capable of running in more than 256 colours, requiring any scanned images to be ‘dithered’ down from their native millions of colours to an artificial 256. The maximum screen resolution in most home display units was 640x480. In order to obtain high quality images, photographs were often taken on slide film. The slides were then scanned using a dedicated slide scanner and then reduced to 256 colours using specialised software.

Digital video has been briefly mentioned, for it was the least developed of all of the supported technologies leading up to the late 1990s. Uncompressed video consumes far more space and processing power than is willingly utilised for fear of performance issues on computers and so video needs to be compressed, or encoded. To play encoded video, a computer needs to decode it and display it in a very short time. In 1994, the processing power to do this was restricted to high-end machines, often with video encoding or decoding hardware to assist the main processor. Video codecs often used to compress the video were Cinepak and Indeo because the processing power to run the superior MPEG video smoothly from CDs was not yet available. In order to record the video segments desired for education, one had to record it on analogue tape, and play the tape into a specially configured computer with AV inputs. Encoding this video could take two hours to complete a one-minute segment running at 320x240 resolution. As processing power increased, the ability to encode and decode video quickly grew, so that eventually computers or digital cameras could do the job efficiently. Even in the modern era, video is still an area where caution needs to be taken as a long or poorly encoded video may still not play efficiently.

In the early 1990s lectures in pathology were replaced in part by computer-based modules with good outcomes for the students (Trelstad & Raskova, 1992). There were applications available to assist students in learning oral anatomy (Bachman, Lua, Clay, & Rudney, 1998), neuroscience (Teyler & Voneida, 1992), lung disease (Bresnitz, Gracely, & Rubenstein, 1992), cardiac anatomy (Stanford, et al., 1994), oral surgery (Matthew, Pollard, & Frame, 1998), ophthalmology (Kaufman & Lee, 1993) and geriatrics (Andrews, Schwarz, & Helme, 1992). The descriptions of evaluated tools showed a great interest in the field by the authors and showed how the preferred software development tools changed towards applications such as Authorware and Director partially because supported the development of CD-ROMs. Where these tools were evaluated they showed equivalence between the computer-assisted learning and traditional methods.

The use of CD-ROMS gave students the ability to actively participate in the educational process away from their institution. They gained the educational benefit of using learning material, but were able to utilise it in their own time, at their own pace and in a comfortable, familiar environment. Those students who did not have computers also benefited, as there were fewer students requiring the use of university computers, thus reducing waiting times and the pressure from other students waiting to use a machine. This caused a newer paradigm of learning to occur as visualised in Figure 3.18.

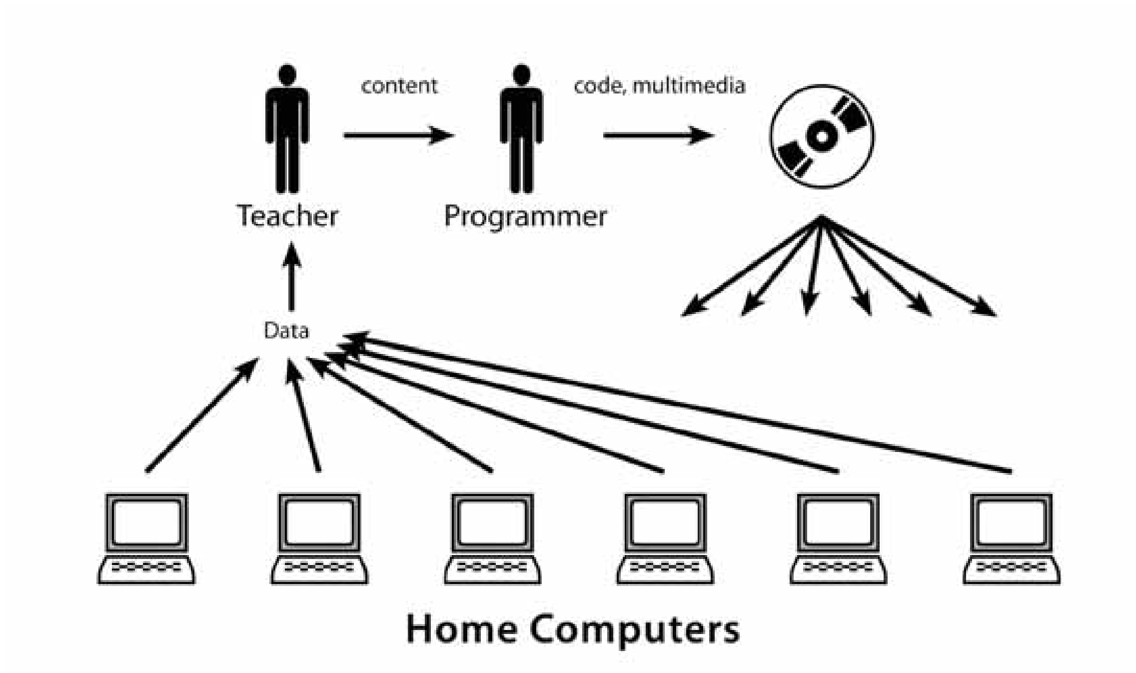


Figure 3.18 Schema illustrating learning using a CD-based methodology

In the late 90s and early 2000s CD-ROM applications were readily available for teaching. Once again, many areas in the health field were represented by the new technology. The interpretation of ultrasounds was supported (Valley, Cardenas, Spangler, & Folstad, 1997), as was the teaching of dietetics (Herriot, Bishop, & Truby, 2004), laparoscopic surgery (Rosser, et al., 2000) and auscultation (Finley, Nanton, & Roy, 1998). Professional CDs were developed and sold addressing fields such as gastroenterology (Valori, 2002), pathology (Spak, 2000), phlebotomy (Littleton, 2000) and anatomy. The latter was represented by the Digital Anatomist program (Rosse, 1995), which could run on CD, laserdisc or over a network thus bridging technologies existing at the time but undoubtedly at a significant cost. Many other programs were developed to run within an institution, over networks or on local workstations, allowing students to learn endocrinology (Holt, et al., 2001), epidemiology (Dean, et al., 1998), urology (Phillips, et al., 1998), neuroanatomy (Lamperti & Sodicoff, 1997) and surgery (Rogers, et al., 1998). There was even one CD-based program that was designed to help paediatric residents to answer the phone correctly (Ottolini & Greenberg, 1998).

When compared with face-to-face teaching, individual studies provided mixed results. Some showed parity with traditional methods (Bissell, McKerlie, Kinane, & McHugh, 2003; Holt, et al., 2001; Williams, Aubin, Harkin, & Cottrell, 2001) and others reported increased benefits for students (Devitt, Smith, & Palmer, 2001; Devitt, Worthley, Palmer, & Cehic, 1998; Elizondo-Omana, et al., 2004; Lowe, Wright, & Bearn, 2001; McGrath, Kucera, & Smith, 2003).

CD-ROM-based learning had its drawbacks. Firstly, there was no way to update the content of a CD. Errors in educational material had to wait until the next release to be corrected, and it was impossible to guarantee which version of a particular resource a student might have on their machine. As was eloquently put by Meyrick (1973) when referring to cassettes, but which applies equally well to CD-ROM-based content

> 'You can produce 100, 1000, 2000 cassettes, every cassette is the same. And what a rotten, stinking mediocrity of uniformity are we going to end up with if that first one isn't jolly good.'

There were also issues with troubleshooting problems that developed on students' home machines. These problems may have involved installation issues, and could be particularly difficult to track down considering every student was likely to have a different configuration of hardware and software on their predominantly Windows-based machines.

With the ability to play video, audio and text resources (often called multimedia), as well as print out any relevant material, computers had become extremely versatile devices. However if they cost over a million dollars as some of the earlier machines did their usefulness would have been to a large extent irrelevant. It could be argued that the biggest change in computers over the last 50 years is the cost of ownership.

Costs of Technology

One of the biggest factors in adoption of computer technology has been the cost. Table 3.2 shows how the prices of computers and key peripheral items have varied over the last 44 years. The cost of supporting this technology has dropped as dramatically as the capabilities of the devices have increased. Prices in Table 3.1 have been converted to US dollars using the average exchange rate at the time (Officer, 2009). Prices were then converted to 2008 equivalent prices in US dollars using the Consumer Price Index (CPI) as an adjustment for time elapsed (Williamson, 2009).

How have the costs changed? In 1984, the basic computer equipment required for a student to learn would have a cost of over $10,000 (in equivalent 2008 $US). In 1994, the equivalent computer would have cost approximately $6000 including vastly improved storage, display and processing power. In 1994, most of the multimedia technologies were in their infancy. CD-ROM drives were new, and single speed drives had only been considered standard in desktop machines since 1990. Hard drives of up to 200MB capacity were appearing in desktop machines. Data transfer was via floppy disk. In 2009, an even more improved version would have cost around $1500.

This dramatic drop in price of a tool, which could be used for education and links all of the media previously discussed together for learning, is one of the key reasons why computers are now one of the most accessible educational devices. Furthermore, any computer purchased since 2000 has one additional feature, which extends its power even further by allowing users to communicate with teachers, students and resources from across the world, and that is the ability to access the Internet.

Table 3.2 Changing costs in multimedia in the last four decades in US dollars.

	1965	1976	1984	1994	2004	2009
Computer	PDP-8 Approx 1Mhz, 4KB RAM $116,000	Altair8800 Microcomputer $6700 2Mhz , 256 Bytes RAM	(IBM AT) 256K RAM, 6 Mhz, $7100	486 33 Mhz, 4MB RAM, 200MB HDD $2470	Pentium 4 (2.8 GHz) 40GB HDD, 256 MB RAM $1340	HP Pavillion 2.5GHz Intel Dual core 8GB RAM, 1TB HDD $998
Hard Drive	32kB (1967) $51415 88MB $1,770,000	635 MB, $265,000	$1440 for 10MB	$1000 for 500MB	$280 250GB	1 TB $97
monitor (viewing size)	12 inch monochrome $443,000	12 inch monochrome $5700	9 inch Monochrome $1350	$1300 (17 inch CRT)	$1108 (LCD 17 inch) $414 (CRT 17 inch)	$189 (23 inch LCD)
RAM	N/A	16K $2890	Approx $8300 for 1MB	Approx $1280 for 16MB	1GB $4108	4GB $84
Portable Storage	N/A	Floppy disk controller $5600	Bubble Drive ($6200 per MB) Floppy disk drive $1700 (360k disks)	SCSI Zip Drive (1995) 100MB disks $400, $30 a disk SCSI Jaz Drive (1996) 1GB disks $1000, $200 a disk.	USB Handy Drive 2GB $800 USB 2.0 hard drive 40GB $256	USB Drive 32GB $110 160 GB $110 1TB portable hard drive
Videodisc	N/A	N/A	$2400 (1985) $24 a disk $400 for a master copy	$1230 player	N/A	N/A
CD-ROM Drive (Data)	N/A	N/A	N/A	$1450 (4x speed)*18	$25 (52x)8	N/A
CD Burner	N/A	N/A	N/A	$20000 (1st cost $280000)16 $30 a disk	$55 $0.30 a disk	N/A $0.20 a disk
DVD Burner	N/A	N/A	N/A	N/A	$98 8 $1.00 a disk	$70 $0.23 a disk, $4 dual layer
Blu-Ray Burner	N/A	N/A	N/A	N/A	N/A	$209 $2.90 a disk (25 GB) $16 a disk (50 GB)
Digitising Still Images	N/A	N/A	Black and white MS300 $4600 (1985)	400dpi scanner $1712 Digital camera $1815 (1 Mpixel)	2400 dpi scanner $186 Digital camera (4MB, 3x optical zoom) $559	4800dpi Scanner $40 Digital camera $200 (8 Mpixel, 12x optical)
Digitising Video	N/A	N/A	N/A	$2760 (analogue)	$400 (digital)	$355

References for prices

1965: Pdp8: (Ceruzzi, 2003), Hard drive (UvA Computer Museum, 2009; Wikipedia, 2009), Monitor (Weisberg, 2008)

1974: Altair 8800 (Wilson, 2003), Hard Drive (Sumney, 2007), monitor terminal (Allan, 2001) p17/12, RAM (Wilson, 2003)

1984: Computer (Borrett, 1984), Hard drive (Smith, 2008; Weyhrich, 2009), Monitor (Borrett, 1984), RAM (Borrett, 1984), Portable Storage (Borrett, 1984; Polsson, 2009), Videodisc (Florell & Nugent, 1985), Image Scanner (Machover & Dill, 1985)

1994: Compaq computer (Hewlett-Packard Development Company, 2009), Hard Drive (Webb, 2003), Monitor (Davis, 1998), RAM (Zisman, 1994), Portable Storage (Johns, 2001), Videodisc (L. Johnson, 1994), Scanner (Business Wire, 1994), Digital Camera (personal purchase), Video Camera (Alldrin, 1994)

2004: Computer (A-Z Computing, 2004), Hard Drive (Data Recovery Resources, 2004), Monitor (A-Z Computing, 2004), RAM (A-Z Computing, 2004), Portable Storage (Jaffe & France, 2004; Nemerovski, 2004), Scanner (IT Reviews, 2004), Digital Camera (Casio, 2004), Video Camera (Bassett, 2004)

2009: All items except Blu-ray burner and media (Wal-Mart, 2009) Bluray burner (Pioneer Electronics, 2009), Blu-Ray blanks (Nextag, 2009)

Computers today are heavily used for communication either for business or personal reasons. The medium for this has become the World Wide Web, which facilitates interactions regardless of distance. A user in Nairobi can purchase items from Sydney or talk with friends in New York, Tokyo or London as easily as visiting with friends who live on his or her street. The Internet has changed the way in which the world works and the computer, in one form or another is the tool we use to access it. As can be seen from Figure 3.19 the number of domain hosts in the world, a measure of the use of the Internet, has increased dramatically since its invention in 1969.

NOTE:
This figure is included on page 130 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.19 Internet domain hosts from 1969-2009 (Zakon)

There have been a large number of online educational tools reported and one study showed that there were over 170 that had been evaluated in some manner (Cook, et al., 2008). These online resources cover a wide range of disciplines including a course in oral surgery (Corrêa, De Campos, Souza, & Novelli, 2003), and assessments in nutrition (Henly, 2003). Surgical online modules were developed in 2004 using software that bypassed the need for any browser but still required online access. (Triola, et al., 2004). Some applications have made the journey through many generations of software development. The pharmacology Computer-Assisted Teaching System (CATS) was developed in the 1970s using Coursewriter and had been rewritten in PILOT amongst 4 other languages (Kerecsen & Pazdernik, 2002). It is now currently written in Java and runs online. The authors emphasise that the development of good content is the key to a successful and useful program for learning.

Educational software has been used regularly from the early 1960s with good reported outcomes (Hansen, 1966). By 1960 there were well over 100 reported studies of the outcomes of the use of teaching machines (Fry, Bryan, & Rigney, 1960) most of them showing benefit to students. Meta analyses of the value of computers in education provide conflicting results. A meta-analysis of studies comparing distance education tools with classroom instruction showed wide variability and almost no effect size (Bernard, et al., 2004). The authors commented on the quality of the studies they had to work with by highlighting key components of studies that they felt were lacking. These included measures of attitude, retention and achievement. They did report findings suggesting that the use of a problem-based learning strategy was a good predictor of achievement and of positive attitudes towards distance education learning. They also reported that the effect sizes measured in the included studies varied significantly and concluded that this was due to methodological flaws in the studies. Overall, they found the results so varied and inconsistent that they felt unable to make strong recommendations about appropriate distance education practices to provide good learning outcomes. A meta-analysis in orthodontics showed minimal benefit for online learning whilst raising the issues of cost, the quality of studies in that field and time effectiveness (Al-Jewair, Azarpazhooh, Suri, & Shah, 2009). The cost and the quality of studies have also been raised by others including Letterie (2003) and Greenhalgh (2001). This suggests that studies with a sound methodological basis are important to this area of research.

A meta-analysis of 47 comparative studies carried out in 1992 (Cohen & Dacanay, 1992) showed a significant positive result for computer-assisted instruction over traditional methods, with an effect size of 0.41. This meant that the computer groups performed at approximately the 66th percentile of the control groups. A similar conclusion was supported by later studies (Kadiyala & Crynes, 2002; Khalili & Shashaani, 1994; Kulik & Kulik, 1991; Liao, 2007). Results of other meta-analyses showed similar positive outcomes for elementary students (Kulik, Kulik, & Bangert-Drowns, 1985), statistics students (Sosa, Berger, Saw, & Mary, 2010) and in nursing (Roh & Park, 2010).

A literature review published in 1999 by Campbell (Campbell & Johnson, 1999) showed only 46 articles reviewing multimedia computer-assisted learning, some 18% of the total number of articles retrieved by the review. This is consistent with a more rigorous review showing 2840 articles on computer-assisted learning in the medical field by Adler (Adler & Johnson, 2000). This review illustrated a strong growth in the literature over 32 years, but predominantly of the demonstration type of article, where authors described how they built their educational tool. Comparative studies made up only 11% of all articles. Other studies have shown significant positive effect for computer-assisted learning. A comprehensive analysis of 201 studies in 2009 showed that Internet-based medical education was significantly better than no education (effect size of 1.0) although produced only moderate effect size differences when compared to traditional methodologies (Cook, et al., 2008). A similar result was obtained by online interventions aimed at K-12 year students (Means, Toyama, Murphy, Bakia, & Jones, 2009).

Overall the results from interactive methods of education (e.g.. computers and interactive video) appear to provide real benefit to students where there is no other educational pathway. Results show greater benefit to students than passive modes (i.e. books, radio, phonograph, video) and this is often attributed to greater interactivity facilitated perhaps by laserdisc (Morgan, et al., 2000), computer (Cohen & Dacanay, 1992) or the online environment (Krishna, et al., 2003). Modern technologies such as MP3 players, personal digital assistants (PDAs) and mobile phones can also deliver such interactivity, but that type of educational content is only now being developed.

When compared to modern teaching methods, which tend to have a lot of audiovisual content as a normal component of course lectures, overall there appears to be little difference in learning outcomes for students. This needs to be considered in the context that many studies report very high effect sizes corresponding to vastly improved student performance. This may indicate that some implementations of online or computer-based learning are far more effective than others, dependent on the strategy used, or the content and that there is untapped potential in online learning.

Online education provides a forum for communication to other people or computers, and combined with a computing tool, be it a desktop, laptop or portable device allows students to potentially learn anywhere and anytime. Education is at a point where the tools that assist students to learn are the cheapest, fastest and most integrated in history. With a combination of Internet use, growth in ownership of computing devices and ever reducing cost, education is being provided with an opportunity that is as revolutionary as the printing press was over 500 years ago. We know that these devices can provide an interactive, formative learning experience for our students and the rest of this chapter will discuss the methods that are typically used to do this.

Given that interactivity may be a powerful learning tool, it appears reasonable to make use of the technologies that are having the largest impact on the world since the book, and furthermore encompasses all of them. The computer and the Internet provide, for the moment, a very economical, exciting and efficient way to learn for today's students. Content development for these tools is very important and it has been shown that there are many options in how computer-assisted learning can be implemented. Experiences with previous technological tools has shown that in many cases, the technology has progressed before it can have long lasting influence in the educational realm and it has not been easy to transfer content developed for one system to another.

From the above it might appear as if there is an accepted understanding of how computers can be used to greatest effect in education. This is not so, and although there does not necessarily need to be only one 'right' way to use computers, a series of clear guidelines, tools, methodologies and training is required. There is an understanding that it has value, but the extent of that value and the understanding of how to effectively implement it is lacking. Observations of teachers have shown that although their beliefs in teaching are quite modern, their use of computers in the classroom revolves solely around knowledge acquisition (Lim & Chai, 2008), although this is not a universally held viewpoint (Niederhauser & Lindstrom, 2006) and it is possible it is culturally dependant. A recent, unpublished study of medical educators (Russell, Devitt, & Palmer, 2009) showed only 11% of respondents reported ever having used a computer-assisted learning tool. Moreover, over 90% of respondents were willing to use these tools and over 50% were prepared to assist in authoring material or contribute to CAL implementation, but stated that time constraints were a significant personal barrier to development of CAL material. Similarly, 40% of respondents cited concerns regarding the technical complexity of CAL development and 32% believed developing such material fell outside their current job description or clinical role. Laurillard (2007) argues that technology has been used typically to support traditional ways of teaching and the historical analysis in this chapter supports this suggestion. Blin and Munro (2008, p. 476) state:

> ‘Although technology is now common place in most higher education institutions there is little evidence of significant impact on teaching practices’

and provide examples from their institution suggesting computer use in learning is based primarily around static or traditional methods such as putting PowerPoint slides and lecture notes online. They cite training as an area that required addressing but suggest the social and cultural environment also requires change. This view is supported by others (MacKeogh & Fox, 2008; Selwyn, 2007). Selwyn (2007) believes sustained change can only come about by focusing:

> ‘on a radical overall and wholesale restructuring of universities and university education. (p. 96)’

Given this is an unlikely scenario, influence by example as has been demonstrated for the last 100 years, is the option usually selected and will be the method chosen for the research in this thesis. Even if one persons work cannot benefit an entire institution if it can be shown to benefit a significant group of students it is worthwhile. So, what might be the best method to provide substantial content to students efficiently using online, computer-based tools and engaging as many educators as possible?

Modes of CAL Instruction

By the middle of the 1970s there were prime examples of using computer-assisted learning in medical education. Each of them addressed modes of CAL as defined by Zinn (1967). These modes have been reiterated (Brock, 1994; Dev, Hoffer, & Barnett, 2001; Seltzer, 1971; Smith, 1974), and modified slightly but the following five modes describe most instances of CAL.

1. Drill and Practice is a repetitive approach to learning designed to develop and improve a skill to high levels. This could include quizzes, or practicing suturing. The use of phonographs and some audio and video approaches to education relied on this method in the past.
2. Tutorials where information was provided and students answered a series of questions. There is variation here as to whether questions or some form of self

assessment is required in this mode, but for this research didactic information with no quizzing is included as part of this category, including an electronic text book or a television show. Immediate evaluation distinguished this passive mode from its more interactive sibling. An audiocassette with instructive booklet follows this mode of learning.

3. Inquiry or Dialogue where students engage in some form of dialogue with a computer in order to elicit information and come to some understanding of a problem. Computers were the first real tools that would allow this type of learning to take place.
4. Simulation where a model of a system (such as a mannequin) provides feedback to the user based on their actions.
5. Problem solving where the student is faced with a problem needing to be solved. It may be purely numerical; the computer could calculate data such as the balance of drugs in a human based on selections by the student or the problem could be a full clinical management scenario.

Drill and Practice

The type of learning approach used by Pressey in 1926 was the Drill and Practice approach. The drill and practice methodology relies on the user to work repetitively on a task in order to develop and improve a skill or skills to high levels and could be included as part of a model of online learning proposed by Anderson (Figure 3.20), much of which could be applied to any sort of learning activity. Drill and practice could be included in the content side of the figure, with Tutorials, games and e-books. The skills developed using drill and practice may range from assembling machinery or carrying out laboratory procedures to performing a problem solving exercise that relies on similar procedural or intellectual tasks to be repeated. Familiar approaches to this would be memorisation of ‘times tables’ for multiplication or the repetition of scales on a musical instrument.

NOTE:
This figure is included on page 136 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.20 A model of online learning (Anderson, 2004, p. 61 (Figure 2-4))

The repetition aims to make the nature of the task second nature or maybe even instinctive. In the medical field this type of approach could be used to assist students in their learning about different dermatological symptoms by studying slides of people with a variety of conditions. Another example of this would be calculating the requirements of patients suffering from fluid loss. Even though the circumstances of each task may be different, the goals of understanding fluid balance can be met. By the student repeating fluid calculations until they achieve a level of knowledge and understanding they may be able to demonstrate mastery of the task. In a study amongst Dutch higher education institutions, the drill and practice approach was used in 20% of cases in medical schools (de Jong, Van Andel, Leiblum, & Mirande, 1992).

Drill and practice tasks may be repetitive but they need not be boring and nor are they necessarily regarded as not stimulating learning (Skinner, 1984). Technologies already discussed which were often used in this way were phonographs, audiocassettes and static image formats such as slides. Using technology, language mastery may come from listening to a variety of language specific audio either on cassette, CD, DVD or other auditory media. The content may well be entertaining and the student can become familiar with accent and word formation from listening to a series of these learning tools. Similarly, a series of engaging case presentations ranging from burn victims to travellers returning from overseas with cholera may impart learning objectives related to fluid balance without appearing repetitive. The drill and practice approach was used to identify parts of nails using NailTutor (Fleckman, Lee, & Astion, 1997), to learn to interpret ECGs (Devitt, et al., 1998), to improve cardiac auscultation (Stern, et al., 2001) and to identify structures of the teeth (Bachman, et al., 1998). The CATS pharmacology online tool discussed previously (Kerecsen & Pazdernik, 2002) also relies on a drill and practice methodology, which the developers have supported for 30 years. It relies on good visual cues and multiple-choice questions to fulfil its function.

Drill and practice tasks may rely heavily on multiple-choice questions as part of formative assessment and learning to achieve sufficient volume of repetition with a minimum, although still substantial, amount of effort from the developers of the educational material. Although this type of style could be considered more traditional than a more modern constructivist approach (Lim & Chai, 2008), non-analytical processes such as pattern recognition are regarded as an essential components to making effective clinical diagnoses (Eva, 2005) as well as important factors in interpreting pathologies (Crowley, Naus, Stewart, & Friedman, 2003) and mammograms (Nodine, Mello-Thoms, Kundel, & Weinstein, 2002) and thus modes of learning which promote this approach should not be summarily dismissed.

Drill and practice exercises form part of traditional teaching, but a certain level of knowledge is often required before this style of learning is usually used. This can be obtained from other forms of instruction. Some of this may be self-directed or may come in the form of the lecture, which is an efficient way of transmitting knowledge to a large group of students (Fowell & Bligh, 1998). That type of transference can also be facilitated using another type of methodology often employed in computer-based education; the tutorial method.

Tutorial

Tutorial learning is used primarily to introduce new material to students and effectively mimics the classroom tutorial or lecture. It can be an audio capture of a lecture, a podcast or something more elaborate such as a multimedia presentation using software such as Articulate or PowerPoint or, in previous eras, educational TV, video cassettes or computer multimedia using software such as Hypercard, Authorware or Toolbook.

Although primarily didactic, the tutorial method provides the student with the opportunity to relive a teaching experience, such as a lecture or a didactic form of tutorial, with some control over the pacing and volume of content being delivered. This allows students to study areas they felt they did not understand clearly or even to learn material they had not encountered in traditional teaching modes. It also allows for a more thorough uptake of the information imparted. It has a stop/start ability so can be accessed over multiple time periods. This can be useful, for often, there is more information imparted in a lecture or some types of tutorials than is able to be absorbed at the pace it is presented.

The didactic nature of presenting material does not necessarily mean that students fail to learn from this style of teaching. In fact, a study by Devitt and Palmer (1999) showed that didactic material may provide just as much effective opportunity for learning as more complex methods to develop such as case-based scenarios with interaction built in. Furthermore, this study also showed that these methods were efficient means of learning.

The tutorial method often applied in CAL relies on some form of self-assessment after the tutorial. This may involve a drill and practice type of component following the didactic material. The previously mentioned nail-tutor program contained a tutorial component (Fleckman, et al., 1997) as did the ECG software (Devitt, et al., 1998).

Inquiry

Inquiry learning revolves around a number of tenets, some of which are to acknowledge that the objective of the inquiry is often a process rather than a solution to a problem with answers not often found in textbooks (Orlich, Harder, Callahan, Trevisan, & Brown, 2009). Learners are able to pose questions, utilise a process of discovery and then test the discoveries in order to obtain greater understanding (De Jong, 2006). Providing an inquiry framework for medical students was the goal of some educators who worked with early computers, (De Dombal, Hartley, & Sleeman, 1969b; Thies, Harless, Lucas, & Jacobson, 1969) and much of the early reported effort focused on simulating the doctor-patient interaction with dialogue. An example of such an interaction is reproduced below and it can be seen that it was designed to develop higher order skills in the clinical management of patients, particularly the taking of a patient history.

NOTE:
This figure is included on page 139 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.21 Paraphrased from a case presented by De Dombal (De Dombal, et al., 1969a)

Early programs were constrained by limited graphical displays, primitive interfaces and low capacity storage options. An early system, *The Leeds Computer Assisted Learning System* could only store 8000 words in its memory (8kB). Students typed in questions via Teletype (da Cruz, 2003), acting as the physician in a clinical situation. The input text was matched against 'keywords' in the systems memory and if a match occurred, the computer would prompt the student to continue questioning. When no match occurred the computer would provide a different prompt. Support materials, such as essential reading, provided the student with appropriate resources to answer questions and the program had sufficient sophistication to preclude students from continuing until they had asked questions central to the management of the patient (De Dombal, et al., 1969a, 1969b). Other programs existed, which worked under similar principles (Gilchrest, et al., 1972; Starkweather, 1967).

While functioning inquiry mode programs with dialogue have been designed and proven to function effectively they are quite complex in nature (Dillenbourg & Self, 1992; Gable & Page, 1980) and vary considerably in their potential for learning (Gorsky & Caspi, 2005). The complexity and the need to anticipate how students might respond to a question is a great challenge for this type of learning tool and the development of content using this mode places a 'considerable burden placed on the author' (Zinn, 1967).

Despite being one of the earliest kinds of CAL modes to be used for medical education the inquiry mode was considered to be still developing in 1974 (Smith, 1974) and although some sophisticated artificial intelligence systems such as GUIDON were developed in the late 1970s to assist in this type of interaction in medicine (Clancey, Shortliffe, & Buchanan, 1979). Only a handful of these kinds of software have been developed (Crowley & Medvedeva, 2006), providing an indication of the difficulty in their development. This is illustrated by an extremely sophisticated system developed by Crowley that effectively separates the actions of problem solving from the pedagogy of teaching, but even then, the authors note (Crowley & Medvedeva, 2006, p. 116) 'clearly an important aspect of medical reasoning - the ability to reason under uncertainty - is not addressed in our current system.'

The cases used in this mode were not necessarily linear. Educators at the time felt that the ability to provide a 'branching' situation, where the 'patient' would respond differently depending on the actions of the doctor, was important to the learning process (Brown, Groome, Niehoff, & Cleaveland, 1968; Starkweather, 1967). Some programs, designed to instruct students on the financial cost, time spent and suitability of investigations such as x-rays and blood tests responded specifically to requests for those tests and provided prompting such as information on how much time was passing and key events occurring at those times (Friedman, Bruce, MacLowry, & Brenner, 1973). Branching was also used as a method to measure the ability of the user and 'jump forward' if questions were too easy (Brown, et al., 1968; Gilchrest, et al., 1972), thus tailoring the user's experience to content, with which they may not have been as familiar (Reed, et al., 1972)

Other programs used multiple-choice questions to progress users through content (Meyer & Beaton, 1974) some of it appearing to be purely knowledge-based rather than testing higher order cognitive skills (Brown, et al., 1968) although some examples provided a branching option dependent on student decisions (Brown, et al., 1968; Hoffer, et al., 1972). Another variation on the type of program available asked students to input text, which would later be read by a human to look for issues that required attention (Reed, et al., 1972). These types of program encouraged inquiry but not through dialogue and could be considered to be a low level of inquiry learning.

Simulation

Simulation involves the active simulation of a system usually by a physical representation (Cooper & Taqueti, 2004) although a patient management scenario could be considered to be a low-level simulation, potentially made more authentic by the use of appropriate media. The main factor involved in a simulation is that there is a sense imposed on the user that the situation is real. This is created by realistic cueing from equipment (real or virtual) causing them to react as they would in a real life situation (Issenberg, Mcgaghie, Petrusa, Gordon, & Scalese, 2005). The goal of simulations is to provide students with the ability to learn skills in an authentic setting. As an example, if one were to be teaching the taking of a patient history a lecturer could choose, if resources permitted, to have the student interview an actor trained to respond like a real patient (Barrows, 1968). This type of simulation requires human support, but it is possible that the scenario can be enacted via discussion boards or role-playing scenarios with students rather than by trained actors.

Other types of simulation may involve the mathematical calculation of outcomes based on student input. An example of this would be a student observing the effect on cardiac performance based on their altering some variables, such as heart rate or preload (Davis & Gore, 2001; Peterson & Campbell, 1984). These types of effects could not be carried out on a patient for fear of causing home, but the computer is able to solve complex equations modelling the heart and present the output to the student. Other examples of this numerical simulation are the calculation of drug concentration versus time in pharmacokinetics (Bolger, 1995), and the effect on blood glucose levels of simulated patients based on diet and insulin levels (Palacio, Lehmann, & Olson, 2007).

Simulations usually go beyond tutorial and inquiry learning by providing a greater level of realism in the feedback to the learner and can encompass virtual patients, simulated patients, mannequins or role-playing. Some procedural simulators provide haptic or force feedback; providing real sensation when using these devices. These types of simulators include laparoscopic training tools that not only look realistic, but also through the haptic feedback provide an accurate impression of what it is like to operate on a real patient. Other examples include mannequins used to simulate anaesthesiology situations (Abrahamson, Denson, & Wolf, 2004; Doyle, 2009; Price, Price, Pratt, Collins, & McDonald, 2010). Although tremendously powerful as learning tools, there is a substantial cost to these devices (Haque & Srinivasan, 2006) and as such, they have not made significant impact into higher education as yet (Cooper & Taqueti, 2004). There is also doubt as to whether these tools provide sufficient educational outcomes for the cost. One study at Wayne State University, USA (Schwartz, et al., 2007) showed no difference in outcomes for students who used computer-based interactive mannequins including those tests measuring the ability to manage a patient under cardiac arrest when compared to a case-based learning approach. This is potentially due to the lack of a clear shared pedagogy revolving around the use of simulators as suggested by Cooper (2004). Certainly laparoscopic simulators, even those without haptic devices have demonstrated learning as measured by reduced time to competency (Hyltander, Liljegren, Rhodin, & Lonroth, 2002; Sherman, Feldman, Stanbridge, Kazmi, & Fried, 2005) and a reported fourfold increase in proficiency in laparoscopic surgery reported by Larsen (2009). A simulated system for training in ultrasounds using virtual patients showed the system developed was capable of training students to the level of practicing gynaecologists (Heer, Middendorf, Miller-Egloff, Dugas, & Strauss, 2004). A meta-analysis of games and simulations from many disciplines has shown that students have improved learning outcomes when they control progress through the simulation as opposed to methods that use some kind of decision making (Vogel, et al., 2006). The study also suggested that greater realism in the media provided to students made the learning process more efficient. This may make virtual worlds a possibility for high fidelity simulations, such as one developed to teach cardiopulmonary resuscitation

(CPR) (Creutzfeldt, 2010). These types of approaches to simulation are relatively new, but promise a different type of authentic environment in which to learn.

Problem solving

Problem solving is a high level task where the synthesis of a multitude of relevant data and making conclusions is a key component. It has been defined as 'the inferential steps that lead from a given state of affairs to a desired goal state.' (Barbey & Barsalou, 2009, p. 35) Formulating differential diagnoses for a patient given a history and symptoms is an example of such a process. This category includes numerical problem solving as well as far more sophisticated tasks. Technologies that can address this issue are mostly beyond the scope of this thesis, but discussion groups, wikis and online role-playing including 3D virtual worlds are all relevant and useful tools for developing full problem-based learning approaches. Within the scope of this research, case-based learning, which tends to not allow students to participate in defining the problem or developing the solution (Savery, 2006), can encourage some of the high level cognitive reasoning required to solve problems. It is acknowledged that cueing from selected response questions and lack of true interaction from entering free responses will restrict some of the learning possible via this mode of learning. Numerical problem solving examples include software designed to simulate the effect of drugs on plasma concentration for clinical pharmacology (Kahn & Bigger Jr, 1974)

Modes of instruction that are beyond the scope of the majority of educators are the use of simulators, which are expensive and tend to not support self-paced learning. This is due to the expense of the resources, which restricts access to them for security purposes, even though they strongly supporting authentic learning. Numerical problem solving is relevant where appropriate, but it isn't possible to develop something generic to solve all medical numerical issues. Full inquiry via dialogue is a path that has been considered by the author and whilst it does give added authenticity, it is at the expense of the amount of content that can be developed. It is very challenging to set a problem to cover the naturally spoken language and have it respond accurately. There are the hazards of spelling errors, missed responses to student comments as well as the massive development time to create a problem in that kind of environment. The biggest issue though is one that is entirely pragmatic and isn't due to technical issues. Clinicians need to write material for students, and their time is rarely given over to such tasks. Furthermore, answers to essay questions written by groups of clinicians have been shown to not be comprehensive enough to anticipate all reasonable student responses (Palmer, et al., 2010). This is precisely the skill required to write inquiry type problems and for this reason it is quite likely that the inquiry with dialogue type of instruction is best carried out with live patients or actors (Barrows, 1968; Bokken, et al., 2009) or perhaps in a virtual reality environment such as second life (Boulos, Hetherington, & Wheeler, 2007).

Experienced practitioners with some effort can develop simulations at a low level, such as case-based scenarios from real-life case studies. This would provide an authentic environment to learn, especially if supplemented by images and video, that is both manageable and cost-effective. A well-written case scenario could provide a purely didactic experience, thus supporting the tutorial mode of learning and it could also be more ‘question and answer’ supporting the drill and practice mode, both of which may be appropriate at different times. More importantly though is that well written case-based scenarios are capable of providing interaction and engage the students in decision making and problem solving. It allows for limited inquiry via cueing and potentially by free text responses from students, thus providing some higher-level learning. Rather than having free range to inquire about all aspects of a case, the student will be constrained, but have sufficient scope to learn. Furthermore, with substantial feedback provided the consequences of making poor decisions, can be made clear.

Summary

This chapter has shown that technology has had a strong influence over education over the last 500 years. It has provided greater levels of authenticity to learning by providing visual and aural support. Furthermore, the computer provides a means of combining all of these media types into one interactive experience for students, which has been shown to provide the greatest effect on learning outcomes. It has also been shown that the technology required to support this type of learning is readily available to students at a low cost. It is therefore reasonable to focus efforts on providing formative assessment on using online educational tools that

1. Support text, still images, sound and video as alone they provide educational outcomes similar to traditional means.
2. Ensure there is some form of interactivity between student and content
3. Combine with computer-based interactivity as there is evidence to suggest that the outcomes are even better in some cases
4. Provide support for drill and practice, tutorial, low-level inquiry and simulation and problem solving forms of learning for both technical reasons and to be realistic about the amount of time clinicians are prepared to devote to the development of these tools. Support free text entry where greater levels of inquiry are required.
5. Attempt to involve as many educators as possible developing content to try and develop a sustainable tool for learning.

Chapter Four: The Design and Development of an Online Learning Tool

Introduction

The research questions posed at the beginning of this thesis were

- Can the use of an online formative assessment tool improve student learning outcomes for medical students as measured by traditional examinations? and
- If the use of an online formative tool did improve student learning outcomes, what strategies will encourage its use?

Before being able to answer these questions it was necessary to examine the characteristics of formative assessment that would need to be considered when developing such a tool. This was considered in Chapter Two, where a learning template for formative assessment was defined. The template's main features were to

- provide material in a case-based environment;
- allow scoring if the author desires. Scoring must be accompanied by immediate feedback;
- support the use of constructed and selected response items;
- provide substantial feedback to the user on both correct and incorrect decisions;
- allow students to investigate alternative choices without penalty;
- allow lecturers to monitor student progress and isolate areas of weakness in both student performance and in the learning material itself;
- allow students to revisit cases as often as desired.

Chapter Three summarised the support required by technology to implement formative assessment as an interactive, educational tool for medical students. This included

- supporting still images, sound and video;
- ensuring there is some form of interactivity between student and content;
- combining with computer-based interactivity as there is evidence to suggest that the outcomes are enhanced;

- providing support for drill and practice, low-level inquiry or simulation and problem solving forms of learning;
- attempting to involve as many educators as possible developing content.

There are some additional recommendations for encouraging students to use, and continue to use, online learning tools proposed by Liao and Lu (2008) and based on a questionnaire measuring perceptions of adopting an information technology innovation (Moore & Benbasat, 1991). These include

- making the software intuitive and easy to access and use;
- providing advantage to users of the material (this might be by providing relevant material and new content on a regular basis, making the work time efficient);
- making the use of the tool compatible with other components of learning;
- result demonstrability. Is the student convinced that the tool is of benefit?
- enhancing the perceived image of the user. This may be culturally dependent.

In this chapter these points are examined with the goal of incorporating them into a model for an online learning tool that will enable the thesis' research questions to be addressed. The implementation of this model will then be described.

Some of this section has been published previously (Palmer & Devitt, 2007a, 2007c).

Content construction

Case-based or scenario-based approaches to teaching rely on providing students with enough data and evidence to develop and make informed judgements, evaluations and decisions. Clinicians take a history from a patient, using their communication skills to obtain information in order to support or refute a number of differential diagnoses. History taking is usually followed by examination of the patient and a provisional diagnosis is often supported by the results of requested investigations, such as blood tests or x-rays. In real-life, this evidence is provided to medical practitioners upon request. Every action regarding patient management is part of a chain of logical clinical reasoning.

Example of Case-based Scenario: Mrs Tebaldi

As an example, consider the patient who has approached their clinician complaining of abdominal pain. There are many factors a clinician would consider in the first few moments of this patient's consultation. General demographics such as age and gender will be important, as will social behaviours such as the extent of smoking, eating and drinking. Information such as any recent travel may be an important factor to consider. Over a series of questions and answers, the clinician will build up a picture of the patient's general health and lifestyle. In a step-wise case-based scenario, however, the author of the scenario may decide that the communication process can be practised and assessed elsewhere (perhaps in another case, or with the aid of fellow students) and may summarise 5-10 minutes of conversation in one block.

The following is an example from a real teaching case of how one might represent the above situation of a patient presenting with acute abdominal pain in a case-based scenario. The author of the scenario has chosen to summarise the findings of history taking in four paragraphs.

> *Mrs Tebaldi is a 35-year-old woman who has been referred to you by her general practitioner. She is normally fit and well and has never had any serious illnesses in her life, but six days ago suffered with abdominal pain of such severity that she went to her local medical practitioner for his help.*
>
> *In this case, you are a resident medical officer. You ask for more details. Until six days ago the patient was well. Midway through the morning she began to notice a discomfort in her upper abdomen. After an hour or so, this had developed into a severe pain and she needed to sit down. Mrs Tebaldi thought that this might be indigestion and remembers that she had eaten rather a large meal the previous evening.*
>
> *She tried some milk, but this did not ease the pain. Things were now getting worse and she was nauseated. She went to bed, but did not feel any better. The nausea continued and she vomited once. She could only bring up a small quantity of bitter tasting fluid, even though she retched on several occasions.*

Mrs Tebaldi was now convinced she had food poisoning and thought she would take it easy for the rest of the day. She had an early night, and although the pain eased off, it did not completely settle and she needed some Panadeine to help her get to sleep.

The following day was Sunday and although the pain had eased, she still did not feel well enough to eat anything. On Monday morning she went to her general practitioner, and here she is now, sitting opposite you.

As Mrs Tebaldi tells her story, you wonder what might have caused her illness…

The author of the above scenario may want to focus the student on determining differential diagnoses and in this example those diagnoses are amongst a list of eight possible alternatives which the student will need to prioritise before deciding on a course of action.

What conditions are reasonable to consider?

1. *An over-indulgence of alcohol*
2. *Gallstones*
3. *Gastroenteritis*
4. *Acute appendicitis*
5. *Myocardial infarction*
6. *Shingles*
7. *Obstructed para-oesophageal hernia*
8. *Peptic ulceration*

After using the history to determine likely diagnoses, the clinician would examine the patient to find evidence (signs) that support or refute the clinician’s hypotheses as to the cause of the patient’s current problem (the diagnosis). The nature of the examination and the expectations of the outcomes could form the basis of another stage in a case-based problem. The goal here is to focus student attention on why they are making a decision and what they expect to be the outcomes. Cognitively this is a high level task. Previously learned knowledge must be applied, data analysed (the physical examination) and a judgement (diagnosis) made.

The next stage of managing such a case would be to consider a diagnosis or list of possible diagnoses. To confirm or exclude a diagnosis the clinician might order some investigations. The next realistic step in the scenario could be to offer a list of investigations for the student to consider and select appropriately. A relative weighting might be put on the selections in order to focus the choices. The weighting may depend on issues including potential adverse outcomes of an action, the cost of an action balanced against benefit, and the evidence supporting an action. Feedback is provided on each choice and the author of the scenario might provide the investigation. In Mrs Tebaldi's case, an ultrasound might be ordered and the student asked to interpret it. Their interpretation may lead to some course of action, with resultant outcomes and further management questions before it has reached its conclusion.

As established in Chapter Three, feedback on any decisions made should be immediate and comprehensive. Any model designed to incorporate case-based learning must include feedback as an integral element. In the above example, feedback on decisions made by a student on reasonable differential diagnoses needs to encompass more than 'correct' or 'incorrect'. An explanation of why the answer is appropriate or not is essential; otherwise the feedback is basically a score or weighting with no supporting evidence.

If a student selected myocardial infarction, gallstones and gastroenteritis as reasonable causes for Mrs Tebaldi's symptoms, the feedback could be:

> *Myocardial infarction (-1)*
> *Incorrect. She is in the wrong age group and the pain is not that of cardiac infarction. Patients who have suffered a myocardial infarction seek medical attention early in the illness—usually within minutes of the onset of pain.*
>
> *Gallstones (1)*
> *Correct. Gallstones are common in western communities and whilst most stones remain asymptomatic, it is quite possible that Mrs Tebaldi's current problem is biliary in origin.*

Gastroenteritis (1)

This is certainly a possibility. You will want to know more about any other associated symptoms. She has not had any diarrhoea. The symptoms described are not supportive of a diagnosis of gastroenteritis.

To support the learning approach highlighted in the example above, a model such as that illustrated in Figure 4.1 is required. An introductory screen displays information to verify the credibility of the case's author. A summary screen provides an overview of the case, summarising key points and providing information on additional learning resources. These two screens act as the start and finish of a case, although if desired it is possible to include a summary screen within the body of a case. This may be beneficial when the case is branching (different decisions take the student down distinct pathways in a case) and the author wishes to highlight different outcomes to students.

In between the introductory and summary screens are the interactive learning materials, presented as a series of information screens, which would be selected from a single selection screen, multiple selection screen or free response screen. Each case has its learning objectives clearly defined at the start, which links clearly to the content and the summary. In this way the author is provided with a structure to develop learning material in a pedagogically sound way and the student is provided with a well-defined learning case.

The structure described in Figure 4.1 provides support for authors to develop material in a case-based environment. It will allow scoring if the author desires accompanied by immediate and substantial feedback on all possible options. It will support the use of constructed and selected response items. Repetition of the information screens structure (step wise, case-based content in Figure 4.1) would allow for the generation of any case-based problem, including the Tebaldi example discussed earlier.

