JIAMSE

Journal of the International Association of Medical Science Educators

Volume 17 | Number 1 | 2007

IAMSE

Digital Classroom - Podcasting

Expert and Non-expert PBL Tutors

Transition to Computer Assessment

An Adaptable Web-Based Atlas

Medical School Student Attitudes in Nepal

IAMSE on the Web
www.iamse.org

ISSN: 1550-8897

JIAMSE

The Journal of the International Association of Medical Science Educators

Volume 17	Number 1	2007

Message from the President

Edward P. Finnerty, Ph.D.
President, IAMSE

For the past 90 and more years the Flexnerian model has dominated medical education. The single greatest influence Flexner had was to call for the integration of science to the study and practice of medicine. This model has served society well. We have seen a better perception by society of our health care practitioners, great improvements in the quality of our health care providers and significant advances in our understanding and treatment/management of health and disease.

Today, in medical education we are faced with great pressures to add more information and skill training to our curriculum. The expectations for us, as educators, and our students have increased as we have amassed greater understanding and knowledge of biomedical sciences. In addition to the 'hard science' knowledge, we also need to recognize and incorporate the behavioral, social and business aspects that are required for a complete medical education.

The end result is a very full and demanding curriculum of study. Holding to the traditional model of two years pre-clinical and two years of clinical study has led to a state of near exhaustion for our students. Over the decades since Flexner's report we have revised and enhanced the clinical experiences and reformed *ad nauseam* the pre-clinical years. We have spent most of our efforts on the pre-clinical years with the greatest changes involving scheduling and use of various instructional strategies.

The roots of IAMSE are in the basic sciences and their relation to medical education. As medical educators we are closest to and the ones who can make a difference in the content and structure of the sciences in the medical education curriculum. IAMSE is embarking on a project to define what sciences should constitute the sciences basic to medical education and practice, the added value of the sciences in the medical education process and where and how best to provide instruction and learning of these sciences.

Obviously, this project cannot be completed without collaboration of our colleagues in the clinical sphere. To determine what sciences are needed as foundational requisites for clinical medicine, we must first define the scope and requirements for our students as they enter their clinical experiences.

To this end, IAMSE in collaboration with the Alliance for Clinical Education (ACE) and the Generalists in Medical Education will be conducting a series of meetings and workshops to address these topics. IAMSE will also be recruiting and collaborating with other professional medical education groups and associations to accomplish this project. It is our intent to present a comprehensive report at the 2010 IAMSE meeting, tentatively planned for Louisville, Kentucky, the professional home of Abraham Flexner on the 100th anniversary of his landmark report.

With the strength of IAMSE, its members, and their collective wisdom regarding the foundational sciences in medical education, we can and should take the lead in this issue. I would ask all to begin a dialogue and conduct research to address these issues. We should not limit ourselves to the role of the sciences simply in the pre-clinical years, but look to the pre-medical curriculum as well as integration in the clinical undergraduate and even the graduate clinical years. The Journal, JIAMSE, would then be the ideal place to publish your research on this topic. This July, at the annual IAMSE meeting in Cleveland, we will have a focus session workshop to begin a discussion on these issues.

As you peruse the agenda for this year's Annual Meeting (http://www.iamse.org/conf/conf11/index.htm) meeting, you will find some very useful sessions addressing everything from instructional strategies, to assessment techniques and even a debate on the topic of medicine as a science. The session topics and posters are interesting and intriguing. This meeting is looking to be an exciting adventure with the opportunity to meet and talk with friends and colleagues from throughout the world.

With regard to other IAMSE business; after a number of years of faithful service, **Regia Kreisle, Tom Schmidt and Deb Vaughn** will be retiring from their IAMSE Board positions. We thank them for their contributions. Replacing them on the Board are **Frazier Stevenson, John Szarek and Mark A.W. Andrews.** Congratulations to these newest additions to the IAMSE Board who will assume their duties in July.

If you have not registered for the 2007 meeting in Cleveland, it is not too late. Neither is planning for the 2008 meeting in Salt Lake City. Have a relaxing summer and looking forward to seeing you all in Cleveland in July.

Message from Editor-in-Chief

Uldis N. Streips, Ph.D.
Editor-in-Chief

It is most gratifying to note the increased numbers and quality of contributions by the membership, both of articles, as well as Innovations and Monographs. The launching of the first Supplement was a success. The response to the call for reviewers has also yielded many new and capable professionals for the future review of all contributions. We hope to hear positive news from listing services this year. In sum, things are going very well with the journal and the editorial team is working hard.

That said, if we hope to reach the first tier of medical education publications, which is a very reachable goal, we require cooperation from the membership. We need for all of you to consider JIAMSE first when you submit papers for review. We need even more and the highest level submissions. Secondly, we need to reference the work published in JIAMSE far more. We were reviewed recently by Current Contents and their view was "While JIAMSE has left a favorable impression with the evaluators, the group notes that the journal is not demonstrating adequate levels of citation activity." The only way in which this can be improved is if the authors who publish in our Journal also reference papers previously published in the JIAMSE. We have the volumes archived on the web and it shouldn't be hard to peruse them to determine if any articles there are relevant to your work sent in for consideration. If we all work together, good things will happen for the Journal.

If you have educational research that needs publishing, look to the JIAMSE. You will be in good company. I look forward to hearing from you. All best,

Uldis N. Streips, Ph.D.
Editor-in-Chief, JIAMSE

The Medical Educator's Resource Guide

John R. Cotter, Ph.D.

It is my guess that the teaching materials and tools prepared by the faculty of some research universities are seldom or minimally recognized for their instructional value and quality and that the faculty that produce them are even less often credited or rewarded for the effort and time spent producing them. That, however, may be changing for faculty working in the field of medical education.