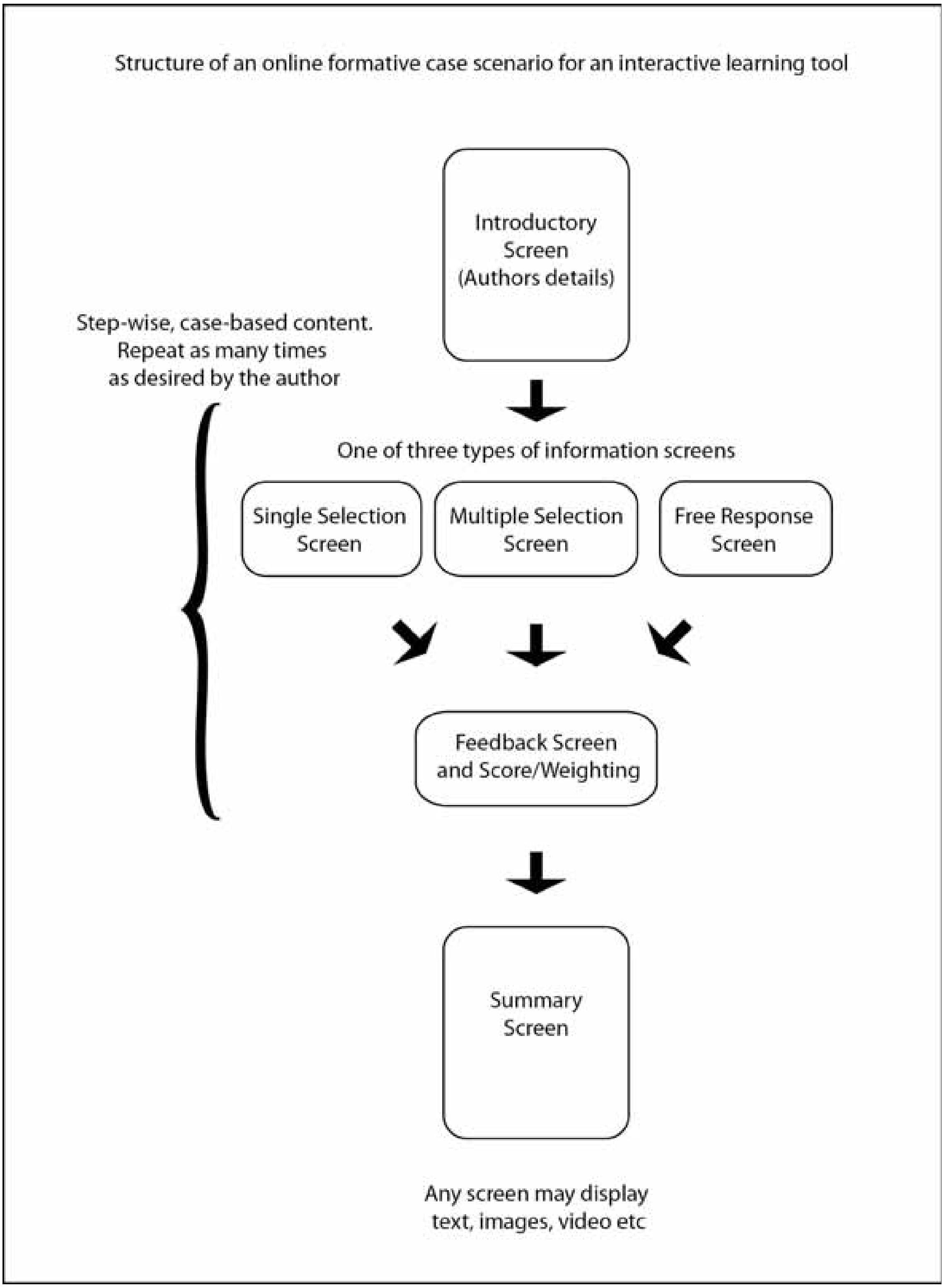


Figure 4.1 Model for interactive learning tool based on the learning template

Figure 4.1 shows the three types of information screens that are required as part of the case-based content and Figures 4.2 and 4.3 illustrate a means of providing these types of screens. Each screen presents some content, poses a question and a means of answering that question. A feedback screen partners any of the information screens and is the core component of much of the formative learning offered by the learning package. Authors need to be able to choose as few or as many of the information-feedback screen pairs as they wish, to develop their case.

Whilst this structure may appear to imply that cases must be linear, it does support the ability to construct multiple pathway cases that allow the students to travel along alternative management lines. This style of case requires substantial additional effort to develop, as each pathway needs to be created in an appropriate methodological and chronological manner. It is unclear whether this type of approach provides sufficient benefit to the student to make the extra development time worthwhile.

Figure 4.2 illustrates the components of an information screen that could be provided to a student for learning. The case scenario needs to be established (Section number 1 in Figure 4.2), a question must be posed (Section 2) and space for a number of choices to choose from, or space to record an answer is required (Section 3). Any images or video linked to the content provided need to be displayed without requiring further navigation (Section 4). The image should be able to be enlarged, especially for images that contain fine detail such as ECGs. Finally, Section 5 provides an onscreen means of navigating through the case without requiring menus. The navigation required would be simple controls such as *Go Forward*, *Go Back*, *Start Again*, *Start a New Case* and *Quit*. In the case of a summary screen, which would usually be presented at the end of the case, boxes 2 and 3 would not be visible.

The structure described by Figure 4.2 supports still images and video and provides for the viewing of these media types simultaneously with the text allowing for a more complete and immersive experience for the student. Interaction is provided by encouraging student decision making in the context of the case via selecting choices or writing responses to questions or examining images in detail. This structure also allows for the use of different learning modes including drill and practice, tutorial, guided inquiry, low-level simulation and problem solving modes and only requires simple navigation for the student to access the learning environment.

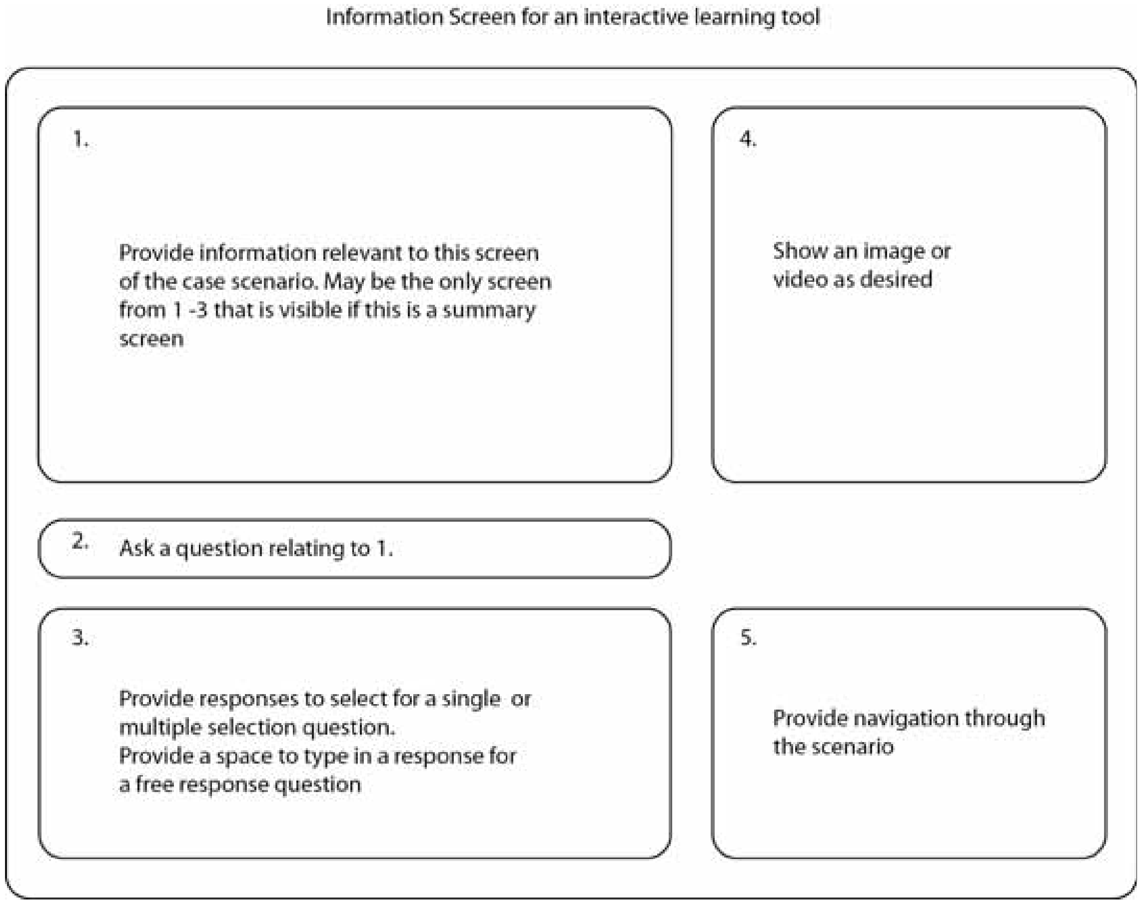


Figure 4.2 Model for an Information Screen

Having made choices or written an answer to the posed question, students could receive instant feedback using the structure proposed in Figure 4.3. In order to encourage as much feedback as possible from authors of the cases, a large amount of space is devoted to it on choices students have selected (Section 1). Scores or weightings indicating the appropriateness of the selection(s) are returned with the feedback. The scoring/weighting system would be optional and the weighting allocated to each option would be a reflection of the author's opinion of the appropriateness of that particular selection. Section 2 would display feedback on options the user did not select but were nevertheless correct. In the case of a written response to a question, the students' answer would be repeated in Section 1, whilst the author's ideal answer would be provided in Section 2 for comparison purposes. Section 3 would use the same navigation as the information screen in Figure 4.2 for consistency.

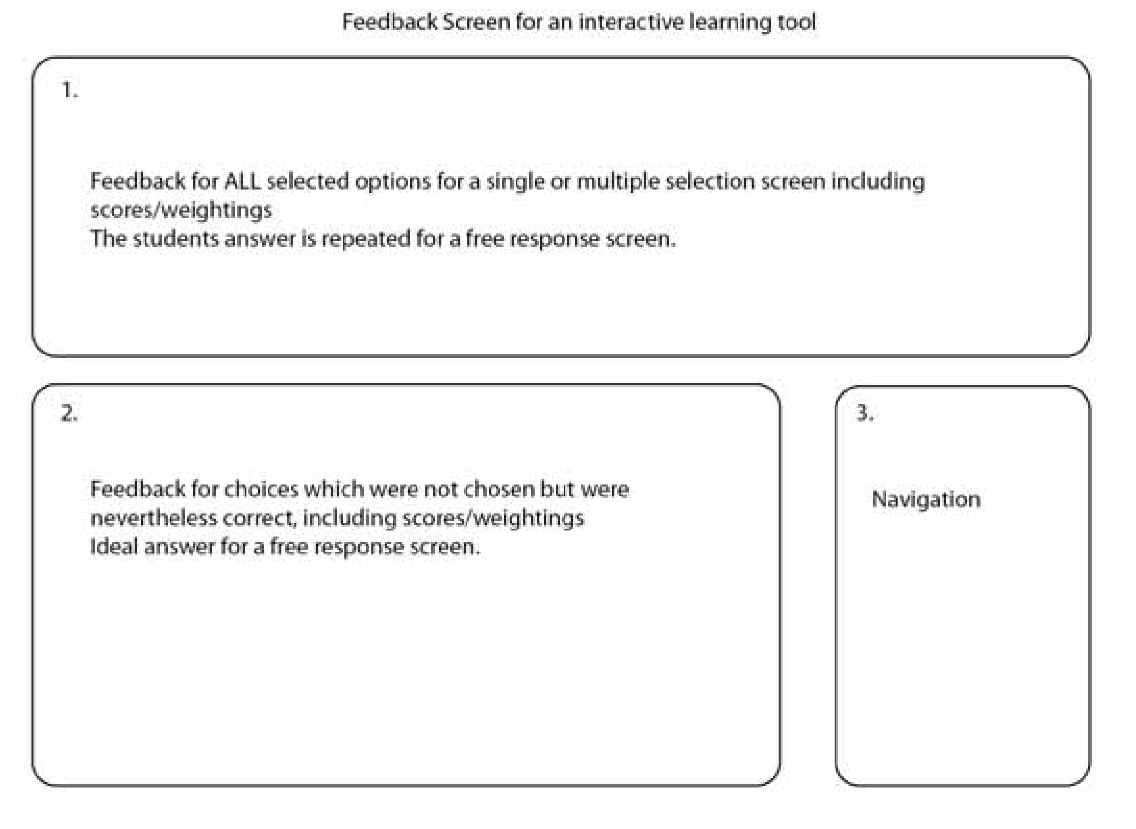


Figure 4.3 Model for providing feedback

The feedback structure described in Figure 4.3 would allow the student to go backwards and examine the choices they had made and change them if they desired, although the benefit of providing feedback for all options is likely to make this unnecessary. If the tool were to be used for summative purposes a score would only be recorded for the first attempt at each screen.

Having established a model for an interactive learning tool, an implementation of it was developed. This online tool was given the name of *Medici* by Peter Devitt; after the famous Medici family who were major influences in the arts and learning in the renaissance period in Florence. They were patrons of some of the greatest artists and humanists of the period and spent vast sums on libraries and learning academies (Gilbert, 2005). The development of *Medici* is described below.

Software Development

It is possible to have applications run as standalone programs, online or on mobile devices using a variety of methods. Programs can be installed on a personal computer, or run directly from CD or DVD; can be used on iPods and iPhones, or other MP3 players and smartphones, and online via websites such as *eMedici* (Palmer & Devitt, 2010). It was important to select a development tool that would encourage access through a wide diversity of devices.

The software used to develop the formative assessment program, called *Medici*, was *Adobe Director*, formerly *Macromedia Director* (Adobe, 2010b). This software was chosen for two main reasons. First, it is cross-platform; able to be used on, and produce content for Apple and Windows-based operating systems. Second, the software could be used to develop *Shockwave* content, which could run online in any browser that used the *Shockwave* plug-in (Adobe, 2010a). The *Shockwave* plug-in is a readily available plug-in, with a high installed base, and is made available by Adobe for a variety of popular Internet browsers. Although not having the market penetration of Adobe's other software platform, *Flash* (Adobe, 2008b; Millward Brown, 2008), *Director* is a versatile platform for software development. *Director* can interact easily with coding technologies such as PHP and ASP. These technologies allow software to transfer content in and out of databases, which was considered to be potentially useful for gathering user data and retrieving content. *Director* also allows for the development of stand-alone CD-based programs. This was important in situations where students were unable to access the Internet.

Director and the other software development components used for *Medici* were used on an Intel-based Macintosh with 2GB of RAM and a 250GB hard drive. *Parallels* (Parallels Holdings Ltd, 2010), a Windows emulation program, was used to test prototype Windows versions of all software, before full testing on a range of Windows-based machines occurred. *Medici* was designed to run on Windows 2000, XP and Vista as well as OS X for Macintosh. A version of *Medici* that would run on some iPods was also created. The software was not developed for other systems such as Linux because Director does not support those environments.

Developing Multimedia Content

Learning resources such as *Medici* require the ability to develop media elements to support the case scenario. This may be audio, video or still images. Each of these media needs to be handled differently, depending on how the user is going to access it. For example, images prepared for the CD version of an application need to be a different size and quality to those to be put on an iPod.

Still Images

Still images can be taken on any digital camera, scanned from slides or paper photos or from other sources, provided the owner grants permission. Images selected for incorporation into Medici usually required editing using *Adobe Photoshop* (Adobe, 2008a). Cropping, rotation, colour correction and level adjustment were all carried out as appropriate for the images. Images that could divulge a patient's identity, had identifiers removed (for example patient names on x-rays). Images for use on CD were prepared in the Joint Photographic Experts Group (JPEG) format, using the highest quality setting of 12 with a size of 800 x 640 pixels and resolution of 72 dots per inch (dpi). Images appropriate for the iPod had a minimum size of 320 x 240 pixels in JPEG Baseline format at a quality sufficient to display the required medical details. For images to be used for online applications, the images prepared for CD were compressed to a lower setting (ranging between 4 and 7, depending on the image) and reduced in size to 640 x 480 pixels.

Video and Audio

Video was filmed on a range of low to mid-end digital video recorders. A direct computer interface, using Firewire or USB 2.0 facilitated fast transfer of video to a computer. *iMovie* (Apple Inc, 2008a) and *QuickTime Pro* (Apple Inc, 2010c) provided sufficient functionality to edit the content. Editing was not elaborate, but rather consisted of cutting out unwanted segments of recorded video and occasionally adding in transitions to provide a smooth viewing experience. Video size on CD ranged up to 640 x 480, but was much smaller for other formats of delivery. On the iPod, the video was encoded to mp4 or H.264 standard with a size of 320 x 240. There is a setting in *QuickTime Pro* specifically for this purpose (Export to iPod). Similar sizes and compression were used for online delivery of video, although as a rule video was considered a poor option for the Internet due to concerns about student ability to download it effectively. The audio component of the video for iPod and CD was encoded using Advanced Audio Coding (AAC) (44.1 or 48 kHz, stereo, up to 160kbit/s on the iPod (Apple Inc, 2007a) and online, 320kbit/s on CD). *QuickTime Pro* was the tool most often used for this function.

When developing content for any educational system, care must be taken regarding copyright and patient consent. Case authors should not use any images from sources such as textbooks or websites without permission and should not use images taken of patients without their permission. Patients are usually happy to consent to have clinical pictures taken, as well as images of test results, including radiological investigations, in order to help educate students. All images and video used have been provided to students only after appropriate permissions have been obtained as certified by the provider.

Programming Methodology

Medici was programmed using *Director's* language, *Lingo*, and an Object Oriented Programming (OOP) approach was used to develop the software. Figure 4.4 illustrates a brief description of how the OOP approach was used to programmatically define the formative assessment tool. As discussed in Chapter Three, multiple methods of user interaction were deemed necessary. These include the free response screen, where students could type their answers to questions posed in *Medici*, and two variations on multiple-choice questions (MCQs). These variations were single selection, which allows the user to select one from many options in response to a question, and multiple selection, where the user can select more than one response from a list of choices.

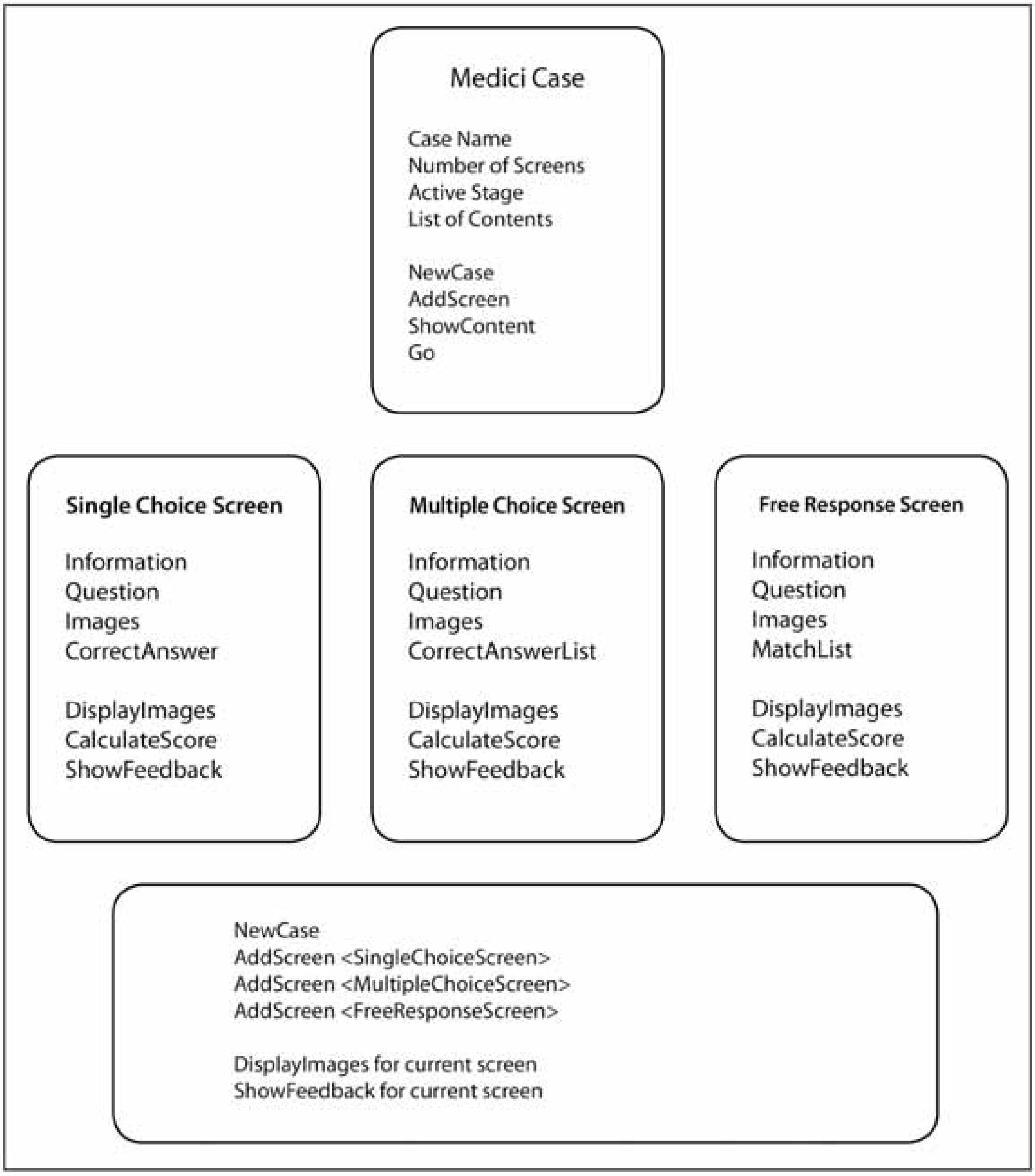


Figure 4.4 Design architecture for *Medici*

In the definition of each of the screen types shown in Figure 4.4, there are some strong similarities in their properties. All of the screens must present some 'information' to students, they must pose a 'question', they must be able to display 'images' or video when required, and they must provide feedback once a student has made a selection.

While these screens have many similarities, they also have many differences. They have specific features that only their class of screen can use. For example a single choice screen (SingleSelectionScreen in Figure 4.4) must only have one correct answer, which is represented by CorrectAnswer (Figure 4.4). The multiple selection screen will have a list of correct answers, and this is reflected by the variable CorrectAnswerList (Figure 4.4). The free response screen, where students can type in their answer may have a list of keywords that the student's answer can be scored against and an ideal answer, which will be displayed as feedback. These are represented by the variables CorrectAnswer, CorrectAnswerList, MatchList and ShowFeedback in Figure 4.4.

Each of these blocks (SingleSelectionScreen, MultipleSelectionScreen etc) are called *classes* and they have all of their common functions made into a separate class and have their own special abilities added as an extension of this basic class. The screens can be managed together in another type of class, which is called MediciCase (Figure 4.4). This class allows screens to be added, displayed and manipulated using a variety of methods. In the *Medici* computer program a temporary manifestation of a class can be created, called an *instance*. Each instance can interact with other parts of the program, and be manipulated according to the code or the user of the code, dependent upon the application. In one example, the author might create a new case and begin populating it with screens. When the user chooses to add a screen, the MediciCase instance has the type of screen selected added to it. If they want to view an image for that screen, they might click on it. The program tells MediciCase to DisplayImages, which then runs the DisplayImages code for that screen. The *Medici* program adds these, displays them or interacts with them without needing to know what type of screen it is. Each screen has its own way of responding to directions from the main program, which is transparent to the user.

The object-oriented approach was chosen in order to facilitate simple modification of the code by future developers. If there was a desire to change the types of screens available, such as a sliding scale indicating agreement or non-agreement with a statement, developers need not be concerned with working on any code but that of their particular module. This is of particular importance, because *Medici* forms part of an Australian Learning and Teaching Council research grant (Australian Learning and Teaching Council, 2009), where it and the content developed for it will be made available to institutions across Australia. As an example, let us consider the Critique screen. A summary screen would be of educational benefit to the users. In nature it is different from any of the screens described in Figure 4.1. It does not need to respond to any choices or input from the user. It is primarily a passive means of summarising a particular case. It does, however, need to provide information to the student and will need to be able to display images. A programmer wishing to implement this type of screen, could use the basic object, which displays information and images without needing to know how the object does its job. The programmer could just focus on developing the components of the new screen, thus reducing development time.

In its entirety, *Medici* consists of 13 programming objects, including the screens described above, 19 multimedia elements, which can link with thousands of others and has over 4500 lines of code.

Programs on CD

Despite comments to the contrary (Kirkwood & Price, 2005; Reisman, 1995), the use of CD-ROMs in education is still supported. They are used to address a wide range of topics such as distributing strategies on how to cope with bullying (Tomei & Burkey Piecka, 2005) and supplements to textbooks (Farrar, 2006) to multimedia projects looking at human dissection (Inwood & Ahmad, 2005), educational CD-ROMs still have a role to play, and they provide significant benefits in important areas:

- Students can take CDs anywhere and use them on most computers, without having to worry about internet connections or payment;
- CDs can hold significant amounts of data, including video, images, video and sound that can often be stored at high quality.
- Downloading the equivalent amount of data that can be stored on a CD (700 MB) would likely use up much of a student's monthly allowance of downloads; and
- Accessing data from a CD is relatively fast.

  A CD-ROM can transfer data at 7.6MB per second, thus ensuring a smooth playback of video and fast display of images. When this is compared with the best speeds available for dial-up (7KB per second), asymmetric digital subscriber line (ADSL) (1MB per second) and ADSL 2+ (3MB per second) (Connectus Pty Ltd, 2008), the most common Internet connection technologies, the CD has a clear advantage. This is most important if it is acknowledged that future clinical scenarios will make generous use of images and video, thus making the need for CD/DVD even more critical. Medicine is an image-rich environment and if, as is hoped, materials are developed to help with clinical skills, then video will play an increasingly important role.

Although the CD/DVD medium can be beneficial in the above areas, the medium is not without its disadvantages, including

- no ability to update the content automatically. There are clear problems with allowing clinical students to have access to out-of-date management advice, so a method must be in place to ensure that content presented on CDs are kept current. Although rewriteable CDs or DVDs would allow for existing content to be wiped and replaced with more current material, there is substantial administrative overhead in ensuring that this would be a viable strategy;
- there are no guarantees as to which version of a case a student may have on their machine, even when updates are provided;
- providing assistance to users can be complex, considering every student may have a different configuration of hardware and software on their (predominantly Windows-based) machines; and
- difficulties automating the connection to a CD for data retrieval, in order to monitor student progress. The administrative overhead in collecting and recording students' use of the program needs to be considered.

Student feedback has suggested that the CD format is still viable, with more than two thirds of University of Adelaide students still preferring CD-based content over iPod or web-based material (Palmer & Devitt, 2007a). This was reflected by a request from students to provide programs in CD format. A *Medici* CD distributed to students in 2007 consisted of 68 cases scenarios supported by 25 high quality videos, lasting over 13 minutes, and over 850 images. The *Medici* program was 6.4MB in size for Mac OSX and 5.4MB for the Windows operating systems. Overall, the entire set of contents required 620 MB of space. This satisfaction with the CD format may have changed in the last three years especially with the existence of online versions of a lot of software, including *Medici* but this has not been measured.

Figure 4.5 shows the model used to provide CD-based content to students using *Medici*. The teacher developed the content, the programmer prepared it for CD production, and the CDs were distributed to students. In order to be able to monitor student progress, *Medici* wrote student results into a data file as they worked on their cases and students supplied the file to the lecturer via email or on a portable data device such as a USB drive. To prevent interference with the data file, it was encrypted and the lecturer used a decryption tool to read the data.

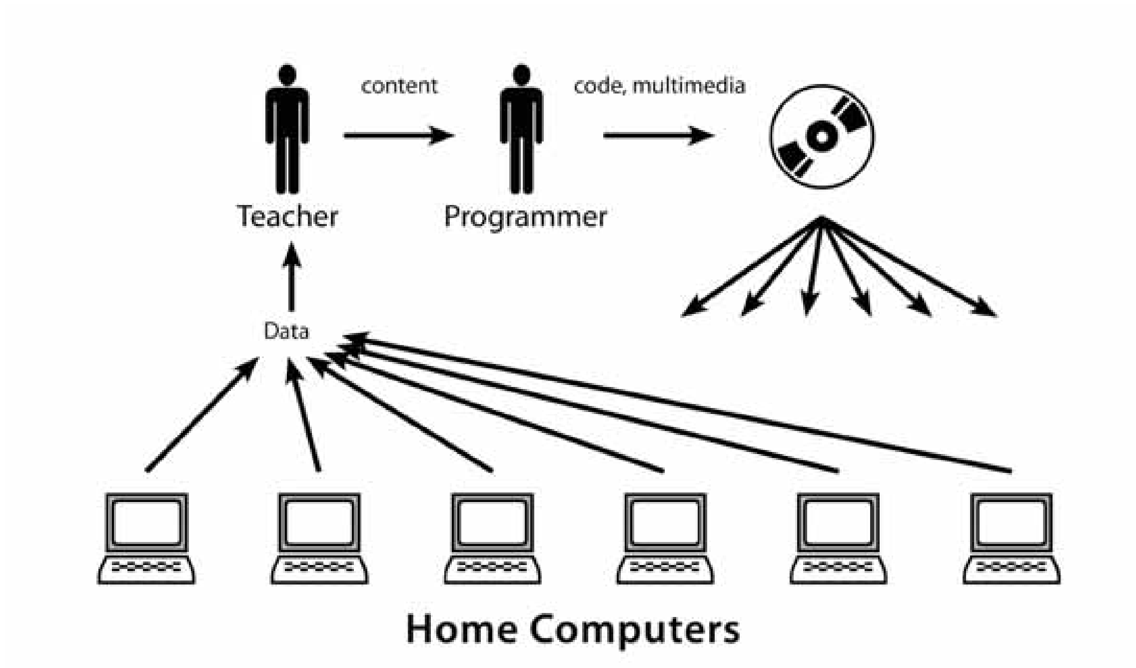


Figure 4.5 Schema illustrating function of *Medici* using a CD-based methodology

As described earlier in this chapter, there are four different types of screens that should be implemented as components of the formative assessment tool template described. These are the single selection screen, multiple selection screen, free text screen and the critique or summary screen. The following pages use a case study to illustrate how *Medici* implements these screens on CD and interacts with the user.

For the CD-based version, the size of images and video is of lesser concern, and an opening screen with musical support was provided (Figure 4.6).

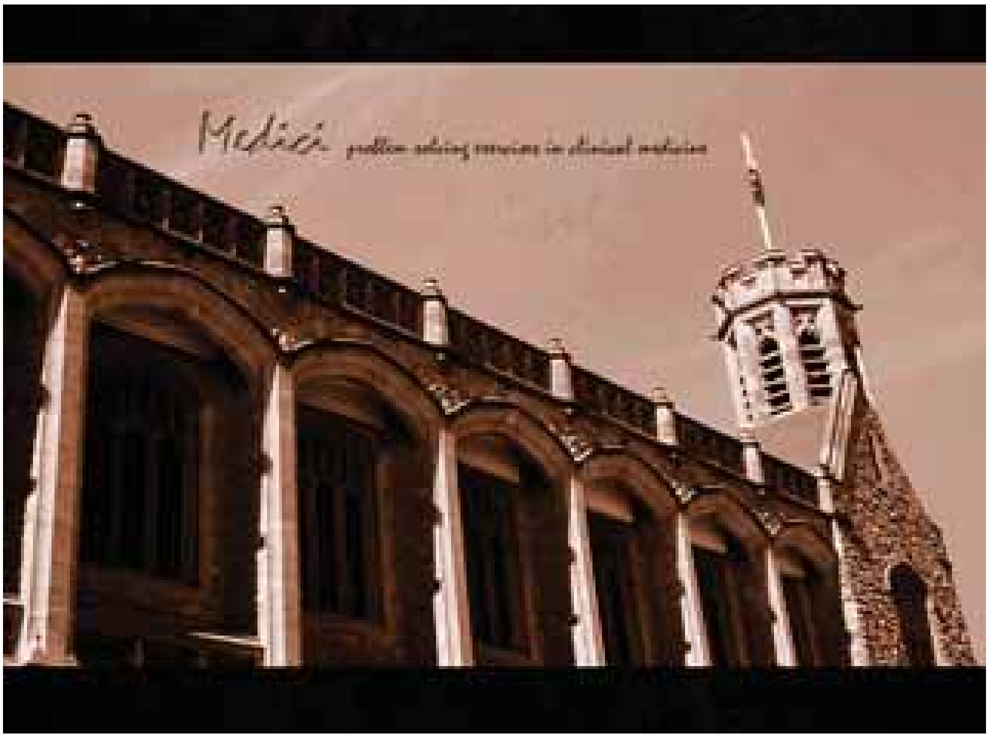


Figure 4.6 Opening screen for one of the CD versions of *Medici*

Figure 4.7 depicts the way in which *Medici* cases are provided to the user. Cases are split into modules and are selectable using either the Proceed navigation on the right of the screen or by double clicking on the case name. A typical problem solving exercise begins with an introductory screen (Figure 4.8), which displays the case title and details of the case author. This provides the user with a direct measure of the pedigree of the author and thus some indication of the potential quality of the case they are about to study. It is an important component in illustrating the appropriate levels of evidence a serious educational tool should possess.

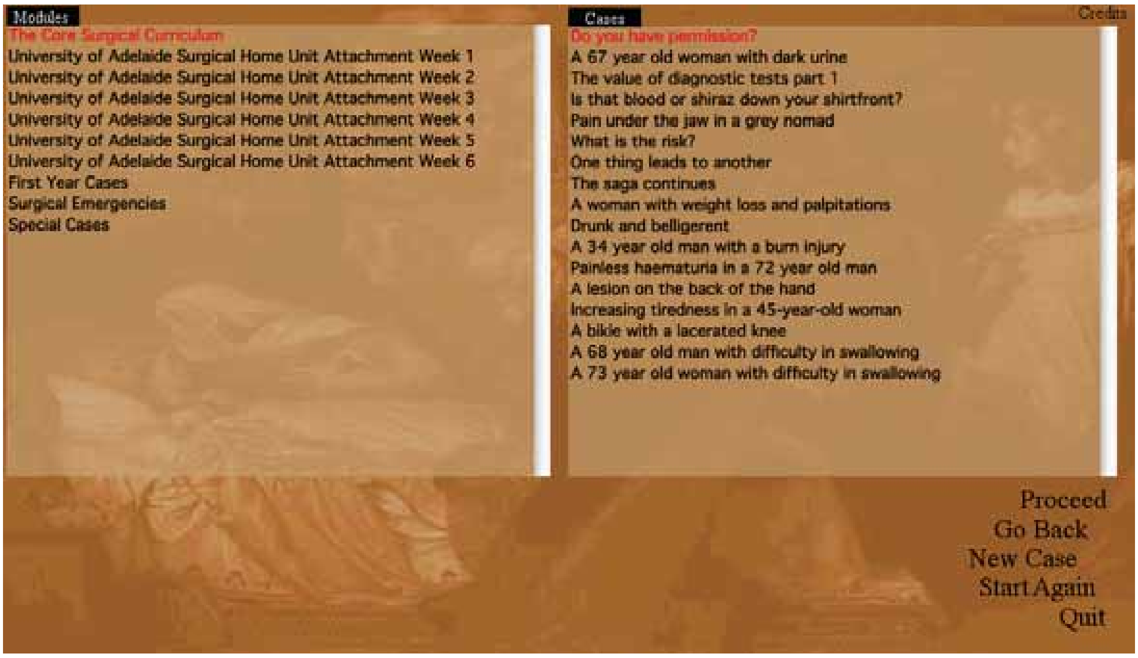


Figure 4.7 Selection of modules and cases in *Medici*

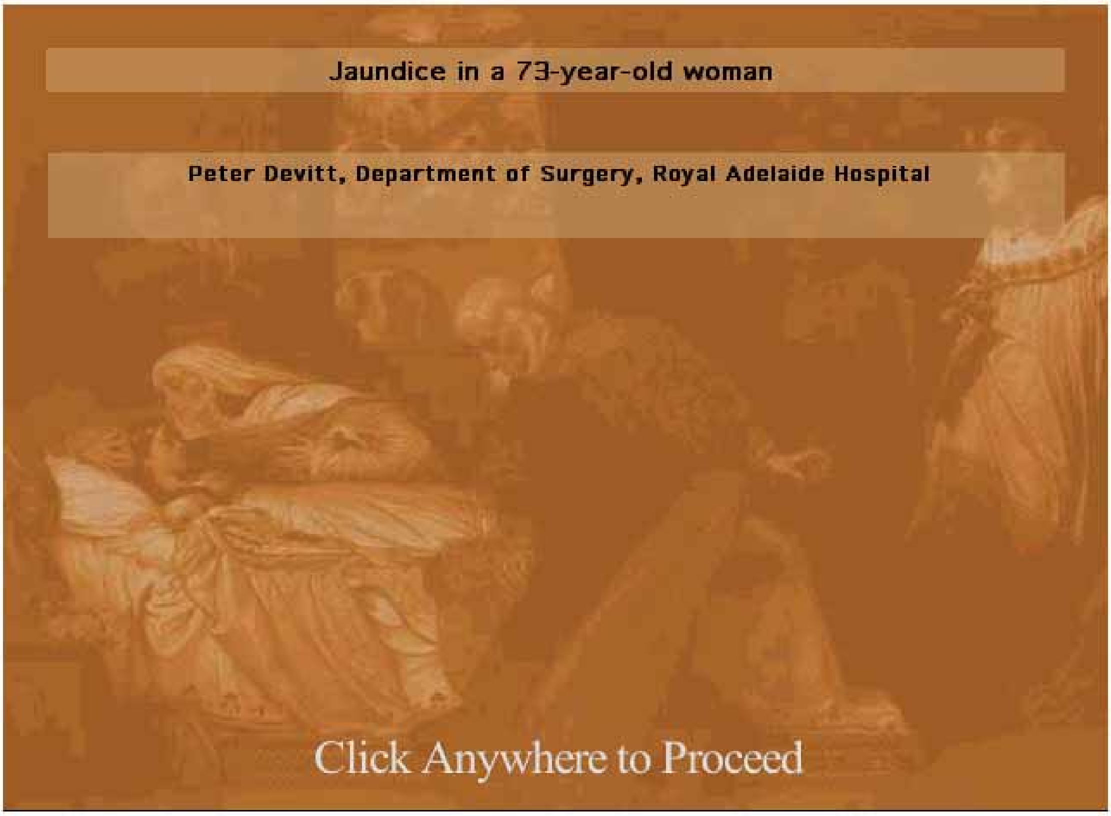


Figure 4.8 Introductory screen for a *Medici* case

Some defined learning objectives or an overview to provide the student with some information as to the likely content may precede the case. These may be particularly important if the author has chosen to construct their case as a tutorial rather than a problem solving exercise. Once the case has begun, each section (or screen) of the case provides some information relating to the current status of a patient, a question about the management of that patient and a series of choices, which turn red when selected. On the single selection screen (Figure 4.9), the user can only choose what they believe is the best of the options provided. For a multiple selection screen (Figure 4.10), the user may select as many options as they wish. In the free response screen, users need to answer the question in their own words (Figure 4.11). All of these screens can be combined into one case scenario, should the author wish as proposed in Figure 4.1.

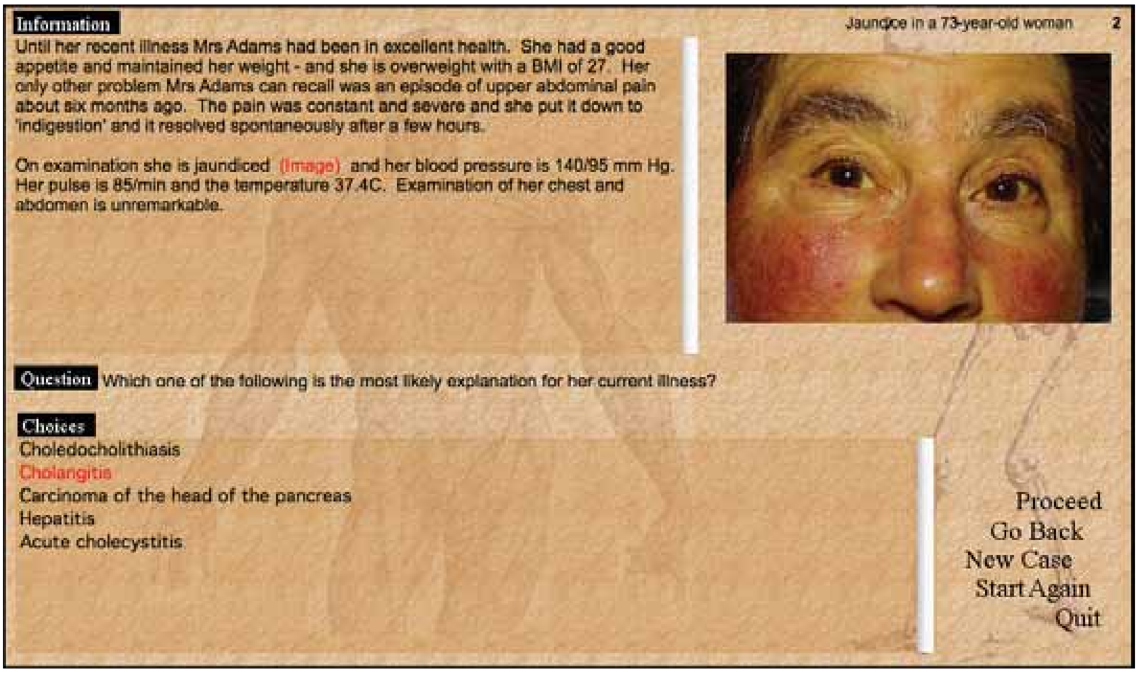


Figure 4.9 *Medici* information and question screen for a single choice screen.
User’s choice is displayed in red

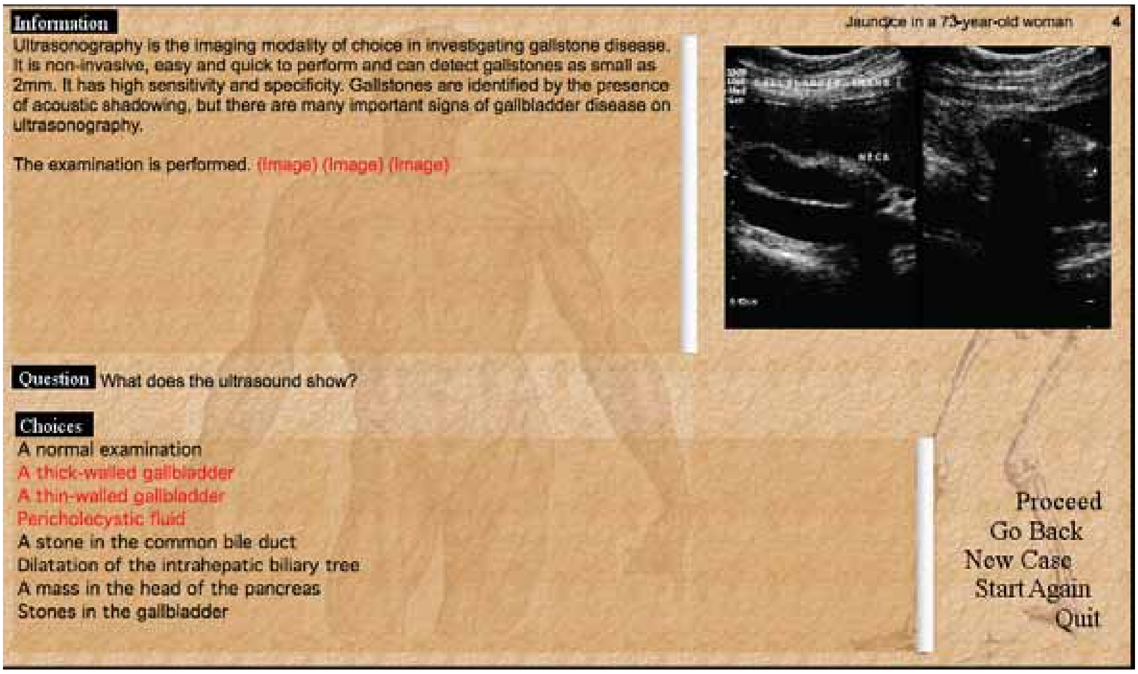


Figure 4.10 *Medici* information and question screen for a multiple-choice screen.

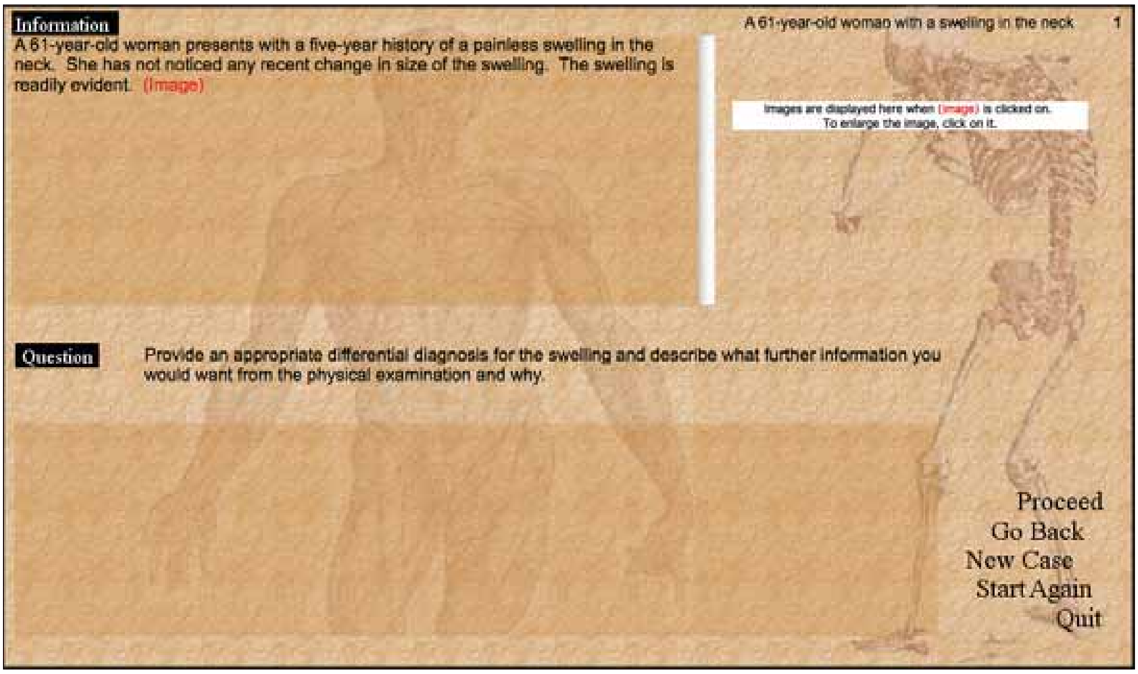


Figure 4.11 *Medici* information and question screen for a free text response question

In the information component of any screen, the word (Image) may appear. This is a cue to the user that there is some form of multimedia assistance provided for this part of the scenario. By clicking on (Image), the appropriate video or still image will appear to the right of screen (Figures 4.9-4.10). The image may be viewed in greater detail by clicking on it (Figure 4.12). Playing video within *Medici* is illustrated in Figure 4.13.

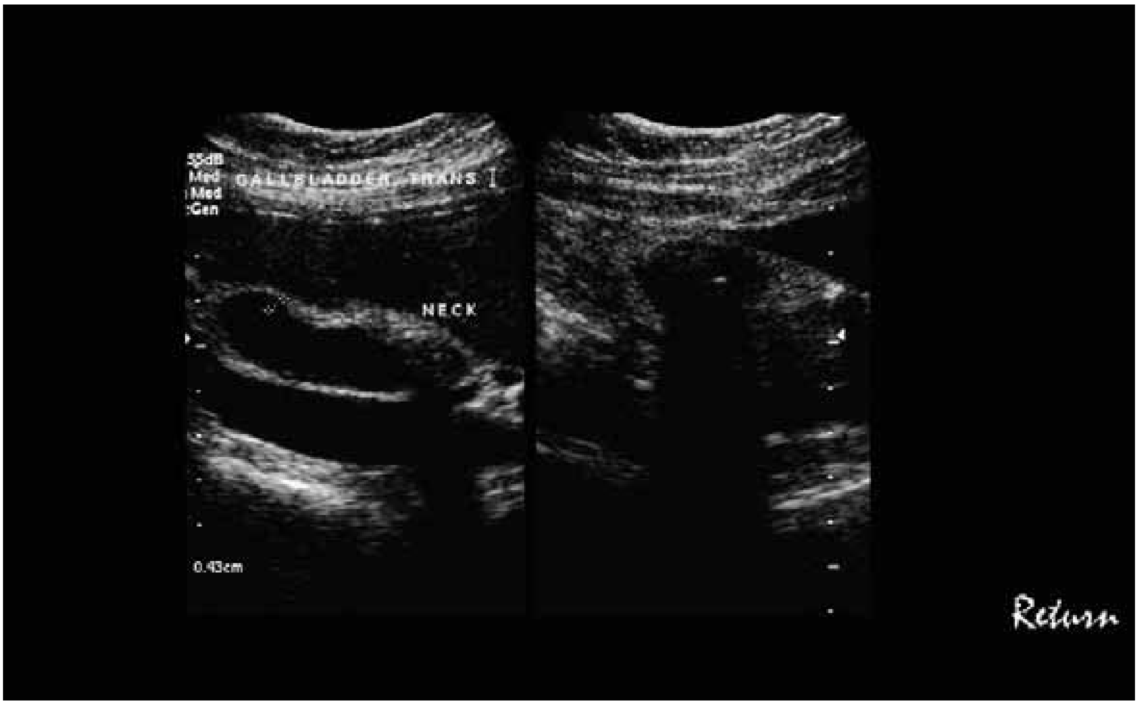


Figure 4.12 Enlarged image from Figure 4.10

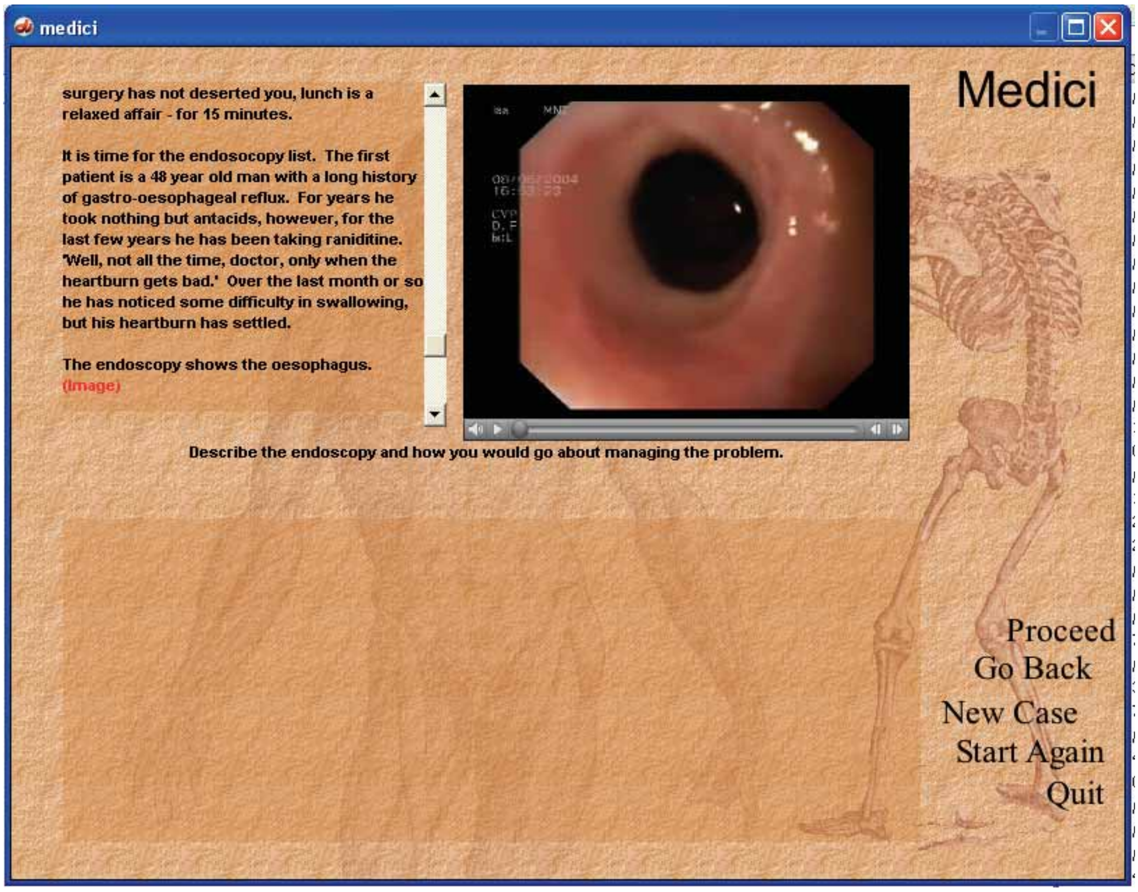


Figure 4.13 Video shown in a free response screen in the Windows version of *Medici*

At every step of the way through a case, feedback is provided about the choices made and also about choices that should have been made and were not (Figure 4.14).

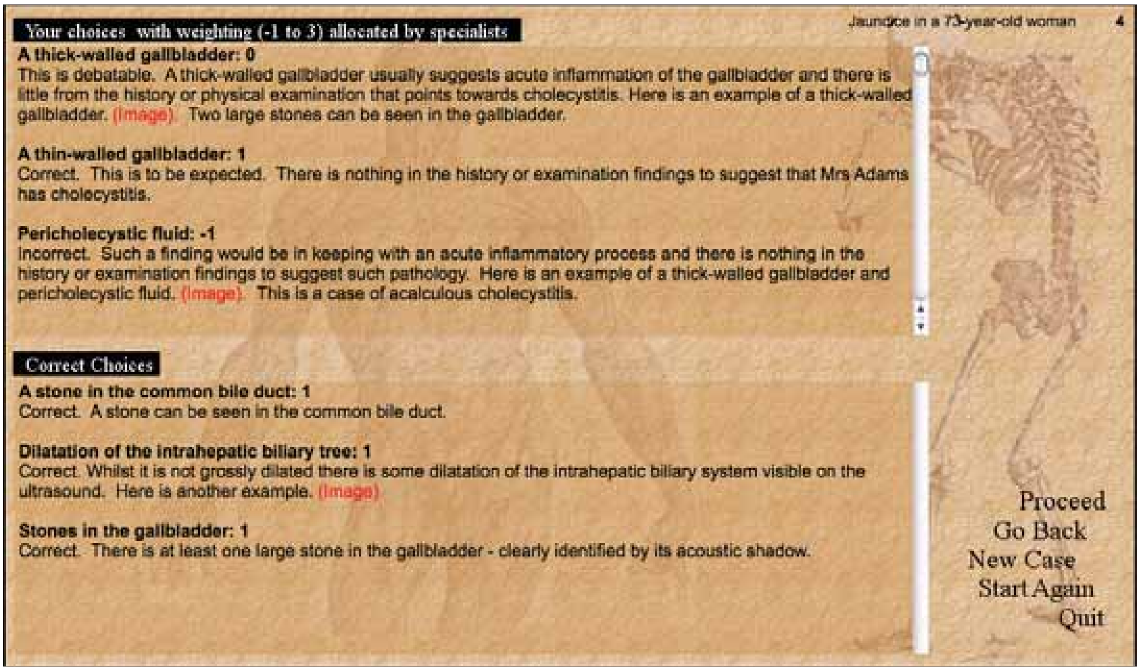


Figure 4.14 Feedback provided to user of a multiple selection screen

For free response screens, feedback is given in terms of an ideal answer and does not specifically address the input of the user (Figure 4.15). Automated scoring is possible for the free response questions, but has not been included in cases used for this thesis due to their complexity and difficulty in engaging clinical staff in their development, thus limiting the number of such cases existing.

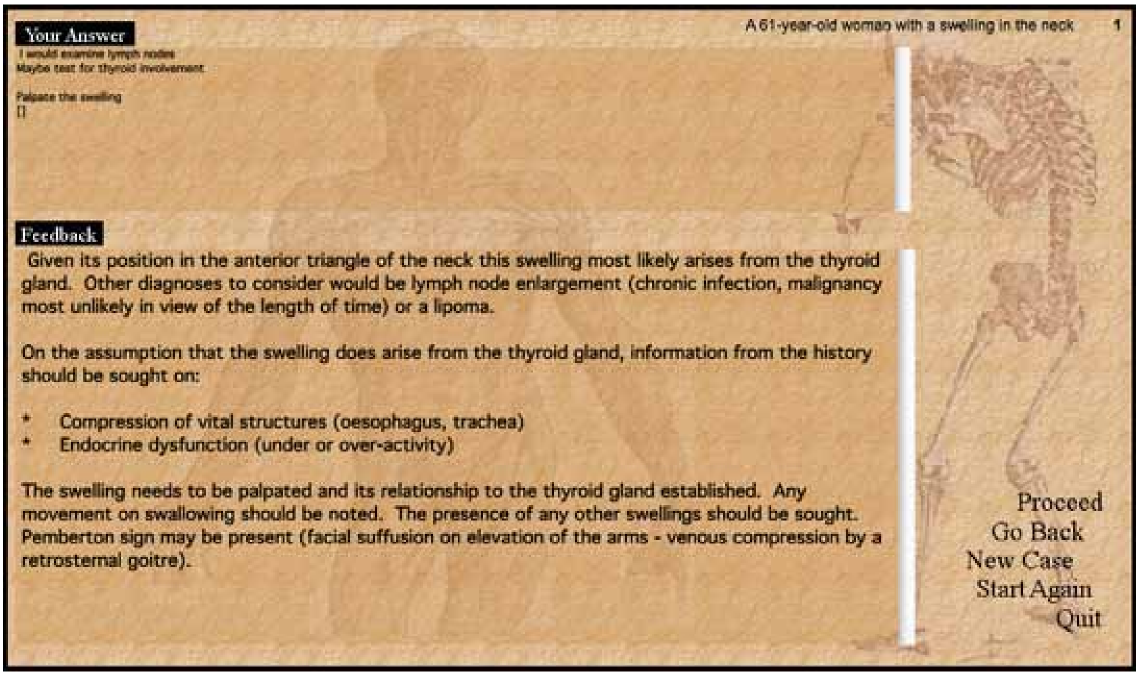


Figure 4.15 Feedback provided to a user of a free response screen

When a *Medici* case has been completed, a score or weighting is calculated as well as a summary screen, which summarises the material covered in the case, provides some guidance to students as to whether repeating the case may be worthwhile and often highlights useful references (Figure 4.16).

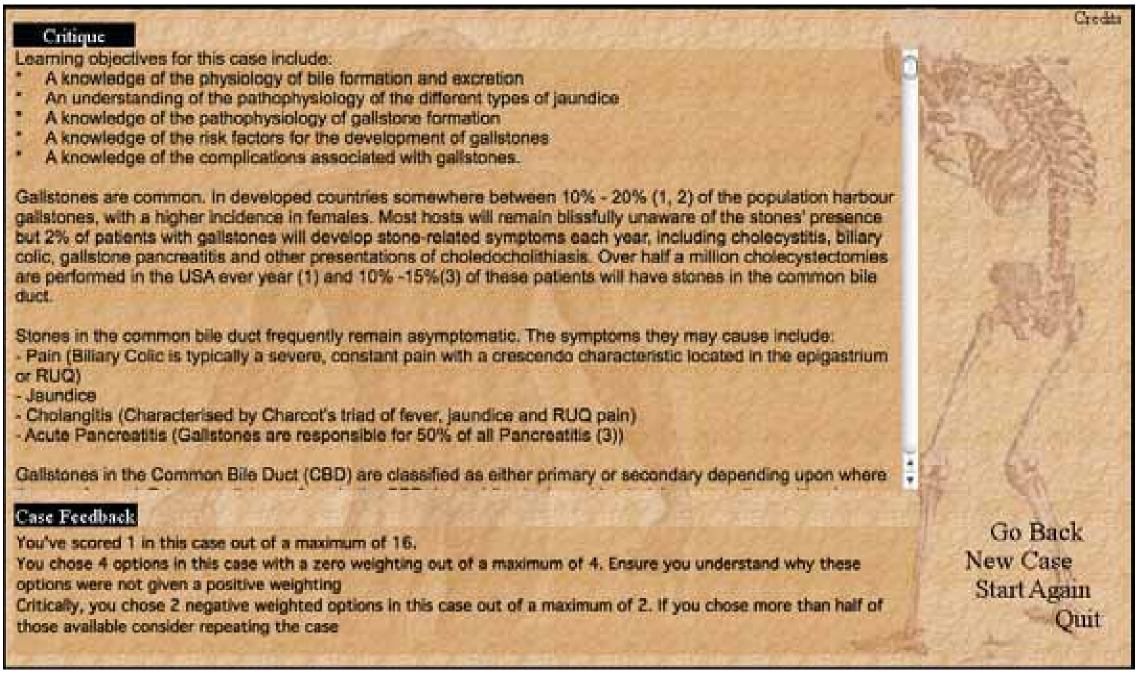


Figure 4.16 Summary or Critique screen

iPod Applications

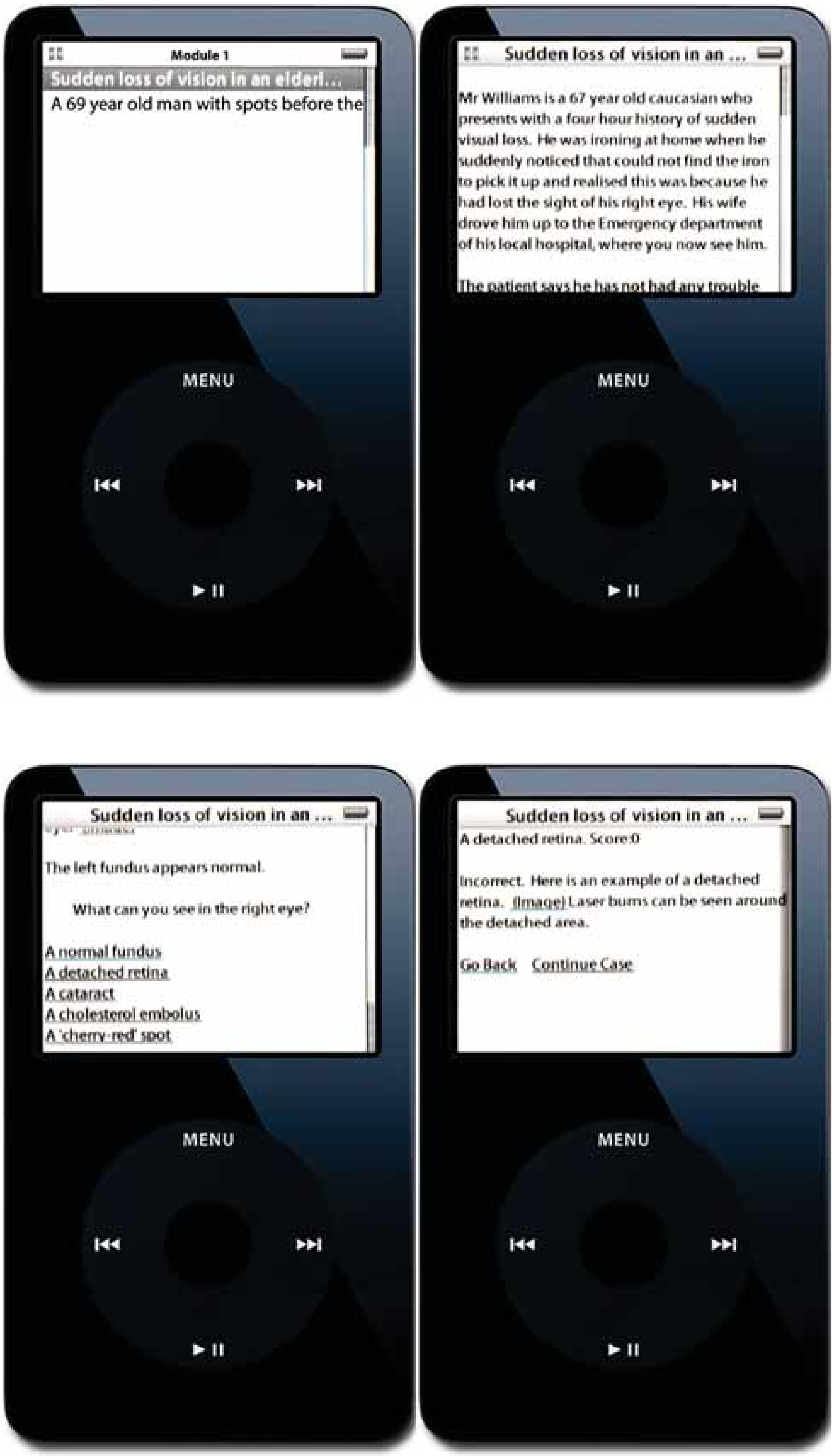
The arrival of new technology, particularly in the entertainment field, is often followed by efforts to leverage this technology for educational purposes. The introduction of records, radio, audio tape, video tape, CDs and laserdiscs have all been followed by educators taking up each of these tools and using them for their students' benefit (Cipriani, 1912; Devitt, et al., 1998; Mir, et al., 1986; Reed, 1941; Thomson & Bryce, 1987; Tomlinson, 1979). The introduction of MP3 players such as Apple's iPod, and Smartphones are proving to be no exception. Educational podcasts, which are audio recordings of items such as lectures, interviews and book readings, have been shown to be increasing rapidly (Cajiao, 2006; Nesbitt, 2005).

In the majority of cases, a podcast is essentially a passive learning experience focused on an audio facility alone. The video equivalent, the vodcast (video podcast), uses additional media and is therefore likely to be more versatile and of interest to the student but is still a passive learning tool. The main benefits of this type of learning material include the ease with which it can be created (Savel, Goldstein, Perencevich, & Angood, 2007), downloaded and used (Boulos, Maramba, & Wheeler, 2006), and its portable nature. This could be beneficial for students and clinicians separated from their peers by distance (Boulos, et al., 2006). It is possible to create medical curriculum content in a short space of time and have it downloaded by many people shortly afterward (Mathieu, 2007). This makes it feasible for students on public transport to choose to listen to, or watch educational material, in addition to viewing videos or listening to their favourite music, as can be done with the iPod *Medici* cases. *Medici* material created for CD-ROM can be transferred onto the iPod rapidly with no additional development time required.

It has been suggested that non-interactive audio and video are less effective educational tools than computer-assisted learning (Cohen, et al., 1981), where the latter has additional scope for interactivity. Some material is available for the iPod that provides a degree of interaction through links to an external website (Whitaker, 2010), but it might be considered more beneficial to provide a portable, self-contained, interactive experience. A version of *Medici* was developed for the iPod (Palmer & Devitt, 2010), specifically the 30GB and second generation nano versions, as they supported images and video (Apple Inc, 2007a) and had high quality displays. The same methodology will work on many of the earlier generations of iPods with screens, but without support for images and video. There are no current versions of Medici for the current generation of iPhones or the iPad, although these are planned.

There are some features in the formative style of case-based learning described, which cannot be transferred into the iPod. These include interactive scoring and the ability of students to select multiple-choices or enter free text. The little-used ‘Notes’ function on the iPod is the key to providing interactive learning material. Within the Notes folder, documents supporting the HTML coding system, the code behind the World Wide Web, can be run. This is used in train schedules (San Francisco Bay Area Rapid Transit District, 2007), and games (Malinche Entertainment, 2007) that are run on the iPod and this function forms the base for interactive learning on the device.

Figure 4.17 demonstrates the stages of a *Medici* case as realised on an iPod using part of an ophthalmology case in *Medici*. The educational content can be delivered to students in two ways. It can be collected together, compressed and be made available for downloading. Students would then need to load images and text files into their Notes folder and drop videos into their Movies library within *iTunes*. Alternatively students can download the material as a series of podcasts. Setting up a podcast server is a relatively straightforward task for those with IT skills, and the process has been described in detail in other forums (Apple Inc, 2007c).


Module 1
Sudden loss of vision in an elderl...
A 69 year old man with spots before the
MENU
Sudden loss of vision in an ...
Mr Williams is a 67 year old caucasian who presents with a four hour history of sudden visual loss. He was ironing at home when he suddenly noticed that could not find the iron to pick it up and realised this was because he had lost the sight of his right eye. His wife drove him up to the Emergency department of his local hospital, where you now see him.
The patient says he has not had any trouble
MENU
Sudden loss of vision in an ...
The left fundus appears normal.
What can you see in the right eye?
A normal fundus
A detached retina
A cataract
A cholesterol embolus
A 'cherry-red' spot
MENU
Sudden loss of vision in an ...
A detached retina. Score:0
Incorrect. Here is an example of a detached retina. (Image) Laser burns can be seen around the detached area.
Go Back Continue Case
MENU

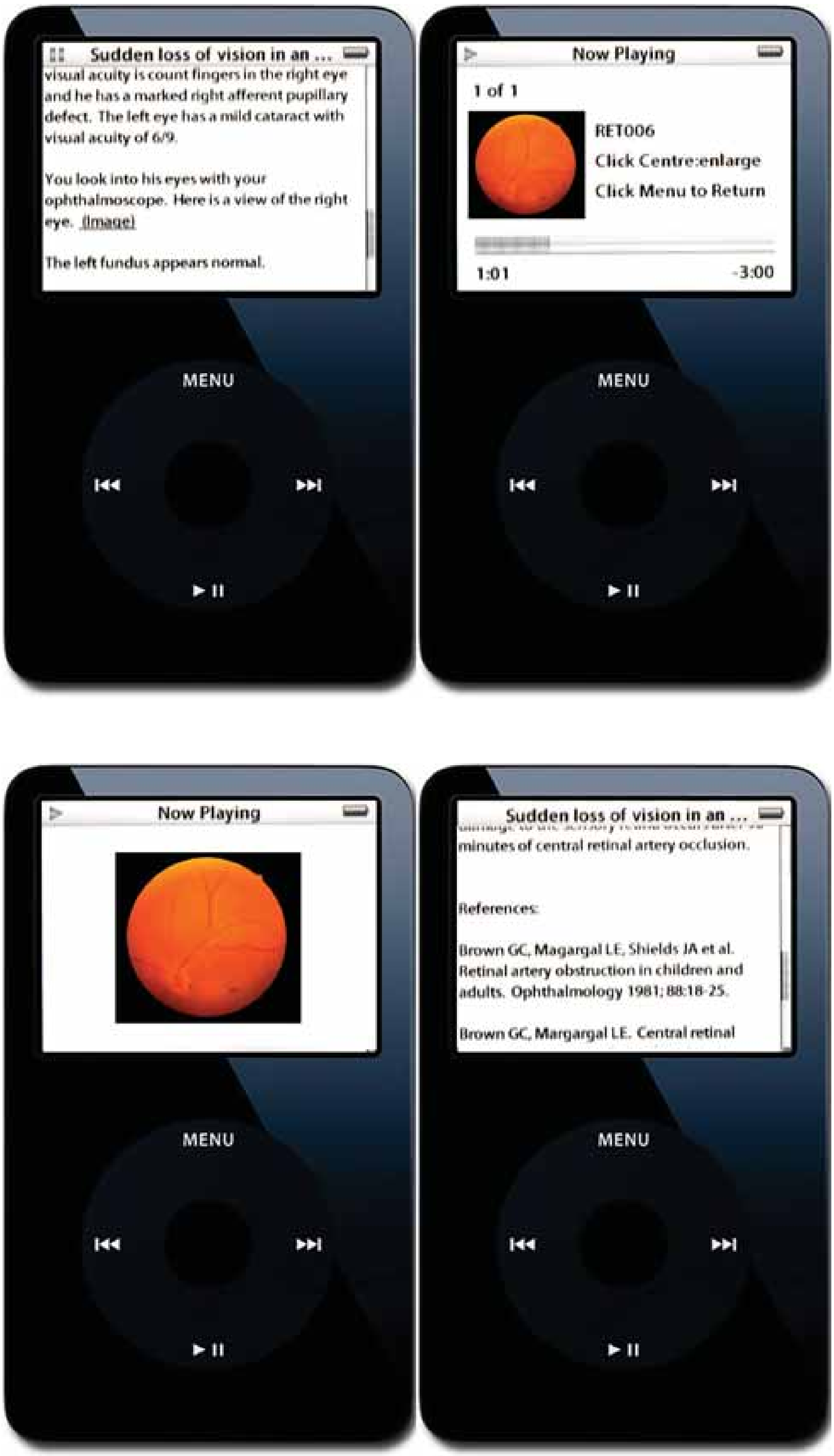


Figure 4.17 Stages of a Medici case on an iPod

iPods are not the only MP3 players. The iRiver *Clix* (iRiver), Sansa's *SanDisk* (sansa) and Microsoft's *Zune* (Microsoft Corporation) are alternative competitors for this market. Not all MP3 players have the capability to run HTML, and the methodology described may not be suitable for other MP3 players, but there are a number of PDAs and mobile phones that can run this content and they can retrieve it from a remote database wirelessly. This latter point is of great importance. The iPod is not an internet active device, so any medical content needs to be downloaded, and may therefore become obsolete without the student's awareness. A 'live' system, where content can be controlled and regularly updated has a great deal of appeal from a quality control perspective. The ability of some MP3 players and other devices to run *Flash* (Adobe, 2008b) applications greatly enhances the opportunities for meaningful interaction on a portable device and is worthy of future investigation. Additionally, the release of newer technologies, such as Apple's iPod Touch and iPhone (Apple Inc, 2008b) will have a large impact on this type of interactive content. At this stage, Medici cannot be used as an iPhone or iPad (Apple Inc, 2010b) application, but this is a logical extension to the work already completed.

Internet Applications

The schema used for *Medici* on the Internet is significantly different from the CD-based version shown as Figure 4.5. In this schema, *Medici* resides on a remote web server, where it is downloaded to a user's computer on request. The large files discussed earlier for the CD version are not downloaded, but a shell of the program, containing essential software code and media, significantly smaller than the 6MB required for the CD version is streamed to the users machine via their web browser. The core engine of *Medici* on the Internet is 192KB in size, thus not prohibiting access even at dial-up speeds. This engine contains no cases, images or video used in cases, but acts as the control system for the user, interacting with both the user and the repository of case information available, which is stored in a MySQL database (Sun Microsystems). When a user logs onto the website, *Medici* requests information on cases available to the user from the database. It then displays this information to the user. Every time the user chooses to run a case, makes a management decision within a case, or views an image or runs a video, *Medici* talks to the database, retrieves the appropriate information and displays it. Additionally, *Medici* monitors every decision made by the user within a case and this data is sent back to the server and stored within the database.

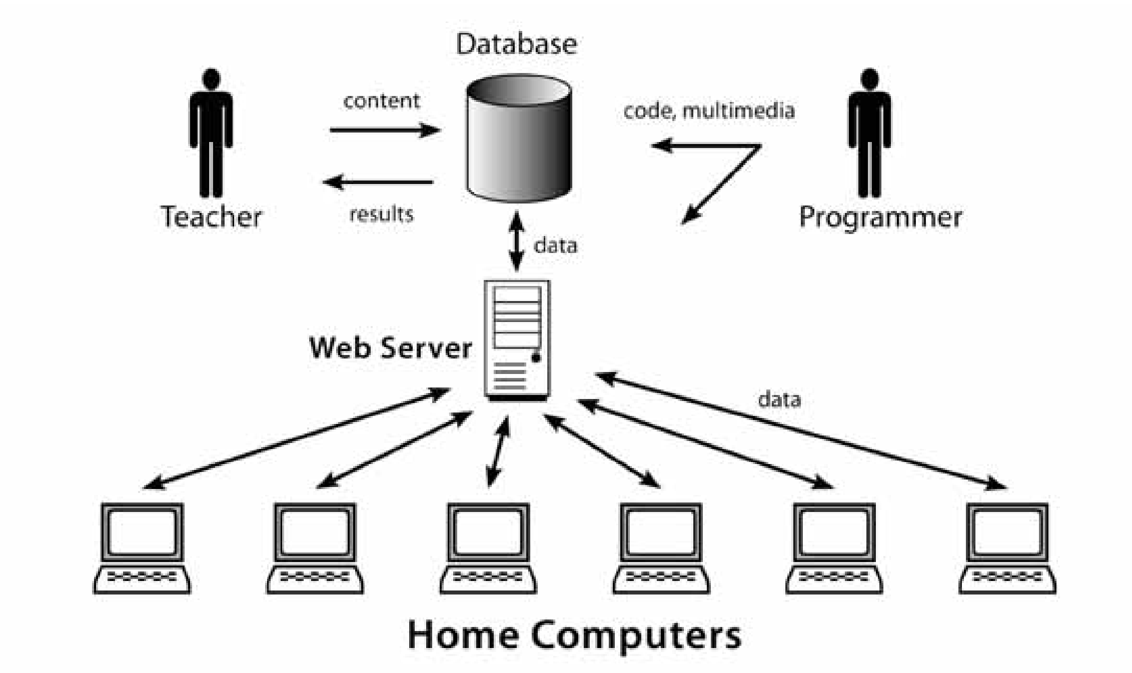


Figure 4.18 Schema illustrating function of *Medici* using a remote access methodology

Medici communicates with the database via a series of scripts; programs that act as a translator between the *Shockwave* program that is *Medici* and the database. This series of scripts handle everything from ensuring the appropriate cases are available to a logged in user, to asking the database to store information on a decision made during a case. There are ten scripts written in the programming language called PHP (The PHP Group), and totalling over 500 lines of code. Complemented by a dozen HTML web pages, they provide the programming and interface for the online version of *Medici*.

Although the CD version of *Medici* could make use of the space afforded by the medium, download time was a paramount design decision for the online version. Image sizes of 10–30k were of sufficient quality to show medically relevant detail and were not prohibitive to download for dial-up users. It was important to not overuse video, as a download of 2MB for a video could potentially take 11 minutes. Shorter segments of video, running at reduced frame rates and compressed as much as possible without compromising quality were utilised.

Access to the online version of *Medici* is via the *eMedici* website (Figure 4.19). The home page describes *Medici* and the necessary requirements to run the software. Students are required to log in, in order to have access to material designed for current learning situation (Figure 4.20). Students on their surgical attachment in fourth year, for example, have access to a different set of cases to first year students. The online version of *Medici* is similar to that for the CD version (Figures 4.6 to 4.16), although it is slightly smaller and simpler to reduce the download size (Figure 4.21).

Figure 4.19 Essential information for users of *Medici* on the *eMedici* website

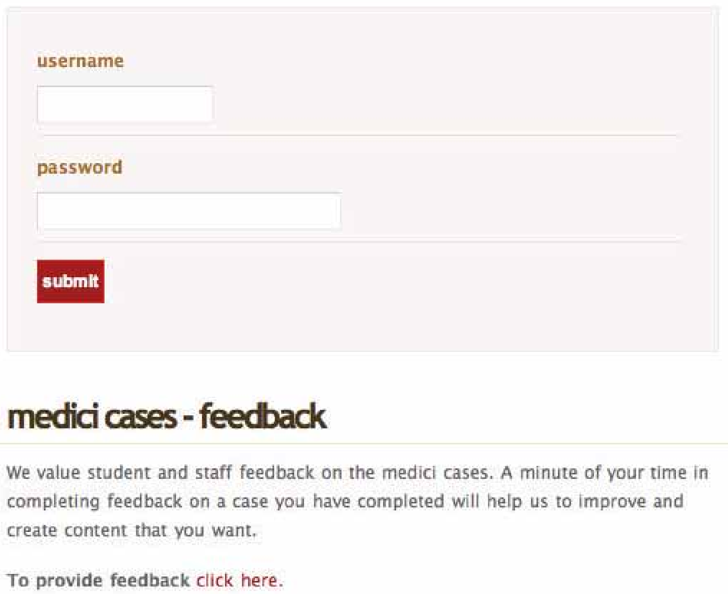


Figure 4.20 Login screen for users of *Medici* on the *eMedici* website

Figure 4.21 *Medici*. Online application written in *Macromedia Director*

Additional Motivation

In order to attract student attention to the website and encourage them to return regularly, additional daily features were added to the site. Milestone of the Day (Figure 4.22) provides a daily medical fact, often presented in a light hearted manner to quietly inform and entertain students. Image of the day (Figure 4.23) provides a new image daily, which students are able to try and identify. There is a large variety of clinical, operative and investigative images present, challenging students to develop the pattern recognition skills for common conditions they will require in their professional lives.

Medical Milestones

Important and occasionally humorous facts about our profession. A new milestone every single day!

Medical facts

In September 2004 Merck & Co announced that it was withdrawing Vioxx from the market. Up until that moment this COX-2 inhibitor had been marketed extremely successfully as an effective agent in the treatment of arthritis with miminal side-effects or complications, particularly gastrointestinal haemorrhage. New data from a clinical trial revealed that Vioxx was associated with an increased risk of myocardial infarction and stroke. A subsequent paper published in the Lancet stated that Vioxx could have caused as many as 140,000 cases of coronary artery disease in the USA since 1999. During that time there were more than 106 million prescriptions for Vioxx in the States.

Figure 4.22 Medical Milestones

Image of the Day

Answer the question about the image of the day. A brand new problem every single day!

Click on the image to open a larger version in a new window in your browser

What does this image show? Select your answer.

Hypopyon | Submit

Figure 4.23 Image of the Day

Developing Clinical Scenarios

The 400-plus cases currently available in *Medici* have been written using *Medici Author*, a program written using the *Director* software. *Medici Author* was written in conjunction with *Medici* to ensure that a programmer would not be necessary to develop clinical teaching scenarios. The writer of the case has full control over every aspect of the case, including its length, the style of question (or Screen), the number of choices students will be able to choose from and select, the nature of the scoring and the use of still images and videos (Figure 4.24).

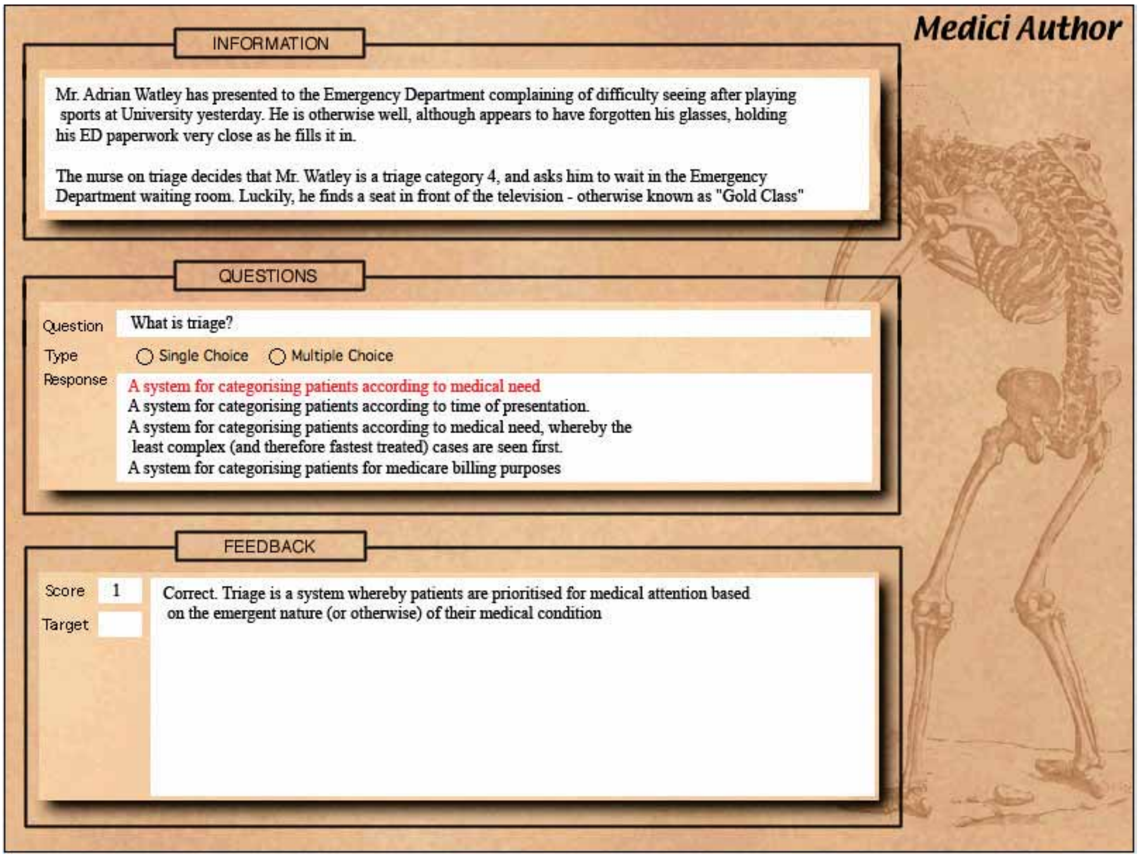


Figure 4.24 *Medici* Author. Image showing interface for development of single and multiple selection problems

The Author program is now integrated with the eMedici website to allow institutions to develop and manage their own content (Figure 4.25). Making the process of content development as simple as possible and with the software well supported encourages educators to develop their own cases to add to a global pool and will assist in ensuring that Medici and its content are sustainable resources for a number of years.

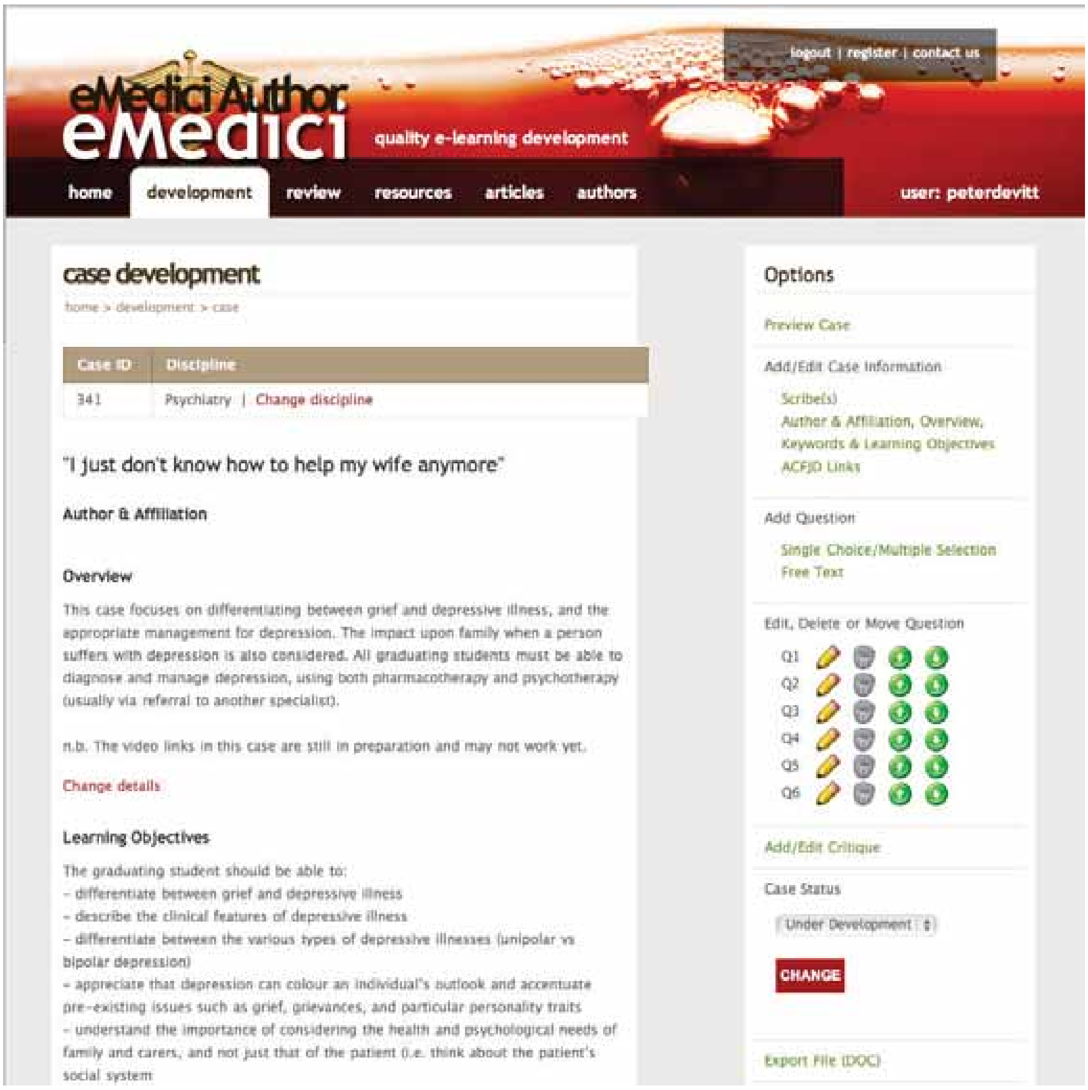


Figure 4.25 *Medici* Author. Image showing interface for development of scenarios on the eMedici web site

This chapter has summarised the construction of an online formative assessment program using criteria and arguments developed in Chapters Two and Three. The following chapters detail the use of *Medici* in learning situations, specifically the use of it in surgical attachments at the University of Adelaide over two years.

Chapter Five: Stage 1 Results

Introduction

Evaluating medical educational interventions can be a challenging task. Like a purely clinical trial, where it is sometimes difficult to isolate participants from treatments that may conflict or complement the trial treatment, it is also technically and ethically difficult to do the same in education (Kember, 2003). Modern education environments are often immersive; students are supported by a variety of educational resources, including multimedia learning programs, learning management systems and online discussion groups, as well as the more traditional clinical attachments, tutorials and lectures. Students in higher education are challenged in many ways in order to encourage learning and it can be difficult to isolate and measure the effect of one particular intervention and exclude other possible influences. Also, a diverse student group makes generalisation of findings difficult.

Without any formal evaluation, it is difficult to justify expenditure on a proposed educational intervention and to encourage other educators to adopt what is essentially a supposedly efficacious, unproven educational strategy. Educational faculties have considerable pressure applied to their budgets, with requests for staff, equipment and building works, and require more than unsubstantiated ideas to fund new and potentially expensive educational initiatives. Although decisions on education should ideally be focused on educational outcomes, and informed by advances in educational pedagogy, it is understandable that those in charge of allocating funds require some evidence to support a new teaching strategy or resource, before supplying development or maintenance funds and encouraging others to participate in and use the educational initiative. It is also reasonable for students to be informed that the style of learning they are undertaking is based on a solid foundation of educational theory and defined outcomes. It is therefore in the best interests of the students, the educator and their faculty to evaluate new educational resources.

Educational Research Methodologies

Research methodologies can be divided into three main categories, quantitative, qualitative and mixed method (Johnson & Onwuegbuzie, 2004). Quantitative research can be regarded as an approach that relies solely on the acquisition of quantitative data. Such research might include measurements of the reduction in morbidity of patients due to a new surgical intervention, or the reduction in road accidents due to the introduction of new safety standards. A qualitative approach to research is typically carried out using tools such as interviews, surveys, and observational data. Whilst this style of research might be considered more subjective than pure quantitative research, it is particularly appropriate for research in social areas. Examples of such research include the effects on children of relocating to another country and the development of strategies that could assist patients in dealing with the emotional stress of having been diagnosed with a serious illness. A mixed method approach to research utilises both quantitative and qualitative approaches and is particularly suitable when the research may be under highly controlled conditions, yet contain significant contextual and social components. Much educational research falls into this category. It is possible to develop and apply educational interventions to a group of students in a controlled and measurable way but students are human beings and, as such, their behaviour is often difficult to predict. An understanding of motivation as well as behaviour is required to fully grasp the impact of a new educational tool.

A Mixed Method Research Approach

The research questions *'Can the use of an online formative assessment tool improve student learning outcomes for medical students as measured by traditional examinations?'* and *'If the use of an online formative tool did improve student learning outcomes, what strategies will encourage its use?'* suggested that a mixed method approach was appropriate for the evaluation of Medici use. The first question was answered by measuring exam results but was also supported by qualitative data such as information on the other types of formative assessment that students encountered. The second question was to be informed by student survey data but also from information on when, and how often, students used the resource.

Quantitative research designs can consist of the pre-test -post-test control-group design, the post-test -only control-group design, the factorial design, the repeated measures design, and the factorial design. These represent strong experimental designs due to the compensation for any confounding factors (Johnson & Christensen, 2004, p. 281). The pre-test-post-test control group design represents the strongest option in terms of being able to generalise to the population. This particular design relies on randomised groups of research participants being selected, some of whom will form an experimental group and the others forming a control group. The control group is not exposed to the experimental conditions and thus provides a baseline for any changes in the experimental group. If the randomisation process is effective, participants in one group should be similar to those in the other group, reducing potentially confounding effects such as gender, culture and age. The next strongest experimental design is the post-test -only control-group design, which has randomized groups, including a control group, but does not provide a pre-test, instead relying on random allocation to ensure equivalence of the two groups at the beginning of the experiment.

It is possible to carry out educational research using one of the above experimental designs but there will still be confounding factors that may influence learning outcomes. These must be acknowledged but often cannot be factored into the research design. Furthermore, true control groups are difficult to create in the educational environment. Students discuss their work with each other and are likely to take advantage of whatever learning tools are available, by sharing resources such as notes, CDs, and even passwords to resources. This has the effect of influencing the control group and experimental group, often unbeknownst to the researcher (Lechner, 2001). Even if this type of interaction did not occur between groups, there are very real ethical concerns about intentionally denying some students access to study aids that may improve their learning outcomes.

For this project pre and post-test evaluations based on student examinations were carried out with randomized groups, including controls. Qualitative data was collected via surveys. Ethical issues were minimised by ensuring all students were given access to the material when it was available and prior to their examinations at the end of the year and the research was approved by the University of Adelaide’s Human Research Ethics Committee.

The evaluation of the *Medici* program took place over a two-year period from 2006 to 2007. Stage 1 of the evaluation was to introduce the program to fourth year medical students as soon as the program and content had been developed. This occurred mid-way through 2006. This allowed for the students in the first half of the year to act as a control group, as they had no access to *Medici*. The students in the second half of the year acted as the experimental group. Stage 2 of the evaluation was based on the outcomes of Stage 1, which are reported below. The results of stage 1 have been previously reported (Palmer & Devitt, 2008).

Stage 1 Evaluation: 2006

The surgical attachment in the fourth (first clinical) year of the Bachelor of Medicine and Bachelor of Surgery (MBBS) program at the University of Adelaide ran four times over a year for a period of nine weeks duration. During that time students were provided with formal tutorials, but primarily they were expected to be out working and studying in a clinical environment where the emphasis is on learning by participation with a major focus on self-directed study. Learning opportunities arose from contact with medical and allied staff and the patients under their care. There were other structured components to the MBBS curriculum, including a common lecture program, which was a weekly three-hour session designed to provide a theoretical background to many aspects of the clinical course. The content of these lectures was common to the entire MBBS program rather than any particular component of the course to which the students may have been attached at that time.

Students on the nine-week surgical attachment were expected to develop skills in clinical and cognitive competency and to demonstrate appropriate professional behaviour. These skills were assessed, based on the results of a mentor report, a written surgical case study and a written summative examination. At the beginning of the attachment, all students were assembled for a briefing to explain the nature of the attachment, the modes and opportunities for learning, their obligations as a student on a surgical clinic, what they could expect of the clinic and what the clinic expected of them. During the attachment, students received feedback on an informal basis from their supervisors and other clinical and academic staff. The students met again at the end of the attachment for a formal debriefing. Students were allocated places in some of the teaching hospitals in Adelaide and in rural attachments, which were up to 600 kilometres away from the major teaching hospitals in Adelaide, which are The Royal Adelaide Hospital, Modbury Hospital, Lyell McEwen Hospital and the Queen Elizabeth Hospital.

Medici was evaluated during the surgical attachments in the fourth year of the program, where students spent nine weeks studying General Surgery in a clinically focused environment. There were opportunities for all the students in the latter half of the course to use this educational initiative. *Medici* was used on a purely voluntary basis for 2006, to ensure equity and to test the effectiveness and uptake of a voluntary use of a recommended teaching resource. The *Medici* program was evaluated using a pre-test-post-test control-group design supported by qualitative data in the form of surveys. The development of *Medici* was scheduled to finish midway through 2006, thus allowing students in the first half of the year to act as a control group.