*In a seminar given in April of 2006 at the Northeastern Ohio Universities Colleges of Medicine and Pharmacy and captured by the school on video, **Robby Reynolds** spoke about the Association of American Medical Colleges (AAMC) initiative that is intended to elevate the standing of outstanding teaching materials in the promotions process.* The website, called **MedEdPORTAL**, publishes peer-reviewed teaching materials for world-wide distribution. In the video, Reynolds, explains that the materials are accepted for publication only after being subjected to a review process that is similar to the one used by **Academic Medicine** in the publication of journal articles. Thus, instructors are afforded the opportunity to receive publication credit for their endeavors, and according to MedEdPORTAL, the publication of the materials by the website "should be considered a compelling scholarly contribution suitable to support promotion and tenure decisions."*

Reynolds continues by saying that MedEdPORTAL will consider any topic dealing with medical education, and that there are no restrictions on the form of the material submitted. For example, case studies, animated artwork, atlases and PowerPoint presentations are published on the website. He explains that the materials need not be web-based and that the site has not as of yet matured into a repository for the materials. As it is currently constituted, the web-based resources are linked to the servers that host them while the authors of the non web-based materials, e.g., paper cases and CD-ROMs must provide the materials to interested parties at no charge (except perhaps for duplication and shipping costs) upon request.

*More information on how to publish teaching materials through MedEdPORTAL can be obtained by visiting the website or linking to the streaming online video seminar. ***

***Google Scholar** and **Wikipedia**, which are also reviewed in this issue of The Medical Educator's Resource Guide (MERG), remind us that users of the Internet are looking for speed and convenience when using websites as sources of educational information, and that they want to know that the information is reasonably complete and accurate. This is the reason the MERG is so useful. In each instance, the sites have been tested in the classroom or office, and have satisfied the needs of faculty and students that use them in their work and studies. If you have experience with a website that offers similar benefits to teachers and students working in the medical sciences please consider sharing your experience with us. You can do so by contacting us by e-mail (jrcotter@buffalo.edu).*

Google Scholar. Google, Inc.

www.scholar.google.com

The team of search-engine gurus at Google(r) has devised a search engine for those interested in finding topics or authors published in scholarly journals. Google Scholar is an astonishingly fast engine capable of searching a word, phrase or name. The most important information concerning the article (title, authors, source, and year) is generated instantly along with several other features which allow the user to quickly obtain the full text or follow its citations to other articles that may be of interest. The importance and position in the list of hits is determined by an articles' citation rate. The key word is returned in bold font (if present in the title) with a brief excerpt from the text. The user can utilize one of many links, some of which include links to the "full text", "citations", "related articles" or "web search" to travel wherever he or she pleases. Perhaps the best feature of Google Scholar is its ability to return not only the full text article, but to return a "grouping" of citations that allows the user to access the same article through many different vectors if the article has been published by more than one source (online journals, online magazines, websites). This feature gives the student, professor, or researcher the ability to obtain the article through their University's library of digital journals and holdings. (*Reviewed by James Marusich, B.S., University at Buffalo*)

MedEdPORTAL. Association of American Medical Colleges.

http://services.aamc.org/jsp/mededportal/goLinkPage.do?link=home

This website is an arm of the Association of American Medical Colleges. The goal of this site is to make instructional materials dealing with any relevant subject available to the medical education community. Given the potential size of the site, users are given four ways to perform a search of the database. The simplest method involves searching with keyword, hot topic, or by discipline. A longer form that asks for more detailed input is used to narrow the outcome. A summary of the results is neatly displayed and organized and the materials can be accessed through links to the identified materials or obtained by contacting the copyright owner. The newness of the site means the coverage of a given subject is likely to be less than ideal. The extensiveness of the coverage can be determined by surveying the number of items attributed to each of the disciplines. The number of items is sure to grow as educators become aware of the service. The site is free, though users must register prior to using it. (*Reviewed by John R. Cotter, Ph.D., University at Buffalo*.)

Wikipedia, The Free Encyclopedia. Wikimedia Foundation, Inc.

http://en.wikipedia.org/wiki/Main_Page

Wikipedia makes finding general information on basic science topics and medicine effortless. Searching the site by using a keyword results in a direct hit or yields a list of potentially relevant articles. The articles are well organized and may be illustrated. The content, which is written and edited by users, is overseen by a group of "trusted users". The methodology has sparked a debate regarding the wisdom of using the website. The issue is acknowledged by the website in articles dealing with the reliability and criticism of the website's content. The ease with which information is obtained while carrying out other tasks on a computer makes the site an attractive alternative to broader searches of the Web. Unfamiliar terms are quickly defined, set within the context of a related subject and linked to pertinent articles. The scale of the site and the organization of information make the site perfect for conducting fast searches and initiating research. It is, as the encyclopedia puts it, "a good starting point for research" but the "articles will, by their nature, vary in standard and maturity." (*Reviewed by Zehra Aftab, B.S., University at Buffalo.*)

* Robby Reynolds is Co-Director of MedEdPORTAL. He has given the MERG permission to use his name and summarize the contents of his seminar.

* * The seminar "MedEdPORTAL: Access and Submission Process" was given by Robby Reynolds on April 13, 2006 and accessed on April 27, 2007. A streaming video of his talk shall be found on the Northeastern Ohio Universities Colleges of Medicine and Pharmacy Faculty Development website at http://www.neoucom.edu/audience/faculty/ProfDev/development/TeachingTalks.html.

COMMENTARY

Delivery of PC-Based Digital Classroom Presentations by Video Podcasting

Charles P. Hutchison, D.V.M.

Department of Structure & Function
School of Veterinary Medicine
Ross University
P.O. Box 334 West Farm
Basseterre, St. Christopher & Nevis, West Indies

Phone: (+) 869 465-41614 Fax: (+) 869 465 1203 Email: chutchison@rossvet.edu.kn

ABSTRACT

The purpose of this article is to heighten awareness concerning an alternative method for delivery of multimedia-rich digital classroom presentations in a manner providing significant ease of access and portability. The method targets Windows® PC-Based presentations and employs syndication by simple internet broadcasts, currently referred to as *podcasts*, of large volumes of such presentations in a video-feed format for viewing on small hand-held devices. The goal is to offer students another alternative format for some of the materials currently used in the classroom or published on the internet in various formats. The methods for producing this content require: 1) special formatting of the initial digital presentation for maximum clarity on the small screen; 2) rendering and translating the video product for podcasting; and 3) production of a web page and appropriate links for the end-user to gain access to and use the materials at large.