In 2006, the cognitive evaluation component of the assessment was extended with the addition of six modified essay questions to complement the 50 multiple-choice questions. For the first time in this surgical attachment, students were provided with a structured formative (self-directed) assessment process. This was undertaken in the setting of a formal evaluation of the initiative. Before commencement of the year, students were randomly allocated to one of four groups with stratification according to sex, origins and academic ability (Figure 5.1). Each group consisted of approximately 30 students. For logistic convenience the first group of the year (Group A) was deemed the control group with no formal provision of any additional learning materials. Groups B, C and D had learning materials provided in the form of clinical scenarios, and intended for use in the *Medici* program. Only Groups C and D had the opportunity to use the newly developed, online formative assessment material. For Group B these resources were printed and provided at the commencement of the attachment. For Groups C and D the material was provided online on a dedicated website (Palmer & Devitt, 2010), also from the beginning of the attachment. The information available in the case-scenarios had been available to students in previous years in the form of appropriate text references and support tutorials but this was the first year the material had been available in a scenario format.

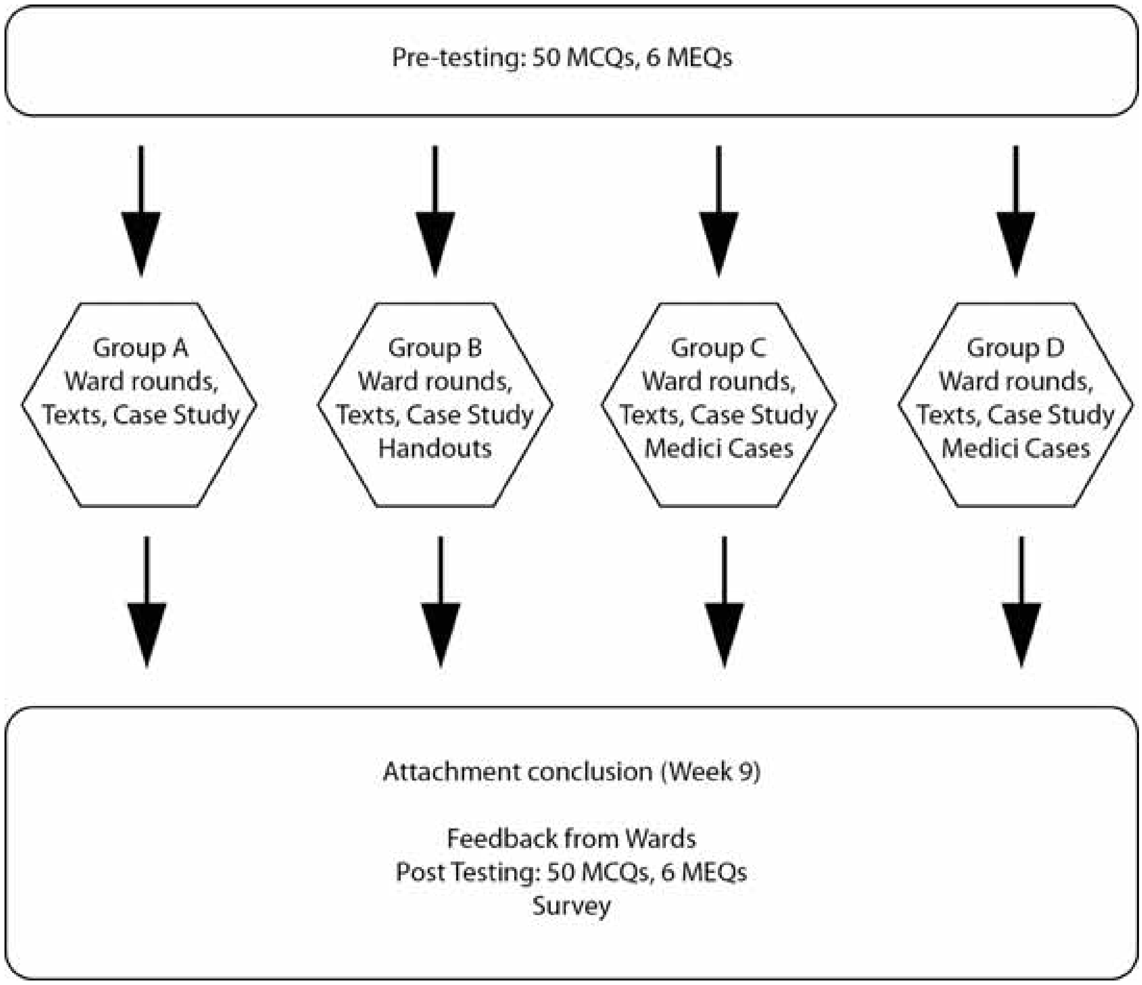


Figure 5.1 Surgical attachment configurations

The aim of this additional material was to stimulate student learning. At the beginning of each attachment, students were provided with the rationale to use either the handouts (Group B) or the online material (Groups C and D) as a formative learning tool. This included describing the problem solving nature of the formative content, which would promote better understanding of the topics covered. It was made clear by the senior surgeon organising the attachment that the topics addressed in the formative material would be assessed in the end-of-attachment exam. Students were encouraged to seek feedback on questions arising from the use of the materials from lecturers, clinicians and their supervisors. Furthermore, supervisors were requested to communicate any feedback to the attachment organisers. Texts recommended for student use included *Essential Surgery* (Burkitt, Quick, Reed, & Deakin, 2007), *Clinical Problems in General Medicine and Surgery* (Devitt, et al., 2003), and *Textbook of Surgery* (Tjandra, Clunie, Kaye, & Smith, 2006).

As part of the formal evaluation of the surgical attachment, and in particular to determine the value of the formative learning materials, the students were pre-tested at the beginning of the nine week period. This test was identical to the final test at the end of the attachment. A resource usage survey was conducted on completion of each attachment (Appendix 1) seeking students' opinions on various components of the course, including any educational activities in which they may have participated, particularly those providing formative assessment. The structure of the study is embedded in the attachment information displayed in Figure 5.1.

Medici cases were made available to all students in the latter part of the year. Groups A and B had access to *Medici* cases after their attachment had finished, and Groups C and D were given access to *Medici* during it. All students were able to access the *Medici* material as a preparative tool for examinations. For those students for whom *Medici* was available during their attachment, they were neither assessed on their use of the material during the study period nor on the scores that they may have achieved in the *Medici* cases. Student use of *Medici* was recorded in detail using an automated logging system. Data was written to a database and included the time the case was attempted, the students' responses to questions, the time spent working on each stage of the case, and the number of attempts required to complete it. Thus, each student's pathway through the *Medici* case screens was monitored for later analysis.

The desired outcome was that exposure to formative assessment material administered in two ways would have a positive effect on student performance in the end-of-attachment assessment. The number of students available in each group restricted the power of the study. Each group consisted of approximately 30 students. A satisfactory mark for the assessment was deemed to be two thirds of the possible responses i.e. a score of 20 out of 30 and the study looked for a 10% difference in outcomes between groups (i.e. a difference of 2 marks). Previous experience had indicated a standard deviation of approximately 3 was expected. This would provide a power for this study of 0.72 (assuming a 2-tailed type 1 error probability of 0.05).

Stage 1 Results

Overall, 129 students completed the pre-test and 136 completed the post-test. Those students who missed the pre-test were already based in rural attachments and could not attend the preliminary session or were absent through illness. The results of the pre- and post-tests are shown in Table 5.1. With a passing grade of 50%, only 30 of the 129 students who sat the pre-test would have passed the MCQ component and no student would have passed the MEQ test. Only six students would have passed the assessment overall. By the end of the attachment the student post-test scores in all groups and for both assessment types had increased significantly. In the final assessment, 126 students passed overall with 124 students passing the MCQ component and 115 passing the MEQ component.

Table 5.1 Test results for Groups A-D.

Group	**MCQ Pre-test** (max score = 46)	**MCQ Post-test** (max score = 46)	**MEQ Pre-test** (max score = 30)	**MEQ Post-test** (max score = 30)
A (n=30)	16.0±0.8	26.7±0.6*	5.9±0.6	17.4±0.8*
B (n=33)	19.4±0.6	27.3±0.7*	5.6±0.5	19.1±0.6*
C (n=33)	21.6±0.7	30.2±0.7*	7.6±0.6	18.8±0.7*
D (n=33)	20.8±0.7	28.8±0.8*	7.8±0.4	19.7±0.7*

Marks are given as score ± standard error, * = significant improvement from pre- to post-test ($p < 0.05$)

The assessment overall was deemed to be an effective tool for measurement. A calibration of the MCQ items with the results of the 136 (post-test) students, using Rasch analysis, yielded an item difficulty mean of 0.00 logits (standardising on item difficulty results in a mean item difficulty of zero) and standard deviation of 1.26 with an item estimate reliability of 0.97 meaning that the item difficulty estimates can be accepted with confidence (Figure 5.2). The mean ability estimate was 0.29 logits and a standard deviation of 0.48 with 0.52 person estimate reliability. There were no items or students with either a perfect or a zero score. The mean ability estimate for the pre-test calibration was -0.63 logits which confirmed the improved performance of 0.92 logits observed. A good match between candidate ability estimates and item difficulty estimates for the post-test on a common scale was observed (Figure 5.2) giving confidence in the examination.

```
--------------------------------------------------------------------------
        High ability               |        Difficult item
                                   |
                                   |    40     48
                                   |
                                   |
  3.0                              |
                                   |
                                   |
                                   |
                                   |
                                   |
                                   |    28
                                   |
                                   |
  2.0                              |
                                   |
                              X    |
                                   |    11     33
                              X    |
                             XX    |    42     50
                           XXXX    |    16     26     30     41
                              X    |
  1.0                        XX    |
                      XXXXXXXXX    |    18
                     XXXXXXXXXX    |     2     19
                   XXXXXXXXXXXX    |    14
                 XXXXXXXXXXXXXX    |     3      8
         XXXXXXXXXXXXXXXXXXXXXX    |
                  XXXXXXXXXXXXX    |    13     21     39
   .0                   XXXXXXX    |    20     25     36     46
                       XXXXXXXX    |    37     45     49
                   XXXXXXXXXXXX    |    43
                       XXXXXXXX    |    23     44
                           XXXX    |    12     35
                           XXXX    |    32
                                   |    47
                             XX    |    31     34
 -1.0                              |     1      5
                                   |     6      7     24
                                   |
                                   |     4     15     22     27     29
                                   |     9
                                   |
                                   |
                                   |    38
 -2.0                              |
                                   |    10     17
                                   |
                                   |
     Low ability                   |        Easy item
--------------------------------------------------------------------------
```

Figure 5.2 Student ability estimates and item difficulty estimates for the post-test.
Each X represents the ability estimate of one student

The four groups (A-D) differed in their pre-test scores. A significant difference in means was found between pre- and post-test results of each group (Table 5.1). Since differences were present in the pre-test results, the pre-test ability estimate was used as a covariate in an Analysis of Covariance (ANCOVA). In the ANCOVA the post-test ability estimate was used as the dependent variable, the group (A-D) was included as the fixed factor and the pre-test ability estimate was included as a covariate. The assumption of equality of variance was not violated (The significance value = 0.9 >> 0.05). There was no significant difference in the post-test ability estimates between the different groups after controlling for pre-test ability estimates prior to the interventions, nor was there any difference in the post-test ability estimates between the students who did the online assessment in Groups C and D and those who did not.

Resource Usage Survey Results

The focus of this survey was to attempt to determine what types of resources students used in their study time and the value they attached to these resources. The survey was completed by 125 students giving a response rate of 92%. Table 5.2 shows the median and the interquartile range of these responses, where 1 corresponded to more than 10 hours, 2 was equivalent to 5–10 hours, 3 was 2–5 hours, 4 was 1–2 hours and 5 was less than 1 hour. Thus a result of 4 [3-4] corresponded to students spending a median of 1-2 hours (4) using a particular resource with 50% of all students spending between one and five hours (3-4). Students reported spending the majority of their time reading the recommended texts, and on the wards (>10 hours, Table 5.2), a moderate amount of time attending lectures or tutorials (5-10 hours), and almost no time on non-journal internet-based resources or interactive learning aids (less than 2 hours).

When the students were asked about the value of each of the resources, they reported valuing texts and tutorials above all other items (Table 5.3). Despite not spending much time on interactive aids, students valued them as highly as paper-based journals, but below lectures, tutorials, ward rounds and web-based journal articles. The overall rating of 2 across all groups (Table 5.3) indicated students believed this type of resource was valuable.

The questionnaires were analysed using the Kruskal-Wallis test of ranks. In order to run *post hoc* contrasts, the data for each significant outcome was ranked and a one-way ANOVA fitted to this ranked data. *Post hoc* contrasts were performed using Fishers LSD. There was no significant difference between groups for the time reported being spent on resources and the value attributed to each resource apart from three instances.

Group A valued texts less than other groups, and valued interactive aids less than Group C. Group B was the group provided with the cases scenarios as handouts, and students in this group reported spending 2–5 hours on using the handouts (median 3, interquartile range 3–4) and found handouts to be valuable (median 2, interquartile range 1–2). Group C reported spending significantly less time on the wards than other groups and Groups C and D valued interactive aids more than Group B.

Table 5.2 Hours spent using various learning resources as reported by students at the end of their attachment.

	Hours spent on learning resources								
Group	Texts	Journal (Paper)	Journal (Web)	Ward	Lectures	Tutorials	Interactive aids	Internet (non-journal)	Other
A	1 [1–1]	4 [2–5]	3 [2–4]	1 [1–1]	2 [1–3]	1.5 [1–2]	4 [3–5]	4 [2–5]	3 [2–5]
B	1 [1–1]	4 [3–5]	2 [1–3]	1 [1–1]	2 [1–4]	2 [1–2]	5 [4–5]	3 [2–5]	4.5 [1–5]
C	1 [1–1]	4 [3–5]	3 [2–4]	1* [1–2]	2 [2–4]	2 [1–2]	4 [3–5]	4 [3–5]	4 [3–5]
D	1 [1–1]	4 [4–5]	3 [2–4]	1* [1–1]	2 [2–3]	1.5 [1–2]	5 [3–5]	3 [2–4]	5 [5–5]
All groups	1 [1–1]	4 [3–5]	3 [2–4]	1 [1–1]	2 [1–3]	2 [1–2]	4 [3–5]	3 [2–5]	5 [3–5]
Sig	ns	ns	ns	0.004	ns	ns	ns	ns	ns

Results are expressed as median [interquartile range]
(1 = more than 10 hours, 2 = 5–10 hours, 3 = 2–5 hours, 4 = 1–2 hours, 5 = less than 1 hour)
* Indicates significant difference between starred groups ($p < 0.05$)

Table 5.3 Perceived value of learning resources as reported by students at the end of their attachment. *

	Perceived value of resources								
Group	Texts	Journal (Paper)	Journal (Web)	Ward	Lectures	Tutorials	Interactive aids	Internet (non-journal)	Other
A	1* [1–1]	2 [2–3]	2 [1–2]	2 [1–2]	2 [2–2]	1 [1–1]	2 [2–3]	3 [2–3]	2 [2–4]
B	1 [1–1]	2 [2–3]	1.5 [1–2]	2 [1–2]	2 [2–2]	1 [1–2]	3* [2–3]	3 [2–3]	2.5 [1–3]
C	1* [1–1]	2 [2–3]	2 [1–2]	2 [2–2]	2 [1–2]	1 [1–1]	2* [1–3]	2 [2–3]	3 [2–3]
D	1 [1–1]	2 [2–3]	2 [1–2]	2 [1–3]	2 [1–2]	1 [1–2]	2 [2–3]	2 [1–3]	3 [3–4]
All groups	1 [1–1]	2 [2–3]	2 [1–2]	2 [1–2]	2 [1–2]	1 [1–2]	2 [2–3]	2 [2–3]	3 [2–3]
Sig	0.001	ns	ns	ns	ns	ns	0.008	ns	ns

Results are expressed as median [interquartile range]
(very valuable = 1, valuable = 2, little value = 3, no value = 4)
* Indicates significant difference between starred groups ($p < 0.05$)

Student Use of Medici

There were 12 cases made available for the students through Medici. Each case was of slightly different length, with a median length of 10.5 screens (range from 4–11 screens per case) (Table 5.4). Overall, students completed 7000 screens, with 30% of all activity occurring after all surgical attachments had been completed. As Groups A and B did not gain access to the material until after the attachments were complete, none of the activity reported for those groups occurred during their attachment.

Table 5.4 Details of cases made available to students

Case ID	Name	Number of Screens
258	Physical signs in general surgery	11
210	Breast disease	11
227	General surgery MCQs – 1	11
251	General surgery MCQs – 2	11
1	Groin hernia	5
2	Acute appendicitis	4
3	Management of hernias	6
4	Management of small bowel obstruction	11
5	Aetiology of neck swellings/management of thyroid lesions	7
6	Goitre	10
7	Chronic arterial occlusion in the leg	9
8	Management of breast cancer	11

Table 5.5 shows the level of activity for each group. Even though Groups C and D had access to *Medici* during their attachment, only 30% of Group C and 20% of Group D made use of the resource. Groups A and B made little use of *Medici*, when they had access after completion of their surgical attachment. Students in Groups A and B spent a little over 20 minutes, attempting 2–4 cases. Group C made the most use of the resource, with a median time spent on it of one and a half hours and attempting the majority of cases (median 11, Table 5.5). Table 5.6 shows the number of students who failed to proceed with a case, once they had commenced it. For Groups C and D this is a low number, but Groups A and B failed to progress in a large number of cases they had begun.

Table 5.5 Student usage of the *Medici* interactive online resource from beginning of attachment to the end of the academic year

Group	**No students attempting cases** (total no. students)	**Number of cases attempted Average, Median [interquartile range]**	**Number of screens** (median of 10 screens per case) **Average, Median [interquartile range]**	**Time spent** (minutes) **Average, Median [interquartile range]**
A	4 (34)	3 2 [1.5–5]	38 40 [12.5–63]	21 28 [9–30]
B	2 (33)	3.5 3.5 [3–4]	37 37 [14–60]	25 26 [9–42]
C	11 (28)	8.6 11 [6–12]	201 129 [59–240]	118 90 [20–143]
D	6 (30)	6.5 7 [2–10.5]	87 82.5 [17–152]	69 53 [25–140]

Results are expressed as average, median [interquartile range]

Table 5.6 Attempts at cases failing to progress past the first screen

Group	Left after 1st screen	Percentage of all attempts
A	5	25%
B	2	20%
C	8	3%
D	2	3.4%

Data illustrating student behaviour by group and by case is provided as Table 5.7. Note that students were able to attempt cases multiple times. Students from Groups A and B chose to complete only half of the available material, whereas Groups C and D attempted all of the available cases. Students in Groups A, B and D rarely attempted a case more than once, whereas students from Group C made 2–3 attempts at each case. Not only did students progress past the first screen of a case (Table 5.6), but a large proportion of students completed the cases. Group C students completed more than three quarters of their cases, whilst Groups B and D completed more than half. Case 258 had a low completion rate in all groups.

Table 5.7 Statistics for the use of *Medici* cases by group and case

CaseID	**Group**	**Number of students attempting case** (number completed)	**Total Attempts at this case** (number completed) Students can attempt cases more than once	**Percentage of all cases completed**
1	A	2 (0)	2 (0)	0
8	A	2 (1)	2 (1)	50
210	A	1 (0)	1 (0)	0
227	A	3 (2)	7 (4)	57
251	A	1 (1)	2 (2)	100
258	A	4 (1)	6 (1)	17
	A	**13 (5)**	**20 (8)**	**40**
1	B	1 (1)	1 (1)	100
2	B	1 (0)	1 (0)	0
8	B	1 (1)	4 (4)	100
210	B	1 (0)	1 (0)	0
227	B	1 (0)	1 (0)	0
258	B	2 (1)	2 (1)	50
	B	**7 (3)**	**10 (6)**	**60**
1	C	8 (7)	22 (16)	73
2	C	8 (6)	20 (16)	80
3	C	7 (6)	18 (16)	89
4	C	9 (7)	21 (16)	76
5	C	8 (6)	19 (16)	84
6	C	7 (6)	15 (11)	73
7	C	8 (6)	21 (14)	67
8	C	6 (6)	12 (11)	92
210	C	9 (9)	24 (20)	83
227	C	10 (9)	26 (20)	77
251	C	9 (7)	27 (19)	70
258	C	9 (6)	26 (15)	58
	C	**98 (71)**	**251 (190)**	**76**
1	D	4 (3)	6 (4)	67
2	D	3 (1)	3 (1)	33
3	D	4 (3)	5 (3)	60
4	D	4 (2)	4 (2)	50
5	D	2 (2)	3 (3)	100
6	D	2 (2)	3 (3)	100
7	D	2 (1)	2 (1)	50
8	D	2 (1)	3 (2)	67
210	D	4 (4_	5 (5)	100
227	D	4 (3)	8 (4)	50
251	D	4 (2)	5 (2)	40
258	D	5 (4)	12 (4)	33
	D	**40 (28)**	**59 (34)**	**58**
Overall		**158 (107)**	**340 (238)**	**70**

Discussion of Stage 1 Results

The students' pre-test scores were uniformly poor. This was not surprising as knowledge in surgery was expected to be limited at the beginning of the attachment. In particular, the first group of the year, which had had no previous clinical attachments and was comprised of students returning from several months break, had a lower pre-test score than other groups. Students in latter groups had the advantage of gaining experience in other clinical attachments before sitting the test.

Although students improved their knowledge and understanding during the nine-week surgical attachment, any advice and help they were offered with regard to self-directed study and formative assessment did not appear to produce any variation in their improvement in cognitive skills or change in study habits. Despite clear guidance at the beginning of the course on the goals of the attachment and how students might help themselves with their learning, little attention appeared to have been paid to this advice. The intention of the course was that students should study as they progressed and, for the appropriate groups, were given guidelines of when and what to study. The intention of the written and computer-based material was that students would be able to see what standards were expected and gain feedback on their individual performance. Teachers would also benefit from examining the performance of students, especially in the online environment, where it would be possible to examine misconceptions and act on them or at least provide prompt feedback. This fits with the generally accepted concept of 'formative assessment' as discussed in Chapter Two.

The survey results from Group C showed small differences in attitude and reported behaviour compared to other groups, but these results were not completely repeated in Group D, making it difficult to generalise. Although Groups C and D did make use of the online material during their attachments, the amount of use was limited. Students failed to utilise the technology as much as had been anticipated. The highest degree of usage came from Group C, yet only 30% of this group attempted any cases. Technically, there were no problems with the *Medici* program and those students who used it tended to complete the cases they began, but the overall result, with no discernable difference in the outcomes of the groups indicated that this method of using *Medici* was not effective.

It was unclear why students failed to utilise the online material. Students may have been using *Medici* out of curiosity, and the cases failed to keep their interest. The cases, on average were 9 screens long and it is possible that this length was considered to be too long by the stduents. Case 258 was the least used case, and in addition to being a long case, it was also the last case in a long list of cases. It is possible that position in the list was a factor in students choosing to attempt the case.

It is also possible that the students felt that they did not need to use *Medici* and they had faith in their own abilities to pass their examination without using it. The majority of students passing ratified this decision to some extent, and although simply passing may not be considered by many to be a sufficient goal for aspiring doctors, students do need to prioritise tasks. In this case it appears use of *Medici* was a low priority especially as it was not directly linked to a summative assessment. Although the use of *Medici* was heavily promoted and regularly encouraged by the lecturer in charge to both students and staff this appears to have had little effect, nor was there any evidence to suggest to students that the use of this formative assessment tool would improve their learning outcomes.

Whilst it could be argued that the aim of formative assessment is not so much to raise the standards of attainment (Black & Wiliam, 1998) but to foster the spirit of learning, it appeared clear that neither goal was achieved in this case by the formative material provided. On the other hand, once students did begin making use of the formative online resource, they did so frequently. It may be that the biggest issue involved in encouraging students to use this type of formative assessment is by discovering a method to encourage the students to begin using it.

The key motivating factors associated with assessments are the perceived relevance of the assessment, the content of the assessment, the enthusiasm of lecturers and group influences (Seale, Chapman, & Davey, 2000). The first two items are related directly to strategic considerations: ‘Can I learn what I need to pass or be good at what I want to do?’ The last two items are external influences and can be controlled to some extent but may be difficult to implement in environments where the teacher is distant from the student. Such situations include the online realm or where the teachers are busy clinicians and may not have the time to enthuse students or build a supportive group structure. This appears to have been reflected in the study, where students apparently placed value on the material provided for them, but in reality made limited use of it.

As well as the methods utilised by the teacher to facilitate learning, there are other factors affecting student motivation, including student goals and interests, creativity and the willingness to learn (Harlen & Crick, 2003; Seale, et al., 2000). Extrinsic motivation is more difficult to apply as some students are easily distracted, tend to take short-cuts, lose interest as they proceed through a course (Crooks, 1988) or adopt defensive approaches where they may reduce effort to protect self-worth (Pintrich & Zusho, 2007). A substantial proportion of students have part-time employment (Krause, et al., 2005) and they may focus on this rather than clinical activities (Crooks, 1988) which may play a part in general exhaustion of medical students reported worldwide (Dahlin, Joneborg, & Runeson, 2007; Dahlin & Runeson, 2007). If students have other concerns and are not motivated to work on clinical activities, the question remains, what would motivate them to work on an online formative assessment tool?

This component of the study was undertaken on the premise that sufficient motivation would be provided to the students by stressing the importance of the formative content and linking the formative assessment material to the learning objectives of the course. Regular feedback, likely to improve students' work ethic (Gibbs & Simpson, 2004), was provided to the students by the formative tool and added incentive was provided by informing students that this content was examinable.

The nature of an assessment can be a key indicator of the effort students will put into assessment tasks (Ramsden, 1997), but students are capable of manipulating their study time to focus on examinations at the expense of understanding subject matter (Gibbs & Simpson, 2004). That strategy appears to have been borne out in this study: where students did appear to use *Medici*, the use was close to the time of the summative examination. In addition, there is doubt as to whether any deep learning was occurring. Deep learning (Marton & Saljo, 1976) is one of the goals of any course and '…if students perceive a need to understand the material in order to successfully negotiate the assessment task, they will engage in deep learning' (MacLellan, 2001, p. 307). This was not reflected in this study where much of the use of *Medici* for exam preparation appeared to be superficial, as judged by the failure of students in Groups A and B to complete many of the individual case studies (Table 5.6). This could be related directly to the perceived relevance of the content to the summative assessment, but the material in the cases was relevant and examinable and the students were aware of that fact.

Many educators will appreciate that an important goal is learning itself for self-improvement and life long learning has been identified as an important element in being a professional clinician (Hojat, Veloski, & Gonnella, 2009). At the University of Adelaide this is reflected by graduate attributes (University of Adelaide, 2009) stating

> 'Being thoroughly versed in the skills and application of adult learning.' and
> 'Deriving enjoyment for the process of learning and the pursuit of knowledge and understanding (where knowledge is defined as information that can be used effectively in familiar and unknown situations)'.

These attributes are particularly important in Medicine where much professional interest is focused on continuing medical education and credentialing, but these values are often difficult to appreciate at the student level, when the barrier of final examinations looms large. This type of behaviour has been observed in medical education, where students learn those elements of the curriculum that are known to be directly assessed and they are more concerned about grades or passing an examination than about using assessments as a learning experience (Norman & Schmidt, 1992), i.e. the overarching goal is to 'know what they need to know' for the examination rather than to improve overall competence. It is this high stakes process that is often considered to have a negative impact on formative assessment (Harlen & Crick, 2003).

A number of criteria were defined to guide assessment practices (Gibbs & Simpson, 2004) as reported in Chapter Two (Table 2.1). Although the written case notes had the potential to meet these criteria, the *Medici* online program dealt with many of these explicitly. The program content provided to students in this study met many of these suggested criteria, including allocating sufficient tasks to utilise available study time, engaging the students in an appropriate activity and providing instant, relevant and complete feedback in sufficient detail, which was focused on student learning. One of the criteria relies on communicating clear and high expectations to the students. Although it was believed that this was done via assessment notes and verbal communication, it may be that the students failed to grasp this fact or that there was a failure to reinforce this message.

The production of educational material for student use is a costly and time-consuming exercise. If educators are required to prepare formative learning materials there must be a clear indication that the effort is worthwhile. It is evident from this part of the study that the strategies and materials provided to students failed to motivate them and make any meaningful difference to their ability to pass a standard summative assessment, thus making the process a failure from the point of view of both parties.

The following chapter discusses the strategy put into place to improve student engagement and motivation with the online formative assessment resource and then reports on the outcomes of the implementation of that strategy.

Chapter Six: Stage 2 Results

Introduction

A review of the 2006 Stage 1 results (Chapter Five) suggested that while the online learning tool, *Medici*, and its content interested those students who used it, it may not have been sufficiently aligned with summative tasks allocated to the students and there was insufficient stimulus to promote its use for self-directed study amongst students as a whole. The strategies implemented were inadequate to motivate students to participate in the formative assessment. It was decided by the researcher to encourage the use of the online *Medici* cases by creating a new learning structure in which they would be used. More case scenarios were produced and were made available on a regular basis throughout the nine-week attachment. The content of these cases were closely linked to the learning objectives provided to all students on commencement of the attachment, as were the cases in Stage 1, but in addition they were tightly linked to a summative assessment. It was decided that a weekly ten question 'mini-test' would be added to the attachment and become a compulsory part of the process. This was despite concerns that the students could be 'over-assessed' (Uden & Beaumont, 2005) and that this mini-test would not represent a true 'achievement' variable' (Sadler, 2007). This was designed to not be particularly onerous and was based on the material produced for that week. Students could complete the test in their own time and at their own pace.

Figure 6.1 illustrates the way in which the *Medici* cases were utilised. Each week for six weeks, students were provided with one module containing an assessment case and six or seven support cases. The assessment case was compulsory and in total they counted for 10% of the grade allocated to students in the surgical attachment. A 'pass' mark could be easily obtained from the assessment case if the other cases in the module were completed. Additional or 'incidental' cases (Module 0 cases in Figure 6.1) were available for use at anytime, but the assessment modules were released on a weekly basis. The last module was due to be completed at least two weeks before the end of attachment summative assessment. Table 6.1 shows the details of the cases in each module. Each case was unique and defined by a unique ID calculated from the order in which the cases had been written. Overall, students had 62 cases to study, although only the six short assessment cases were compulsory.

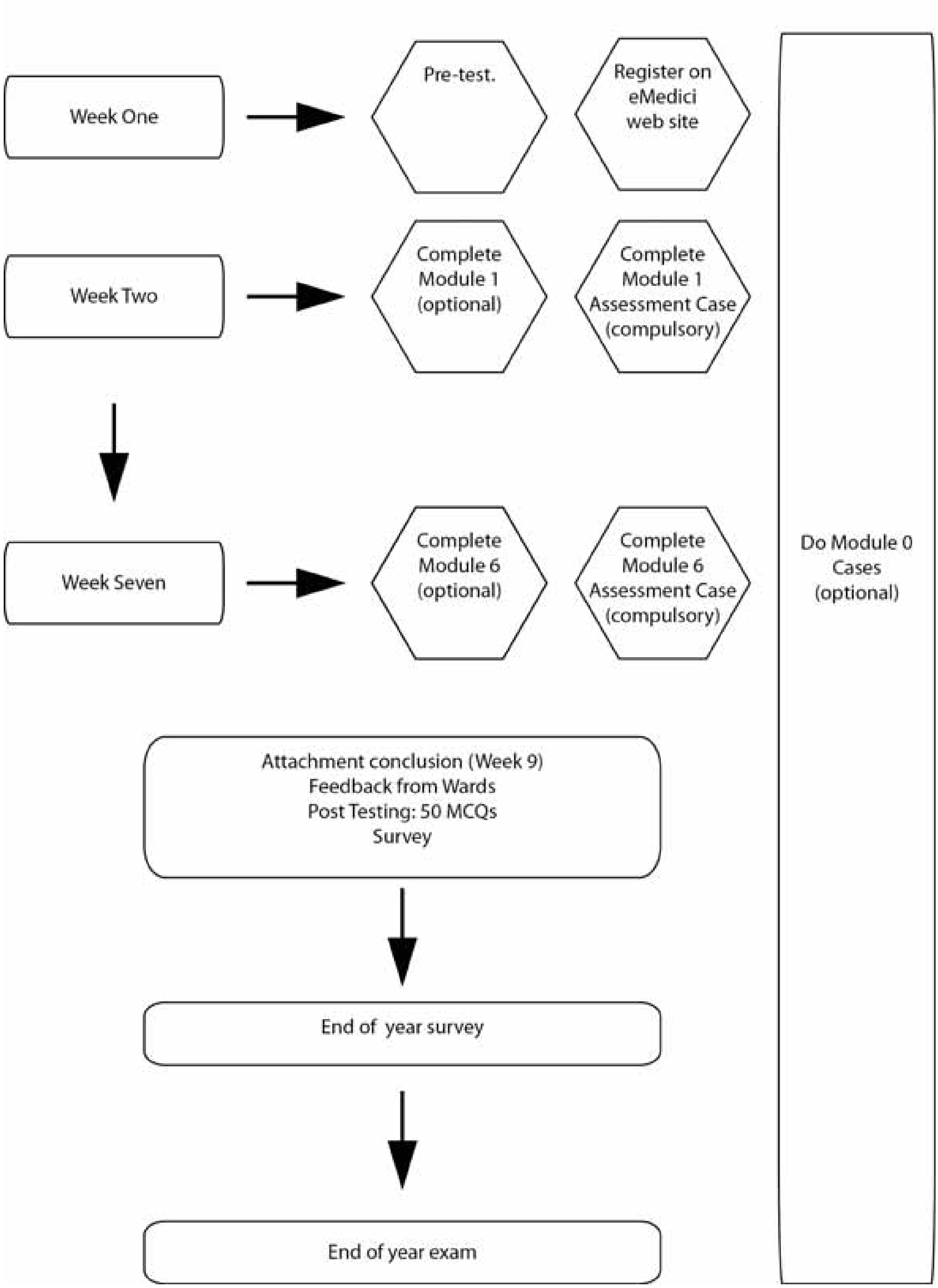


Figure 6.1 Application of formative assessment modules to surgical attachment

Table 6.1 Details of cases provided for surgical attachment students 2007
Where the case name is non-descriptive, a brief description of the content is provided in brackets

Module 0: Incidental cases

CaseID	Case name	Number of Screens
62	Do you have permission? (acute abdomen)	5
1	A 67 year old woman with dark urine (jaundice)	7
2	The value of diagnostic tests part 1	7
3	Is that blood or shiraz down your shirtfront? (upper GI bleeding)	10
4	Pain under the jaw in a grey nomad (neck swelling)	4
5	or are you just happy to see me? (hernia)	5
6	What is the risk? (informed consent)	4
7	One thing leads to another (acute abdomen)	11
8	The saga continues (large bowel malignancy)	10
9	A woman with weight loss and palpitations	13
10	Drunk and belligerent (penetrating wound)	7
11	A 34 year old man with a burn injury	10
12	Painless haematuria in a 72 year old man	7
13	A lesion on the back of the hand	6
14	Increasing tiredness in a 45-year-old woman	5
15	A bikie with a lacerated knee	4
16	A 68 year old man with difficulty in swallowing	6
17	A 73 year old woman with difficulty in swallowing	8

Modules 1-6: Support cases + Assessment cases

Module 1

18	A scrotal swelling in a 72 year old man	5
19	A 65-year-old man with a swelling in the scrotum	5
20	It doesn't feel quite right doctor (testicular swelling)	5
21	A 22-year-old man with a scrotal swelling	5
22	Acute abdominal pain in a 17 year old woman	6
23	A young woman with acute abdominal pain	4
24	I Claudius (atherosclerosis)	9
25	A 67-year-old man with pain in the right leg	8
26	Week 1 Home Unit Assessment	11

Module 2

27	Jaundice in a 73-year-old woman	5
28	A 62-year-old woman with jaundice	5
29	A breast lump in a young woman	10
30	A young woman with a breast lump	5
31	A 69-year-old man with blurring of vision	5
32	Transient weakness in a 74-year-old man	5
33	Week 2 Home Unit Assessment	11

Module 3

34	An elderly woman with a swelling in the groin	6
35	A 72-year-old woman with a lump in the groin	5
36	Intermittent abdominal pain in an elderly woman	6
37	A 77-year-old woman with abdominal pain and vomiting	5
38	A 30 year old man with a lump in the neck	10
39	A 61-year-old woman with a swelling in the neck	5
40	Week 3 Home Unit Assessment	11

Module 4

41	Acute epigastric pain in a middle-aged man	7
42	A 72-year-old man with upper abdominal pain	5
43	A 50-year-old man with a foot ulcer	5
44	A 55-year-old man with a leg ulcer	5
45	A young man with a swelling in the neck	6
46	A 68-year-old woman with a lump in the neck	6
47	Week 4 Home Unit Assessment	11

Module 5

48	Blood on the Floor (gastrointestinal haemorrhage)	5
49	A 36-year-old man with a haematemesis	5
50	A 32 year old woman with a palpable breast lump	6
51	A 26-year-old woman with a breast lump	5
52	A 55 year old woman with a pigmented skin lesion	8
53	A 70-year-old woman with a lip lesion	4
54	Week 5 Home Unit Assessment	11

Module 6

55	Rectal bleeding in a 58 year old man	7
56	A 69-year-old woman with rectal bleeding	4
57	Anorexia and weight loss in a 65 year old woman	8
58	A 41-year-old woman with anorexia and weight loss	7
59	Leg cramps in a 64 year old man	4
60	A 58-year-old man with pain in the right leg	4
61	Week 6 Home Unit Assessment	11

The aims of structuring the formative assessment material in this way were to ensure that:

- Students did not feel unduly pressured about other deadlines during the formative assessment period. The two-week gap before the end of attachment examination was designed so that students would focus on only one assessment at a time;
- Students could feel confident in their ability to complete the weekly assessment provided they had studied the accompanying clinical scenarios in that week's module; and

- The time required to work on the formative assessment exercises was not excessive. Six cases a week was anticipated to take 60-90 minutes to complete in total.

The cases used for this formative assessment process were carefully selected and constructed, in the knowledge that the students who used them would be working in diverse environments and with different clinical materials. For example, students attached to the Colorectal Unit at one of the tertiary referral centres would not see the breadth of clinical material available to a student working in one of the rural settings, but might be able to work in more of a collaborative environment than the isolated rural practices. Thus cases were produced to reflect the broad scope of surgical practice and to encourage the user to study further in each of the designated areas.

No restrictions were made as to how the formative assessment material was to be used and any student collaboration on the cases was deemed to be a positive outcome. The goal of the material was to expose students to scenarios and content they might not encounter on the wards and to engage in the diagnostic and management issues in a constructive and non-threatening manner. Group participation in achieving this was not considered to be a negative result, especially for students who were isolated in small groups in rural communities. It was acknowledged that at some stage, students were likely to collaborate on the weekly assessment case, potentially making the assessment non-discriminatory. This risk was balanced against the potential benefit of students completing the formative material, which was the initial goal of the teaching strategy.

The structure of the nine-week surgical attachment was identical to 2006 (Figure 6.2), but the assessment was slightly altered. Independent of this study it had been shown that well constructed MCQs are capable of testing the higher order cognitive skills of reasoning and judgement and did so more reliably than MEQs (Palmer & Devitt, 2007b). Since MCQs can be objectively marked and done so automatically in contrast with the more laborious and unreliable hand-marking of MEQs, a decision was made to drop MEQs from the summative assessment process that occurred in 2006. The end-of-attachment assessment was delivered fully online. As in 2006, there were four groups, labelled A to D, who completed their surgical attachment in sequence (Group A at the beginning of the calendar year and Group D at the end). In addition to a survey provided to students at the end of their attachment, an extra online survey was administered at the end of the year to all students, including the students who had access to Medici in 2006 (Chapter Five). Note that Group A was not provided a pre-test due to scheduling difficulties. Group A also had fewer cases available for study, as they were still under development. Allocation to the groups was via a randomized method and independent of the researcher and teaching staff.

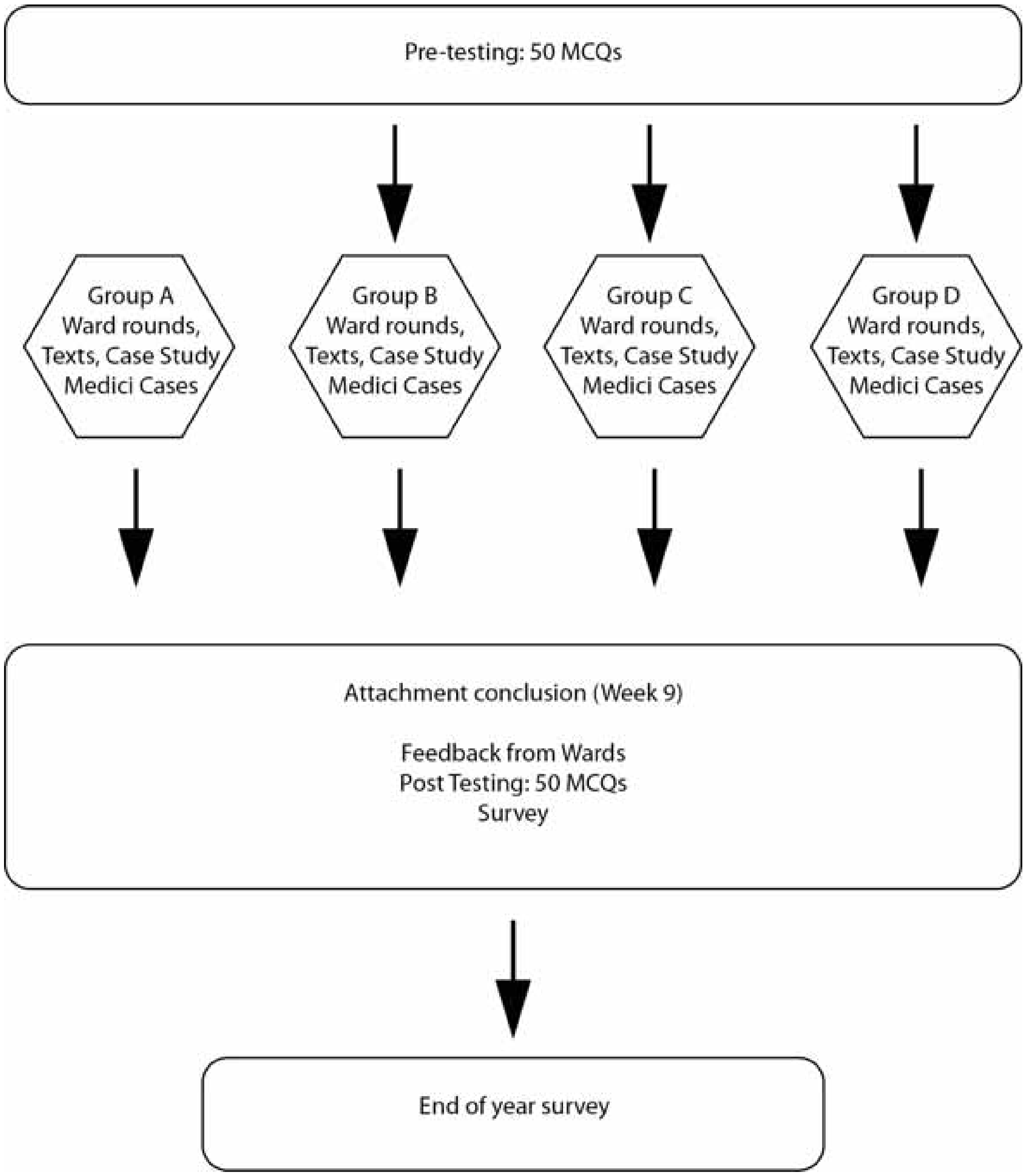


Figure 6.2 Course templates for surgical attachment 2007

Comparison of Test Scores 2006-2007

The raw percentages for the pre- and post-test are provided as Table 6.2. All groups in 2007 obtained a higher score in the final end of attachment assessment than the comparable group in 2006. The pre-test scores were similar, but Group C in 2007 performed significantly worse than the equivalent group in 2006. No explanation is available for this difference. In 2006 the pre- to post-test improvement ranged from 40% to 66%. In 2007 the improvement ranged from 77% to 117%. Group D in 2007 performed better than any of the other three groups.

Table 6.2 Test results for Groups A-D in 2007 and 2006.

Group 2007	MCQ Pre-test percentage	MCQ Post-test	Percent Improvement
A (n=33)	N/A	75±2	N/A
B (n=34)	43±2	76±1	77%
C (n=32)	35±2	76±2	117%
D (n=31)	39±3	83±2	112%

Group 2006	MCQ Pre-test	MCQ Post-test	Percent Improvement
A (n=30)	35±2	58±1	66%
B (n=33)	42±1	59±2	40%
C (n=33)	47±2	66±2	40%
D (n=33)	45.2±0.9	62.5±0.8	37%

Marks are given as score (percent) ± standard error

To compare improvement scores between 2006 and 2007 while adjusting for group, a two-way Analysis of Variance (ANOVA) model was fitted to the data. In the model, the improvement score (difference between the post-test and pre-test result) was entered as the outcome variable, while year (2006, 2007) and Group (A, B, C or D) were entered as predictor variables. The ANOVA model showed that after adjusting for group there was a significant difference in improvement scores between students in 2006 and 2007 (mean improvement 8.8 vs. 17.9, $p < 0.0001$). Note that Group A in 2007 was not included in this analysis as the improvement scores were undefined.

Given that improvement scores were undefined in Group A in 2007 (as they had missing pre-test scores), a second set of analyses comparing only post-test results was conducted. To compare post-test scores between 2006 and 2007 while adjusting for group, a two-way ANOVA model was fitted to the data. In the model the post-test score was entered as the outcome variable, while year (2006, 2007) and Group (A, B, C or D) were entered as the predictor variables. The ANOVA model showed that after adjusting for group there was a significant difference in post-test scores between students in 2006 and 2007 (mean score 28vs. 36 $p = 0.0004$).

Use of Modules Containing Support Material versus Module 0 (Incidental Cases)

The statistical data for each group and case are summarised in Table 6.3. Apart from the optional Module 0, the behaviour in the groups was similar. The majority of students attempted and completed cases in each Module (1–6) with a 90% completion rate. Often, there was more than one attempt on each case, but the completion rate on later attempts fell to 65% overall. Module 0 had fewer students attempting cases and lower completion rates for the cases that were attempted. Fewer cases were completed from Module 0 when compared with all other modules (Table 6.4).

Table 6.3 Case statistics by module and group

Group	Module	Av. no. students attempting cases in this module	Av. no. students completing cases in this module	% complete	Av. no. attempts at cases in this module	Av. no. attempts completed for the cases in this module	% complete
A	0	18±1	15±1	80±3	42±6	19±2	49±4
B	0	26.6±0.8	22±1	83±2	64±4	32±2	51±3
C	0	14±1	11±1	77±5	24±3	13±2	57±5
D	0	9.9±0.8	7.6±0.8	74±3	16±2	9±1	52±4
		17.2±0.9	**13.8±0.9**	**78±2**	**36±3**	**18±2**	**52±2**
A	1	21.2±0.5	17±1	80±4	42±6	23±2	59±6
B	1	32.0±0.4	28.8±0.8	90±3	107±11	63±5	61±4
C	1	25.6±0.7	23.4±0.7	92±3	66±6	40±3	62±4
D	1	23.2±0.8	19±1	83±4	53±5	31±3	60±4
		25.5±0.7	**22.2±0.9**	**86±2**	**67±6**	**39±3**	**60±2**
A	2	27±1	24.0±0.9	89±3	71±7	38±2	57±5
B	2	31.7±0.6	30±1	94±3	98±13	60±6	63±3
C	2	24.9±0.8	23±1	93±2	64±7	41±3	67±5
D	2	21±1	18±1	87±4	44±6	28±4	64±3
		25±1	**23±1**	**87±4**	**62±6**	**39±3**	**66±2**
A	3	27±1	24.0±0.9	89±3	71±7	38±2	57±5
B	3	32.1±0.9	29±1	90±5	91±9	52±5	59±5
C	3	26.1±0.7	25.0±0.8	96±2	67±8	42±3	66±6
D	3	21±2	18±2	84±5	42±6	24±4	57±5
		27±1	**24±1**	**90±2**	**68±5**	**39±3**	**60±3**
A	4	28±1	27±1	95±2	68±10	45±5	69±4
B	4	30.7±0.7	28±1	92±3	86±9	55±6	65±2
C	4	27.1±0.5	26.3±0.5	97±1	63±7	42±3	68±2
D	4	22±2	20±2	90±2	46±6	30±5	64±3
		27.0±0.8	**25.3±0.9**	**93±1**	**66±5**	**43±3**	**67±1**
A	5	27.0±0.8	92±3	72±8	47±3	68±6	27.0±0.8
B	5	29±1	25±2	87±4	85±12	49±6	61±5
C	5	25.4±0.8	24±1	93±3	60±8	37±4	65±6
D	5	22±2	20±2	90±2	46±6	30±5	64±3
		26.4±0.8	**23.9±0.8**	**91±2**	**66±5**	**40±3**	**64±3**
A	6	28±1	26±1	92±3	64±11	40±4	69±7
B	6	29.9±0.8	27±2	91±4	79±11	47±6	64±7
C	6	23±1	21±2	92±3	53±8	34±4	67±6
D	6	21±1	19±2	88±4	40±6	25±3	64±7
		25.6±0.9	**23.3±0.9**	**91±2**	**59±5**	**37±3**	**66±3**

Statistical data presented as average ± standard error

Table 6.4 Case statistics by module and group. Modules 1–6 collected together

Group	Module	Av. no. students attempting cases in this module	Av. no. students completing cases in this module	% complete	Av. no. attempts at cases in this module	Av. no. attempts completed for the cases in this module	% complete
A	0	18±1	15±1	80±3	42±6	19±2	49±4
B	0	26.6±0.8	22±1	83±2	64±4	32±2	51±3
C	0	14±1	11±1	77±5	24±3	13±2	57±5
D	0	9.9±0.8	7.6±0.8	74±3	16±2	9±1	52±4
		17.1±0.9	**13.8±0.9**	**78±2**	**36±3**	**18±2**	**52±2**
A	1–6	25.7±0.7	23.0±0.8	87±3	59±4	36±2	65±2
B	1–6	30.9±03	28.1±0.5	91±1	92±5	55±2	62±2
C	1–6	25.4±0.4	23.8±0.5	94±1	63±3	39±1	66±2
D	1–6	21.8±0.6	19.0±0.6	87±2	46±2	28±2	62±2
		25.9±0.4	**23.5±0.4**	**89.5±0.9**	**65±2**	**40±1**	**64±1**

Statistical data presented as average ± standard error

To compare the percentage of support cases completed with the percentage of incidental (Module 0) cases completed a negative binomial GEE model was fitted to the data. In the model, the number of cases completed was entered as the outcome variable, while type of case (support/incident), Group (A, B, C or D) and the interaction between type of case and group were entered as predictor variables. The total number of cases available was included as an offset variable to account for the slightly fewer cases available to Group A.

The interaction term was statistically significant and there was no need to interpret the other two terms. The interaction term showed that the difference between the proportion of support and incident cases completed varied across the four groups ($p = 0.035$). The proportion of completed cases was higher for support cases than incident cases in all groups, particularly in Group D (Table 6.5). In all groups, the proportion of completed incident cases was found to be lower than the proportion of completed support cases ($p < 0.05$). The column labelled rate ratio expresses the ratio of the two proportions. The proportion of completed incident to support cases was highest in Group B (rate ratio = 0.859) and lowest in Group D (rate ratio = 0.554).

Table 6.5 Differences of adjusted means

Effect	Group	Type of Case	_Group	_ Type of Case	Estimate	Standard Error	DF	Chi-Square	P-value	Rate ratio
Type of Case*Group	A	Incident	A	Support	-0.3008	0.1072	1	7.87	0.0050	0.740
Type of Case*Group	B	Incident	B	Support	-0.1525	0.0607	1	6.31	0.0120	0.859
Type of Case*Group	C	Incident	C	Support	-0.4058	0.1169	1	12.05	0.0005	0.666
Type of Case*Group	D	Incident	D	Support	-0.5901	0.1691	1	12.17	0.0005	0.554

In order to compare the number of attempts per completed case between support and incident cases, a negative binomial GEE model was fitted to the data. In the model, the number of cases attempted was entered as the outcome (dependent variable), while type of case (support/incident), Group (A, B, C or D) and the interaction between type of case and group were entered as predictor variables. The total number of completed cases was entered into the model as an offset variable. A significant interaction effect was found ($p < 0.0001$) suggesting that the difference in the number of attempts per completed case between incident and support cases varied across the four groups. The rate of attempts per completed case was highest for support cases in Group B (2.9 attempts per completed case) and lowest for incident cases in Group D (1.6 attempts per completed case). The *post hoc* comparisons revealed that in Groups B, C and D, the rate of attempts/completed case was lower for incident cases than support cases (as the rate ratios were less than 1). In Group A the rate of attempts to completed cases was higher for incident cases than support cases (rate ratio = 1.153, $p = 0.048$). Students in Group D attempted and completed significantly fewer cases than other groups. Group B behaved in an opposing manner, attempting and completing more cases than other groups.

Students did not devote their time solely to the assessment cases. Regardless of group or module, the cases designed to support students doing the summative assessment case were often attempted and completed. Students failed to go past the first screen of a case on an average of 7% of all attempts (Table 6.6). There was no difference between groups with this statistic.

Table 6.6 Students failing to continue past the first screen of a case.

Group	Left after 1st screen	Percentage of all attempts
A	194	6%
B	391	8%
C	234	7%
D	175	8%

Student Commitment of Time Towards Online Learning

When compared with 2006 data, it is clear that students attempted more cases in 2007 (Table 6.7). In the module containing incidental material, on average twice as many students attempted cases than in 2006, where all cases were optional. In the modules where the assessment case was present, the number of students attempting cases was approximately three times higher. Students spent between 10 and 20 hours on the *Medici* cases (Table 6.7). When considering the extra cases involved, this is comparable to the best performing group from 2006 (Group C) who spent an average of two hours on case studies using *Medici*.

Table 6.7 Student usage of the *Medici* interactive online resource from beginning of attachment to the end of the academic year

Group	**No students attempting cases** (total no. students)	**Number of cases attempted Average, Median [interquartile range]**	**Number of screens** (median of 10 screens per case) **Average, Median [interquartile range]**	**Time spent** (minutes) **Average, Median [interquartile range]**
A	33 (33)	40 44 [29.5–54]	420 410 [230–560]	640 506 [340–890]
B	34 (34)	54 57.5 [48.25–60]	660 560 [450–800]	1100 1045 [670–1510]
C	30 (32)	46 46.5 [41.75–56.25]	490 450 [350–580]	800 765 [490–1060]
D	29 (31)	39 42 [35–51]	350 360 [230–440]	600 560 [370–770]

Results are expressed as average, median [interquartile range]
Students who did not complete cases in groups C and D withdrew during the attachment.

To compare the time spent online between the four groups in 2007, a one-way ANOVA model was fitted to the data. Time spent online by students was entered as the outcome variable in the model, while Group (A, B, C or D) was entered as the predictor variable. The ANOVA model demonstrated that there was a significant difference between the four groups in the time spent online ($p < 0.0001$). The mean time spent online by students was highest in Group B (mean = 1115 minutes) and lowest in Group D (mean = 595 minutes).

The *post hoc* tests revealed that students in group B spent more time online than students in Groups A, C and D ($p < 0.0001$, $p = 0.0070$ and $p < 0.0001$ respectively). No significant differences were found between Groups A, C and D.

Predictors of Success

To test whether improvement scores were related to Group (B, C or D), total time spent online, number of support cases completed or number of incident cases completed, a linear regression model was fitted to the data. Only the student group was found to be a significant predictor of improvement scores ($p = 0.046$). There was no evidence that the number of incident cases completed, the number of support cases completed or time spent online had an influence on improvement scores. After adjusting for other variables in the model, students in Group B were found to have significantly lower improvement scores than students in Group C (15.3 vs. 18.9, $p = 0.023$).

To test whether post-test scores were related to Group (A, B, C or D), total time spent online, number of support cases completed or number of incident cases completed, a linear regression model was fitted to the data. Both group ($p = 0.0012$) and the number of support cases completed ($p = 0.013$) were found to be predictive of post-test scores. For every single unit increase in the number of support cases completed, the post-test score increased by an average of 0.13 units ($p = 0.013$). After adjusting for other variables in the model, Group D was found to have higher post-test scores than Groups A, B and C ($p = 0.0003$, $p = 0.0013$ and $p = 0.0020$ respectively).

Survey Responses

Two anonymous surveys were implemented in 2007. An Attachment Survey (Appendix 2) was the first of these and was carried out at the end of each attachment. This survey focused on student understanding and use of formative material. A *Medici* Usage Survey (Appendix 3) was administered at the end of the year and focused on their use of and attitudes towards *eMedici*. The *Medici* Survey participants included the students from 2006 as well as the students from 2007.

Attachment Survey

The attachment survey consisted of four short Yes/No questions, eight Likert scale questions and four open-ended questions. The Likert questions used 5-point scales where 1 corresponded to the respondent strongly agreeing with a statement and 5 corresponded to strongly disagreeing. The total number of responses received was 91, corresponding to a response rate of 71%. Most students believed they had encountered some form of formative assessment during their attachment (Table 6.8), highlighting case presentations either in the ward or at a special session for case presentations (Table 6.9). Students also identified *eMedici* as a formative assessment tool (Table 6.9 and Figure 6.3) but some students identified the end-of-attachment summative examination as a formative exercise also, even though formative assessment was clearly defined for the students before they completed the survey.

Table 6.8 Results of yes/no questions from 2007 cohort

Question	Percentage agreeing
Did you have any formative assessment during the attachment	69%
Did you have tutorials with individual appraisal	45%
Did you have peer review groups	7%
Did you have case presentations with immediate feedback	79%

Table 6.9 Results of question ‘What formative assessment did you have during your attachment?’

Type of assessment	Percentage of students identifying this as a formative experience
Case presentation	25
Ward round presentations	20
Presented to consultant	16
eMedici	15
On ward tutorials	5
Spot quizzing	5
Feedback form	5
Viva/Outpatients/Theatre	5
End of attachment test	4

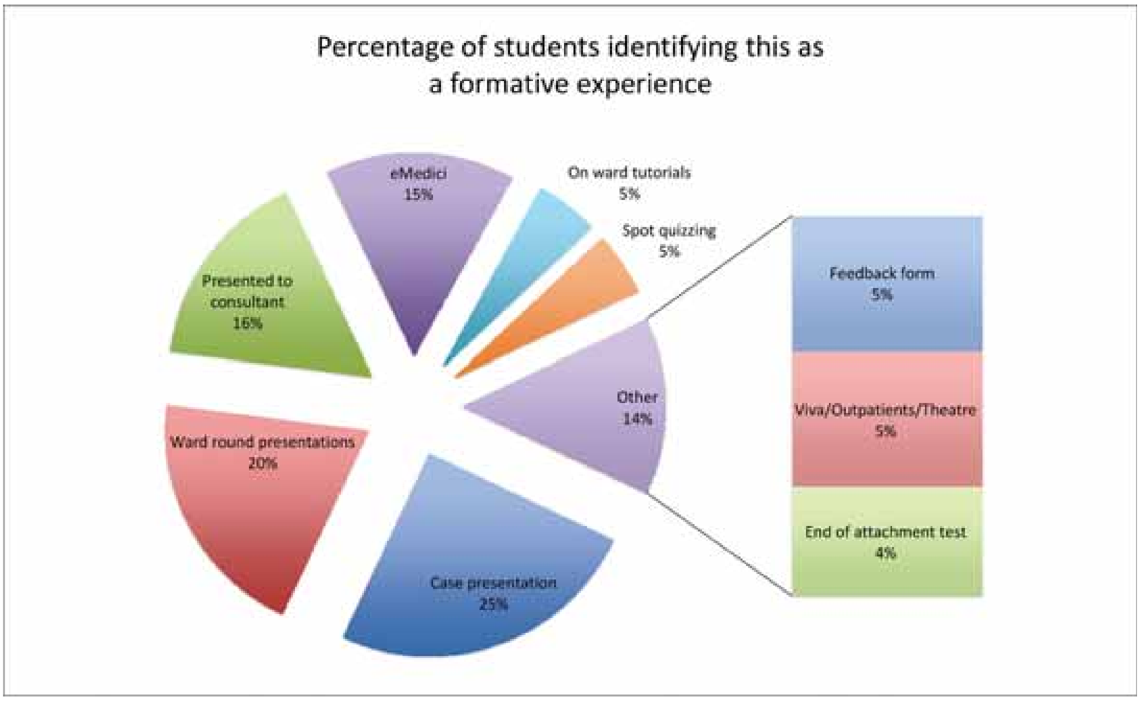


Figure 6.3 Graph of formative assessments identified by students in use during their attachment

When students answered the questions on the provided online material, they reported that it aided their learning, was relevant and they would have liked to have had more of it. Many students only accessed the online material at the end of the week. It is unclear whether this was because it was a more convenient time for students to do the exercises or because they put off the work until only a short time before the deadline. The students used the material for revision, would have liked access to the online case material via CD, and believed that they would have completed the online material without it being compulsory (Table 6.10, and Figures 6.4 to 6.10).

Table 6.10 Results of Attachment Survey Likert scale questions from 2007 cohort

Question (1 = strongly agree, 5 = strongly disagree)	Median (Interquartile range)
The online material aided my learning	1 (1-2)
I only accessed the online material at the end of the week	2 (2-4)
I would have completed the online material without it being compulsory	2 (2-3)
The online material was relevant to my attachment	1 (1-2)
I used the online material for revision	1 (1-2)
I would like to have the learning material on CD	1 (1-2)
I would like to have more online material	2 (1-2)

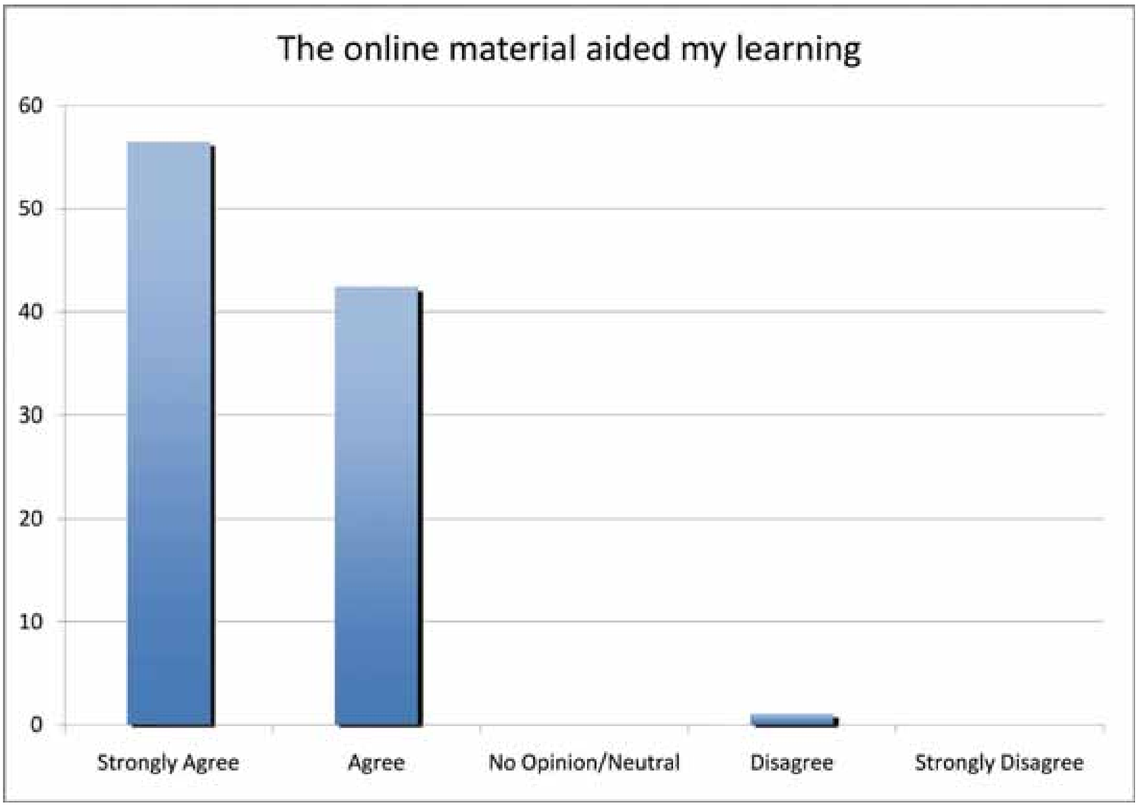


Figure 6.4 The online material aided my learning

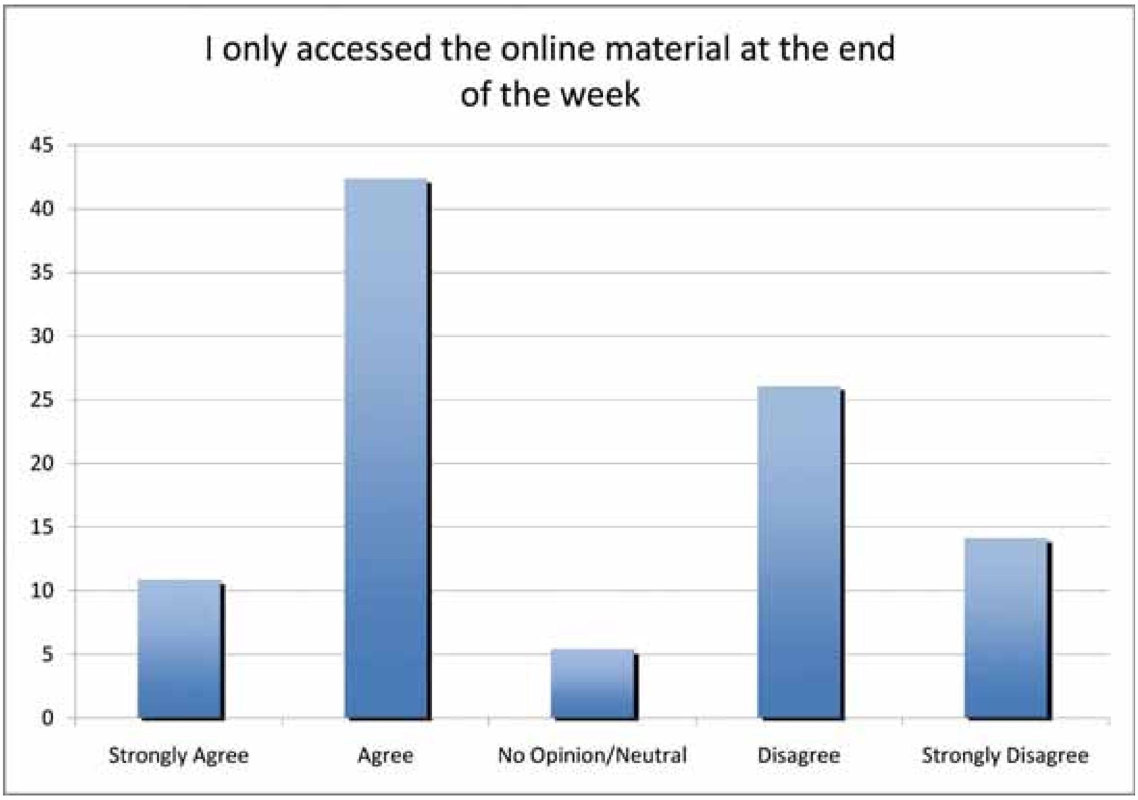


Figure 6.5 I only accessed the online material at the end of the week

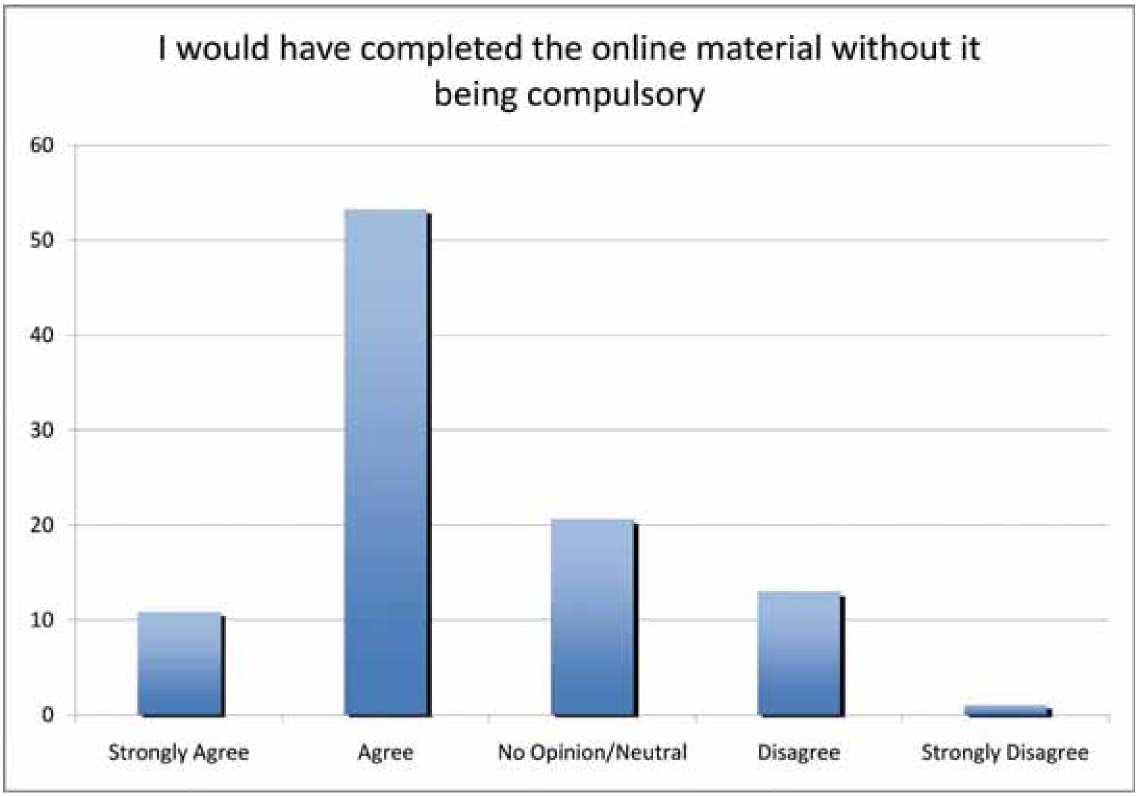


Figure 6.6 I would have completed the online material without it being compulsory

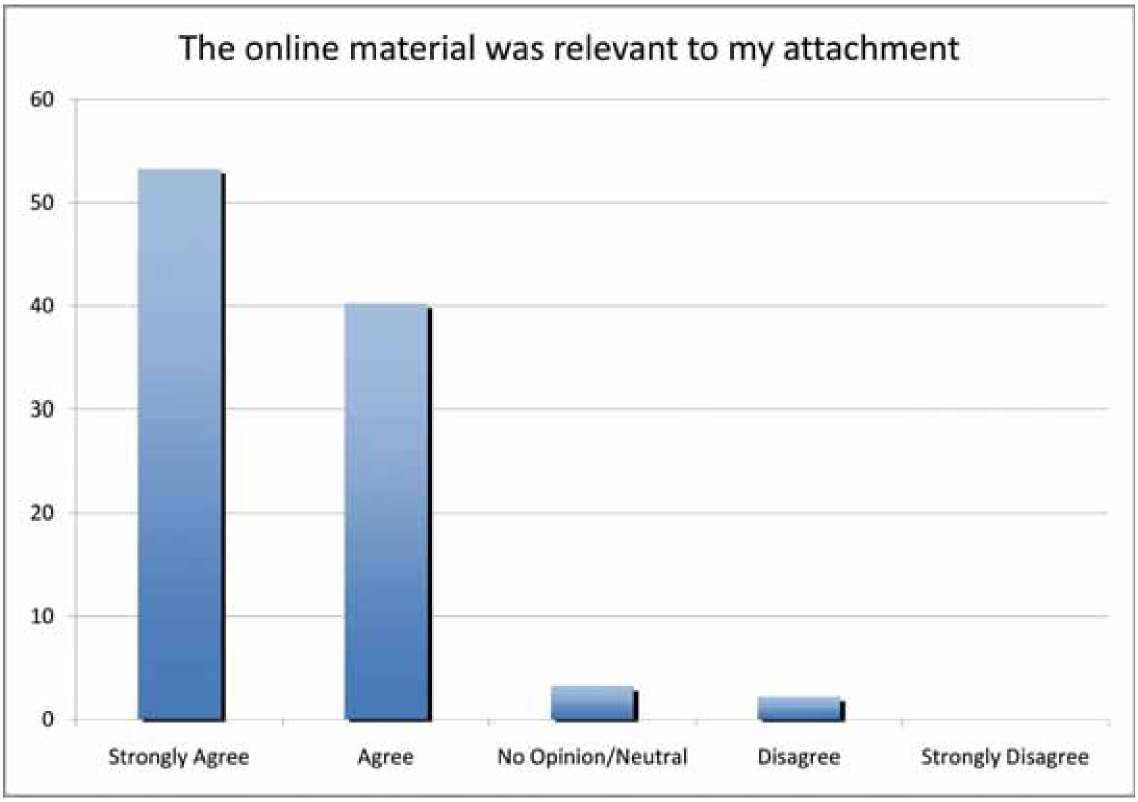


Figure 6.7 The online material was relevant to my attachment

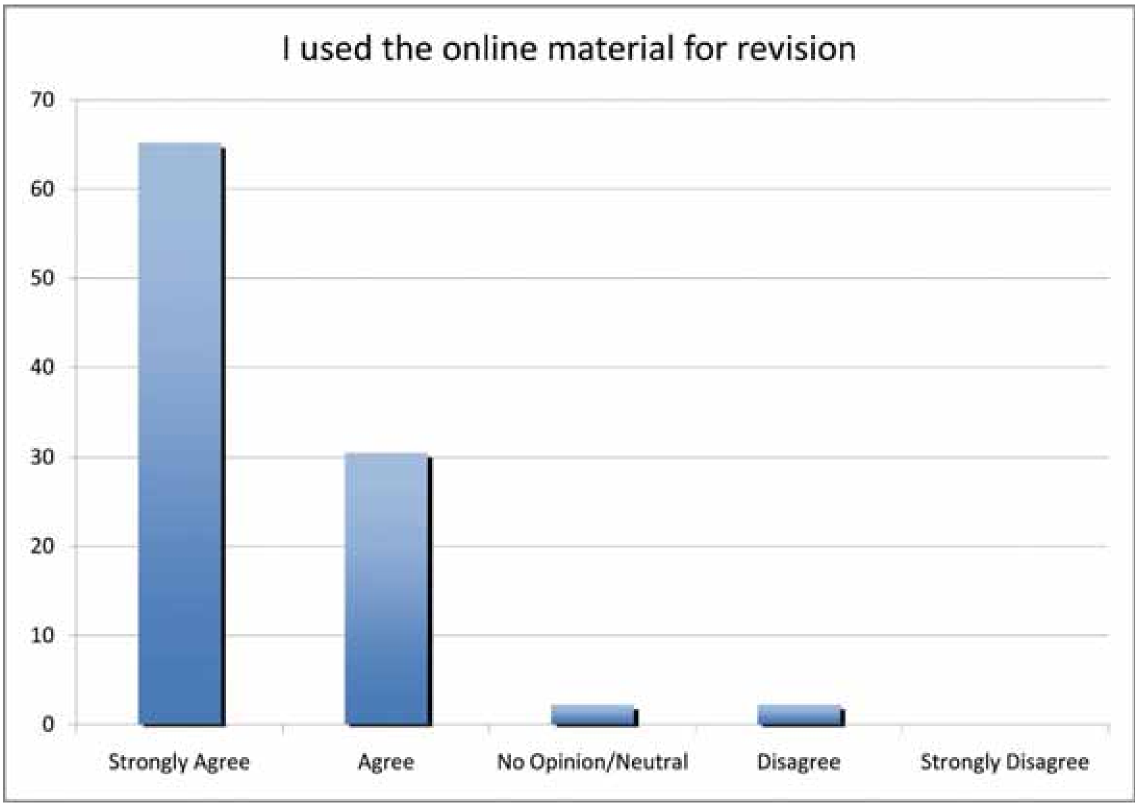


Figure 6.8 I used the online material for revision

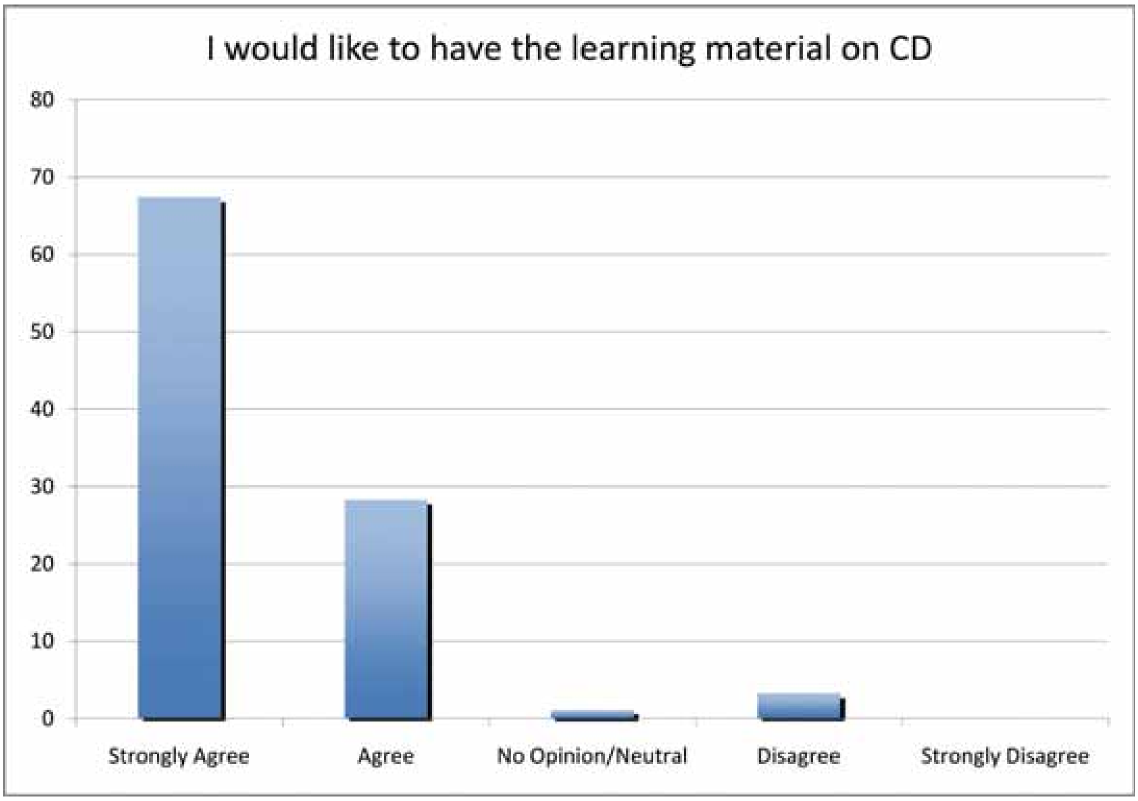


Figure 6.9 I would like to have the learning material on CD

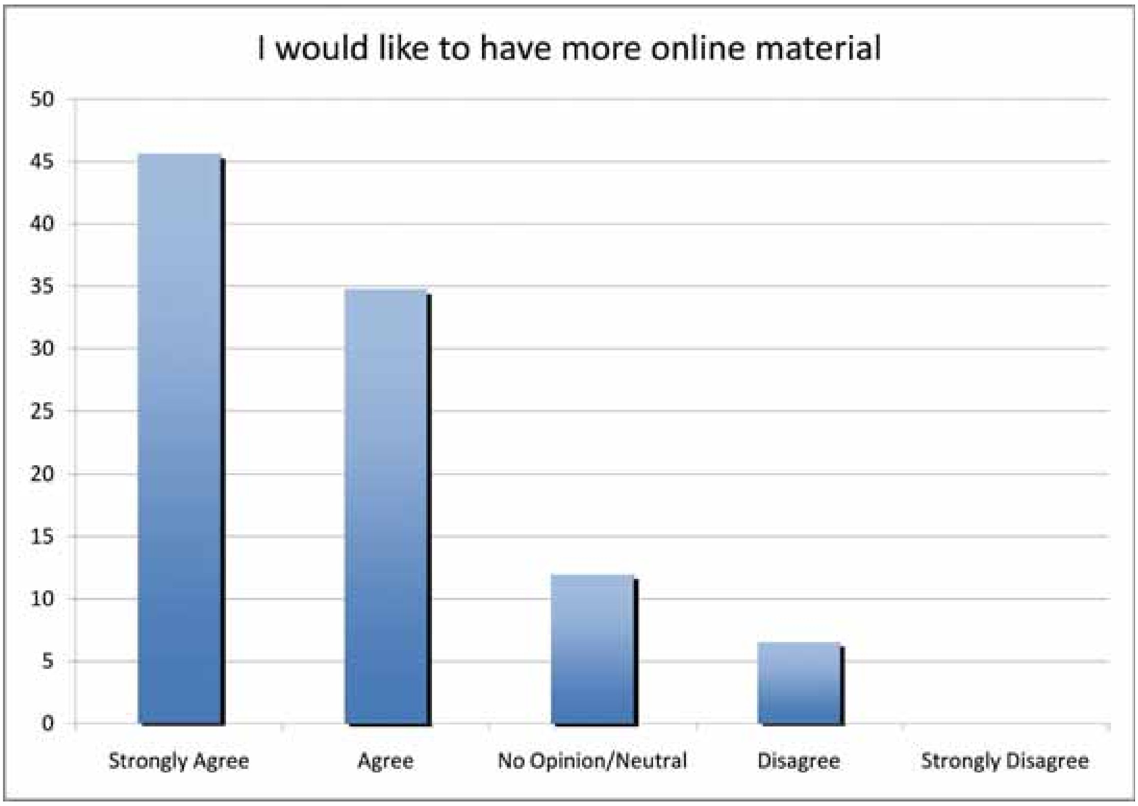


Figure 6.10 I would like to have more online material

The free response questions provided additional information (Table 6.11). Students felt that there should be more online material, there should be feedback provided on the summative assessment case provided for each module and efforts should be made to ensure consistency with texts and other clinicians. They also felt that some of the questions were ambiguous and this should be addressed (Table 6.11). The benefits of *eMedici* as perceived by the students were many. They appreciated the case-based scenario approach and the structure designed to assist in developing clinical reasoning. The instant and comprehensive feedback, as well as whether the answer was right or wrong, was well regarded. Students felt the cases benefited from the inclusion of images and they appreciated the ease of use of the program and its easy accessibility (Table 6.12).

Table 6.11 Responses to the question 'How would you improve the online learning material'

Comment	Number of students suggesting this theme
More of it	20
Explanations for the summative assessment case	11
Nothing	6
Consistency with textbook-other sources	6
Some questions ambiguous	4
Grouping cases better by topic	3
Avoid random MCQs that don't relate to the cases	3
Allow printing so don't have to be online	3
Some images didn't come up	3
More prac MCQs	2
More images	3
More detail	2
More basic science	2
Keep it up to date	2
Build in discussion group	2
Learning objectives for every case	1
Give out a CD	1
More videos	1
If cases could be downloaded	1
Feels unrealistic	1
Different people have different opinions on how to manage a case	1
Videos take a while to download	1
More relevant cases	1
Identify areas where controversy may exist	1
Links to other resources	1
Put up previous weeks scores	1
Make it all MCQ (bit unsure)	1

Table 6.12 Responses to the question 'What did you like about the online material'

Comment	Number of students suggesting this theme
Useful info about why answers were correct/incorrect	23
Case-based/clinical reasoning	18
Instant feedback/explanation	17
Images...helped understanding	13
Relevant	13
Easy to use	11
Online (more convenient than handouts)	11 (1)
Provided a basis to guide learning	8
Good way to test myself	6
I learnt from it/informative	6
Summaries at end of cases	5
Interactive	4
Some form of assessment every week	3
Good variety (common and uncommon)	3
Do it in your own time at own pace	3
The layout	3
Learning objectives	3
Thought stimulating	2
Could view it repeatedly	2
Thoughtful questions	2
Variety of multiple-choice and short answer questions	2
Exposure to topics not necessarily on wards	2
comprehensive	2
Instant results as to how we were going	1
Concise	1
Grouping of subject matter	1
Fast	1
Liked the structure of cases to support summative case	1

Students were asked about what had motivated them to study, in general and specifically for their surgical attachment. Students identified assessment as the main motivating factor. Other motivators were an interest in the topic; working on the wards; a desire to learn; and fear of failure or embarrassment. Good teachers were also identified as a key motivator for some students (Table 6.13). One student commented

'It was good there was an assessment case or I might have left it till the end of the rotation'

Table 6.13 Responses to the question ‘In general, what are the things that motivate you to study’

Comment	Number of students
Exams/assessment	36
Interesting topics	19
Real world experiences/on the ward	19
Desire to learn	16
Fear of not knowing when asked/failure..feel comfortable being asked questions	11
Good teachers	8
Patients presenting	7
Tutorials	5
Regular Questions and Answers	5
deadlines	4
Presentations	4
Being prepared to talk to patients	4
Supplied learning objectives	3
Interactive learning	2
Passion for Medicine	2
What I’m doing is relevant to the real world	2
Regular feedback	2
Realistic summative assessment	1
Having small specific areas to focus on (even if multiple)	1
Not much actually	1
discussions	1
Expectations of staff	1
Positive feedback	1
marks	1
challenge	1
stress	1
Peer support	1

For the surgical attachment, the main motivators for learning were reported as interesting case studies, *eMedici* and good quality teaching staff (Table 6.14 and Figure 6.11). One student lamented that they were ‘Motivated by interesting cases on the wards but they were mostly not examinable’.

Table 6.14 Responses to the question 'Please provide examples of things that motivated you during this attachment'

Comment	Number of students
Interesting case studies (new, weird)	17
EMedici	17
Good supervisors and staff, consultants	16
Relating theory to practice	8
Consultants quizzing	7
Seeing patients/cases on wards then reading about them in depth after	6
Case presentation	6
Exposure to a large number of patients and pathologies	5
Exam	5
In theatre	3
I enjoyed the topic	3
Fear	1
Want to be a surgeon and wanted to make a good impression	1
Study environment	1
Intra hospital talks	1
outpatients	1
peers	1
Lots of study materials	1

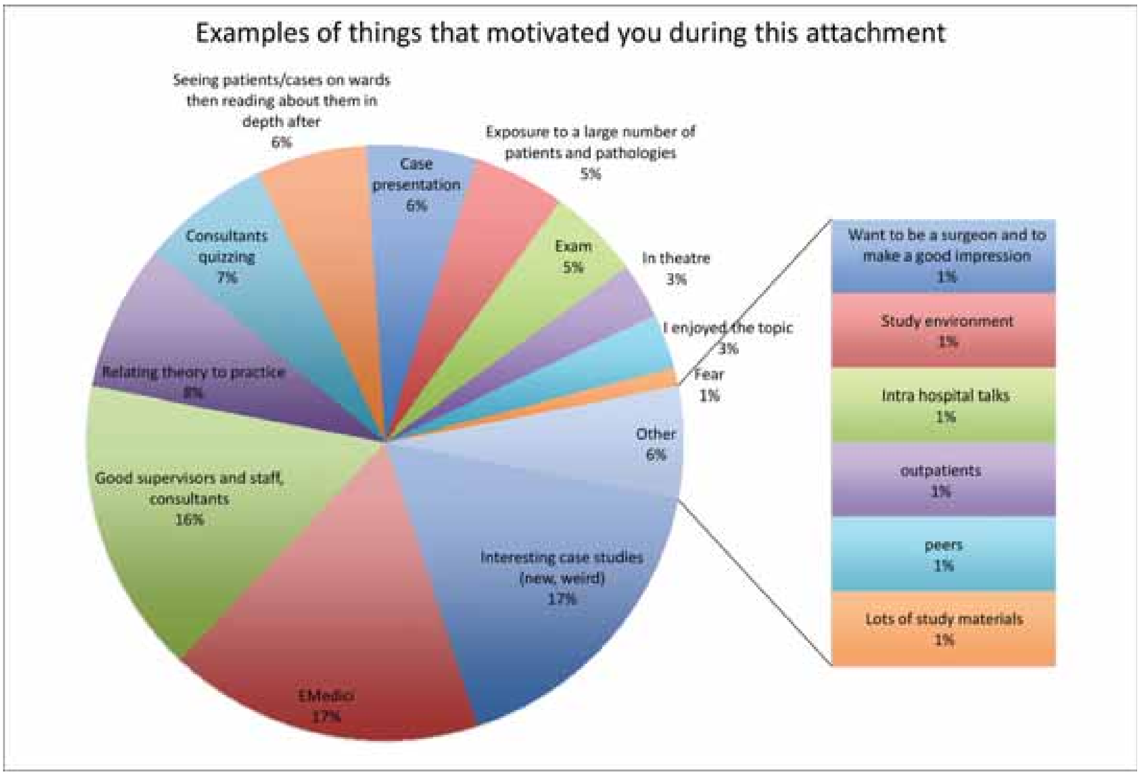


Figure 6.11 Student motivators

Medici Usage Survey

The *Medici* Usage Survey was delivered online at the end of the year to 130 fourth year students of which 85 responded (response rate 65%). The survey also had 66 responses from a group of 128 fifth year students (response rate 52%), who had experience with *Medici* in 2006. These students had been provided with new cases for their use in 2007. The survey contained nine Likert questions on *eMedici* and two open-ended questions. The Likert questions used 7-point scales where 1 corresponded to the respondent strongly disagreeing with a statement and 7 corresponded to strongly agreeing. This is in contrast to the previous survey due to this online survey needing to conform to University standards, which differed from those used locally. Students responded that they found the quality of the content of *Medici* to be good; they understood the concepts presented using *Medici* and felt motivated to learn using *Medici* (Table 6.15, Figures 6.11–6.13). They found the online format assisted in balancing their life-study commitments, were satisfied with the number of cases available, and wanted to have *Medici* available on CD. Students believed *Medici* helped them develop their problem solving skills.

Table 6.15 Summary of results of Likert questions of *Medici* survey:

Question (1 = strongly disagree, 7 = strongly agree)	Year 4	Year 5
Overall, I am satisfied with the quality of the *Medici* cases.	6 (6-6)	6 (6-7)
Overall, I am satisfied with the number of *Medici* cases available ONLINE	6 (5-7)	5 (4-6)
I would like to have *Medici* available to me on CD	7 (7-7)	7 (7-7)
I feel part of a group committed to learning.	6 (5-7)	6 (5-6)
The feedback provided to me from *Medici* assists my learning.	6 (6-7)	6 (6-7)
I am motivated to learn using *Medici*.	7 (6-7)	6 (6-7)
Medici helps me develop my thinking skills (eg. Problem solving, analysis).	6 (6-7)	6 (6-7)
I understand the concepts presented using *Medici*.	6 (6-7)	6 (6-7)
The flexibility of the on-line format helps me balance study and other commitments.	7 (6-7)	6 (6-7)

Results reported as median (interquartile range)

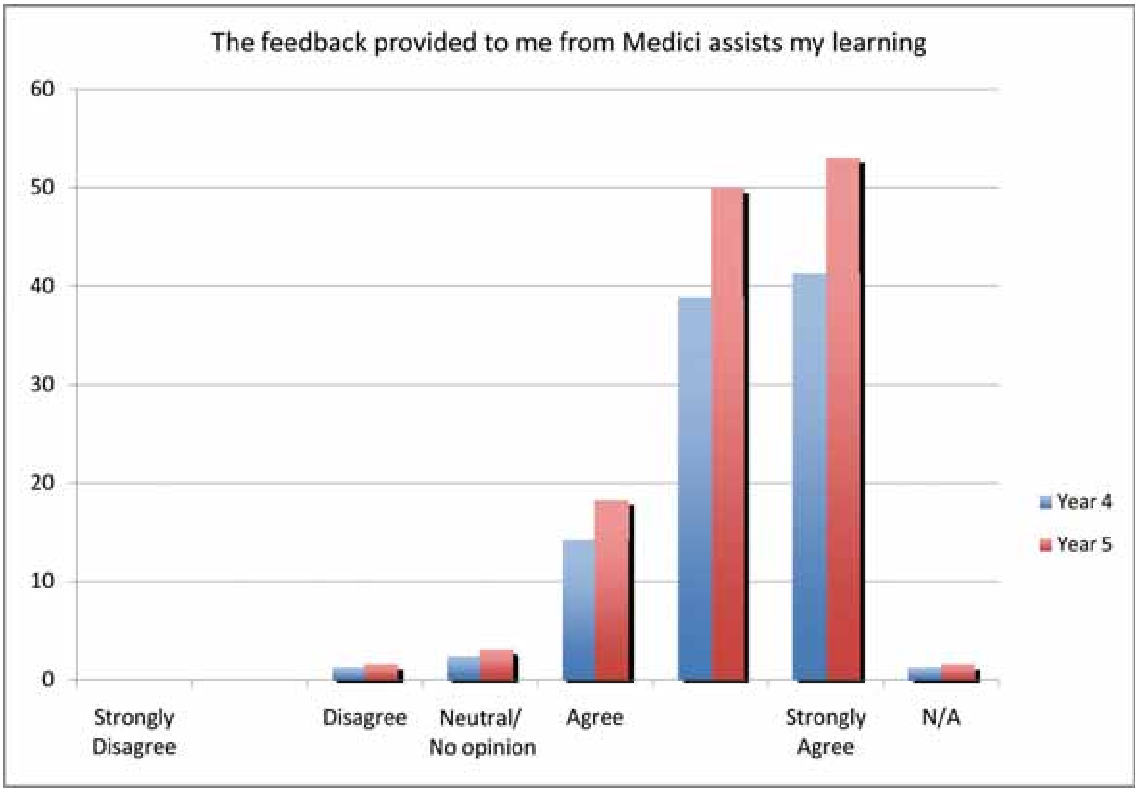


Figure 6.12 The feedback provided to me from *Medici* assists my learning

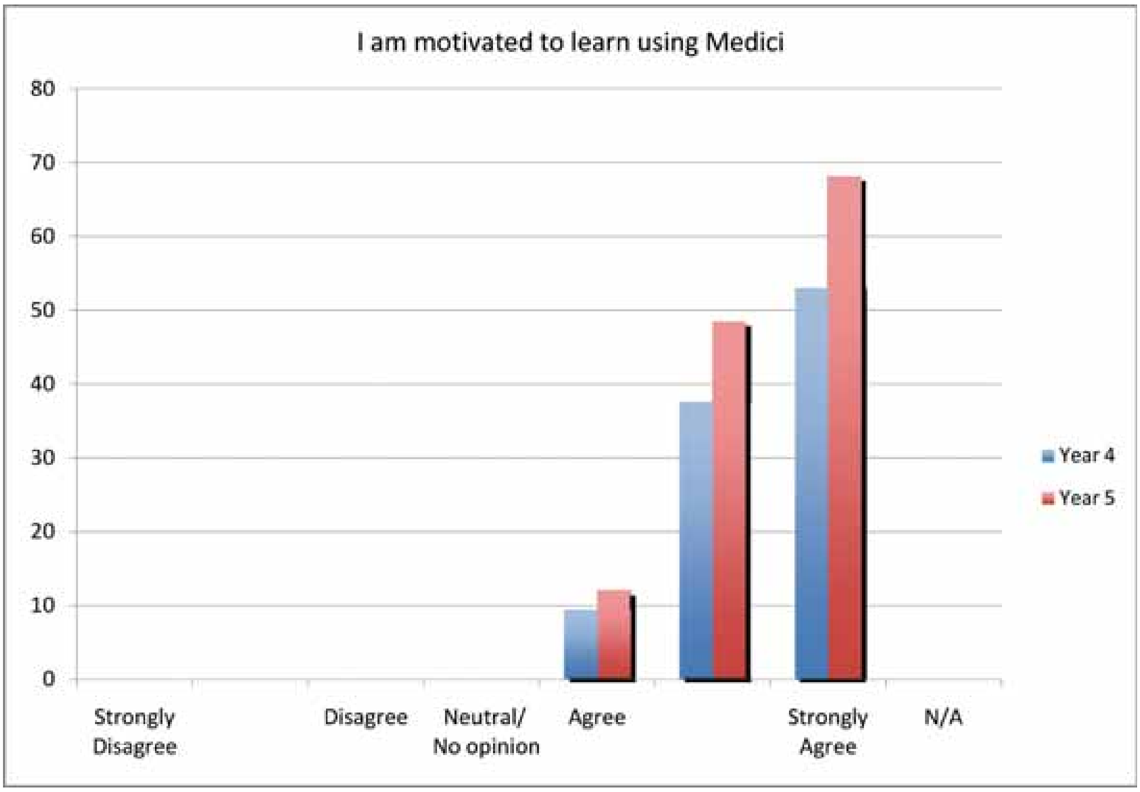


Figure 6.13 I am motivated to learn using *Medici*

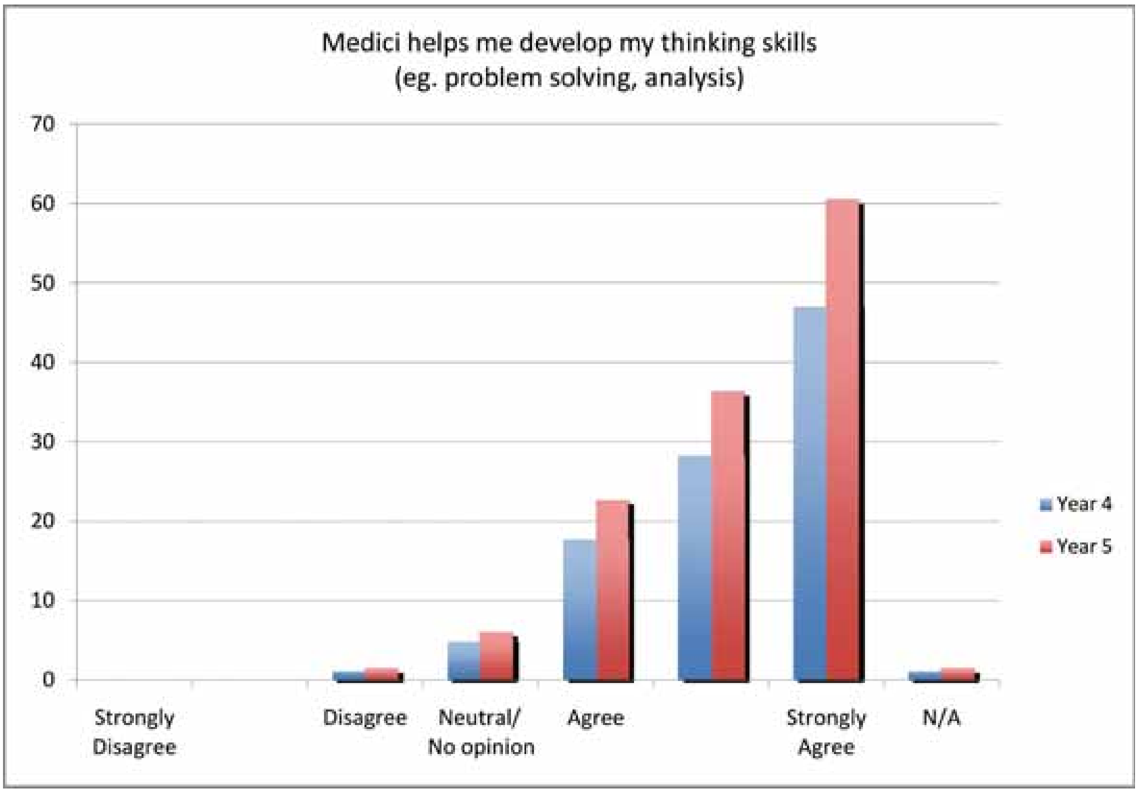


Figure 6.14 *Medici* helps me develop my thinking skills (e.g. problem solving, analysis)

The qualitative responses to the survey supported the contention that feedback on authentic case material delivered when the students want it is highly regarded by students. This was consistent in the responses provided by both fourth and fifth year students, which are summarised thematically in Table 6.16.