To date, podcasting has primarily been used to feed simple text, still photos or MP3 audio. What makes this method somewhat novel is the production of the next level of podcast output: video MP4 files. Currently, video podcasts are being offered to first- and second-year students in gross anatomy and neuroscience classes at Ross University School of Veterinary Medicine to offer an alternative learning format to facilitate the needs of students with varied learning styles.

INTRODUCTION

As new technology emerges, it is the responsibility of faculty at all levels to keep abreast of these advancements and to remain open-minded, concerning the applications of such technology in their own teaching endeavors. Software applications that were once mind-boggling to utilize have become more user-friendly as methods toward their application and frequency of their use expands. The introduction and advancement of IT resources in schools and colleges has sparked more interest and facilitated many educators to apply techniques for learning beyond the typical classroom environment; one such application target being digital distance learning.

Distance learning itself is not a new idea, having beginnings in the mid 1800s with correspondence courses via the mail being conceived by Sir Isaac Pitman[1], but the availability of the World Wide Web and advancements in computer technology have made the digital application of these practices mainstream in contemporary education.
This article explores a novel method for the application of distance learning and introduces educators to the techniques necessary to produce podcasts of their lectures in a multimedia format using RSS (Really Simple Syndication) feeds.

The term "podcast" is a relatively new term combining the words "iPod" (Apple, Inc.) and "broadcasting" first coined by British journalist Ben Hammersley in 2004.[2-4] The technology for such productions emerged and developed in 2000-2001 and podcasts started hitting the web commonly after the first transmission in 2003.[5] Currently, most podcasts are related to news and audio feeds and exhaustive web searches for digital multimedia presentation podcast feeds of higher education content will yield few, if any, published links.

In the practical application of podcasting, the multimedia content will be broadcast via the internet and loaded into one or more devices for viewing. In reality, such devices can be anything that can play video content of the specific type produced (computers, hand-held devices, DVD players), but the idea presented here is to target a small, portable device for ease and frequency of access to the information podcasted.

An alternative use for this format targets the presenter. Persons that must deliver lectures or seminars utilizing digital presentations can use podcasts on their own devices for study and rehearsal prior to delivery of their presentation. The ease of access of small portable devices makes this a favorable choice to consider.

TECHNIQUES

The sample podcasts produced and published for this tutorial were designed for an end-user with a handheld-type multimedia device capable of playing the specific type of media, a computer to dock the device to and a broadband internet connection. The specific device utilized to view the final video podcasts produced is an iPod® with color video capability. Herein, we will explore the techniques and necessary hardware and software to complete such podcast content.

Four major steps to accomplish the podcasts will be outlined below: 1) Proper formatting of the digital presentation; 2) rendering of the digital presentation to a continuous-feed video format; 3) translating the initial video to device-compatible video format; and 4) electronic publication of the final product with instructions to the end-user for their application. The details of each of these steps is completely outlined in an online presentation which provides complete instructions including examples and links to any and all software necessary to implement your own podcasts of similar content at www.rossvet.edu.kn/podcasts/test/tutorial.htm.

The first steps employ the use of PowerPoint® (Microsoft Corporation) as a standard digital classroom presentation medium. The presentation content must be conformed to high-resolution and contrast standards and specifically formatted for viewing on a small-screen as a continuous video feed. Poor formatting includes the use of incorrect fonts or font sizes, too much text per slide, poor color schemes, inadequate contrast between backgrounds and foregrounds, inadequate or excess animation. Basically, if it presents poorly on the large screen, it may appear even worse on the small screen.