Table 6.16 Responses to the question 'What are the best aspects of *Medici*, and why?'

Aspect	**Year 4**	**Year 5**
Clinical case approach/problem solving/real life	28	19
Explanation why correct or incorrect	26	24
Available 24 hours a day anywhere/flexible	22	11
Images	16	6
Instant Feedback	13	3
End of case summaries	10	0
Relevant	8	5
Easy to use/	8	4
Helps with developing clinical thinking	7	4
Good revision tool/testing yourself/reflection	7	12
Wide variety	7	6
Helps me learn	5	5
Like being able to go back and make other choices	5	1
Test higher order skills	5	1
Cases are a good length	4	0
Think about differentials	4	1
Step by step approach through case/well structured	4	7
More interesting/fun	4	3
Assists learning	3	0
Helped me address weaknesses	3	2
Video	3	0
Good to know authors can be easily contacted	3	0
Interactive	3	9
Covers things not met earlier or on attachment	3	0
Cases released weekly on themes	2	0
Good mix of types of questions	2	0
Fills in gaps in rotations	2	0
Real time marking	1	0
Language is colourful and exciting	1	0
Accurate and up to date	1	0
Great across all disciplines	1	0
Helped me set objectives for each week	1	0
Cant find this sort of info in texts	1	0
Guides study	1	0
No pressure/anxiety	1	0
Based on common conditions	1	3
Encourages you to study broadly. Not get bogged down in one area	1	0
Ensures students are aware of what is relevant and necessary	1	1
Learning objectives	1	0
Cases well written…at good level	0	3
Work at own pace	0	1
Different format to reading	0	2
Better than textbooks	0	3
Like it on CD	0	1

Student Comments

Student comments consistently praised the online learning resource. The module structure and the scenario-based approach were also highly regarded. The following quotations support this assertion and are typical of the comments received.

> ‘In a course with very little student direction in the pre-clinical years and the expectation that the student will teach themselves the right things it is reassuring to have access to the *Medici* material to know what areas are important and what the correct way to learn around a disease is’ and

> ‘cases described explore all facets of the disease involved (clinical - investigation - diagnosis - management)—allows you to have an appreciation of clinical problem solving. Questions regarding differentials not only prompt a correct answer but jog your memory regarding diagnosis not previously considered’.

The students made insightful comments on the feedback provided with comments such as

> ‘from my perspective, the best aspect of *Medici* is the thorough and comprehensive discussion of all responses. it is from these responses that the most understanding is gained’ and

> ‘feedback, particularly in response to incorrect responses the staged answering and response to questions—it allows my thought process to continue and respond to the different aspects of the question e.g. investigation, examination, history etc. (i.e. it facilitates active thought process rather than answer-response, answer-response).’

The focus on clinical scenarios as a learning tool provided some students with an immersive experience.

> ‘Clinically relevant information, it is an exciting learning tool, because as a student I feel that I am active in the learning process because I am making decisions and diagnosing. Furthermore, it enables me to have a break from book work, while still learning new information’ and

> 'Having cases to study from helps me retain medical knowledge that is applicable to what I may see on the wards/ in my career. Instead of just seeing a page in a texbook when I think about a topic, I see a patient'.

Students were asked how the *Medici* program could be changed to improve their learning. The majority of comments focused on content specific areas such as enhanced feedback and greater amounts of content covering broader areas (Table 6.17). Specific comments of note relate to the nature of the questions asked:

> 'the cases where you type in your answer aren't as good as the multiple-choice ones. just because it's too easy to not type it in, and click next.'
>
> 'Simply MCQ. Some of the weekly non-assessed cases have space to write answers but these are never checked.' and
>
> 'If the incorrect answer is selected at first, it would be more beneficial if the correct answer did not appear on the screen straight away, but rather the student was given the opportunity to try a different response after learning why their answer was not correct.'

Table 6.17 Responses to the question '*Medici* could be changed in the following ways to improve my learning:'

Comment	**Year 4**	**Year 5**
More cases in other areas	10	7
Some answers are contestable	9	0
Feedback on weekly assessment	9	0
Better wording of some questions/answers	9	0
More cases	7	10
More detail in feedback	6	4
Images didn't come up	6	0
Nothing	5	5
More pictures	4	3
Provide up to date info with references on changes in management	4	0
CD version	4	7
More MCQs	3	0
Consistency in answers	3	0
Categorise cases by system/area	2	1
Downloadable cases/images	2	0
Be able to cut/past text	2	0
Test questions don't reflect content in cases	2	0
More learning objectives	2	0
Make assessments more like cases	2	0
Simply MCQ	1	0
Focus on which are important rather than most important-subjective	1	0
Some cases are a bit long	1	0
More path	1	0
A means of letting students provide feedback on each case for changes etc	1	0
Don't provide correct answer straight away	1	0
Video on examination	1	0
Print summary	1	0
Some cases very similar	1	0
Videos take too long to load	1	0
More interactive	1	0
View cases by diagnosis	1	0
More video	1	0
Specify software requirements	1	1
Randomise questions at each log on	1	0
Larger images	0	1
More detail in summary page	0	1
Automatic restart for critical errors	0	1
Have a view all feedback option	0	2
Branching cases	0	1
Topic Search Function	0	1
Restart cases where I left off	0	1
More investigations	0	1

Discussion of Results

It is apparent that in every outcome measured in this study, the fourth year group in 2007 improved on the outcomes of the previous 2006 group. They made greater use of the online material (Table 5.5 and Table 6.7) and performed better in the end-of-attachment exam. It is tempting to suggest that the former caused the latter and given that the comparison group in 2006 was in every respect identical to the 2007 group but without the support of a structured formative assessment strategy it may be a valid conclusion. Perhaps the question should be how did this result eventuate? The exam material addressed the learning objectives of the cases, but did not mirror it, so rote learning is not a reasonable conclusion to draw. It is certainly possible that the cases themselves taught the students all they needed to do well in the exams, but that may be limiting the effect of the learning tool. Considering the students regarded Medici as one of the key motivators of their attachment it could be that the use of the cases inspired students to seek more information in related fields, visit a greater number of patients or examine old case notes. It is also likely that they would have discussed management decisions suggested within the case scenarios with clinicians, especially if they disagreed with some of those decisions. It is worth remembering that there were not necessarily any correct answers in many cases, only decisions based on evidence and clinical acumen. Different clinicians are likely to make different, but not incorrect decisions about how a patient is managed. Learning that these differences exist and the reasons behind them may help students build up a databank of knowledge, perhaps like the scripts suggested by those proponents of script concordance methods of learning and assessing.

An observation of interest is that there was no direct link between time spent on the online cases and examination outcome. It might be supposed that the more work the better the outcome, but it is worth examining the time spent on the cases. The interquartile range for group A was 340-890 minutes and for group D it was 370-770 minutes. These are non-trivial times. Six hours for the student represented by the lowest bound in these ranges is a significant amount of study time on a formative exercise. It is possible that the student achieved all that could be gained from the exercises by spending that amount of time on them. In other words, it is possible that there is a threshold where nothing further can be gained by repeating case scenarios. Intuitively this makes sense. There comes a point where there is a need to learn in a different way or to recognise that what something has to offer has been learned to an appropriate level. If that is the case then students that spent twice this time studying formative saes may have not been using there time effectively. Perhaps it is important to tell students not to study too hard using one type of learning?

The behaviour of different groups was interesting. Groups B and C dedicated more time to the Medici cases than groups A and D. The reason behind this is likely due to acclimatisation and examinations. Group A was the first group to do the cases and the surgical attachment was their first clinical attachment. In this situation, the students would not have the benefit of advice from other students in their cohort about the value of tasks and they would have been spending significant amounts of time familiarising themselves with hospital structures as well as the appropriate protocols when dealing with patient, interns and senior clinicians. This process takes significant time and may well have taken it away from formative assessment exercises. Group D was the last attachment and although they would have understood the requirements of an attachment at that stage, they would have also been focussing on examinations and preparing for them. In this instance, the formative assessment cases, although useful for one discipline were not going to be of assistance in others and thus were perhaps given a lower priority. Groups B and C were likely the most relaxed about their attachments with no exam pressure and thus they could devote more time to formative assessment tasks. This does not explain the lower level of performance by Group B and the cause of this result is unclear.

The strategy used to encourage students to use Medici could be considered to be a misuse of assessment. Certainly, there was an understanding that by providing the summative assessment quizzes at the end of each week online, there was likely to be collusion or circulation of the answers. Collusion, or preferably collaboration was not considered to be a poor result because students would discuss the problems before coming up with answers. Circulation of answers was only going to be possible if one student ‘took a bullet’ for the others by sitting the assessment and reporting the outcomes. This requires a strange version of altruism and it is equally likely that a student, having once completed the assessment and realised his or her mistakes would not allow others to gain an advantage by providing answers. Nonetheless there is anecdotal evidence that this did occur and there are certain strategies that would allow students to share the risk (and benefit) if they so desired.

Moreover, the question arises should assessment be used as a lure to encourage students to do what they should do anyway. Certainly some would argue strongly against this approach, just as they would argue against marks being provided for participating in tutorials or going on field trips and the point made is valid. However, there is evidence here that such a strategy is not only effective but also essential. Students in the 2007 group clearly stated that they would do the Medici cases even if they weren't compulsory, but there is evidence from the 2006 group that this is not what would, in reality occur. Students prioritise tasks and they tend to prioritise to assessments. In the case of an online formative assessment task and perhaps any formative task, they chose to make it a low priority, potentially to their detriment. The 2007 group were aware that they would have to complete the cases provided online in order to perform well in the weekly assessments, however the 2006 group were also informed that the online cases linked tightly to the curriculum and it was strongly recommended that they complete the cases in order to perform well on the wards and in the end-of-attachment assessment. They chose not to do so. In all likelihood, it was the immediacy of the weekly assessment for the 2007 students, which motivated them to complete so many of the cases provided as support material. This was supported by the survey results that highlight that one of the key motivators for this group of students were the *eMedici* cases. Assessment drives learning but formative assessment may not unless it is tightly coupled to a summative component. In the end the greatest formative assessment tools in the world are useless if nobody completes them. So a pragmatic approach may be to recognise that the reality of student life is that 'if it doesn't count for marks it doesn't count at all'. This may feel a hollow victory if one believes that students should complete formative assessments because they are beneficial to their learning, but it is nonetheless a victory.

Medici was one of the main motivators for learning for the 2007 cohort of students. Gratifyingly, of equal value as motivators were good supervision and interesting case material. The issue of motivation for students is at the heart of the success of any educational initiative, particularly one focusing on a formative approach to learning. In this case, the combination of the *Medici* software and an effective formative assessment strategy motivated students to do well as specified by the indicators for success such as the use of the program and improved assessment results.

As discussed in Chapter Two, motivation can be regarded broadly as being either intrinsic or extrinsic. According to Deci (1991) an intrinsically motivated behaviour is based on a person's interest and enjoyment of a task with any reward being satisfaction at achieving the task or enhancing ones knowledge. Extrinsically motivated behaviour may be based on factors outside the person's control, such as a reward for achieving the task or may also be based on internal factors, such as recognition of the value of the task. Considering the group of students from 2006, the majority presumably had little intrinsic motivation to complete the formative *Medici* cases and this behaviour was replicated to some degree by the behaviour of the 2007 students and the Module 0 cases (Table 6.3). When an extrinsic motivator was applied to the students (assessment cases in each of Modules 1–6), this external regulation may have suggested to the students that the cases associated with the assessment in each module were important, and worth completing; a conscious valuing of the activity.

Not surprisingly, the majority of students found exams and assessment to be the main motivators in their study. Assessment is considered to have a substantial influence on student motivation (Leach, Neutze, & Zepke, 2001; Race, Learning, & Centre, 2001; Regan, 2003; Seale, et al., 2000) and this study supports this observation. The students' comments from the survey are reflected by their actions in the use of *eMedici*. The cases, which were designed to support the assessment cases were attempted and completed in greater numbers than those which were clearly delineated as non-assessment (although certainly core material that would be useful for students in examinations as well as on the wards). More students (52%) attempted cases in the assessable modules of cases (Modules 1–6) than those in Module 0 and 90% of cases in Modules 1–6 were completed compared to 78% of those in Module 0. Overall students who attempted cases in Modules 1–6, attempted the cases 2.5 times compared with 2.1 times in Module 0.

Chapter Seven: Conclusion

Modern day Medicine involves fast turnaround of patients with a large proportion of medical procedures taking place as part of day clinics. The benefits of this are that costs for overnight stays in hospital are kept down, the patients can go home on the same day of their procedure and waiting lists are kept as short as possible. Unfortunately, the historical strength of teaching hospitals relied upon patients being *in situ* for several days, allowing students to observe patients, take history information and suggest management options. Students no longer see a huge variety of patients and other strategies must be developed to compensate for the reduction in the teaching and learning experience imposed by the changes in today's healthcare environment. One of the solutions is to develop computer-based materials and using examples from real life, supplement the student's exposure to clinical Medicine.

The evidence supporting interactive educational use for computers is compelling, and is further supported by the use of media that history has shown to be effective replacements or adjuncts for traditional teaching. Combined with a formative assessment strategy, which is also supported by evidence that it is helpful for student learning, we are left with a resource that could guide learning well into the 21st century.

The research questions posed at the beginning of this thesis were

- Can the use of an online formative assessment tool improve student learning outcomes for medical students as measured by traditional examinations? and
- If the use of an online formative tool did improve student learning outcomes, what strategies will encourage its use?

Both questions have been answered to satisfaction. An effective strategy to encourage student learning was developed and implemented. Evidence showed substantial student engagement with the online material and students saw the resource as a key motivator for their learning.

Student outcomes as measured by standard examinations were improved over the period in which they used the online resource. Although it cannot be established that this is a direct causal relationship, comparisons with a previous year, which was in all other aspects identical, showed marked improvement in exam results. The conclusion drawn is that the online resource was the key factor. This may not be due solely to knowledge or understanding gained directly from the educational material directly but also from the motivation that may have been gained to investigate information not in the online cases by utilising textbooks, specialists or by finding patients that would provide deeper understanding of the concepts being taught. Most students provided positive comments on the learning material, as is nearly always the case with many types of educational interventions, and expressed a desire to see CAL introduced into the curriculum. Students saw it as a key motivator in their learning.

Medici has been shown to be an effective implementation of the model developed in Chapter Four and has been established as an effective online formative tool in surgery. There is little reason to suspect it would be less effective in other medical disciplines such as cardiology or ophthalmology and it is quite likely that it will be effective in related fields such as psychiatry and health psychology. This is currently being investigated as part of a competitive research grant awarded to the author from the Australian Learning and Teaching Council and the outcomes will inform future development of cases. The content is intended for students who are experiencing clinical attachments, but lower level content can and has been developed for other year levels, including first years, with a good reception from those students. In addition, *Medici* has been used as a tool to prepare students for experimental teaching sessions in molecular biology with encouraging results so far. It is likely that the application will be transferable to other areas due to its design and underlying pedagogical principles.

The use of *Medici* in the learning environment has been effective, but there are potential areas for expansion of its use. *Medici* has been used as an adjunct to other forms of learning in this research, but it is possible to consider a more integrated role for its use in the future. With a sufficient library of cases, specialists could encourage students to exam specific cases to reinforce concepts and cases seen on the wards. Lecturers could use the tool in the classroom in isolation (Palmer & Devitt, 2006) or in conjunction with voting systems such as Votapedia (urVoting, 2009) or clickers. Indeed this has been done with reported success (Palmer, Devitt, De Young, & Morris, 2005). To encourage greater discussion of concepts introduced in *Medici* cases, discussion boards, blogs or wikis could be used to engage the lecturer, clinicians and the students in a wider dialogue. If the appropriate mechanism for engagement was discovered, this could be a significant enhancement to a useful learning tool.

Other considerations for improving the use of *Medici* might be to examine the use of scoring and weighting systems in greater detail. This is an area that could be of significant interest to medical educators. The nature of feedback delivery is also worthy of investigation. Immediate feedback has been used in this design, where students are supplied with feedback on all of the options they have selected, plus feedback on correct options that they did not consider. An alternative design might provide feedback only on the chosen options, or on all options. There are a wide variety of permutations and an appropriate research design and implementation could be enlightening. Another area of interest would be in providing more personalised feedback to students who answer free response questions. There is little doubt that this additional complexity would create additional pressure on case authors, but it would be useful to determine if the effort would improve learning outcomes for students.

How will changing technology affect this learning tool? Computer processing power will continue to increase, as will storage capacity and memory size and as long as we have enough resources, costs are likely to continue to fall. Educational software has evolved from mainframe-terminal structures to the one used for Medici where cases are retrieved over the web in xml form via a back-end database. It is likely that this structure will be effective for some time to come; yet the main application may well evolve into something significantly different. Programs such as Second Life (Rymaszewski, et al., 2007) may host real-life virtual ward rounds, which would provide substantial interactivity and more realistic levels of simulation. This type of virtual world program has been used to learn in areas such as nutrition, genetics and cardiac auscultation (Boulos, et al., 2007). There is no reason to suspect that it would not be useful in other areas of medical education, including surgery, but one should not underestimate the amount of work required to develop such elaborate scenarios, nor the potentially steep learning curve for students (Wang & Braman, 2009).

Mobile computing is growing as evidenced by the massive sales of Apple's iPad. The screens of these devices are large and allow for the display of media, whereas those of smartphones do not. Portable devices also provide substantial computing power to run interactive applications. It is worth noting that some of these devices do not support the technologies used in this project. Potential further research leading from this project would be to examine student usage of a portable version of *Medici* available for iPad or other equivalent device. It would also be useful to examine a structure of learning combining free response and multiple-choice questions as a method to remove cueing. Similarly a series of Medici cases written to support learning according to script concordance theory would be intriguing. It might also be worthwhile examining the concept of a 'threshold of learning' where too much time can be spent by students learning in one way with no return in learning outcomes.

What's next? There is little reason to doubt that technological change will continue. Greater bandwidth at home and the use of wireless technology combined with larger and more flexible displays will allow many of our students in the future to learn on the move. Overall, the largest benefit gained by technology has been the ability to get the information to students on their terms. Greater access to computers at universities and a larger proportion of machines in student homes or bags means that students have the opportunity to use the teaching resources they choose at a time and place of their convenience. It is quite possible to foresee students completing many problem solving learning exercises in medicine on their portable devices, gathering real life data from remote locations and analysing it on the train or bus. Interaction with fellow students and colleagues will be more remote, but may be more frequent. Our interactions with fellow humans are changing and it will be fascinating to see the effect this has on our portable education in the coming decade.

Bates (2005, p. 48) makes the point that is far easier for individuals to make the potentially expensive leap towards using new technology for learning than institutions which need to have a strategic and more global viewpoint . Yet, pioneers who experimented with new technologies may have made history but had little impact on educational development due to the conservative nature of education (Saettler, 1968). History has shown that this is a likely path for our current innovations, and yet we sit in a moment in time where everything appears to be aligned. Cheap resources, powerful computers and sound pedagogical theories combined with a massive amount of supportive evidence suggests that maybe, this time there *will* be an impact.

Appendices

Appendix 1: Resource Usage Survey

1. What types of resources did you use to help you learn during this attachment and how much did you use them?

The times refer to total time over the whole attachment

Type of resource	More than 10 hours	5-10 hours	2-5 hours	1-2 hours	Less than 1 hour
Text Books					
Paper-based journal articles					
Web-based journal articles					
Ward Rounds					
Lectures					
Tutorials					
Interactive computer aided learning tools					
Non-journal, web-based material					
Other					

2. What value do you attach to these resources?

Type of resource	Very Valuable	Valuable	Of little value	Of no value
Text Books				
Paper-based journal articles				
Web-based journal articles				
Ward Rounds				
Lectures				
Tutorials				
Interactive Computer aided learning tools				
Non-journal, web-based material				
Other				

Appendix 2: Attachment Survey

Yes/No Questions

Did you have any formative assessment during the attachment? (Yes =1, No = 0)

Did you have tutorials with individual appraisal?

Did you have peer review groups?

Did you have case presentations with immediate feedback?

Likert Questions

The online material aided my learning

I only accessed the online material at the end of the week

I would have completed the online material without it being compulsory

The online material was relevant to my attachment

I used the online material for revision

I would like to have more online material

I would like to have the learning material on CD

I accessed Medici every week

Open-ended Questions

How would you improve the online learning material?

What did you like about the online material?

In general, what are the things that motivate you to study?

Please provide examples of things that motivated you during this attachment

Appendix 3: Medici Usage Survey

Likert Questions

Overall, I am satisfied with the quality of the *Medici* cases.

Overall, I am satisfied with the number of *Medici* cases available ONLINE

I would like to have *Medici* available to me on CD

I feel part of a group committed to learning.

The feedback provided to me from *Medici* assists my learning.

I am motivated to learn using *Medici*.

Medici helps me develop my thinking skills (eg. Problem solving, analysis).

I understand the concepts presented using *Medici*.

The flexibility of the on-line format helps me balance study and other commitments.

Open-ended questions

What are the best aspects of *Medici*, and why

Medici could be changed in the following ways to improve my learning

References

A-Z Computing (2004). A- Computer Services - June 2004 Price List Retrieved August 23, 2009, from http://www.azcomputers.com.au/AZ06PL1.pdf

Abdulla, A. M., Watkins, L. O., & Henke, J. S. (1984). The use of natural language entry and laser videodisk technology in CAI. *Academic Medicine, 59*(9), 739-745.

Abrahamson, S., Denson, J., & Wolf, R. (2004). Effectiveness of a simulator in training anesthesiology residents. *British Medical Journal, 13*(5), 395-397.

Ackers, G. W., & Oosthoek, J. K. (1972). The evaluation of an audio-tape mediated course-I. *British Journal of Educational Technology, 3*(2), 136-146.

Adamo, G. (2003). Simulated and standardized patients in OSCEs: achievements and challenges 1992-2003. *Med Teach, 25*(3), 262-270.

Adamson, R. (1915). Fire Administration. *Proceedings of the Academy of Political Science in the City of New York, 5*(3), 66-78.

Adler, M., & Johnson, K. (2000). Quantifying the literature of computer-aided instruction in medical education. *Academic Medicine, 75*(10), 1025-1028.

Adobe (2008a). Adobe - Photoshop CS3 editions Retrieved August, 2007, from http://www.adobe.com/products/photoshop/index.html

Adobe (2008b). Adobe Flash CS3 Professional, Interactive Multimedia, Interactive Design Retrieved August, 2009, from http://www.adobe.com/products/flash/

Adobe (2010a). Adobe - Adobe Shockwave Download Center Retrieved July, 2010, from http://www.adobe.com/shockwave/download/

Adobe (2010b). Adobe - Director 11 : Multimedia Authoring Software, Multimedia Authoring Tool Retrieved November, 2010, from http://www.adobe.com/products/director/

Agard, F. B., Clements, R. J., Hendrix, W. S., Hocking, E., Pitcher, S. L., van Eerden, A., et al. (1945). A survey of language classes in the Army Specialized Training Program. *Modern Language Journal*, 155-160.

Al-Jewair, T., Azarpazhooh, A., Suri, S., & Shah, P. (2009). Computer-assisted learning in orthodontic education: a systematic review and meta-analysis. *Journal of Dental Education, 73*(6), 730-739.

Alavi, C., Loh, S. H., & Reilly, D. (1991). Reality basis for teaching psychomotor skills in a tertiary nursing curriculum. *Journal of Advanced Nursing, 16*(8), 957-965.

Albanese, M., & Mitchell, S. (1993). Problem-based learning: a review of literature on its outcomes and implementation issues. *Acad Med, 68*(1), 52-81.

Albanese, M., & Sabers, D. L. (1988). Multiple true-false items: A study of interitem correlations, scoring alternatives, and reliability estimation. *Journal of Educational Measurement*, 111-123.

Alcott, W. (1843). *Slate and black board exercises*. New York: Mark H. Newman.

Allahbadia, N. K. (1959). Auscultation on the Gramophone. *British Medical Journal, 2*(5162), 1336.

Allan, D. (1991). Tape journals: Bridging the gap between communication and correction. *ELT Journal, 45*(1), 61-66.

Allan, R. A. (2001). *A history of the personal computer: the people and the technology*. London, Ontario: Allan Publishing.

Alldrin, L. (1994, September 2009). Camcorder Reviews. *Videomaker, 1994*.

Alma Mater Studiorum - Università di Bologna (2009). Our History - Università di Bologna Retrieved November, 2009, from http://www.eng.unibo.it/PortaleEn/University/Our+History/default.htm

Alpert, D., & Bitzer, D. (1970). Advances in Computer-Based Education. *Science, 167*(3925), 1582-1590.

Amazon.com (2010). Amazon.com: Kindle Wireless Reading Device, Wi-Fi, 6" Display, Graphite - Latest Generation: Kindle Store Retrieved October, 2010, from http://www.amazon.com/Kindle-Wireless-Reader-Wifi-Graphite/dp/B002Y27P3M/ref=amb_link_354202202_3?pf_rd_m=ATVPDKIKX0DER&pf_rd_s=center-1&pf_rd_r=1R4P0HAHHAQTVW0FG2SX&pf_rd_t=101&pf_rd_p=1277208842&pf_rd_i=507846

Anderson, C. (1962). Technology in American education 1650-1900. New media for instruction 1, OE 34018. *Bulletin*, 19-53.

Anderson, L., Krathwohl, D. R., Airasian, P. W., & Samuel, B. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Anderson, T. (2004). Foundations of educational theory for online learning. In T. Anderson (Ed.), *Theory and practice of online learning* (pp. 45-75). Edmonton: AU Press.

Andrews, P., Schwarz, J., & Helme, R. (1992). Students can learn medicine with computers. Evaluation of an interactive computer learning package in geriatric medicine. *The Medical Journal of Australia, 157*(10), 693-695.

Anfara, J. V., Anfara, V. A., Anfara, A., Andrews, G., & Mertens, S. B. (2005). *The encyclopedia of middle grades education*. Greenwich, CN: Information Age Pub Inc.

Apple Inc (2007a). Apple - iPod classic - Technical Specifications Retrieved August, 2007, from http://www.apple.com/ipod/specs.html

Apple Inc (2007b). Apple-iTunes-iTunes Overview Retrieved June, 2007, from http://www.apple.com/itunes/

Apple Inc (2007c). Apple-iTunes-iTunes Store-Pocasts-Technical Specification Retrieved June, 2007, from http://www.apple.com/itunes/store/podcaststechspecs.html

Apple Inc (2008a). Apple - iLife - iMovie Retrieved August, 2008, from http://www.apple.com/ilife/imovie/

Apple Inc (2008b). Apple - iPhone Retrieved August, 2008, from http://www.apple.com/iphone/

Apple Inc (2010a). Apple - iPad - Buy and read books like never before. Retrieved October, 2010, from http://www.apple.com/ipad/features/ibooks.html

Apple Inc (2010b). Apple - iPad - See the web, email, and photos like never before Retrieved October, 2010, from http://www.apple.com/ipad/

Apple Inc (2010c). Apple - QuickTime - QuickTime Pro Retrieved November, 2010, from http://www.apple.com/quicktime/pro/

Archbald, D. A., & Newmann, F. M. (1988). Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School Available from http://eric.ed.gov:80/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED301587&ERICExtSearch_SearchType_0=no&accno=ED301587

Armes, R. (1988). *On video*. London: Routledge.

Asher, H. (1962). Teaching on tape. *Medical & biological illustration, 12*, 253.

Ashwal, S., Russman, B. S., Blasco, P. A., Miller, G., Sandler, A., Shevell, M., et al. (2004). Practice parameter: diagnostic assessment of the child with cerebral palsy: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology, 62*(6), 851-863.

Association for the Advancement of Artificial Intelligence (2009). AITopics / NaturalLanguage Retrieved August 7, 2009, from http://www.aaai.org/aitopics/html/natlang.html

Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater v. 2. *Journal of Technology, Learning, and Assessment, 4*(3), 1-31.

Australian Learning and Teaching Council (2009). The Medici project: developing a multi-disciplinary, sustainable resource for blended learning initiatives in tertiary medical education | Australian Learning and Teaching Council Retrieved November, 2010, from http://www.altc.edu.au/project-medici-project-adelaide-2009

Australian Private Hospitals Association (2007). Australian Private Hospitals Association - 2007 Facts & Figures Retrieved January, 2009, from http://www.apha.org.au/media_files/2375386209.html

Bachman, M., Lua, M., Clay, D., & Rudney, J. (1998). Comparing traditional lecture vs. computer-based instruction for oral anatomy. *Journal of Dental Education, 62*(8), 587-591.

Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short-answer questions in a marketing context. *Journal of Marketing Education, 25*(1), 31-36.

Ballard, P. B. (1944). *The New Examiner*. London: Habell, Watson and Vinny Ltd.

Baptiste, S. (2003). *Problem-based learning : a self-directed journey*. Thorofare, NJ: SLACK.

Barabas, A. (1965). Blackboard drawing in medical teaching. *British Medical Journal, 1*(5437), 782-784.

Barbey, A., & Barsalou, L. (2009). Reasoning and problem solving: models. *Encyclopedia of Neuroscience, 8*, 35-43.

Barrett, M. J., Kuzma, M. A., Seto, T. C., Richards, P., Mason, D., Barrett, D. M., et al. (2006). The power of repetition in mastering cardiac auscultation. *The American journal of medicine, 119*(1), 73-75.

Barrows, H. (1968). Simulated Patients in Medical Teaching. *Canadian medical Association Journal, 98*, 674-676.

Barrows, H., & Tamblyn, R. (1980). *Problem-based learning: an approach to medical education.* (Vol. 1). New York: Springer Publishing Company.

Basdogan, C., Ho, C. H., & Srinivasan, M. A. (2001). Virtual environments for medical training: Graphical and haptic simulation of laparoscopic common bile duct exploration. *IEEE/ASME Transactions On Mechatronics, 6*(3), 269-285.

Bassett, M. (2004). JVC GR-D33 Camcorder Review Retrieved August 23, 2009, from http://www.camcorderinfo.com/content/jvc-gr-d33-camcorder-review.htm

Bates, A., & Bates, T. (2005). *Technology, e-learning and distance education* (2nd ed.). New York, NY: Routledge.

Bauer, W. (1935). Concerning health education by radio. *California and Western Medicine, 43*(6), 457.

BBC, & The British Museum (1957). Telcan Domestic Video Recorder. *A History of the World* Retrieved October, 2010, from http://www.labguysworld.com/VTR_BirthOf.htm

BBC News (2005). Wordsmiths hail podcast success. *BBC News, Wednesday, 7 December 2005*. Retrieved from http://news.bbc.co.uk/2/hi/technology/4504256.stm

Beard, J. (1860). *The rational primer; or, First reader*. Manchester, UK: J Heywood.

Beck, A. H. (2004). The Flexner report and the standardization of American medical education. *JAMA, 291*(17), 2139-2140.

Bengtsson, M., & Ohlsson, B. (2010). The nursing and medical students motivation to attain knowledge. *Nurse education today, 30*(2), 150-156.

Berci, G., & Davids, J. (1962). Endoscopy and television. *British Medical Journal, 1*(5292), 1610-1613.

Berg, G. (2002). *Why distance learning?: higher education administrative practices*. Westport, CT: Praeger Publishers.

Bernard, E. (1936). Silent Films and Lantern Slides in Teaching French. *Modern Language Journal, 21*(2), 109-115.

Bernard, R., Abrami, P., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., et al. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379-439.

Beun, A. M., van Emde Boas, W., & Dekker, E. (1998). Sharp transients in the sleep EEG of healthy adults: a possible pitfall in the diagnostic assessment of seizure disorders. *Electroencephalogr Clin Neurophysiol, 106*(1), 44-51.

Biggs, J. (1998). Assessment and Classroom Learning: a role for summative assessment? *Assessment in Education: Principles, Policy & Practice, 5*(1), 103-110.

Biggs, J., & Tang, C. (1999). Teaching for quality learning at university Available from http://74.125.155.132/scholar?q=cache:PWEdVVcwSq0J:scholar.google.com/+biggs+1999&hl=en&as_sdt=2001

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.

Biran, L., Biran, L., Dunn, W., & Harden, R. (1985). Using the overhead projector to present patient management problems to groups. *Medical Teacher, 7*(3), 257-269.

Bissell, V., McKerlie, R. A., Kinane, D. F., & McHugh, S. (2003). Teaching periodontal pocket charting to dental students: a comparison of computer assisted learning and traditional tutorials. *British dental journal, 195*(6), 333-336.

Bitzer, D., Lyman, E., & Easley, J. (1965). The Uses of PLATO: A Computer-Controlled Teaching System Available from http://www.ideals.illinois.edu/bitstream/handle/2142/1760/Bitzer3446.pdf?sequence=2

Bitzer, D., Sherwood, B., & Tenczar, P. (1973). Computer-based science education (CERL Report X-37). Urbana: University of Illinois. *Computer Education Research Laboratory*.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan, 86*(1).

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

Black, P., & Wiliam, D. (2003). 'In praise of educational research': formative assessment. *British Educational Research Journal, 29*(5), 623-637.

Blackman, J. A., & Huntley, J. S. (1984). The pediatrician's guide to computer videodiscs. *Pediatrics, 74*(4), 543-547.

Blanchard, A. (1935). Health Information on the Air. *American Journal of Public Health, 25*(10), 1081-1088.

Blank, J. W., Kirk, K. W., & Weinswig, M. H. (1975). Evaluation of an audio cassette tape lecture course. *American journal of pharmaceutical education, 39*(3), 275-278.

Blenheim, L. (1969). Television teaching by professional performers? *Educational technology research and development, 17*(3), 322-326.

Blin, F., & Munro, M. (2008). Why hasn't technology disrupted academics' teaching practices Understanding resistance to change through the lens of activity theory. *Computers & Education, 50*(2), 475-490.

Block, A. (1975). Continuing medical education in small community hospitals. *Western Journal of Medicine, 123*(2), 154-158.

Bloom, B. S. (1956). Taxonomy of educational objectives, handbook I: Cognitive domain. *New York: David McKay*.

Bloom, B. S. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill.

Bloom, S. (1988). Structure and ideology in medical education: an analysis of resistance to change. *J Health Soc Behav, 29*(4), 294-306.

Bock, R. D. (1997). A Brief History of Item Response Theory. *Educational measurement: issues and practice, 16*(4), 21-33.

Bokken, L., Rethans, J., van Heurn, L., Duvivier, R., Scherpbier, A., & van der Vleuten, C. (2009). Students' views on the use of real patients and simulated patients in undergraduate medical education. *Academic Medicine, 84*(7), 958-963.

Bolger, M. (1995). Cyber Patient (TM): A Multimedia Pharmacokinetic Simulation Program for Case Study Generation in a Problem-Solving Curriculum. *American Journal of Pharmaceutical Education, 59*(4), 409-416.

Bond, W., Deitrick, L., Arnold, D., Kostenbader, M., Barr, G., Kimmel, S., et al. (2004). Using simulation to instruct emergency medicine residents in cognitive forcing strategies. *Academic Medicine, 79*(5), 438-446.

Bonner, T. N. (1995). *Becoming a Physician: Medical education in Britain, France, Germany, and the United States, 1750–1945*. New York: Oxford University Press.

Borrett, L. (1984). Your IBM Computer, Jan-1984. *Your Computer Magazine* Retrieved August 20, 2009, from http://www.borrett.id.au/computing/art-1984-01-01.htm

Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education, 22*(2), 151-167.

Boud, D., & Feletti, G. (1997). Changing Problem-Based Learning. Introduction to the Second Edition. In D. Boud & G. Feletti (Eds.), *The Challenge of Problem-Based Learning* London: Kogan Page.

Boulos, M., Hetherington, L., & Wheeler, S. (2007). Second Life: an overview of the potential of 3 D virtual worlds in medical and health education. *Health Information & Libraries Journal, 24*(4), 233-245.

Boulos, M. N., Maramba, I., & Wheeler, S. (2006). Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education. *BMC Med Educ, 6*, 41. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16911779

Bowman, R. C. (2002). Osler and Rural Practice Retrieved April, 2009, from http://www.unmc.edu/Community/ruralmeded/osler.htm

Bramwell, B., & Murray, R. M. (1888). A Method of Graphically Recording the Exact Time-Relations of Cardiac Sounds and Murmurs. *British Medical Journal, 1*(1410), 10-16.

Brancolini, K. R. (1998). New and emerging video technologies: A status report. *October 29*. Retrieved from http://www.dlib.indiana.edu/~brancoli/videostatus.html

Brayton, D. (1967). Postgraduate education in medicine by television. *Biomedical sciences instrumentation, 3*, 81-88.

Bresnitz, E., Gracely, E., & Rubenstein, H. (1992). A randomized trial to evaluate a computer-based learning program in occupational lung disease. *Journal of Occupational and Environmental Medicine, 34*(4), 422-427.

Briggs, A., & Burke, P. (2009). *Social history of the media: from Gutenberg to the Internet* (3rd ed.). Cambridge, UK: Polity Press.

Brigham, C., & Kamp, M. (1974). The current status of computer-assisted instruction in the health sciences. *Academic Medicine, 49*(3), 278-279.

Brock, P. (1994). *Educational technology in the classroom*. New Jersey, USA: Educational Technology Pubns.

Brong, R. (1964). Is Television The Answer? *AJN The American Journal of Nursing, 64*(1), 77-79.

Brooks, V. (2004). Double marking revisited. *British Journal of Education Studies, 11*, 29-46.

Brown, A., & Green, T. (2007). Video podcasting in perspective: The history, technology, aesthetics, and instructional uses of a new medium. *Journal of educational technology systems, 36*(1), 3-17.

Brown, D. W., Groome, D. S., Niehoff, R. D., & Cleaveland, J. D. (1968). Computer-assisted instruction in nuclear medicine. *Jama, 206*(5), 1059-1062.

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

Brown, G., Manogue, M., & Martin, M. (1999). The validity and reliability of an OSCE in dentistry. *European journal of dental education: official journal of the Association for Dental Education in Europe, 3*(3), 117-125.

Brownlow, K. (1968). *The parade's gone by*. Los Angeles: University of California Press.

Bryant, P., Hartley, S., Coppola, W., Berlin, A., Modell, M., & Murray, E. (2003). Clinical exposure during clinical method attachments in general practice. *Med Educ, 37*(9), 790-793.

Burge, E., & Howard, J. (1990). Audio-conferencing in graduate education: A case study. *American Journal of Distance Education, 4*(2), 3-13. Retrieved from http://www.ed.psu.edu/acsde/deos/deosnews/deosnews1_3.asp

Burkitt, H. G., Quick, C. R. G., Reed, J. B., & Deakin, P. J. (2007). *Essential Surgery: Problems, Diagnosis and Management*. Oxford, England: Churchill Livingstone.

Burton, R. F. (2002). Misinformation, partial knowledge and guessing in true/false tests. *Med Educ, 36*(9), 805-811.

Business Wire (1994). HP offers $150 rebate on color scanners. *Business Wire* Retrieved August 23, 2009, from http://findarticles.com/p/articles/mi_m0EIN/is_1994_Nov_1/ai_15873017/

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology, 79*(4), 474-482.

Cajiao, V. (2006). ITunes Podcasts Grow 10 Fold In Just Over A Year. Let's Leverage The Power. *Typical Mac User*. Retrieved from http://typicalmacuser.com/wordpress/2006/09/24/itunes-podcasts-grow-10-fold-in-just-over-a-year/

Cambridge Esol (2007). Cambridge IELTS 6 Audio CDs: Examination papers from University of Cambridge ESOL Examinations (Ielts Practice Tests) (No. 6): Cambridge University Press.

Campbell, J. (1994). *The Roman Army, 31 BC-AD 337: A Sourcebook*. New York: Routledge.

Campbell, J., & Johnson, C. (1999). Trend spotting: fashions in medical education. *BMJ, 318*(7193), 1272.

Carle, W. M. (1959). Needed Research in School Broadcasting. *Educational Research Bulletin*, 120-130.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Thousand Oaks: Sage Publications.

Carroll, K., Iedema, R., & Kerridge, R. (2008). Reshaping ICU ward round practices using video-reflexive ethnography. *Qualitative Health Research, 18*(3), 380-390.

Case, S., & Swanson, D. (1996). Constructing written test questions for the basic and clincial sciences. Retrieved August, 2008, from http://www.nbme.org/PDF/ItemWriting_2003/2003IWGwhole.pdf

Casio (2004). Digital Camera Reviews, Ratings, Tips and Comparisons Retrieved August 23, 2009, from http://www.livingroom.org.au/photolog/2004/10/index.php

Cavazzi, F. (nd). The Roman Empire Retrieved August, 2008, from http://www.roman-empire.net/army/

Ceruzzi, P. (2003). *A history of modern computing*. Salisbury, UK: The MIT press.

Chambers, R. (1840). *First book of drawing: being exercises for children and young persons on the slate and black board*. Edinburgh: Sorin & Ball.

Chan, C. C., Tsui, M. S., Chan, M. Y. C., & Hong, J. H. (2002). Applying the Structure of the Observed Learning Outcomes (SOLO) Taxonomy on Student s Learning Outcomes: an empirical study. *Assessment & Evaluation in Higher Education, 27*(6), 511-527.

Chapman, M. (2006). Podcasting: Who's Tuning In? *eMarketer*. Retrieved from http://www.emarketer.com/Report.aspx?code=podcasting_mar06

Chaput De Saintonge, D. M., & Pavlovic, A. (2004). Cheating. *Medical education(Oxford. Print), 38*(1), 8-9.

Charlin, B., Roy, L., Brailovsky, C., Goulet, F., & van der Vleuten, C. (2000). The Script Concordance test: a tool to assess the reflective clinician. *Teach Learn Med, 12*(4), 189-195.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE bulletin, 39*(7), 3-7.

Chinn, M. D., & Fairlie, R. W. (2010). ICT Use in the Developing World: An Analysis of Differences in Computer and Internet Penetration. *Review of International Economics, 18*(1), 153-167. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9396.2009.00861.x/full

Cipriani, C. (1912). The Use of Phonetics and the Phonograph in the Teaching of Elementary French. *The School Review, 20*, 516-525.

Clancey, W., Shortliffe, E., & Buchanan, B. (1979, October 17). *Intelligent computer-aided instruction for medical diagnosis.* Paper presented at the AMWA.

Clark, D. (2008). Nursing students to pay for new lab - Las Vegas Sun Retrieved April, 2009, from http://www.lasvegassun.com/news/2008/dec/18/nursing-students-pay-new-lab/

Clark, R., & Raffin, T. (1992). Efficacy of computers in teaching arterial blood gas analysis. *Academic Medicine, 67*(6), 365-366.

Clarke, C. C. (1918). The phonograph in modern language teaching. *Modern Language Journal*, 116-122.

Clifton, J. L. (1930). Education by Radio. *Educational Research Bulletin*, 199-201.

Cockshott, W. P. (1973). Audio-visual systems in support of radiology education for medical students. *Clinical radiology, 24*(1), 25-27.

Cohen, L., Manion, L., & Morrison, K. (1996). *A guide to teaching practice*. Evanston, IL: Routledge.

Cohen, P., & Dacanay, L. S. (1992). Computer-based instruction and health professions education: a meta-analysis of outcomes. *Evaluation & the Health Professions, 15*(3), 259-281.