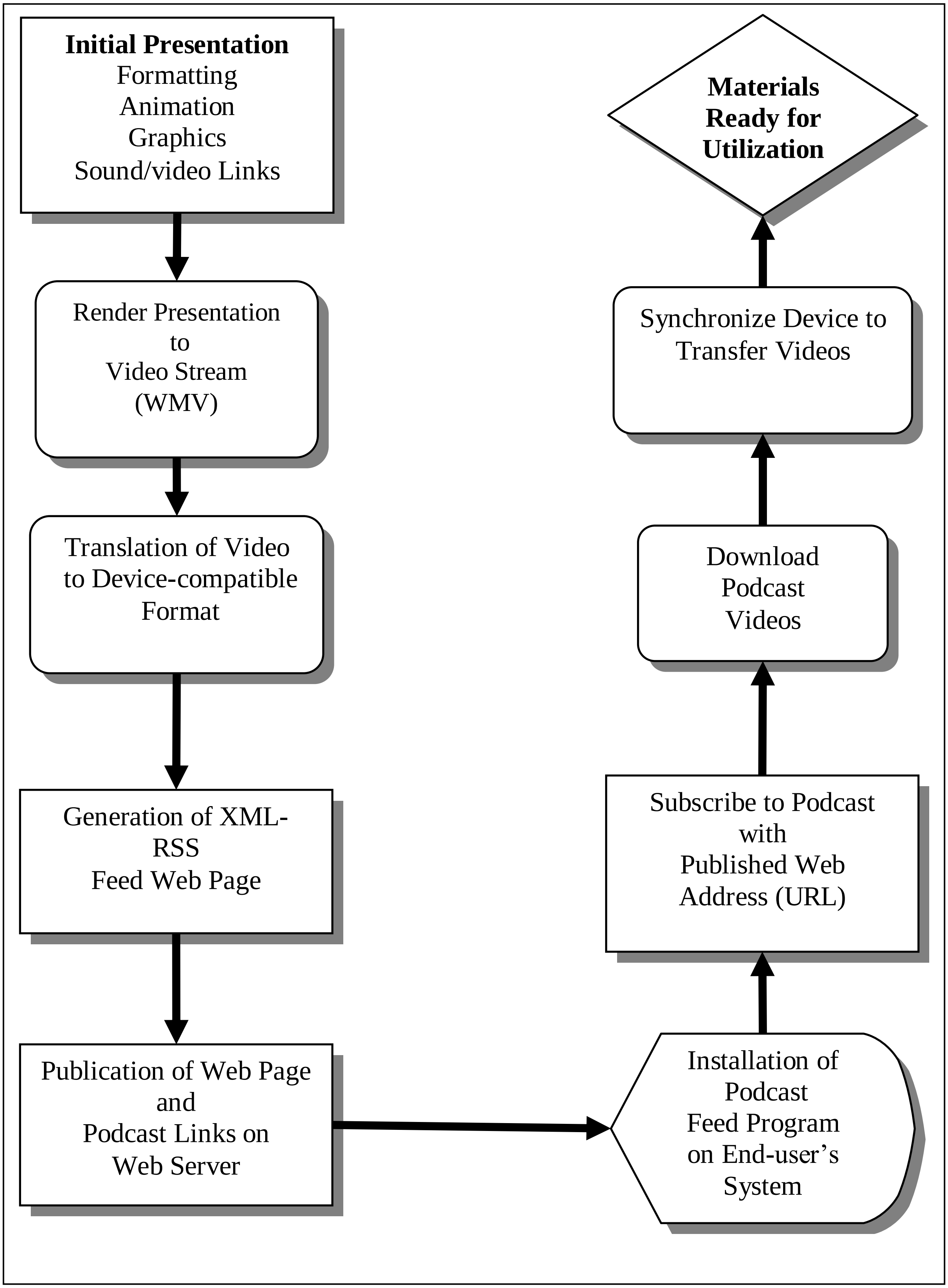
It is recommended to employ the default presentation ratio and dimensions (4:3 width:height ratio: 10x7.5 inches) with a dark-colored gradient background and off-white colored text. For readability, the font should be a sans-serif face (Arial, Comic Sans MS) with an optimal size of 32-40 point and a minimal size of 28 point. To maximize the esthetic quality of the presentation, text should be animated so that it does more than simply appear and lie on the screen. Text that is animated to appear word by word or paragraph by paragraph allows the viewer to read along as text appears. Further formatting is done so as to ensure that the final presentation podcast video flows from slide to slide with some calculated delay time; otherwise slides progress before the viewer is finished with them. This is accomplished by using a "pause indicator" to let the viewer know that all content for that slide is finished. Once they see the "pause indicator" appear, they can then hit the "pause" key on their device to temporarily stop the slide sequence and allow for more time on the slide. For the "pause indicator," the word: "pause" was animated to materialize at the end of each slide after a slight delay from the emergence of the final slide content.

The next considerations involve numbers of slides and details of content that affect file size. First, the initial size of the presentation is critical as it dictates the final size of the podcast. It is imperative to minimize the size of the podcast for timely internet downloading without deteriorating its quality and effectiveness. To accomplish this, keep podcasts under 20 megabytes (MB) in size by sectioning presentations into parts if podcast file sizes become too large.

Further, any images embedded into the presentation should be no larger than the presentation screen size. The type of image used, for example JPEG, GIF, PNG, is not as critical as the clarity of the base image. All images used in presentations in this project are one of these three types. If the base images are larger than the presentation screen size, they are rendered at no greater display resolution, but result in a higher presentation file size. For example, two one-slide presentations were made with the same photograph placed into each one. In the first presentation, the photograph had a base size of 3072x2304 pixels and in the second presentation it was resized to 800x600 pixels (6.8% of the total square pixels of the original image). The presentation file sizes were 1573 kilobytes (kB) and 95 kB, respectively, and the resolution of both photos is the same on screen and the two presentations could not be visibly distinguished.

Likewise, if the original images are too small, they may be difficult to see on a small-screen device and enlarging them within the presentation may result in significant "pixelation." Optimal image size is between 800x600 and 960x720 pixels. Many simple programs can be utilized to resize base images including the presentation software itself. The file size for a simple podcast of a one-hour lecture, with still images, exclusive of embedded videos and sound, was 23Mb from an initial 60-slide PowerPoint presentation size of 8.9Mb. See this link to view this sample podcast in your browser: http://www.rossvet.edu.kn/podcasts/test/test1.mp4.

Figure 1. Flowchart Demonstrating Step-by-step Procedure for Production of Video Podcasts from Digital Classroom Presentations



Sound files such as clips of the lecturer's voice can be linked into presentations. However, when the podcast is rendered from the presentation, it will call the sound file and embed it into the final video product, thus increasing output file size significantly. For this reason, it is imperative to keep linked sound files as small as possible and use them only when necessary. In this project, short audio clips (less than 10 seconds) were recorded and linked from certain slides which, in the author's opinion, needed further explanation. All audio clips were recorded and saved in standard Windows "wav" file format. A podcast of the presentation with embedded audio clips rendered at 14Mb from an initial 18-slide PowerPoint presentation size of 1Mb. See this link to view this sample podcast in your browser: http://www.rossvet.edu.kn/podcasts/test/test2.mp4.

Finally, video files linked within presentations can be rendered into podcasts in a similar manner as audio. However, video clips have much larger file sizes than audio clips of similar length and when they are rendered in this process they become fully embedded. To preserve video quality, while maximizing compression to minimize video file size, the Windows Media Video (WMV) compression format was utilized. A podcast of a short presentation with embedded video rendered at 15Mb from an initial 10-slide PowerPoint presentation size of 1Mb with one embedded video per slide. This link shows this sample podcast in your browser:
http://www.rossvet.edu.kn/podcasts/test/test3.mp4.

The next step involves rendering the presentation into a video file. Inexpensive software packages are available to perform this task with excellent results. For example:
PowerVideoMaker Professional V2.2.0, PresenterSoft,Inc., www.presentersoft.com. Such software is capable of producing feeds in WMV, MPEG or AVI video formats. The format chosen for maximal size compression and quality was WMV at 8000Kbps and a resolution of 640x480 pixels. Conversion of the presentation files to video format is a time-consuming process and typical programs must run uninterrupted on a computer system during the conversion. Conversion of presentation number 1 (test1.mp4 above) took about 42 minutes on a Pentium®-4 system with 1MB RAM and a 2.8GHz processor.

Then, to render the podcast involves translation of the video to a device-compatible mode. Our target devices utilize video in DVD-quality MPEG-4 with Advance Audio Coding (AAC) audio compression. The podcast videos were produced from the raw WMV videos using a simple freeware package (Videora iPod Converter V 0.91, Videora Holdings, www.videora.com). The program default output format was chosen for all podcasts produced: 320x240 pixel video resolution at 768kbps and stereo audio at 128kbps.
The final step is to publish the podcast in some manner. This involves using Extensible Markup Language (XML) and Really Simple Syndication (RSS) to construct the podcast feeds. The student can then use their RSS capable browser or other software to subscribe to the podcast. In this project, a freeware program (RSSReader, V 1.0.88.0, www.rssreader.com) was tested and utilized to subscribe to podcasts and download the videos. Using this podcast subscription program allows the podcasts to be automatically updated so that changes to existing podcasts as well as podcasts added or deleted is routine and inherent to the RSS feed program's notification mode. Once the video podcasts are located, they can be downloaded to your computer system and then migrated into your handheld device with the appropriate device synchronization program (iTunes 7, Apple Computer, Inc.). Figure 1 demonstrates a summary flowchart of the steps of this publication process.

Assuming the use of presentations of typical type and average length and in comparison with results demonstrated in the test examples published above, it would take somewhere between 25 to 50 MB of video podcast feeds with video runtimes of at least 15 minutes to deliver a typical one-hour lecture. This equates to nearly 4000 MB (4.0 gigabytes [GB]) of video with a total runtime of 20 or more hours for 5 hours of lecture per course weekly over an entire semester. With current handheld devices having storage capacities in the tens of gigabytes, there is adequate room for storage of video podcasts for multiple courses on a single hand-held device.

REFERENCES

1. Matthews, D. The Origins of Distance Education and its Use in the United States. *The Journal* 1999. Available at http://www.thejournal.com/articles/14278. Accessed September 25, 2006.
2. Wikipedia.org Podcast. *Wikipedia Foundation, Inc.* 2006. Available at http://en.wikipedia.org/wiki/Podcast. Accessed September 26, 2006.
3. Hammersley, B. Audible Revolution. *The Guardian Unlimited, Digital Edition*. 2004. Available at http://technology.guardian.co.uk/online/story/0,,1145689,00.html. Accessed October 11, 2006.
4. Meyers, M. 'Podcast,' the Word of the Year. *CNET Networks*. 2005. Available at http://news.com.com/2061-11199_3-5986198.html Accessed October 11, 2006.
5. Doyle, B. The First Podcast. *Information Today*. 2005. Available at http://www.econtentmag.com/Articles/ArticleReader.aspx?ArticleID+13515. Accessed October 10, 2006.

REVIEW

Expert and Non-expert Tutors: Role in Problem-Based Learning

Ihab A.M. Ahmed, MBBS, FRCS, CME, MD, FRCS (Gen) [1], Maha A.M. Elseed, MBBS, MRCPCH [2], and Mohamed. A. El-Sheikh, MBBS, FRCS, MSc, FRCS (Gen) [3]

[1] Bishop Auckland General Hospital
Cockton Hill Road
Bishop Auckland
Co. Durham DL14 6AD
UK
[2]Newcastle General Hospital
Westgate Road
Newcastle upon Tyne NE4 6BE
[3]Nottingham University Hospital NHS Trust
City Hospital Campus
Nottingham Renal Unit
Nottingham NG5 1PB
UK

Phone: (+)1-44-1388455180 Fax: (+)1-44-1388455054 Email: ihabahmed@doctors.org.uk

ABSTRACT

Problem-based learning (PBL) is considered by many to be an important innovation in medical education. Integral to the success of PBL is the role of the tutor, which is considered a prime determinant of how the tutorial group functions. Recently, there has been considerable debate about the role of subject expert and non-expert tutors in facilitating PBL session and who will be more effective and efficient in facilitating the learning experience. The aim of this review is to evaluate the role of expert and non-expert tutors in facilitating PBL sessions.

An electronic search of the Medline database was undertaken for the articles published on tutoring in PBL with specific reference to expert and non-expert tutors in PBL. Relevant articles were chosen for review and analysis.

Tutoring in PBL is a multifaceted process, with the issue of the tutor expertise forming only one part of this complex process. Published literature was equivocal on this issue because of the inconsistency in defining expert and non-expert tutors and the inconsistency in evaluation tools used. There is no consistent evidence to suggest that groups facilitated by expert tutors do better, in terms of student academic achievement and student perception, than those facilitated by non-expert tutors.

INTRODUCTION

Problem-based learning (PBL) was first developed at McMaster University in the 1960s and is regarded by some as the most important innovation in medical education. Definitions of PBL vary, but a comprehensive example would be “an educational method characterized by the use of patient problems as a context for students to learn problem-solving skills and acquire knowledge about the basic and clinical sciences”.[1]

The PBL approach is based on principles of adult education and cognitive psychology. It differs fundamentally from traditional curricula, in which students acquire "background" knowledge of the basic sciences in the early years of the course and in the later years apply this knowledge to the diagnosis and management of clinical problems. The traditional approach has been criticized for several reasons including the fact that it creates an artificial divide between the basic and clinical sciences, wastes time in acquiring knowledge that is subsequently forgotten or found not to be useful, and for the difficulty in the application of the acquired knowledge. In addition, in traditional learning, the acquisition and retention of information may seem irrelevant and can even be perceived as boring to the students. Problem-based learning, with its educational objectives, can avoid many of these problems. Various disciplines, particularly the basic and clinical sciences, are integrated throughout the curriculum. As the students attempt to understand and solve clinical problems, they learn about normal bodily structure and function, and apply this knowledge for their search for a solution. Learning occurs in this context and builds on what students already know. PBL can aid retention, add interest and increase motivation to learn. Students, with initial help from tutors, determine their own learning needs and the strategies they need for learning e.g., the efficient accessing of library resources or the formation of study groups.[1]

Integral to the success of a PBL program is the role of the tutor. The tutor performance is considered (along with the student prior knowledge and the quality of the cases discussed) to be a prime determinant of how the tutorial group functions.[2] The main role of the tutor in PBL is to facilitate the proceedings, to ensure that the group achieves appropriate learning objectives in line with those set by the curriculum design team, to ensure that all the students have done the appropriate work and to help students self assess their understanding of the material by encouraging them to ask questions.[3] Some authors have argued that subject-expert tutors would facilitate the tutorial process more effectively and more efficiently than non-experts, because such tutors were more comfortable tutoring in areas of their expertise.[4] Recently, there has been considerable debate about this theory and about the role of expert and non-expert tutors.