Cohen, P., Ebeling, B. J., & Kulik, J. A. (1981). A meta-analysis of outcome studies of visual-based instruction. *Educational Technology Research and Development, 29*(1), 26-36.

Cohen, R., Reznick, R. K., Taylor, B. R., Provan, J., & Rothman, A. (1990). Reliability and validity of the objective structured clinical examination in assessing surgical residents. *American journal of surgery, 160*(3), 302-305.

Collins, J. (2006). Education techniques for lifelong learning: writing multiple-choice questions for continuing medical education activities and self-assessment modules. *Radiographics, 26*(2), 543-551.

Colliver, J. (2000). Effectiveness of problem-based learning curricula: research and theory. *Acad Med, 75*(3), 259-266.

Colliver, J. (2001). Are qualitative studies of the PBL tutorial process indicated? *Acad Med, 76*(3), 215-216.

Connectus Pty Ltd (2008). Need Broadband Help - Broadband Guide can help Retrieved August, 2008, from http://broadbandguide.com.au/broadbandhelp/broadbandhelp

Connor, D. V. (1970). Educational Technology in Australia. *British Journal of Educational Technology, 1*(3), 207-216.

Cook, D., Levinson, A., Garside, S., Dupras, D., Erwin, P., & Montori, V. (2008). Internet-based learning in the health professions: a meta-analysis. *Jama, 300*(10), 1181-1196.

Cooper, J., & Taqueti, V. (2004). A brief history of the development of mannequin simulators for clinical education and training. *British Medical Journal, 13*(Supplement 1), i11-i18.

Copley, J. (2007). Audio and video podcasts of lectures for campus-based students: production and evaluation of student use. *Innovations in education and teaching international, 44*(4), 387-399.

Corliss, R. (2001). That Old Feeling: Mercury, God of Radio - TIME. *Time*. Retrieved from http://www.time.com/time/arts/article/0,8599,172672,00.html

Corrêa, L., De Campos, A., Souza, S., & Novelli, M. (2003). Teaching oral surgery to undergraduate students: a pilot study using a Web based practical course. *European Journal of Dental Education, 7*(3), 111-115.

Cosmann, R. (1996). The Evolution of Educational Computer Software. *Education, 116*(4), 619-623.

Crampton, A., Vanniasinkam, T., & Ragusa, A. (2008). *Microbial vodcasting-supplementing laboratory time with vodcasts of key microbial skills.* Paper presented at the UniServe Science Proceedings Visualisation, University of Sydney.

Creutzfeldt, J. (2010). Exploring Virtual Worlds for Scenario-Based Repeated Team Training of Cardiopulmonary Resuscitation in Medical Students. *J Med Internet Res, 12*(3), e38. Retrieved from http://www.jmir.org/2010/3/e38/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%253A+JMedInternetRes+%2528Journal+of+Medical+Internet+Research+%2528atom%2529%2529&utm_content=Google+Reader

Crisp, G. T. (2009). *HERDSA Guide. Designing and using e-assessments*. Milperra, N.S.W: Higher Education Research and Development Society of Australasia.

Crisp, G. T., & Palmer, E. J. (2009). Engaging academics with a simplified analysis of their multiple-choice question (MCQ) assessment results. *Journal of University Teaching and Learning Practice, 4*(2), 88-106.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Cronbach, L. (1984). Elements of psychological testing: New York: Harper & Row.

Cronbach, L., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychol. Bull., 52*, 281-302.

Cronin, J. F. (1993). Four Misconceptions about Authentic Learning. *Educational Leadership, 50*(7), 78-80.

Croninger, W., Tumiel, J., & Sowa, T. (1995). The effects of a Hypercard tutorial on student learning of biomechanics. *The American journal of occupational therapy.: official publication of the American Occupational Therapy Association, 49*(5), 456-460.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Crowley, R., & Medvedeva, O. (2006). An intelligent tutoring system for visual classification problem solving. *Arificial Intelligence in Medicine, 36*(1), 85-117.

Crowley, R., Naus, G., Stewart, J., & Friedman, C. (2003). Development of visual diagnostic expertise in pathology-An information-processing study. *British Medical Journal, 10*(1), 39-51.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. New York: Teachers College Press.

Culbert, D. (1979). Education Unit in World War II: An Interview with Erik Barnouw. *Journal of Popular Culture, 12*(2), 275-284.

Curtis, S., & Lucas, R. (2001). A coincidence of needs?: Employers and full-time students. *Employee Relations, 23*(1), 38-54

da Cruz, F. (2003). Teletype Machines. *Comumbia University Computing History* Retrieved October 29, 2010, from http://www.columbia.edu/acis/history/teletype.html

Dahlin, M., Joneborg, N., & Runeson, B. (2007). Performance-based self-esteem and burnout in a cross-sectional study of medical students. *Medical Teacher, 29*(1), 43-48.

Dahlin, M., & Runeson, B. (2007). Burnout and psychiatric morbidity among medical students entering clinical training: a three year prospective questionnaire and interview-based study. *BMC Medical Education, 7*(1), 6. Retrieved from http://www.biomedcentral.com/1472-6920/7/6

Dashiell, J. F. (1928). *Fundamentals of Objective Psychology*. Boston: Houghton Mifflin Company.

Data Recovery Resources (2004). Hard Drive Pricing Prices of Disk Drives, from http://www.dtidata.com/resourcecenter/hard-drive-pricing/

Davies, M., Gale, J., & Clarke, W. (1977). Audiotape and booklet self-instructional materials in physiology: An evaluation of their effectiveness and acceptability in the pre-clinical curriculum. *Medical Education, 11*, 370-373.

Davis, J. (1998). High monitor prices buck PC trend. *CNET News* Retrieved August 23, 2009, from http://news.cnet.com/High-monitor-prices-buck-PC-trend/2100-1001_3-208528.html

Davis, M., & Gore, R. (2001). Determinants of cardiac function: simulation of a dynamic cardiac pump for physiology instruction. *Advances in Physiology Education, 25*(1), 13-35.

De Dombal, F. T., Hartley, J. R., & Sleeman, D. H. (1969a). A computer-assisted system for learning clinical diagnosis. *Lancet, 1*(7586), 145-148.

De Dombal, F. T., Hartley, J. R., & Sleeman, D. H. (1969b). Teaching surgical diagnosis with the aid of a computer. *Br J Surg, 56*(10), 754-757.

De Garay, A. (1910). Teaching of Hygiene in Normal Schools. *American journal of public hygiene, 20*(3), 477-481.

De Jong, T. (2006). Computer simulations: Technological advances in inquiry learning. *Science, 312*(5773), 532-533.

de Jong, T., Van Andel, J., Leiblum, M., & Mirande, M. (1992). Computer assisted learning in higher education in The Netherlands: a review of findings. *Computers & Education, 19*(4), 381-386.

De Muth, J., & Bruskiewitz, R. (2006). A comparison of the acceptability and effectiveness of two methods of distance education: CD-ROM and audio teleconferencing. *American Journal of Pharmaceutical Education, 70*(1), 1-8.

Dean, M., Andrew, G., Shah, M., Svati, P., Churchill, M., & Jeanetta, E. (1998). DoEpi:: Computer-Assisted Instruction in Epidemiology and Computing and a Framework for Creating New Exercises. *American journal of preventive medicine, 14*(4), 367-371.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26*(3), 325-346.

DenBeste, M. (2003). Power Point, Technology and the Web: More Than Just an Overhead Projector for the New Century? *History Teacher, 36*(4), 491-504.

Denniston, C. (2008). Reflections of a mid-career clinician educator--fanning the flames for teaching. *Fam Med, 40*(2), 134-135.

Dent, E. C. (1947). The Audio-Visual Aids of Tomorrow. *Music Educators Journal*, 22-38.

Denton, L. F., Doran, M. V., & McKinney, D. (2002). *Integrated Use of Bloom and Maslow for Instructional Success in Technical and Scientific Fields* Paper presented at the 2002 American Society for Engineering Education Annual Conference & Exposition from http://www.cis.usouthal.edu/~mckinney/session2793.PDF

Department of English, U. o. T. (1997). Edmund Coote's English School-maister (1596) Retrieved August 10, 2009, from http://www.library.utoronto.ca/utel/ret/coote/ret2.html

Detweiler, D. K. (1964). Genetic aspects of cardiovascular diseases in animals. *Circulation, 30*(1), 114-127.

Dev, P., Hoffer, E., & Barnett, G. (2001). Computers in Medical Education. In E. Shortliffe & L. Perreault (Eds.), *Biomedical Informatics Computers in Medical Education* (pp. 737-762). New York: Springer.

Devitt, P. G., Barker, J., Mitchell, J., & Hamilton-Craig, C. (2003). *Clinical Problems in General Medicine and Surgery*. Edinburgh : New York: Churchill Livingstone.

Devitt, P. G., & Palmer, E. J. (1998). Computers in medical education 1: evaluation of a problem-orientated learning package. *ANZ Journal of Surgery, 68*(4), 284-287.

Devitt, P. G., & Palmer, E. J. (1999). Computer-aided learning: an overvalued educational resource? *Medical Education, 33*(2), 136-139.

Devitt, P. G., Smith, J. R., & Palmer, E. J. (2001). Improved student learning in ophthalmology with computer-aided instruction. *Eye, 15*(Pt 5), 635-639.

Devitt, P. G., Worthley, S., Palmer, E., & Cehic, D. (1998). Evaluation of a computer based package on electrocardiography. *Aust N Z J Med, 28*(4), 432-435.

Diekhoff, G. M., Labeff, E. E., Shinohara, K., & Yasukawa, H. (1999). College cheating in Japan and the United States. *Research in Higher Education, 40*(3), 343-353.

Dillenbourg, P., & Self, J. (1992). People power: A human-computer collaborative learning system. *Lecture Notes in Computer Science, 608*, 651-660.

Dja65 (nd). Vintage Gramophone Isolated. Clipping Path Included. Stock Photo 60486508 : Shutterstock. In shutterstock_60486508.jpg (Ed.) (Vol. 4328x4636, pp. Phonograph): Shutterstock Images.

Dochy, F., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation, 23*(4), 279-298.

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.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction, 13*(5), 533–568.

Donsky, H. J. (1971). Lecture tapes in medical education. *International Journal of Dermatology, 10*(2), 104-106.

Downing, S. M. (2003). Guessing on selected-response examinations. *Med Educ, 37*(8), 670-671.

Downing, S. M. (2004). Reliability: on the reproducibility of assessment data. *Medical Education, 38*(9), 1006-1012.

Downing, S. M., Baranowski, R. A., Grosso, L. J., & Norcini, J. J. (1995). Item type and cognitive ability measured: the validity evidence for multiple true-false items in medical specialty certification. *Applied measurement in Education, 8*(2), 187-197.

Doyle, D. (2009). Simulation in medical education: Focus on Anesthesiology. *Medical Education Online, 7*(16), 1-15. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.7.1011&rep=rep1&type=pdf

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, California: Corwin Press.

Ebel, R. L. (1970). The Case for True-False Test Items *The School Review, 78*(3), 373-389.

Educational Testing Service (2009). Research: Automated Scoring Capabilities Retrieved August, 2009, from http://www.ets.org/research/erater.html

Educause (2007). Educause: Education Learning Initiative. *Educause*. Retrieved from www.educause.edu/eli/sessions

Eisenstein, E. L. (1979). *The printing press as an agent of change: Communications and cultural transformations in early modern Europe*. Cambridge, UK: Cambridge University Press.

Eisenstein, E. L. (2005). *The printing revolution in early modern Europe*. Cambridge, UK: Cambridge University Press.

Elizondo-Omana, R. E., Morales-Gomez, J. A., Guzman, S. L., Hernandez, I. L., Ibarra, R. P., & Vilchez, F. C. (2004). Traditional teaching supported by computer-assisted learning for macroscopic anatomy. *The Anatomical Record*(1), 18-22.

Ellenbogen, E., Encyclopedia of education, Second edition. New York: Macmillan. (2002). *Museum as an educational institution*.

Encarta, M. U.S. Army Alpha Intelligence Test: Sample Questions - Table - MSN Encarta Retrieved August, 2008, from http://encarta.msn.com/media_461550852_761570026_-1_1/u_s_army_alpha_intelligence_test_sample_questions.html

Engel, C. (2000). Medical Education in Australia, New Zealand and Great Britain in the 21st Century. In L. H. Distlehorst, G. L. Dunnington & J. R. Folse (Eds.), *Teaching and Learning in Medical and Surgical Education: Lessons Learned for the 21st Century*. Mahwah, NJ: Lawrence Erlbaum Associates.

Engel, C., & Clarke, R. (1979). Medical Education with a Difference. *Innovations in Education and Teaching International, 16*(1), 70-87.

Entwistle, N. J., & Brennan, T. (1971). The Academic Performance of Students. 2. Types of Successful Students. *British Journal of Educational Psychology, 41*(3), 268-276.

Epstein, R. (2004). Learning from the problems of problem-based learning. *BMC Med Educ, 4*, 1. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14713320

Epstein, R. (2007). Assessment in medical education. *New England Journal of Medicine, 356*(4), 387-396.

Erwin, T., & Wise, S. (2002). A scholar-practitioner model for assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment*. San Francisco, CA: Jossey-Bass.

Essex, D., & Sorlie, W. (1979). Effectiveness of instructional computers in teaching basic medical sciences. *Medical Education, 13*(3), 189-193.

Essex-Lopresti, M. (1979). Use an Overhead Projector. *Medical Teacher, 1*(1), 9-15.

Eva, K. (2005). What every teacher needs to know about clinical reasoning. *Medical Education, 39*(1), 98-106.

Ewbank, H. L. (1930). A Wisconsin Experiment in Radio Teaching. *Music Supervisors Journal*, 29-29.

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and Psychological Measurement, 58*(3), 357-381.

Farber, S. (1966). Continuing Medical Education: Medical Radio Conferences. *California Medicine, 104*(2), 141.

Farrar, J. (2006). Book Reviews. Tropical Infectious Diseases: Principles, Pathogens, and Practice. *The New England Journal of Medicine, 355*(22), 2379-2380.

Fawver, A., Branch, C., Trentham, L., Robertson, B., & Beckett, S. (1990). A comparison of interactive videodisc instruction with live animal laboratories. *Advances in Physiology Education, 259*(6), S11-S14.

Feldhusen, J. F. (2003). Lewis M. Terman: a Pioneer in the Development of Ability Tests. In B. J. Zimmerman & D. H. Schunk (Eds.), *Educational Psychology: A Century of Contributions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Feletti, G. I., & Smith, E. K. M. (1986). Modified Essay Questions: are they worth the effort? *Medical education, 20*(2), 126-132.

Fenderson, B. A., Damjanov, I., Robeson, M. R., Veloski, J. J., & Rubin, E. (1997). The virtues of extended matching and uncued tests as alternatives to multiple choice questions. *Human pathology, 28*(5), 526-532.

Ferguson, K. J. (2006). Beyond multiple-choice questions: using case-based learning patient questions to assess clinical reasoning. *Medical education, 40*(11), 1143-1143.

Fietze, S. (2009). Podcast in higher education: Students usage behaviour. *Same places, different spaces. Proceedings ascilite Auckland 2009*.

Fildes, R. (1935). Blackboards and Their Use. *The Elementary School Journal, 35*(10), 760-767.

Finley, S., Nanton, C., & Roy, P. (1998). Auscultation of the heart: a trial of classroom teaching versus computer-based independent learning. *Medical Education, 32*(4), 357-361.

Finucane, P. M., Johnson, S. M., & Prideaux, D. J. (1998). Problem-based learning: its rationale and efficacy. *Med J Aust*(168), 445-448.

Flanagan, J., & Berger, D. (1948). Research for or by the Armed Forces. *Review of Educational Research, 18*(6), 615-641.

Fleckman, P., Lee, J., & Astion, M. (1997). Nail-Tutor: an image-based personal computer program that teaches the anatomy, patterns of pathology, and disorders of the nails. *Dermatology online journal, 3*(2), 2. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9452368

Flexner, A. (1910). *Medical Education in the United States and Canada*. New York, NY: Carnegie Foundation for the Advancement of Teaching.

Florell, R. J., & Nugent, R. W. (1985). The Videodisc: More Than a Toy. *Journal of Extension, 23*(1). Retrieved from http://www.joe.org/joe/1985spring/a5.php

Ford, J., Bosworth, D., & Wilson, R. (1995). Part-time work and full-time higher education. *Studies in Higher Education, 20*(2), 187-202.

Forrest, D., Ryan, J., Glavin, R., & Merritt, H. (1974). Through the viewing tube: Videocassette psychiatry. *American Journal of Psychiatry, 131*(1), 90-94.

Foster, D., & Poppers, A. (2009). Using Formative Assessment to Drive Learning.

Fournier, J. P., Demeester, A., & Charlin, B. (2008). Script Concordance Tests: Guidelines for Construction. *BMC Medical Informatics and Decision Making, 8*(1), 18.

Fowell, S. L., & Bligh, J. G. (1998). Recent developments in assessing medical students. *Postgrad Med J, 74*(867), 18-24.

Friedman, R., Bruce, D., MacLowry, J., & Brenner, V. (1973). Computer-assisted identification of bacteria. *American journal of clinical pathology, 60*(3), 395-403.

Frisbie, D. A. (1973). Multiple Choice versus True-False: A Comparison of Reliabilities and Concurrent Validities *Journal of Educational Measurement, 10*(4), 297-304.

Frisbie, D. A., & Becker, D. F. (1991). An Analysis of Textbook Advice About True-False Tests. *Applied Measurement in Education, 4*(1), 67-84.

Frisbie, D. A., & Sweeney, D. C. (1982). The relative merits of multiple true-false achievement tests. *Journal of Educational Measurement*, 29-35.

Frontline (2008). frontline: secrets of the sat: where did the test come from?: the 1901 college board Retrieved August, 2008, from http://www.pbs.org/wgbh/pages/frontline/shows/sats/where/1901.html

Fry, E., Bryan, G., & Rigney, J. (1960). Teaching machines: An annotated bibliography. *Audio Visual Communication Review, 8*(2), 1-2,5-79.

Furness, E. L. (1952). Techniques for the Teaching of Listening. *Modern Language Journal*, 124-128.

Gable, A., & Page, C. (1980). The use of artificial intelligence techniques in computer-assisted instruction: an overview. *International Journal of Man-Machine Studies, 12*(3), 259-282.

Gale, M., & Clarke, W. D. (1977). Audiotape and booklet self-instructional materials in physiology: an evaluation of their effectiveness and acceptability in the pre-clinical curriculum. *Med Educ, 11*, 370-373.

Galloway, D., Cantlie, J., Graham, J., & Sambon, L. (1905). The Popular Lecture. *The British Medical Journal, 2*(2327), 283-284.

Garfinkel, A. (1972). Teaching languages via radio: A review of resources. *Modern Language Journal*, 158-162.

Gaskins, R. (2010). Home Page for Robert Gaskins Retrieved November 25, 2010, from http://www.robertgaskins.com/#powerpoint-history

Gates, A. I. (1921). The True-False Test as a Measure of Achievement in College Courses. *Journal of Educational Psychology, 12*(5), 276-287.

Gayeski, D. (1990). Why information technologies fail. In L. Lipsitz (Ed.), *Telecommunications for learning* (Vol. 3, pp. 159-167). Englewood Cliffs, New Jersey: Educational Technology Publications Inc.

Gelula, M. (1997). Working with slides and transparencies for presentations. *Surgical neurology, 47*(3), 308-312.

General Medical Council (1993). *Tomorrow's Doctors: Recommendations on undergraduate medical education*. London: General Medical Council.

Genereux, R. L., & McLeod, B. A. (1995). Circumstances surrounding cheating: A questionnaire study of college students. *Research in Higher Education, 36*(6), 687-704.

Genova, T. (nd). Television History - The First 75 Years. Retrieved from http://www.tvhistory.tv/index.html

George, J. H., & Doto, F. X. (2001). A Simple Five-step Method for Teaching Clinical Skills *Family Medicine, 33*(8), 577-578.

Gervais, R. (2010). Ricky Gervais... Obvioulsy. *rickygervais.com*. Retrieved from http://www.rickygervais.com/podcast.php

Gibbs, G., & Simpson, C. (2004). Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education, 1*, 3-31.

Gibbs, G., Simpson, C., & Macdonald, R. (2003). *Improving student learning through changing assessment–a conceptual and practical framework.* Paper presented at the European Association for Research into Learning and Instruction Conference, Padova, Italy.

Gibson, D., & Campbell, R. M. (2000). Promoting effective teaching and learning: hospital consultants identify their needs. *Med Educ, 34*(2), 126-130.

Gibson, S., Fyock, C., Grimson, E., Kanade, T., Kikinis, R., Lauer, H., et al. (1997). *Simulating arthroscopic knee surgery using volumetric object representations, real-time volume rendering and haptic feedback.*

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. . *Review of Educational Research, 75*, 27-61.

Gilbert, K. (2005). *Medici power and patronage under Cosimo the elder and Lorenzo the magnificent.* Unpublished Open Access, Eastern Michigan University.

Gilchrest, B. A., Deykin, D., & Bleich, H. L. (1972). A computer-based teaching program in hemostasis. *Blood, 40*(4), 560-567.

Ginsburg, C. (1957). The Birth of Video Recording Retrieved October 28, 2010, from http://www.labguysworld.com/VTR_BirthOf.htm

Gizmo Highway (2005). The history of the Laserdisc. *GizmoHighway*. Retrieved from http://www.gizmohighway.com/history/laserdisc.htm

Gleason, C. W. (1926). An experiment in testing engine lathe aptitude. *Journal of Applied Psychology, 10*(2), 171-179.

Gorsky, P., & Caspi, A. (2005). Dialogue: A theoretical framework for distance education instructional systems. *British Journal of Educational Technology, 36*(2), 137-144.

Gosch, M. (1940). Foreign Languages in Junior Colleges. *The Modern Language Journal, 25*(1), 44-47.

Grant, J. (1907). Experiments and Results in Wireless Telephony. *The American Telephone Journal, 26*, 49–51.

Graves, J. (1965). The use of tape in medical teaching. *British Medical Journal, 41*(474), 208-212.

Greenhalgh, T. (2001). Computer assisted learning in undergraduate medical education (Vol. 322, pp. 40-44).

Griffith, J. (1988). Measurement of Group Cohesion in U.S. Army Units. *BASIC AND APPLIED SOCIAL PSYCHOLOGY, 9*(2), 149-171.

Gronow, P. (1983). The Record Industry: The Growth of a Mass Medium. *Popular Music, 3*, 53-75.

Gross, J., & Dobozy, E. (2009). *Adding value to first year student learning with embedded library pod/vodcasts*. Paper presented at the Ascilite. Retrieved December 2010, from http://www.ascilite.org.au/conferences/auckland09/procs/gross.pdf

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*(3), 67-86.

Haladyna, T. M., & Downing, S. M. (1989a). A Taxonomy of Multiple-Choice Item-Writing Rules. *Applied Measurement in Education, 2*(1), 37-50.

Haladyna, T. M., & Downing, S. M. (1989b). Comparative Review: Validity of a Taxonomy of Multiple-Choice Item-Writing Rules. . *Applied Measurement in Education, 2*(1), 51-78.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. . *Applied Measurement in Education, 15*(3), 309-334.

Hall, J. (1947). Solving verbal arithmetic problems without pencil and paper. *The Elementary School Journal, 48*(4), 212-217.

Hall, R. (1969). Audio visual aids in medical education. *British Medical Journal, 4*(5674), 40-41.

Hall, V. C. (2003). Educational Psychology from 1890 to 1920. In B. J. Zimmerman & D. H. Schunk (Eds.), *Educational Psychology: A Century of Contributions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hammer, C. (1945). Enlivening the Scientific German Class. *German Quarterly, 18*(3), 109-115.

Hammond, A. (1972). Computer-Assisted Instruction: Two Major Demonstrations. *Science, 176*(4039), 1110-1112.

Hancock, G. R. (1994). Cognitive Complexity and the Comparability of Multiple-Choice and Constructed-Response Test Formats. *Journal of Experimental Education, 62*(2), 143-158.

Hands, W. J., Scholes, P. A., & Whittaker, W. G. (1934). The Gramophone in Schools. *The Musical Times, 75*(1092), 164-165.

Hansen, D. (1966). Chapter VII: Computer Assistance with the Educational Process. *Review of Educational Research, 36*(5), 588-603. Retrieved from http://www.jstor.org/stable/1169483

Haque, S., & Srinivasan, S. (2006). A meta-analysis of the training effectiveness of virtual reality surgical simulators. *Information Technology in Biomedicine, IEEE Transactions on, 10*(1), 51-58.

Harden, R. M. (1999). Stress, pressure and burnout in teachers: Is the swan exhausted? *Med Teach, 21*(3), 245-247.

Harden, R. M., Stevenson, M., Downie, W. W., & Wilson, G. M. (1975). Assessment of clinical competence using objective structured examination. *British Medical Journal, 1*(5955), 447-451.

Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning*. London, UK: Sage Pubns Ltd.

Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education, 10*(2), 169-207.

Harper, D. (2001). Online Etymology Dictionary Retrieved August, 2008, from http://www.etymonline.com/index.php

Harrower-Erickson, M. R. (1943). A Multiple Choice Test for Screening Purposes *Psychosom Med*(5), 331-341.

Harshbarger, H. C. (1941). Education Speaks. *The Journal of Educational Sociology*, 513-520.

Hawkins, H. (1960). *Pioneer: a History of the Johns Hopkins University, 1874-1899*. Ithaca, NY: Cornell University Press.

Heer, I., Middendorf, K., Miller-Egloff, S., Dugas, M., & Strauss, A. (2004). Ultrasound training: the virtual patient. *Ultrasound in Obstetrics and Gynecology, 24*(4), 440-444.

Henderson, E. S. (1980). The Essay in Continuous Assessment *Studies in Higher Education, 5*(2), 197-203.

Henly, D. (2003). Use of Web based formative assessment to support student learning in a metabolism/nutrition unit. *European Journal of Dental Education, 7*(3), 116-122.

Herriot, A., Bishop, J., & Truby, H. (2004). The development and evaluation of Student Training, Education and Practice for Dietetics CD ROM: a computer assisted instruction programme for dietetic students. *Journal of Human Nutrition and Dietetics, 17*(1), 35-41.

Hewlett-Packard Development Company (2009). Compaq Presario 600 Desktop PC series Retrieved August 20, 2009, from http://h20000.www2.hp.com/bizsupport/TechSupport/Document.jsp?lang=en&cc=us&taskId=120&prodSeriesId=96395&prodTypeId=12454&objectID=c00009599

Highley, S. (1863). The Application of Photography to the Magic Lantern, Educationally Considered. *Journal of the Society of Arts, 11*(530), 141-147.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99-107.

Hoban, C., & Van Ormer, E. (1970). *Instructional film research 1918-1950*. USA: Arno press Inc.

Hodges, B., Regehr, G., McNaughton, N., Tiberius, R., & Hanson, M. (1999). OSCE checklists do not capture increasing levels of expertise. *Academic Medicine, 74*(10), 1129-1134.

Hoffer, E. P., Barnett, G. O., & Farquhar, B. B. (1972). Computer simulation model for teaching cardiopulmonary resuscitation. *Academic Medicine, 47*(5), 343-347.

Hofstadter, R., & Hardy, a. C. D. (1952). *The Development and Scope of Higher Education in the United States*. New York: Columbia University Press.

Hojat, M., Veloski, J., & Gonnella, J. (2009). Measurement and correlates of physicians' lifelong learning. *Academic Medicine, 84*(8), 1066-1074.

Holmes, G. (1983). Creating CAL courseware: Some possibilities. *System, 11*(1), 21-32.

Holt, R. I. G., Miklaszewicz, P., Cranston, I. C., Russell-Jones, D., John Rees, P., & Sonksen, P. H. (2001). Computer assisted learning is an effective way of teaching endocrinology. *Clinical Endocrinology, 55*(4), 537-542.

Holzl, J. (1997). Twelve tips for effective PowerPoint presentations for the technologically challenged. *Medical Teacher, 19*(3), 175-179.

Hood, J. (1968). An audio visual teaching unit for radiology. *British Journal of Radiology, 41*(482), 150-151.

Horner, P. (1973). The teaching of reading in England up to 1870. *Reading, 7*(3), 3-10.

Howard, C., Rogers, P. L., Boettcher, J., Berg, G. A., Justice, L., & Schenk, K. (2005). *Encyclopedia of distance learning*. Hershey, PA: Idea Group Reference.

Hrabak, M., Vujaklija, A., Vodopivec, I., Hren, D., Marusic, M., & Marusic, A. (2004). Academic misconduct among medical students in a post-communist country. *Medical education, 38*(3), 276-285.

Hull, C. L. A. T. (1928). *Aptitude Testing*. Yonkers-on-Hudson, NY: World Book Company.

Hyltander, A., Liljegren, E., Rhodin, P., & Lonroth, H. (2002). The transfer of basic skills learned in a laparoscopic simulator to the operating room. *Surgical endoscopy, 16*(9), 1324-1328.

Ice, P., Curtis, R., Phillips, P., & Wells, J. (2007). Using asynchronous audio feedback to enhance teaching presence and students' sense of community. *Journal of Asynchronous Learning Networks, 11*(2), 3-25.

IEEE Global History Network (2010a). Image:0014 - Ampex Mark III video recorder.jpg. In Image:0014_-_Ampex_Mark_III_video_recorder.jpg (Ed.) (Vol. 1440x1110): GHN: IEEE Global History Network.

IEEE Global History Network (2010b). Image:0025 - ENIAC.jpg - GHN: IEEE Global History Network. In I.-. ENIAC.jpg (Ed.) (Vol. 1381x1080): GHN: IEEE Global History Network.

Immersion Corporation (2003). CathLabVR System Virtual Reality Medical Simulator Retrieved April, 2009, from http://www.immersion.com/markets/medical/products/endovascular/index.html

Imrie, B. W. (1995). Assessment for learning: quality and taxonomies. *Assessment & Evaluation in Higher Education, 20*(2), 175-189.

Indvik, L. (2010). Sony Retires the Cassette Walkman After 30 Years. *Mashable*. Retrieved from http://mashable.com/2010/10/24/sony-walkman-rip/

Intel (2005). Moore's Law, The Future - Technology & Research at Intel Retrieved November 5, 2010, from http://www.intel.com/cd/corporate/techtrends/EMEA/ENG/209729.htm

Inwood, M. J., & Ahmad, J. (2005). Development of instructional, interactive, multimedia anatomy dissection software: a student-led initiative. *Clin Anat, 18*(8), 613-617.

Irish College of General Practitioners (2009). Short Essay Question (SEQ) - ICGP Web Site Retrieved August 7th, 2009, from http://www.rcog.org.uk/education-and-exams/examinations

iRiver (2007). www.iRiver.com Retrieved June, 2007, from www.iriver.com

Isaacs, G. (1996). Bloom's taxonomy of educational objectives Retrieved August, 2008, from http://www.tedi.uq.edu.au/downloads/Bloom.pdf

Issenberg, S., Mcgaghie, W., Petrusa, E., Gordon, D., & Scalese, R. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher, 27*(1), 10-28.

IT Reviews (2004). HP - Scanjet 3970 review Retrieved August 23, 2009, from http://www.itreviews.co.uk/hardware/h568.htm

Jacoby, C., Smith, W., & Albanese, M. (1984). An evaluation of computer-assisted instruction in radiology. *American Journal of Roentgenology, 143*(3), 675-677.

Jaffe, J., & France, J. (2004). Stoarge in a flash. *cnet reviews* Retrieved August 23, 2009, from http://reviews.cnet.com/4520-3240_7-5128560-1.html

James, K., Burke, L., & Hutchins, H. (2006). Powerful or pointless? Faculty versus student perceptions of PowerPoint use in business education. *Business Communication Quarterly, 69*(4), 374-396.

James, S. (1959). Auscultation on the Gramophone. *British Medical Journal, 2*(5162), 1336.

Januszewski, A. (2001). *Educational Technology: The Development of a Concept*. Englewood, CO: Libraries Unlimited.

Jennett, P., & Swanson, R. (1994). Lifelong, self-directed learning: why physicians and educators should be interested. *Journal of Continuing Education in the Health Professions, 14*(2), 69-74.

Jham, B., Duraes, G., Strassler, H., & Sensi, L. (2008). Joining the podcast revolution. *Journal of Dental Education, 72*(3), 278-281.

Johns, K. (2001). Which SyQuest Should I Buy? *The Information Booth* Retrieved August 23, 2009, from http://www.kassj.com/articles/sqzip_jaz.html

Johnson, B., & Christensen, L. B. (2004). *Educational Research: Quantitative, Qualitative, and Mixed Approaches*. Boston, MA: Allyn and Bacon.

Johnson, L. (1994). Advantage, videodisc. (enjoying opera productions on laserdisc) (includes names and addresses of laserdisc manufacturers) (Buyers Guide). *Opera News*. Retrieved from http://www.accessmylibrary.com/coms2/summary_0286-5517993_ITM

Johnson, R., & Onwuegbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher, 33*(7), 14-26.

Jones, B. (1968). Audio-visual aids in medical teaching. *Proceedings of the Royal Society of Medicine, 61*(1), 89-94.

Judd, C. (1923). Education and the Movies. *The School Review, 31*(3), 173-178.

Kadiyala, M., & Crynes, B. (2002). *Where's the proof? A review of literature on effectiveness of information technology in education.* Paper presented at the Frontiers in Education Conference, Tempe, AZ, USA.

Kahn, N., & Bigger Jr, J. (1974). Instruction in pharmacokinetics: a computer-assisted demonstration system. *Academic Medicine, 49*(3), 292-295.

Kanellos, M. (2006). ENIAC: A computer is born. *CNET News*. Retrieved from http://news.cnet.com/2009-1006_3-6037980.html

Kanner, J., & Rosenstein, A. (1960). Television in army training: Color vs black and white. *Educational technology research and development, 8*(5), 243-252.

Katzman, N., & Nyenhuis, J. (1972). Color vs. black-and-white effects on learning, opinion, and attention. *Educational technology research and development, 20*(1), 16-28.

Kaufman, D., & Lee, S. (1993). Formative evaluation of a multimedia CAL program in an ophthalmology clerkship. *Medical Teacher, 15*(4), 327-340.

Keckley, P. (2004). Evidence-based medicine in managed care: a survey of current and emerging strategies. *Med Gen Med, 6*(2), 56. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1395794/

Kellogg, W., & Payne, B. (1938). The true-false question as an aid in studying. *Journal of Educational Psychology, 29*(8), 581-589.

Kember, D. (2003). To Control or Not to Control: the question of whether experimental designs are appropriate for evaluating teaching innovations in higher education. *Assessment & Evaluation in Higher Education, 28*(1), 89-102.

Kennedy, T., & Preminger, G. (1987). Impact of video on endourology. *Journal of Endourology, 1*(2), 75-79.

Kenshole, G. (1969). An experiment in undergraduate teaching using audio visual aids. *Physics Education, 4*, 157-161.

Kerecsen, L., & Pazdernik, T. (2002). From mainframe to Web-based: 30 years of experience in computer-aided instruction of pharmacology. *Naunyn-Schmiedeberg's archives of pharmacology, 366*(1), 83-89.

Khaled, M. A., Ellis, A. G., & Buck, P. (1999). *Short essay questions for MRCOG Part 2*. Leith Walk, Edinburgh: Churchill Livingstone.

Khalili, A., & Shashaani, L. (1994). The effectiveness of computer applications: A meta-analysis. *Journal of Research on Computing in Education, 27*, 48-61.

Khoo, H. (2003). Implementation of problem based learning in Asian medical schools and students' perceptions of their experience. *Medical Education, 37*(5), 401-409.

Kilpatrick, J. (1993). The chain and the arrow: from the history of mathematics assessment. In M. Niss (Ed.), *Investigations Into Assessment in Mathematics Education: An Icmi Study*. Dordrecht, Netherlands: Luwer Academic Publishers Group.

Kim, L. (2004). Online technologies for teaching writing: Students react to teacher response in voice and written modalities. *Research in the Teaching of English, 38*(3), 304-337.

Kirk, R. (2007). Medical Education Changes Should Be Evidence Based. *Kuwait Medical Journal, 39*(3), 213-215.

Kirkwood, A., & Price, L. (2005). Learners and learning in the twenty-first century: what do we know about students' attitudes towards and experiences of information and communication technologies that will help us design courses? *Studies in Higher Education, 30*(3), 257-274.

Kirschner, P., Brink, H., & Meester, M. (1991). Audiotape feedback for essays in distance education. *Innovative Higher Education, 15*(2), 185-195.

Knight, F. (1922). Data on the True-False Test as a device for college examination. *Journal of Educational Psychology, 13*(2), 75-80.

Knight, P. (2002). Summative assessment in higher education: practices in disarray. *Studies in Higher Education, 27*(3), 275-286.

Koh, G. C., Khoo, H. E., Wong, M. L., & Koh, D. (2008). The effects of problem-based learning during medical school on physician competency: a systematic review. *Cmaj, 178*(1), 34-41.

Koperski, M. (1989). Health education using video recordings in a general practice waiting area: an evaluation. *The Journal of the Royal College of General Practitioners, 39*(325), 328-330.

Kordas, M., & Lajovic, J. (1987). Personal computers in teaching quantitative aspects in undergraduate physiology. *Medical Education, 21*(6), 468-470.

Krause, K., Hartley, R., James, R., & McInnes, C. (2005). *The first year experience in Australian universities: Findings from a decade of national studies.* Canberra, Australia: Department of Education, Science and Training.

Kreiter, C. D., & Bergus, G. R. (2007). A study of Two Clinical Performance Scores: Assessing the Psychometric Characteristics of a Combined Score Derived from Clinical Evaluation Forms and OSCEs. *Med Educ Online [serial online], 12*(10), 1-5.

Krishna, S., Francisco, B. D., Balas, E. A., Konig, P., Graff, G. R., & Madsen, R. W. (2003). Internet-enabled interactive multimedia asthma education program: a randomized trial. *Pediatrics, 111*(3), 503-510.

Kuhn, W. (1954). Trends in Phonographs. *Music Educators Journal, 40*(3), 37-37.

Kulik, C., & Kulik, J. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior, 7*(1-2), 75-94.

Kulik, J., Kulik, C., & Bangert-Drowns, R. (1985). Effectiveness of computer-based education in elementary schools. *Computers in Human Behavior, 1*(1), 59-74.

Kunze, E. G. (1929). The Use of the Phonograph in Language Work. *The Modern Language Journal, 13*(7), 538-549.

Kwan, C. (2002). Problem-based learning and teaching of medical pharmacology. *Naunyn Schmiedebergs Arch Pharmacol, 366*(1), 10-17.

Kwan, R., Naqvi, S. H., Narayanan, G., Mir, M. A., Care, D., & Briscoe, M. H. (1988). Use of interactive video in teaching history-taking to medical students: a pilot project. *The Journal of audiovisual media in medicine, 11*(4), 121-124.

Lackey, L. W., & Lackey, W. J. (2002). Influence of True/False Tests and First Language on Engineering Students' Test Scores. *Journal Of Engineering Education, 91*(1), 25-32.

Laidlaw, J. (1987). Twelve Tips for the Users of the Overhead Projector. *Medical Teacher, 9*(3), 247-252.

Lamperti, A., & Sodicoff, M. (1997). Computer-based neuroanatomy laboratory for medical students. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology, 249*(3), 422-428.

Lancy, D. F. (1980). Becoming a Blacksmith in Gbarngasuakwelle. *Anthropology & Education Quarterly, 11*(4), 266-274.

Landman, J. H. (1927). Phonograph Records as an Aid in the Teaching of American History. *The School Review, 35*(9), 681-685.

Landow, G. (1996). Newman and the Idea of an Electronic University, The Idea of a University. New Haven: Yale University Press.

Lane, J. L., & Gottlieb, R. P. (2000). Structured clinical observations: a method to teach clinical skills with limited time and financial resources. *Pediatrics, 105*(4), 973-977.

Lanius, C. (2004). PowerPoint, not your grandmotherís presentations, but is it evil. *Cell Biol. Educ, 3*, 158-160.

Larsen, C., Soerensen, J., Grantcharov, T., Dalsgaard, T., Schouenborg, L., Ottosen, C., et al. (2009). Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *British Medical Journal, 338*(14), b1802.

Laurillard, D. (2007). Forward. In H. Beetham & R. Sharpe (Eds.), *Rethinking pedagogy for a digital age: Designing and delivering e-learning*. New York: Psychology Press.

Lawrence, P. F. (2005). *Essentials of general surgery* (4th ed.). Baltimore,MD: Lippincott Williams & Wilkins.

Laycock, S. D., & Day, A. M. (2003). *Recent developments and applications of haptic devices*.

Le Bohec, Y., & Bate, R. (2000). *The Imperial Roman Army*. London: Routledge.

Leach, L., Neutze, G., & Zepke, N. (2001). Assessment and empowerment: Some critical questions. *Assessment & Evaluation in Higher Education, 26*(4), 293-305.

Lechner, S. (2001). Evaluation of teaching and learning strategies. *Med Educ Online [serial online], 6*(4).

Leighton, J., & Gierl, M. (2007). *Cognitive diagnostic assessment for education*. Cambridge: Cambridge University Press.

Lemieux, C. (1949). Audio-Visual Aids in the Teaching of Russian. *Modern Language Journal, 33*(8), 594-602.

Letterie, G. S. (2003). Medical education as a science: The quality of evidence for computer-assisted instruction. *American Journal of Obstetrics & Gynecology, 188*(3), 849-853.

Levine, R., Jones, J., & Morgan, C. (1987). Comparison of computer assisted learning with tutorial teaching in a group of first year dental students. *Medical Education, 21*(4), 305-309.

Levy, A., & Thursh, D. (1989). The implementation of a knowledge-based Pathology Hypertext under HyperCard. *Journal of Medical Systems, 13*(6), 321-329.

Lewis, J. (1946). Educational Radio Rides the Range. *Hollywood Quarterly, 1*(3), 312-316.

Lewy, A. (1971). Affective Outcomes of Musical Education. *Journal of Research in Music Education, 19*(3), 361-365.

Liao, H., & Lu, H. (2008). The role of experience and innovation characteristics in the adoption and continued use of e-learning websites. *Computers & Education, 51*(4), 1405-1416.

Liao, Y. C. (2007). Effects of computer-assisted instruction on studentsí achievement in Taiwan: A meta-analysis. *Computers & Education, 48*(2), 216-233.

Lim, C., & Chai, C. (2008). Teachersí pedagogical beliefs and their planning and conduct of computer mediated classroom lessons. *British Journal of Educational Technology, 39*(5), 807-828.

Lim, S., & Ratnam, S. (1979). The use of video in a teaching unit. *Singapore Medical Journal, 20*(2), 347-350.

Linn, R. L. (2001). A Century of Standardized Testing: Controversies and Pendulum Swings. *Educational Assessment, 7*(1), 29-38.

Lipsett, A. (2008, July 29). Dyslexic student fights exam system. *Guardian.co.uk*, from http://www.guardian.co.uk/education/2008/jul/29/dyslexia.exams

Littleton, D. (2000). PhlebotomyTUTOR. *Bulletin of the Medical Library Association, 88*(2), 205.

Lowe, C. I., Wright, J. L., & Bearn, D. R. (2001). Computer-aided Learning (CAL): an effective way to teach the Index of Orthodontic Treatment Need (IOTN)? *Journal of Orthodontics, 28*(4), 307-311.

Lowenstein, S. R., Fernandez, G., & Crane, L. A. (2007). Medical school faculty discontent: prevalence and predictors of intent to leave academic careers. *BMC Med Educ, 7*, 37. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17935631

Lyman, P., Varian, H., & Swearingen, K. (2003). How much information? Retrieved from http://www.attitudeweb.be/doc/resources/studies/how_much_information_produced_world_year_en.pdf

Maag, M. (2006). Podcasting and MP3 players: Emerging education technologies. *Computers Informatics Nursing, 24*(1), 9-13.

Macdonald, P., & Paunonen, S. V. (2002). A Monte Carlo comparison of item and person statistics based on item response theory versus classical test theory. *Educational and Psychological Measurement, 62*(6), 921-943.

Machover, C., & Dill, J. (1985). New Products. *IEEE Computer Graphics and Applications, 5*(11), 67-75.

MacKeogh, K., & Fox, S. (2008). *Opening access to higher education to all? What motivates academic staff in traditional universities to adopt elearning?* Paper presented at the European Distance & eLearning Network 5th Research Conference. from http://doras.dcu.ie/2099/1/eden_2008.pdf

MacLellan, E. (2001). Assessment for learning: the different perceptions of tutors and students. *Assessment and Evaluation in Higher Education, 26*(4), 307-318.

Magnuson, W. (2000). Educational Television Progress Report. *Congressional Record, Oct. 27, 1965*. Retrieved from http://www.current.org/pbpb/legislation/ETVProgressRept1965.html

Malinche Entertainment (2007). Fiction Books Transformed by Interactive Fiction Retrieved March, 2007, from http://www.malinche.net/index.html

Mamary, E., & Charles, P. (2003). Promoting self-directed learning for continuing medical education. *Med Teach, 25*(2), 188-190.

Mangione, S., Nieman, L., & Gracely, E. (1992). Comparison of computer-based learning and seminar teaching of pulmonary auscultation to first-year medical students. *Academic Medicine, 67*(10), S63-S65.

Mann, K. (1999). Motivation in medical education: how theory can inform our practice. *Academic Medicine, 74*(3), 237-239.

Margolis, C. Z., Cook, C. D., Barak, N., Adler, A., & Geertsma, A. (1989). Clinical algorithms teach pediatric decisionmaking more effectively than prose. *Medical Care, 27*(6), 576-592.

Martin, J. J., Kulinna, P. H., & Cothran, D. (2002). Motivating Students through Assessment. *JOPERD--The Journal of Physical Education, Recreation & Dance, 73*(8), 18-21.

Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist, 34*(4), 207-218.

Marton, F., & Saljo, R. (1976). On Qualitative Differences in Learning: 1--Outcome and Process. *British Journal of Educational Psychology, 46*(1), 4-11.

Mason, O., & Grove-Stephensen, I. (2002). *Automated free text marking with paperless school.* Paper presented at the 6th CAA Conference, Loughborough: Loughborough University.

Mathews, P. W. (1947). Edison's First Great Gift to Education. *Music Educators Journal*, 25-27.

Mathieu, J. (2007). Blogs, podcasts, and wikis: the new names in information dissemination. *J Am Diet Assoc, 107*(4), 553-555.

Matthew, I., Pollard, D., & Frame, J. (1998). Development and evaluation of a computer aided learning package for minor oral surgery teaching. *Medical Education, 32*(1), 89-94.

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist, 59*(1), 14-19.

McDonald, K. (2004). Points of View: PowerPoint in the Classroom: Examining PowerPointlessness. *Life Sciences Education, 3*(3), 160-161.

McGrath, P., Kucera, R., & Smith, W. (2003). Computer simulation of introductory neurophysiology (Vol. 27, pp. 120-129): Am Physiological Soc.

McInnis, C., James, R., & Hartley, R. (2000). *Trends in the first year experience in Australian universities*. Canberra: AGPS.

McKee, I. H. (1965). The Tape-Recorded Lecture in Undergraduate Medical Education. *Lancet, 39*, 863-864.

McKendrick, J. G. (1895). The Tone and Curves of the Phonograph. *Journal of Anatomy and Physiology, 29*(4), 583-591.

McLoughlin, C., & Lee, M. (2007). *Listen and learn: A systematic review of the evidence that podcasting supports learning in higher education.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunication, Chesapeake, VA.

McNeil, B. J., & Nelson, K. R. (1990). Meta-Analysis of Interactive Video Instruction: A 10 Year Review of Achievement Effects. *Journal of Computer-Based Instruction, 18*(1), 1-6.