EXPERT AND NON-EXPERT TUTORS

There has been no published uniform definition of "tutor expertise" in PBL. Generally, most of the published literature about tutoring in PBL, have divided expertise into two categories:

1. Some authors have defined an expert tutor as a tutor with previous PBL experience.[2] Expertise was further subclassified by these authors according to the number of PBL modules tutored by a particular tutor.
2. Other authors have defined an expert tutor as a subject-matter expert.[4-8] However, different authors have employed different definitions of what constitutes subject-matter expertise in tutoring. Some authors have defined expertise in terms of background training. For example, Schmidt described non-expert tutors as either those whose previous training was unrelated to the problem at hand or tutors who have a general background knowledge regarding the unit, but not specific expertise.[9] He describes expert tutors as those with specific background knowledge relevant to the unit.[9] Bochner and coworkers defined an expert as one who displays specific skills or knowledge derived from training or experience.[2] Dolmans and colleagues have defined expertise as the tutor self-perceived subject matter expertise,[8] whereas Schmidt and Moust defined expertise as a tutor with advanced disciplinary training and/or research experience in the problem under study.[7]
3. In addition, expertise has been further defined by whether the tutor is medically qualified and/or whether he/she holds an academic (compared with non-academic or a student tutor) appointment.

These inconsistencies in the definition of expertise might have contributed to the variable results from different studies, as discussed later.

DIFFERENCES IN OUTCOME BETWEEN EXPERT AND NON-EXPERT PBL TUTORS

Studies that have examined the effect of expertise (with specific regard to subject-matter expertise) have used different methods to measure the outcome. Most studies have employed the academic achievements of students at the end of the unit or module as the major outcome. Other studies have used different end points, including student and tutor perception or time spent on studying. Other studies have used a combination of two or more assessment methods.

STUDENT ACADEMIC ACHIEVEMENT AS A MEASURE OF OUTCOME

Some authors found that students guided by subject-matter experts achieved better than those guided by non-experts. Schmidt and coworkers examined the effect of expertise in PBL and compared students academic achievements relative to expert and non-expert PBL tutors.[5] Achievement was measured from 100-150 true-false items and by short essay questions. They found that students guided by subject-matter experts, defined as tutors who have specific knowledge derived from training or experience, achieved better than did the students guided by non-expert tutors. The effect of subject matter expertise on achievement was strongest in the first curriculum year, suggesting that new students are more dependent on their tutor expertise than are the more advanced students. In another study to assess the effect of subject-matter expertise on outcome, Hay and Katsikitis assessed the student achievement by administering a random set of five questions to each student at the end of the module.[10] Students who were taught by experts scored higher in the end-of-course test in the topic area, suggesting that students guided by an expert have a better learning outcome. Expertise was again defined as professional competence in the subject.

Dolmans and coworkers also compared groups guided by subject-matter expert versus non-expert tutors. Expertise here was defined as the tutor's self-perceived subject matter expertise. In this study, measurement of achievement consisted of three parts for each individual student: a multiple-choice test, short essay questions and a tutor assessment regarding the student performance.[8] No difference emerged from the multiple-choice tests. In one of the two years studied, the students guided by the expert tutors performed significantly better on the essay tests. The authors attributed this finding to the fact that the experts were actually involved in the production of the questions for the tests, which may have contaminated the results.

However, other authors also demonstrated that there was no difference in performance of students in PBL groups led by subject-matter experts or non-expert tutors. Steele and colleagues compared learning outcomes in PBL groups led by students and those led by faculty.[11] Learning outcomes were assessed by performance on objective examinations covering factual materials pertinent to the case. No differences were detected in student performance on the objective evaluation, based on whether the facilitator was a faculty member or peer group member. Similarly, in a retrospective study to establish whether or not tutor subject-matter expertise influences student achievement on content-based examination in a PBL curriculum, no difference was demonstrated between groups led by subject-matter experts or non-expert tutors.[6] Assessment was conducted by end-of-block modified essay questions. In this study, the expert tutor was a medically qualified tutor while the non-expert tutor was from a humanities background. Matthes and colleagues compared the learning outcomes by an end-of-term examination consisting of multiple-choice questions and short-essay questions. No difference was demonstrated between subject-matter experts versus non-experts led groups.[12] In this study, non-expert tutors were peer undergraduate students and junior staff members while expert tutors were senior staff members who had completed post-graduate education. Other studies also did not reveal any difference in the level of achievement of students led by expert or non-expert tutors.[13, 14]

OTHER MEASURES OF OUTCOME

Student perception, tutor perception and student-related factors, e.g. time for preparation, have also been assessed as outcomes to compare expert and non-expert tutors.

Self-report ratings by tutors and students were one of the measures used by Hay and Katsikitis to evaluate outcomes in PBL groups led either by expert or non-expert tutors.[2] There were no significant differences between the student overall ratings of expert or non-expert tutors; however, the non-expert tutor was rated more highly for group management skills. Steele and colleagues compared perceptions of facilitator behaviors and group functioning in PBL groups, led by students or by faculty, using a questionnaire completed at the end of each individual case.[11] No differences were detected in the perception of group process based on whether the facilitator was a faculty member or peer group member. Matthes and colleagues compared the influence of tutor expertise on the process and outcome in a PBL basic medical pharmacology course.[12] Expert tutors were staff members who completed postgraduate education while non-expert tutors were junior staff members. In addition, peer tutors (i.e., undergraduate medical students) were included in the study. The authors found that the tutor performance scores of peer-led groups did not differ from those of staff-led groups. Student weekly preparation time tended to be lower in peer-led groups, while learning time spent specifically on exam preparation seemed to be increased compared to PBL-groups of staff tutors. As part of their study, Matthes and coworkers looked at tutor experience in coaching PBL groups. Groups led by experienced staff, defined as tutors with at least one term of previous PBL tutoring, were found to have significantly higher evaluation scores.[12]

REASONS FOR EQUIVOCALITY OF PUBLISHED STUDIES

The results of the studies reviewed here were generally equivocal. The reasons for that are:

1. These equivocal results might be related to the inconsistent definition of expertise. While few studies looked at expertise as experience in PBL tutoring, most of the studies looked at expertise as subject-matter expertise. There was a wide spectrum in defining subject-matter expertise. For instance, Schmidt described an expert as someone who has specific background knowledge relevant to the unit.[10] Davis described an expert as someone with advanced disciplinary training and/or research experience in the problem under study,[13] while Dolmans employed a self-report measure of expertise by which tutors indicated the extent to which they considered themselves experts.[9]
2. The extent to which students were exposed to PBL may be another confounding factor. It is often observed that students with little to no experience in PBL rely more heavily on their tutors as sources of guidance and information. If these tutors are familiar with the subject matter to be mastered, this may make a difference in learning outcome and level of achievement.[5]
3. The methods of assessment for learning outcomes may confound the results. Some of the studies have assessed the student achievement, in the form of examinations, while others have analyzed the student and tutor perception of the PBL sessions as a measure of outcome.

For the studies that assessed outcomes using an examination, both short-essay type of questions and multiple-choice questions (MCQs) were used. MCQs are probably a better form of assessment because of the consistency of the questions and expected answers. However, short-essay questions might be a subjective form of examination as they can be open to different interpretation.

Another method of assessment was questionnaires. Use of questionnaires can be a very subjective way of assessment (there was no mention of use of a validated questionnaire format in the reviewed studies).

OTHER FACTORS AFFECTING OUTCOME

Schmidt and colleagues demonstrated that the effects of tutor expertise showed up in only some of the first-year units, whereas in other units the effect could not be found.[5] These authors suggested that the effects of tutor expertise on learning must be mediated through an unknown, unit-related factor. Another study by Schmidt concluded that a minimum level of structure is needed for students to benefit from problem-based instruction.[10] This structure can be externally provided, through the structure provided by the learning materials. Alternatively, the structure can be internally provided through prior knowledge available for understanding the new subjects, or offered by the environment in the form of cues of what is relevant and what should be the focus of the activities. If prior knowledge falls short, or if the environment lacks structure, students will turn to their tutors for help and direction. Under these conditions, students who are guided by a subject-matter expert tutor may benefit more than students guided by a non-expert tutor.

PROBLEMS OF EXPERT TUTORS

It is believed by some that subject-expert tutors would facilitate the tutorial process more effectively and more efficiently than non-experts, since such tutors were supposedly more comfortable tutoring in areas of their expertise. Some of the studies we reviewed have challenged this hypothesis. Silver and Wilkerson have shown that subject-matter experts, compared to non-experts, were significantly more directive, spoke more often and for longer, provided more direct answers, suggested more discussion topics and presided over exchange patterns that were predominantly tutor to student compared with student to student.[14] This study suggested that tutors should recognize the potential effects of their authority and knowledge. The study has also raised concerns that students with dominant tutors might miss opportunities to prioritize their learning needs, ask and answer crucial questions and synthesize their learning. The authors concluded that tutor expertise might have deleterious effects on the process of collaborative learning, endangering the development of student skills in active, self-directed learning.

OTHER FACTORS CONTRIBUTING TO TUTORING IN PBL

In addition to the use of subject-matter knowledge, five other sets of behaviors guide tutorial groups. These are: use of authority, achievement orientation, an orientation towards cooperation in the tutorial group, role congruence, and cognitive congruence. Role congruence is defined as the willingness of the tutor to be a “student among the students,” that is, to seek an informal relationship with the students and display an attitude of personal interest and caring. Cognitive congruence was defined as the ability to express oneself in the language of the students, using the concepts they use and explaining things in ways easily grasped by students. This was framed in the context of a theory of problem-based learning.[9] The authors assume that the tutor behavior is one of three factors affecting the way in which small-group tutorials function (the other two being the student prior knowledge and the quality of the problem handled). In turn, the small group functioning would influence time spent on self-directed learning activities and intrinsic interest in the topic studied. Finally, time spent would influence achievement on appropriate tests.[9]
From the above, it can be concluded that tutoring in PBL is a complex process with the issue of tutor expertise forming one part of this complex process. The student prior knowledge, quality of the presented problem and the student-tutor relationship form other parts of this process. Most of the reviewed studies failed to take these factors into consideration, which will limit the interpretation or external validity of these studies.

SUMMARY

The question of the effect of the tutor expertise on student learning has raised considerable controversies. The published literature was equivocal in answering this question. Part of the published literature demonstrated that expert tutors have a positive impact of student learning and educational outcomes, while other studies failed to confirm these findings. Several reasons might explain these inconsistencies, including an inconsistent definition of expertise, differences in the methods used to assess outcome and previous exposure of students to PBL. The authors suggest that the definition of “expert tutor” should be restricted to those who have back-ground professional knowledge and/or expertise in the subject-matter while “non-expert” tutors should be those who have no back-ground knowledge about the subject.

Tutoring in PBL is a complex process with the tutor subject matter expertise forming only part of this process. Future larger studies, taking into consideration factors which resulted in inconsistencies in the published literature and other factors affecting tutoring in PBL, are needed to answer the question of the impact of tutor subject-matter expertise in PBL presentation and the role of PBL in student learning.

References

1. Finucane, P.M., Johnson, S.M., and Prideaux, D.J. Problem based learning: its rationale and efficacy. *Medical Journal of Australia*. 1998; 168: 445-8.