Meade, O., Dianne, B., & Joanne, L. (2009). Pharmacology as a foreign language: A preliminary evaluation of podcasting as a supplementary learning tool for non-medical prescribing students. *BMC Medical Education, 9*. Retrieved from http://www.biomedcentral.com/1472-6920/9/74/

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington DC: US Dept. of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service.

Meggs, P. (1998). *A history of graphic design* (3rd ed.). New York: Wiley and Sons.

Meterissian, S. (2006). A novel method of assessing clinical reasoning in surgical residents. *Surgical Innovation, 13*(2), 115-119.

Meterissian, S., Zabolotny, B., Gagnon, R., & Charlin, B. (2007). Is the script concordance test a valid instrument for assessment of intraoperative decision-making skills? *The American Journal of Surgery, 193*(2), 248-251.

Meyer, J., & Beaton, G. (1974). An evaluation of computer-assisted teaching in physiology. *Academic Medicine, 49*(3), 295-297.

Meyrick, R. (1973). Videocassettes for Medical Education.

Microsoft (2010). Microsoft Office 2010: Make It Great with Office 2010, from http://www.microsoft.com/office/make-it-great/en-au/applications/PowerPoint

Microsoft Corporation (2008). Zune.net | Zune.net Home Retrieved August, 2008, from http://www.zune.net:80/en-US/

Middleton, A., & Nortcliffe, A. (2009). Audio Feedback design: principles and emerging practice. 1-13. Retrieved from http://www.qmu.ac.uk/cap/audiofeedbackdesign.pdf

Miller, K., & Willett, J. (1987). Audio-enhanced computer-assisted learning in science teaching. *Biochemical Society transactions, 15*(3), 351-354.

Miller, M. (1944). To teach and to learn. *Modern Language Journal, 28*(4), 378-380.

Miller, M. (1974). *Plain speaking: an oral biography of Harry S. Truman*. New York: Putnam Adult.

Millward Brown (2008). Adobe - Shockwave Player Adoption Statistics Retrieved July, 2008, from http://www.adobe.com/products/player_census/shockwaveplayer/

Minton, H. (1990). Robert M. Yerkes and the Mental Testing Movement. In M. M. Sokal (Ed.), *Psychological Testing and American Society, 1890-1913*. New Brunswick, NJ: Rutgers University Press.

Minton, H. L. (1905). Introduction. *New Methods for the Diagnosis of the Intellectual Level of Subnormals* Retrieved August, 2008, from http://psychclassics.yorku.ca/Binet/intro.htm

Mir, M., Marshall, R., Evans, R., Hall, R., & Duthie, H. (1984). Comparison between videotape and personal teaching as methods of communicating clinical skills to medical students. *British Medical Journal, 289*(6436), 31-34.

Mir, M., Marshall, R. J., Evans, R. W., Hall, R., & Duthie, H. L. (1986). Teaching medical history-taking: a comparison between the use of audio- and videotapes. *Med Educ, 20*(2), 102-108.

Mitchell, T., Aldridge, N., & Broomhead, P. (2003). *Computerised Marking of Short-Answer Free-Text Responses*.

Molnar, A. (1997). Computers in education: A brief history. *THE Journal (Technological Horizons In Education), 24*(11). Retrieved from http://www.questia.com/googleScholar.qst;jsessionid=BCEC793CEE3C8428AB269E217866418B.inst1_1a?docId=5001519381

Moore, G., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information systems research, 2*(3), 192-222.

Moore, L., Waechter, D., & Aronow, L. (1991). Assessing the effectiveness of computer-assisted instruction in a pharmacology course. *Academic Medicine, 66*(4), 194-196.

Morgan, M., Deber, R., Llewellyn Thomas, H., Gladstone, P., Cusimano, R., O'rourke, K., et al. (2000). Randomized, controlled trial of an interactive videodisc decision aid for patients with ischemic heart disease. *Journal of General Internal Medicine, 15*(10), 685-693.

Morrissey, M. J. (1960). Points from letters: Auscultation on the Gramophone. *British Medical Journal, 1*(5175), 806.

Morwood, J. M. (1953). Relaxation by gramophone in asthma. *The Practitioner, 170*(1018), 400.

Moy, R. H. (2000). Medical Education in the 20th Century. In L. H. Distlehorst, G. L. Dunnington & J. R. Folse (Eds.), *Teaching and Learning in Medical and Surgical Education: Lessons Learned for the 21st Century*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mulholland, H. (1997). 'De facto'assessment: issues of validity and reliability. *Current Obstetrics & Gynaecology, 7*(3), 145-148.

MySchizoBuddy (2010). iPod sales per fiscal quarter. Retrieved from http://en.wikipedia.org/wiki/File:Ipod_sales_per_quarter.svg

Nair, B. R., Alexander, H. G., McGrath, B. P., Parvathy, M. S., Kilsby, E. C., Wenzel, J., et al. (2008). The mini clinical evaluation exercise (mini-CEX) for assessing clinical performance of international medical graduates. *Medical Journal of Australia, 189*(3), 159-161.

Nandi, P. L., Chan, J. N., Chan, C. P., Chan, P., & Chan, L. P. (2000). Undergraduate medical education: comparison of problem-based learning and conventional teaching. *Hong Kong Med J, 6*(3), 301-306.

National Center for Education Statistics (2005). Indicator 30: Employment of College Students, from http://nces.ed.gov/programs/youthindicators/Indicators.asp?PubPageNumber=30&ShowTablePage=TablesHTML/30.asp

Neilson, F. (1946). Some Thoughts on Education. *American Journal of Economics and Sociology, 5*(2), 261-269.

Nemerovski, J. (2004). Rocport 4 Portable FireWire/USB 2.0 Hard Drive - Review. *MyMac Magazine* Retrieved August 23, 2009, from http://www.mymac.com/showarticle.php?id=1710

Nendaz, M. R., & Tekian, A. (1999). Assessment in Problem-Based Learning Medical Schools: A Literature Review. *Teaching and Learning in Medicine, 11*(4), 232–243.

Nesbitt, A. (2005). The Podcast Value Chain Report Retrieved March, 2007, from http://www.digitalpodcast.com/PVCR_3Q_Update.pdf

Newble, D. (2004). Techniques for measuring clinical competence: objective structured clinical examinations. *Med Educ, 38*(2), 199-203.

Newman, M. (2004). *A pilot systematic review and Meta-analysis on the effectiveness of Problem Based Learning*. Newcastle: University of Newcastle.

Newmann, F. M., & Wehlage, G. G. (1993). Five standards of authentic instruction. *Educational Leadership, 50*, 8-8.

Nextag (2009). Blu Ray Blank Disc Retrieved August 23, 2009, from http://www.nextag.com/blu-ray-blank-disc/shop-html

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Nicol, D., & Milligan, C. (2006). Rethinking technology-supported assessment practices in relation to the seven principles of good feedback practice. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp. 64-77). London: Routledge.

Niederhauser, D., & Lindstrom, D. (2006). Addressing the NETS for students through constructivist technology use in K-12 classrooms. *Journal of Educational Computing Research, 34*(1), 91-128.

Nilson, L. B. (2003). Improving Student Peer Feedback. *College Teaching, 51*(1), 34-39.

Niss, M. (1993). *Investigations Into Assessment in Mathematics Education: An Icmi Study*. Dodrecht, Netherlands: Kluwer Academic Publishers.

Nodelman, P. (1988). *Words about pictures: The narrative art of children's picture books*. Georgia, USA: University of Georgia Press.

Nodine, C., Mello-Thoms, C., Kundel, H., & Weinstein, S. (2002). Time course of perception and decision making during mammographic interpretation. *American Journal of Roentgenology, 179*(4), 917-923.

Norcini, J. J., Blank, L. L., Arnold, G. K., & Kimball, H. R. (1995). The mini-CEX (clinical evaluation exercise): a preliminary investigation. *Ann Intern Med, 123*(10), 795-799.

Norcini, J. J., Blank, L. L., Duffy, F. D., & Fortna, G. S. (2003). The mini-CEX: a method for assessing clinical skills. *Ann Intern Med, 138*(6), 476-481.

Norcini, J. J., Swanson, D., Grosso, L., & Webster, G. (1985). Reliability, validity and efficiency of multiple choice question and patient management problem item formats in assessment of clinical competence. *Medical Education, 19*(3), 238-247.

Noreña, C. G. (1970). *Juan Luis Vives*: Springer.

Norman, G. R., & Schmidt, H. G. (1992). The Psychological Basis of Problem-Based Learning: A Review of the Evidence. *Acad. Med, 67*(6), 557-565.

Norman, G. R., & Schmidt, H. G. (2000). Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Med Educ, 34*, 721-728.

Northern Illinois University (2007). Artifacts - NIU - Blackwell Museum Retrieved May 13, 2008, from http://www.cedu.niu.edu/blackwell/artifacts.html

Norton, L. S. (1990). Essay-writing: what really counts? *Higher Education, 20*(4), 411-442.

O'Neill, P. N. (1990). Developing videodisc instructions for health sciences: a consortium approach. *Academic Medicine, 65*(10), 624-627.

O'Connor, V. (2008). Teaching and learning with simulated patients. *O&G Magazine, 10*(1), 41-42.

Oberhelman, H. D. (1960). The Foreign Language Laboratory in Secondary Schools. *Hispania*, 145-148.

Officer, L. H. (2009). Exchange Rate Between US Dollar and 40 Countries Retrieved August 16, 2009, from http://www.measuringworth.com/calculators/uscompare/result.php

Olsen, J., & Bass, V. (1982). The application of performance technology in the military: 1960-1980. *Performance & Instruction, 21*(6), 32-36.

Orlich, D., Harder, R., Callahan, R., Trevisan, M., & Brown, A. (2009). *Teaching strategies: A guide to effective instruction*. Boston,Ma: Wadsworth Pub Co.

Orr, R. G. (1968). Audio tape versus the live lecture for health education. *Medical & biological illustration, 18*(1), 37.

Osipian, A. (2004). Corruption as a legacy of the medieval university: Financial affairs. *Munich Personal RePEc Archive* Retrieved August, 2008, from http://mpra.ub.uni-muenchen.de/8472/

Osterlind, S. J. (1997). *Constructing test items: Multiple-choice, constructed-response, performance, and other formats*: Kluwer Academic Pub.

Ottolini, M., & Greenberg, L. (1998). Development and evaluation of a CD-ROM computer program to teach residents telephone management. *Pediatrics, 101*(3), e2. Retrieved from http://pediatrics.aappublications.org/cgi/content/abstract/101/3/e2

Oxford University Press (2006). Oxford Handbook of Anaesthesia Retrieved July 19, 2009, from http://oha.oxfordonline.com/

Page, E. (2005). Project Essay Grade (PEG) Retrieved May, 2006, from http://134.68.49.185/pegdemo/

Palacio, A., Lehmann, E., & Olson, D. (2007). Teaching Diabetes to Middle-School Students with the www. 2aida. net AIDA Online Diabetes Software Simulator. *Journal of diabetes science and technology (Online), 1*(1), 106-115. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2769599/

Palaniyappan, L. (2008). An elegy to essay writing. *Psychiatric Bulletin, 32*, 152.

Palmer, E., & Devitt, P. (2006). *Evaluation and assessment in the classroom: The computer aided tutorial*. Paper presented at the Evaluation Forum.

Palmer, E., & Devitt, P. (2008). Limitations of student-driven formative assessment in a clinical clerkship. A randomised controlled trial. *BMC Medical Education, 8*(1), 29. Retrieved from http://www.biomedcentral.com/1472-6920/8/29

Palmer, E., & Devitt, P. (2010). Medici - Quality e-learning Retrieved November, 2010, from www.emedici.com

Palmer, E., Devitt, P., De Young, N., & Morris, D. (2005). Assessment of an electronic voting system within the tutorial setting: A randomised controlled trial. *BMC Medical Education, 5*(1), 24. Retrieved from http://www.biomedcentral.com/1472-6920/7/25/

Palmer, E., & Devitt, P. G. (2007a). A method for creating interactive content for the iPod, and its potential use as a learning tool: technical advances. *BMC Med Educ, 7*, 32. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17888168

Palmer, E., & Devitt, P. G. (2007b). Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Med Educ, 7*, 49. Retrieved from http://www.biomedcentral.com/1472-6920/7/49/

Palmer, E., & Devitt, P. G. (2007c). *The Evolution of Interactive Learning Material for Medical Students*. Paper presented at the Apple University Consortium Conference.

Palmer, E., Duggan, P., Devitt, P. G., & Russell, R. (2010). The Modified Essay Question – its exit from the exit examination? *Medical Teacher, 32*(7), e300-e307.

Parallels Holdings Ltd (2010). Automation Solutions for Desktops, Servers, Hosting, SaaS – Parallels Optimized Computing Retrieved November 18, 2010, from http://www.parallels.com

Paris, S. G., & Stahl, S. A. (2005). *Children's Reading Comprehension and Assessment*. Mahwah, New Jersey: Lawrence Erlbaum Publishers.

Pattee, E. B. (1923). The Phonograph as a Medium of Foreign-Language Instruction. *The School Review*, 604-607.

Paul, S., Dawson, K., Lanphear, J., & Cheema, M. (1998). Video recording feedback: a feasible and effective approach to teaching history-taking and physical examination skills in undergraduate paediatric medicine. *Medical Education, 32*(3), 332-336.

Peacock, M. (1997). The effect of authentic materials on the motivation of EFL learners. *ELT Journal, 51*(2), 144-156.

Pearson Education, I. (2009). Pearson's Knowledge Technologies Products - Intelligent Essay Assessor Retrieved August, 2009, from http://www.k-a-t.com/prodIEA.shtml

Pernick, M. (1978). Thomas Edison's Tuberculosis Films: Mass Media and Health Propaganda. *Hastings Center Report, 8*(3), 21-27.

Peters, A. S., Greenberger-Rosovsky, R., Crowder, C., Block, S. D., & Moore, G. T. (2000). Long-term outcomes of the New Pathway Program at Harvard Medical School: a randomized controlled trial. *Acad Med, 75*(5), 470-479.

Peterson, N., & Campbell, K. (1984). The Physiology Teacher: Simulated Laboratory for Teaching Cardiac Mechanics. *Physiologist, 27*(3), 165-169.

Phillips, A. F. (1949). A loud-speaker stethoscope for clinical teaching. *British Heart Journal, 11*(1), 48-54.

Phillips, C., Henderson, P., Mandel, L., Kim, S., Schaad, D., Cooper, M., et al. (1998). Teaching the microscopic examination of urine sediment to second year medical students using the Urinalysis-Tutor computer program. *Clinical chemistry, 44*(8), 1692-1700.

Phillips, K. (1956). Television in operating room nursing. *AJN The American Journal of Nursing, 56*(2), 162-165.

Phillips, R. (1968). An audio-visual teaching unit for dermatology. *British Journal of Dermatology, 80*(6), 406-409.

Piemme, T. E. (1988). Computer-assisted learning and evaluation in medicine. *JAMA, 260*(3), 367-372.

Pinkerton, P. (1961). Teaching by tape. A method of undergraduate instruction in child psychiatry. *Lancet, 2*(7197), 308.

Pinney, J. (1936). The Motion Picture and Social-Hygiene Education. *The Journal of Educational Sociology, 10*(3), 158-167.

Pintrich, P., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. In R. Perry & J. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 731-810). Netherlands: Springer.

Pioneer Electronics (2009). Pioneer USA - DVD/Blu-ray Burners Retrieved August 23, 2009, from http://www.pioneerelectronics.com/PUSA/Shop/DVD-Burners

PLATO Learning Inc (2010). Educational & Instructional Technology News & Events by PLATO Learning Retrieved December 4, 2010, from http://www.plato.com/News-and-Events.aspx

Plumpers, V. (1870). The Magic Lantern in Medicine. *The British Medical Journal, 2*(497), 43-44.

Pollock, R. (2009). miscellaneous factZ Retrieved May, 2009, from http://rufuspollock.org/tags/culture-and-society/

Polsson, K. (2009). Chronology of Personal Computers (1984) Retrieved August 20, 2009, from http://www.islandnet.com/~kpolsson/comphist/comp1984.htm

Popham, W. J. (2003). *Test better, teach better: The instructional role of assessment*: Association for Supervision & Curriculum Development.

Powicke, F. (1934). Presidential Address: Some Problems in the History of the Medieval University. *Transactions of the Royal Historical Society, Fourth Series, 17*, 1-18.

Prentice, J., & Kenny, G. (1986). Microcomputers in medical education. *Medical Teacher, 8*(1), 9-18.

Pressey, S. (1926). A simple device which gives tests and scores and teaches. *School and Society, 23*, 373-376.

Price, J., Price, J., Pratt, D., Collins, J., & McDonald, J. (2010). High-fidelity simulation in anesthesiology training: a survey of Canadian anesthesiology residentsí simulator experience. *Canadian Journal of Anesthesia/Journal canadien d'anesthÈsie, 57*(2), 134-142.

Race, P., Learning, & Centre, T. S. N. G. (2001). *A briefing on self, peer & group assessment* (Vol. 9). York, UK: Learning and Teaching Support Network.

Rademakers, J., Cate, T. J., & Bär, P. R. (2005). Progress testing with short answer questions. *Medical Teacher, 27*(7), 578-582.

Rahmlow, H., Fratini, R., & Ghesquiere, J. (1980). *Plato*. New Jersey: Educational Technology Publications.

Rainsbury, J., & McDonnell, S. (2006). Podcasts: an educational revolution in the making? *Journal of the Royal Society of Medicine, 99*(9), 481-482.

Ramos, R. A., & Stern, J. (1973). Item Behavior Associated with Changes in the Number of Alternatives in Multiple Choice Items *Journal of Educational Measurement,, 10*(4), 305-310.

Ramsden, P. (1997). The Context Of Learning In Academic Departments. In F. Marton, D. Hounsell & N. Entwhistle (Eds.), *The Experience of Learning* (pp. 124-143). Edinburgh: Scottish Academic Press.

Ravitch, D. (2006). The Fall of the Standard-Bearers. *School and College, 52*(27), B44.

Reed, F., Collart, M. E., & Ertel, P. Y. (1972). Computer assisted instruction for continued learning. *The American Journal of Nursing, 72*(11), 2035-2039.

Reed, P. (1941). Radio as an Aid to Learning. *Annals of the American Academy of Political and Social Science, 213*(1), 31-36.

Reed, V., Jernstedt, G., & McCormick, T. (2004). A longitudinal study of determinants of career satisfaction in medical students. *Med Educ Online, 9*(11). Retrieved from http://med-ed-online.net/coaction/index.php/meo/article/viewArticle/4351

Regan, J. A. (2003). Motivating students towards self-directed learning. *Nurse Education Today, 23*(8), 593-599.

Reimer, J. (2005). Total share: 30 years of personal computer market share figu. *ars technica,* (12). Retrieved from http://arstechnica.com/old/content/2005/12/total-share.ars

Reimer, J. (2009). Total Share: Personal Computer Market Share 1975-2005. *Jeremey Reimer's Blog* Retrieved October 29, 2010, from http://jeremyreimer.com/postman/node/329

Reiser, B. J. (2004). Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work. *The Journal of the Learning Sciences, 13*(3), 273-304.

Reiser, R. (2001). A history of instructional design and technology: Part I: A history of instructional media. *Educational technology research and development, 49*(1), 53-64.

Reisman, S. (1995). Why CD-ROM Is A Dead-End Technology. *IEEE Computer Society, 12*(6), 119-120.

RIAA New Shipments Statistics Database (2008). Music Unit Sales Analysis - Stop Music Theft. Retrieved from http://www.stopmusictheft.com/music-sales-analysis

Richardson, R. (1992). The multiple choice true/false question: what does it measure and what could it measure? *Medical Teacher, 14*(2), 201-204.

Rickards, B. (1923). Broadcasting Health. *The American Journal of Public Health*, 417-419.

Roberts, C., Walton, M., Rothnie, I., Crossley, J., Lyon, P., Kumar, K., et al. (2008). Factors affecting the utility of the multiple mini interview in selecting candidates for graduate entry medical school. *Medical Education, 42*(4), 396-404.

Roberts, G. (1994). An Evaluation of the Use Made by Students of the Audio Recording of Lectures. *Australian Journal of Educational Technology, 10*(2), 96-102.

Robertson, L. (2000). Twelve tips for using a computerised interactive audience response system. *Medical Teacher, 22*(3), 237-239.

Rodin, F. (1927). Conservation of hearing and the hard-of-hearing child. *California and western medicine, 27*(5), 643-647.

Rogers, M., David, A., Regehr PhD, G., Yeh, M., Karen, A., Howdieshell, M., et al. (1998). Computer-assisted learning versus a lecture and feedback seminar for teaching a basic surgical technical skill. *The American Journal of Surgery, 175*(6), 508-510.

Roh, K., & Park, H. (2010). A Meta-Analysis on the Effectiveness of Computer-Based Education in Nursing. *Healthcare Informatics Research, 16*(3), 149-157.

Roig, M., & Caso, M. (2005). Lying and cheating: Fraudulent excuse making, cheating, and plagiarism. *The Journal of Psychology: Interdisciplinary and Applied, 139*(6), 485-494.

Rojas, R., & Hashagen, U. (2002). *The first computers: history and architectures* (1st ed.). Massachusetts: The MIT Press.

Rorabaugh, W. (1986). *The Craft Apprentice: From Franklin to the Machine Age in America*. Oxford: Oxford University Press US.

Rose, M. (1976). Instructional Technology in the Department of Defense. *Air University Review,* (May-June). Retrieved from http://www.airpower.maxwell.af.mil/airchronicles/aureview/1976/may-jun/rose.html

Rosse, C. (1995). The potential of computerized representations of anatomy in the training of health care providers. *Academic Medicine, 70*(6), 499-505.

Rosser, J., Herman, B., Risucci, D., Murayama, M., Rosser, L., & Merrell, R. (2000). Effectiveness of a CD-ROM multimedia tutorial in transferring cognitive knowledge essential for laparoscopic skill training. *The American Journal of Surgery, 179*(4), 320-324.

Rotheram, B. (2007). Using an MP3 recorder to give feedback on student assignments. *Developing tribes and territories, 8*, 7-11.

Routzahn, E. (1924). Health Education and Publicity. *American Journal of Public Health, 14*(2), 172.

Royal Australasian College of Surgeons (2007). Fellowship Examination, from http://www.surgeons.org/Content/NavigationMenu/EducationandTraining/Examinations/Fellowshipexam/default.htm

Royal College of Obstetricians and Gynaecologists (2009). Examinations | Royal College of Obstetricians and Gynaecologists Retrieved August 7th, 2009, from http://www.rcog.org.uk/education-and-exams/examinations

Rudner, L. M., Garcia, V., & Welch, C. (2006). An Evaluation of IntelliMetric Essay Scoring System. *Journal of Technology, Learning, and Assessment, 4*, 4.

Rudner, L. M., & Liang, T. (2002). Automated essay scoring using Bayesí theorem. *Journal of Technology, Learning, and Assessment, 1*(2).

Rumble, G. (2001). The Costs and Costing of Networked Learning. *Journal of Asynchronous Learning Networks, 5*, 75-96.

Russell, R., Devitt, P., & Palmer, E. (2009). Clinicians attitude towards computer-assisted learning.

Ryan, K. (2006). *Those who can teach*. Boston New York: Houghton Mifflin Company.

Rymaszewski, M., Au, W., Wallace, M., Winters, C., Ondrejka, C., & Batstone-Cunnungham, B. (2007). *Second life: The official guide*. Indianapolis, Indiana: John Wiley and Sons.

Sadler, D. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education: Principles, Policy & Practice, 14*(3), 387-392.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science, 18*(2), 119-144.

Saettler, P. (1968). *A History of Instructional Technology*. New York, N.Y.: McGraw-Hill Book Company.

San Francisco Bay Area Rapid Transit District (2007). BART Mobile Options Retrieved March, 2007, from http://www.bart.gov/stations/quickplanner/downloadipod.asp

Sand, R. (1923). Health and the Workers. *International Labour Review, 7*(4), 507-514.

sansa (2008). SanDisk | Home Retrieved July, 2008, from www.sandisk.com

Sanson-Fisher, R. W., & Lynagh, M. C. (2005). Problem-based learning: a dissemination success story? *Med J Aust, 183*(5), 258-260.

Savel, R. H., Goldstein, E. B., Perencevich, E. N., & Angood, P. B. (2007). The iCritical care podcast: a novel medium for critical care communication and education. *J Am Med Inform Assoc, 14*(1), 94-99.

Savery, J. (2006). Overview of problem-based learning: Definitions and distinctions. *The Interdisciplinary Journal of Problem-based Learning, 1*(1), 9-20.

Savery, J. R., & Duffy, T. M. (2001). *Problem Based Learning: An Instructional Model and Its Constructivist Framework* (No. CRLT Technical Report No. 16-01). Bloomington, IN: : Indiana University.

Schinnerer, M. C. (1945). Radio: The Fourth R in Education. *Teachers' College Journal, 17*(1), 2.

Schramm, W. (1962). What we know about learning from instructional television. In L. Asheim (Ed.), *Educational television: The next ten years* (pp. 52-76). California: Stanford University.

Schreiber, B., Fukuta, J., & Gordon, F. (2010). Live lecture versus video podcast in undergraduate medical education: A randomised controlled trial. *BMC Medical Education, 10*(1), 68. Retrieved from http://www.biomedcentral.com/content/pdf/1472-6920-10-68.pdf

Schuwirth, L. W., van der Vleuten, C. P., & Donkers, H. H. (1996). A closer look at cueing effects in multiple-choice questions. *Med Educ, 30*(1), 44-49.

Schwartz, L., Fernandez, R., Kouyoumjian, S., Jones, K., & Compton, S. (2007). A Randomized Comparison Trial of Case based Learning versus Human Patient Simulation in Medical Student Education. *Academic Emergency Medicine, 14*(2), 130-137.

Schwartz, R. W., Witzke, D. B., Donnelly, M. B., Stratton, T., Blue, A. V., & Sloan, D. A. (1998). Assessing residents' clinical performance: Cumulative results of a four-year study with the Objective Structured Clinical Examination. *Surgery, 124*(2), 307-312.

Scott, L. K., Brascho, D. J., & Brown, S. (1978). An evaluation of the effectiveness of audio cassettes for CME. *Med Educ, 12*(6), 427-430.

Scutter, S., Stupans, I., Sawyer, T., & King, S. (2010). How do students use podcasts to support learning. *Australasian Journal of Educational Technology, 26*(2), 180-191.

Seale, J. K., Chapman, J., & Davey, C. (2000). The influence of assessments on students' motivation to learn in a therapy degree course. *Med Educ, 34*(8), 614-621.

Seltzer, R. (1971). Computer-assisted instruction: What it can and cannot do. *American Psychologist, 26*(4), 373-377.

Selwyn, N. (2007). The use of computer technology in university teaching and learning: a critical perspective. *Journal of Computer Assisted Learning, 23*(2), 83-94.

Shane, M. L. (1940). The audio-visual library: An acquisition plan. *Peabody Journal of Education*, 420-430.

Shantikumar, S. (2009). From lecture theatre to portable media: students' perceptions of an enhanced podcast for revision. *Medical Teacher, 31*(6), 535-538.

Sherman, V., Feldman, L., Stanbridge, D., Kazmi, R., & Fried, G. (2005). Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. *Surgical endoscopy, 19*(5), 678-682.

Sibert, L., Darmoni, S. J., Dahamna, B., Weber, J., & Charlin, B. (2005). Online clinical reasoning assessment with the Script Concordance test: a feasibility study. *BMC Medical Informatics and Decision Making, 5*(1), 18. Retrieved from http://www.biomedcentral.com/1472-6947/5/18

Siepmann, C. A. (1941). Can Radio Educate? *The Journal of Educational Sociology*, 346-357.

Sigsby, L. (2004). Perioperative Clinical Learning Experiences. *AORN, 80*(3), 476-490.

Simpson, D., Cookson, M., Percy-Robb, I., & Henderson, G. (1987). Application of a frame based computer aided learning system in clinical chemistry. *Medical Education, 21*(5), 399-404.

Sinclair, S. (2004). Evidence-based medicine: a new ritual in medical teaching. *Br Med Bull, 69*(1), 179-196.

Skinner, B. (1984). The shame of American education. *American Psychologist, 39*(9), 947-954.

Skrabec, Q. (2009). *William McGuffey: mentor to American industry*. USA: Algora Publishing.

Slabaugh, W. (1951). The overhead projector and chemical demonstrations. *Journal of Chemical Education, 28*(11), 579-580.

Smith, G., & Wyllie, J. (1965). Use of closed-circuit television in teaching surgery to medical students. *British Medical Journal, 2*(5453), 99-101.

Smith, H. (1934). What Is Progress in Teaching Foreign Languages. *French Review, 7*(4), 269-283.

Smith, I. (2008). Cost of Hard Drive Space Retrieved August 20, 2009, from http://www.littletechshoppe.com/ns1625/winchest.html

Smith, L. (1974). The medical librarian and computer-assisted instruction. *Bulletin of the Medical Library Association, 62*(1), 6-18.

Society of Motion Picture and Television Engineers (nd). SMPTE Sales Chart Retrieved October 30, 2010, from http://www.avsforum.com/avs-vb/showthread.php?t=1146753

Sokal, M. M. (1990). Sokal, Michael M. *Psychological Testing and American Society, 1890-1913*. New Brunswick, NJ: Rutgers University Press.

Sosa, G., Berger, D., Saw, A., & Mary, J. (2010). Effectiveness of Computer-Assisted Instruction in Statistics: A Meta-Analysis. *Review of Educaitonal Research*. Retrieved from http://rer.sagepub.com/content/early/2010/08/10/0034654310378174

Souder, W. (1955). Composition, properties and behavior of ball pens and inks. *The Journal of Criminal Law, Criminology, and Police Science, 45*(6), 743-747.

Sousa, J. (1993). Machine Songs IV: The Menace of Mechanical Music. *Computer Music Journal, 17*(1), 14-18.

Spak, J. (2000). WebPath: The Internet Pathology Laboratory for Medical Education on CD-ROM. *Bulletin of the Medical Library Association, 88*(2), 205-207.

Special Correspondant (1967). Television in Medical Education. *British Medical Journal, 4*(25th November), 479-480.

Spector, J. (2008). *Handbook of research on educational communications and technology*. New York: Routledge.

Spencer, K. (1977). An Evaluation of the Cost-Effectiveness of Audiotaped and Videotaped Self-Study Media Presentations. *British Journal of Educational Technology, 8*(1), 71-77.

Stanford, W., Erkonen, W., Cassell, M., Moran, B., Easley, G., Carris, R., et al. (1994). Evaluation of a computer-based program for teaching cardiac anatomy. *Investigative radiology, 29*(2), 248-252.

Starkweather, J. A. (1967). Computer-assisted learning in medical education. *Canadian Medical Association Journal, 97*(12), 733-742.

Stensrud, B. (2008). Thoughts on the Supply Chain for Recorded Music. *The Business of Classical Music*. Retrieved from http://businessofclassicalmusic.blogspot.com/2008/12/thoughts-on-supply-chain-for-recorded.html

Stern, D., Mangrulkar, R., Gruppen, L., Lang, A., Grum, C., & Judge, R. (2001). Using a multimedia tool to improve cardiac auscultation knowledge and skills. *Journal of General Internal Medicine, 16*(11), 763-769.

Stinson, A., & Smith, E. (1968). Student sets of color slides as an aid in teaching microscopic anatomy. *Academic Medicine, 43*(1), 83-85.

Stipek, D. (1996). Motivation and instruction. *Handbook of educational psychology*, 85-113.

Stratford, P., & Pierce-Fenn, H. (1985). Modified essay question. *Physical Therapy, 65*(7), 1075-1079.

Student Book World (2009). Free Online Medical Textbooks Retrieved April, 2009, from http://www.studentbookworld.com/Browse/Free-Online-Medical-Textbooks.html

Sumney, L. (2007). *Collaborative Research in Semiconductors* (PowerPoint Presentation). Semiconductor Research Corporation.

Sumstine, D. (1918). A comparative study of visual instruction in the high school. *School and Society, 7*, 235-238.

Sun Microsystems (2008). MySQL Retrieved August, 2008, from http://www.sun.com/software/products/mysql/index.jsp

Sunderman, L. F. (1943). The Music Program in Teacher-Education Institutions. *The Elementary School Journal*, 290-297.

Syrjälä, A., Knuuttila, M., & Syrjälä, L. (1990). Suitability of Krathwohl's affective taxonomy for evaluating patient attitudes to dental care. *Community Dent Oral Epidemiol, 18*(6), 299-303.

Taekman, J., Hobbs, G., Barber, L., Phillips-Bute, B., Wright, M., Newman, M., et al. (2004). Preliminary report on the use of high-fidelity simulation in the training of study coordinators conducting a clinical research protocol. *Anesth Analg, 99*(2), 521-527.

Taras, M. (2002). Using assessment for learning and learning from assessment. *Assessment & Evaluation in Higher Education, 27*(6), 501-510.

Tattersall, M., & Butow, P. (2002). Consultation audio tapes: an underused cancer patient information aid and clinical research tool. *The Lancet Oncology, 3*(7), 431-437.

Taylor, J. (1995). Distance education technologies: The fourth generation. *Australian Journal of Educational Technology, 11*, 1-7.

Taylor, R., & Mason, V. (1934). Radio Education or Radio Propaganda. *California and Western Medicine, 41*(5), 355-356.

Terry, P. W. (1933). How students review for objective and essay tests. *The Elementary School Journal*, 592-603.

Teyler, T., & Voneida, T. (1992). Use of computer-assisted courseware in teaching neuroscience: the Graphic Brain. *Advances in Physiology Education, 263*(6), S37-S44.

The Board of Censors of the Royal College of General Practitioners (1971). *The modified essay question.* Paper presented at the Proceedings of the Royal College of General Practitioners.

The PHP Group (2008). PHP: Hypertext Preprocessor Retrieved August, 2008, from http://www.php.net/

The University of Sydney (2006). Assessment tools - Modified essays - Professional educational services - The University of Sydney Retrieved August, 2008, from http://www.medfac.usyd.edu.au/showcase/assessment/meq/index.php

Thelin, J. R. (1994). 1 Campus and Commonwealth: a Historical Interpretation. In P. G. Altbach, R. O. Berdahl & P. J. Gumport (Eds.), *Higher Education in American Society*. Amherst, NY: Prometheus Books.

Thies, R., Harless, W. G., Lucas, N. C., & Jacobson, E. D. (1969). An experiment comparing computer-assisted instruction with lecture presentation in physiology. *Academic Medicine, 44*(12), 1156-1160.

Thistlethwaite, J. E., Jacobs, H., & Rudolphy, S. (2005). Undergraduate general practice attachments--implications and challenges. *Aust Fam Physician, 34*(3), 181-182.

Thompson, E. (1995). Machines, music, and the quest for fidelity: marketing the edison phonograph in America, 1877-1925. *The Musical Quarterly, 79*(1), 131-171.

Thomson, I., & Bryce, C. F. (1987). Interactive videodiscs in the life sciences: the use of 'generic' discs in the teaching of cell biology. *Biochem Soc Trans, 15*(3), 354-356.

Thorndike, E. (1911). *Animal intelligence*. New York: Macmillan.

Thorndike, E. L. (1921). On the Organization of Intellect. *Psychological Review, 28*(2), 141-151.

Thorndike, E. L. (1923). *Education Psychology: Briefer Course* (2nd ed.). London, UK: Routledge.

Thursh, D., Mabry, F., & Levy, A. (1986). Computers and videodiscs in pathology education: ECLIPS as an example of one approach. *Human pathology, 17*(3), 216-218.

Timmins, M. (2010). In The Time of PLATO. *Illinois Alumni Magazine*.

Tjandra, J. J., Clunie, G. J. A., Kaye, A. H., & Smith, J. A. (2006). *Textbook of surgery*. Malden, Mass: Blackwell Publishing.

Tomei, L. A., & Burkey Piecka, D. (2005). Designing, Developing and Delivering a Technology-Based Bullying Prevention Lesson for Parents and Communities. *T.H.E. Journal, 32*. Retrieved from http://www.thejournal.com/articles/17293

Tomlinson, D. R. (1979). Independent learning packages (audio tape/pamphlet) for a course in cell excitation: evidence that they can be at least as effective as conventional lectures. *Med Educ, 13*(4), 257-262.

Towle, A. (1998). Changes in health care and continuing medical education for the 21st century. *BMJ, 316*, 301-304.

Traub, R. E. (1997). Classical test theory in historical perspective. *Educational Measurement: Issues and Practice, 16*(4), 8-14.

Trelstad, R., & Raskova, J. (1992). *Teaching pathology without lectures through computer-based exercises, small-group discussions and reading.* Paper presented at the AMIA.

Trent, C. J. (2008). MU School of Medicine opens clinical center - Columbia Missourian Retrieved April, 2009, from http://www.columbiamissourian.com/stories/2008/05/12/mu-school-medicine-opens-clinical-center/

Triola, M., Hopkins, M., Phil, M., Weiner, M., Holloway, W., Levin, R., et al. (2004). *Surgical interactive multimedia modules: A novel, non-browser based architecture for medical education.* Paper presented at the CBMS.

Trochim, W. M. K. (2006). Reliability & Validity Retrieved July 24, 2009, from http://www.socialresearchmethods.net/kb/relandval.php

Turner, C., Drenckhahn, V., & Bates, M. (1935). Effectiveness of Radio in Health Education. *American Journal of Public Health, 25*(5), 589.

U.S. Department of Education, N. N. P. S. A. S. N. (2002). *Percentage of undergraduates working while enrolled but considering themselves primarily students who reported various effects of working, by student status: 1999&2000*: National Center for Education Statistics.

U.S. Food and Drug Administration (2007). U.S. Food and Drug Administration FDA Food and Drug Podcasts. Retrieved from http://www.fda.gov/cder/drug/podcast/default.htm

U.S. National Library of Medicine (2003, 12 February 2009). The National Library of Medicine's Visible Human Project Retrieved April, 2009, from http://www.nlm.nih.gov/research/visible/visible_human.html

Uden, L., & Beaumont, C. (2005). *Technology and problem-based learning* (Illustrated ed.). Hershey, PA: Idea Group Inc (IGI).

University of Adelaide (2009). Bachelor of Medicine and Bachelor of Surgery | Programs and Courses Retrieved December 5, 2010, from http://www.adelaide.edu.au/programfinder/2009/bmbbs_bmbbs.html#graduate

unknown (1889). Topics of the day. *British Medical Journal, 1490*, 912.

unknown (1890a). A New Use for the Phonograph. *Science, 15*, 43.

unknown (1890b). A new use of the phonograph. *British Medical Journal, 1555*, 912.

unknown (1891). The phonograph in medicine. *British Medical Journal, 1575*, 537.

unknown (1925). The" Talking Film" In Medical Education. *The British Medical Journal, 2*(3387), 1021-1021.

unknown (1932). New Technique of Science Education to Be Introduced. *The Science News-Letter, 21*, 112.

unknown (1942). Review: With the Technicians. *The Journal of Higher Education, 13*(3), 156-160.

urVoting (2009). VotApedia - Free Audience Response by Mobile Phone Retrieved December 5, 2010, from http://www.urvoting.com/

Usova, G. M. (1997). Effective test item discrimination using Bloom's taxonomy. *Education, 118*(1), 100-111.

UvA Computer Museum (2009). PDP-8 Retrieved August 20, 2009, from http://www.science.uva.nl/museum/pdp8.html

Valenti, S., Neri, F., & Cucchiarelli, A. (2003). An overview of current research on automated essay grading. *Journal of Information Technology Education, 2*(19), 319-330.

Valley, V., Cardenas, E., Spangler, H., & Folstad, S. (1997). Review of Ultrasound Educational CD-ROM Software. *Annals of emergency medicine, 29*(3), 375-379.

Valori, R. (2002). Interactive Continuing Education in Gastroenterology CD-ROM. *Gut, 50*(4), 582-583.

van der Issenberg, S. B., & McGaghie, W. C. (2002). Clinical skills training--practice makes perfect. *Med Educ, 36*(3), 210-211.

van der Vleuten, C. (1996). The assessment of professional competence: developments, research and practical implications. *Advances in Health Sciences Education, 1*(1), 41-67.

van der Vleuten, C., Scherpbier, A. J. J. A., Dolmans, D. H. J. M., Schuwirth, L. W. T., Nen, G. M. V., & Wolfhagen, H. A. P. (2000). Clerkship assessment assessed. *Med Teach, 22*(6), 592-600.

Van Merrienboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the Load Off a Learner s Mind: Instructional Design for Complex Learning. *Educational Psychologist, 38*(1), 5-13.

van Reenen, J. (1990). Technology for health care education. *The Canadian nurse, 86*(10), 24-26.

Vantage Learning (2006). Intellimetric Retrieved May, 2006, from www.vantage.com

Verbeek, H. A. (1987). Interactive Clinical Teaching with Computers combined with a Laser Vision Disc in Clinical Teaching, Past and Present. *Clio Medica, 21*(1-4), 169-174.

Verney, E. (1895). The Public Libraries Act of 1892 in a Small Country Parish. *The Library, 1*(1), 353-358.

Vernon, D. T., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Acad Med, 68*(7), 550-563.

Vogel, J., Vogel, D., Cannon-Bowers, J., Bowers, C., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*(3), 229-243.

Vu, N. V., & Barrows, H. S. (1994). Use of Standardized Patients in Clinical Assessments: Recent Developments and Measurement Findings. *Educational Researcher, 23*(3), 23-30.

Wakeford, R. E., & Roberts, S. (1984). Short answer questions in an undergraduate qualifying examination: a study of examiner variability. *Medical Education, 18*(3), 168-173.

Wal-Mart (2009). Walmart.com: Electronics Retrieved August 23, 2009, from http://www.walmart.com/Computer/cp/3951?fromPageCatId=3944&catNavId=3944

Walker, R. (1964). Teaching basic chemistry concepts with the overhead projector. *Journal of Chemical Education, 41*(12), 663-665.

Wang, Y., & Braman, J. (2009). Extending the Classroom through Second Life. *Journal of Information Systems Education, 20*(2), 235-247.

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

Wasserman, J. D., & Tulsky, D. S. (2005). A History of Intelligence Assessment. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues, 2nd ed*. New York: Guilford Press.

Watson, P. (1972). *Using the computer in education: A briefing for school decision makers*. Englewood Cliffs, New Jersey: Educational Technology Pubns.

Watt, M., & Watt, A. (1987). A tape-based system of interactive video for computerised self-instruction. *Medical Teacher, 9*(3), 309-315.

Weatherall, D. J. (2006). Science in the undergraduate curriculum during the 20th century. *Med Educ, 40*(3), 195-201.

Webb, G. K. (2003). A Rule Based Forecast of Hard Disk Drive Costs. *Issues in Information Systems, 4*(1), 337-343.

Webber, R. H. (1992). Structured short-answer questions: an alternative examination method. *Medical Education, 26*(1), 58-62.

Weinberg, A., & Knoerr, H. (2003). Learning French Pronunciation: Audiocassettes or Multimedia? *Calico Journal, 20*(2), 315-336.

Weisberg, D. E. (2008). The Engineering Design Revolution: CadHistory. net.

Weisenthal, J. (2008). Analyst Whacks Entertainment Industry: Major Cannibalization Set To Begin… Now. *paidcontent.org*. Retrieved from http://paidcontent.org/article/419-analyst-whacks-entertainment-industry-digital-too-disuprtive/

Wells, W. (1862). *The graded school: A graded course of instruction for public schools: with copious practical directions to teachers*. New York: AS Barnes & Burr.

Wetzel, C. D., Radtke, P. H., & Stern, H. W. (1994). *Instructional effectiveness of video media.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Weyhrich, S. (2009). Apple II History Timeline Retrieved August 20, 2009, from http://apple2history.org/history/appy/ahb3.html#1984

Whitaker, A. (2010). Instant Anatomy-Podcasts Retrieved June 2007, from http://www.instantanatomy.net/podcasts.html

Whitney, E., & Pessina, M. (2008). Does Availability of Audio Podcasts Enhance the Classroom Experience for First Year Dental Students? *International journal of instructional technology and distance learning, 5*(8). Retrieved from http://www.itdl.org/Journal/Aug_08/article03.htm

Whyte, J. (1931). On the Use of Phonograph Records in the Teaching of German Pronunciation. *The German Quarterly, 4*(4), 170-174.

Wikipedia (2009). Early IBM disk storage Retrieved August 20, 2009, from http://en.wikipedia.org/wiki/Early_IBM_disk_storage

Wilbrink, B. (1997). Assessment in historical perspective. *Studies in Educational Evaluation, 23*, 31-34.

Wiliam, D., & Black, P. (1996). Meanings and consequences: a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal, 22*(5), 537-548.

Williams, B., & Bearman, M. (2008). Podcasting Lectures: the next silver bullet? *Journal of Emergency Primary Health Care (JEPHC), 6*(3). Retrieved from http://www.jephc.com/uploads/990309BW.pdf

Williams, C., Aubin, S., Harkin, P., & Cottrell, D. (2001). A randomized, controlled, single-blind trial of teaching provided by a computer-based multimedia package versus lecture. *Medical Education, 35*(9), 847-854.

Williams, M. (2003). *Deforesting the earth: from prehistory to global crisis*. Chicago, USA: University of Chicago Press.

Williamson, E. G. (1936). The Role of Guidance in Educational Methodology. *The Journal of Higher Education*, 184-190.

Williamson, S. H. (2009). Six Ways to Compute the Relative Value of a U.S. Dollar Amount, 1790 to Present Retrieved August 16, 2009, from http://www.measuringworth.com/calculators/uscompare/result.php

Wilson, R. (2003). MITS 1976 Prices Retrieved August 20, 2009, from http://rwebs.net/micros/altair/1976prices.htm

Wintersgill, P. (1959). Auscultation on the Gramophone. *British Medical Journal, 2*(5162), 1336.

Wisan, J. (1940). Radio Manual: A Compilation of Radio Broadcasts for Mouth Health Education. *American Journal of Public Health, 30*(6), 698.

Wong, K. (1992). The use of the liquid crystal display (LCD) panel as a teaching aid in medical lectures. *Medical Teacher, 14*(1), 33-36.

Wood, B. (2000). Feedback: A Key Feature of Medical Training. *Radiology, 215*(1), 17-19.

Wood, E. (2003). What are Extended Matching Sets Questions? . *Bioscience Education e journal, 1*(1), r2. Retrieved from http://bio.ltsn.ac.uk/journal/vol1/beej-1-2.htm

Woolley, D. (1994). PLATO: The Emergence of Online Community Retrieved December 4, 2010, from http://thinkofit.com/plato/dwplato.htm

Worley, P., Esterman, A., & Prideaux, D. (2004). Cohort study of examination performance of undergraduate medical students learning in community settings. *BMJ, 328*(7433), 207-209.

Wright, A. (2007). *Glut: Mastering information through the ages*. US: National Academy Press.

Yoder, M. E. (1993). Transfer of cognitive learning to a clinical skill: linear versus interactive video. *Western journal of nursing research, 15*(1), 115-117.

Yonge, K. (1965). The use of closed-circuit television for the teaching of psychotherapeutic interviewing to medical students. *Canadian Medical Association Journal, 92*(14), 747-751.

Yusof, K. M., Tasir, Z., Harun, J., & Helmi, S. A. (2005). Promoting Problem-Based Learning (PBL) in Engineering Courses at the Universiti Teknologi Malaysia. *Global J. of Engng. Educ, 9*(2), 175-184.

Zakon, R. (2010). Hobbes' Internet Timeline - the definitive ARPAnet & Internet history Retrieved November, 2010, from http://www.isoc.org/zakon/Internet/History/HIT.html

Zinn, K. (1967). Computer technology for teaching and research on instruction. *Review of Educational Research, 37*(5), 618-634.

Zisman, A. (1994). Buying a computer for school. *Our Computer Player,* (October). Retrieved from http://www.zisman.ca/Articles/1994/School.html