2. Bochner, D., Badovinac, R.L., Howell, T.H., and Karimbux, N.Y. Tutoring in a problem-based curriculum: Expert versus nonexpert. *Journal of Dental Education*. 2002; 66: 1246-1251.

3. Wood, D. ABC of learning and teaching in medicine: Problem based learning. *British Medical Journal*. 2003; 326: 328-330.

4. Schmidt, H.G., Van Der Arend, A., Moust, J.H.C., Kokx, I. and Boon, L. Influence of tutors' subject-matter expertise on student effort and achievement in problem-based learning. *Academic Medicine*. 1993; 68: 784-791.

5. Kwizera, E.N., Dambisya, Y.M. and Aguirre, J.H. Does tutor subject-matter expertise influence student achievement in problem-based learning curriculum at UNITRA Medical School? *South African Medical Journal*. 2001; 91: 514-516.

6. Gilkison, A. Techniques used by 'expert' and 'non-expert' tutors to facilitate problem-based learning tutorials in an undergraduate medical curriculum. *Medical Education*. 2003; 37: 6-14.

7. Schmidt, H.G., and Moust, J.H. What makes a tutor effective? A structural-equations modeling approach to learning in problem-based curricula. *Academic Medicine*. 1995; 70: 708-14.

8. Dolmans, D.H.J.M., Wolfhagen, I.H.A.P., and Schmidt, H.G. Effect of tutor expertise on student performance in relation to prior knowledge and level of curricular structure. *Academic Medicine*. 1996; 71: 1008-1011.

9. Schmidt H.G. Resolving inconsistencies in tutor expertise research: Does lack of structure cause students to seek tutor guidance? *Academic Medicine*. 1994; 69: 656-662.

10. Hay, P.J., and Katsikitis, M. The 'expert' in problem-based and case-based learning: necessary or not? *Medical Education*. 2001; 35: 22-26.

11. Steele, D.J., Medder, J.D., and Turner P. A comparison of learning outcomes and attitudes in student- versus faculty-led problem-based learning: an experimental study. *Medical Education*. 2000; 34: 23-29.

12. Matthes, J., Marxen, B., Linke, R.M., Antepohl, W., Coburger, S., Christ, H., Lehmacher, W., and Herzig, S. The influence of tutor qualification on the process and outcome of learning in a problem-based course of basic medical pharmacology. *Naunyn-Schmiedeberg's Archives of Pharmacology*. 2002; 366:58-63.

13. Swanson, B.D., and Case, S.M. Assessment in basic science instruction: directions for practice and research. *Advances in Health Sciences Education* 1997; 2(1): 71-84.

14. De Volder, M.L., and De Graves, W.S. Approaches to learning in a problem-based medical program: a developmental study. *Medical Education*. 1989; 23: 262-4.

15. Davis, W.K., Nairn, M.E., and Anderson, R.M., Effect of expert and non-expert facilitators on small-group process on a student performance. *Academic Medicine*. 1992; 67:470-4.

16. Silver, M., and Wilkerson, L.A. Effects of tutors with subject matter expertise on the problem-based tutorial process. *Academic Medicine*. 1991; 66: 298-300.

Stages of Concern During Curriculum Change

India Broyles, Ed.D. [1], Mildred Savidge, Ph.D. [2], Evelyn Schwalenberg-Leip, D.O. [3], Kathryn Thompson, Ph.D. [4], Rorie Lee, Ph.D. , M.P.H. [5], and Sarah Sprafka, Ph.D. [2]

[1]Department of Pediatrics
[2]Department of Family Medicine
[3]Office of Faculty Development
[4]Department of Biochemistry and Nutrition
University of New England, College of Osteopathic Medicine
11 Hills Beach Rd, Biddeford, ME 04005
[5]Maine Medical Center, Office of Medical Education
22 Bramhall Street
Portland, Maine 04102

Phone: (+)1-207-602-2694 Fax: (+)1-207-602-5931 Email: ibroyles@une.edu

ABSTRACT

This formative evaluation examines the process of curricular change as a College of Osteopathic Medicine develops a competency-based curriculum. The study explores how faculty perceive the need for change, the process of change, and the results of specific changes to guide the college in the change process and improve the chances of achieving a long-lasting and well-accepted curriculum innovation. Study objectives were to determine status of faculty members regarding**:** a) levels of readiness for change early in the change process, including changes in perceptions after implementation; and b) specific concerns and level of use of the Competency-Based Curriculum. The evaluation is based on well-established principles that focus on change, first in individuals and then in organizations. The Concerns-Based Adoption Model describes seven levels of concern that users experience while adopting a new program or practice. The Stages of Concern (SoC) questionnaire was administered to all basic science and clinical faculty, and to related staff. The results showed an overall profile consistent with an institution beginning a change process. The profile revealed very high awareness concerns, moderately high information and personal concerns, moderate concerns in management, low concerns in impact or consequences, with rising concerns in collaboration and refocusing. The institutional profile shows the typical non-user profile except for some tailing up at Stages 6 and 7, which may indicate resistance to the innovations. Understanding the profile will help in tailoring information and training sessions. Follow-up interviews of a purposeful sample of faculty focused on the five key elements of the Competency-Based Curriculum: overall integration of national competencies; student accountability for learning outcomes; curriculum scope based on clinical relevance; innovative use of curriculum time and instructional modalities, and curriculum expansion allowing student choice. Interviews were analyzed for level of use and stage of concern, bringing to light new themes that will guide faculty development and implementation.

INTRODUCTION

When the Accreditation Council for Graduate Medical Education (ACGME) and the American Osteopathic Association (AOA) set new standards for competency-based curriculum during residency education in the United States, many medical schools replicated the process in an effort to establish a pre-doctoral education foundation for the new core competencies. One medical school organized a committee of 22 faculty and staff to review the current curriculum and implement the competency-based approach to curriculum, instruction, and assessment. Dialogue within the curriculum committee resulted in the identification of five elements as guiding principles for the design of the curriculum:

- Integration of national core competencies
- Curriculum scope based on clinical criteria of commonality and criticality
- Student responsibility for course learning outcomes
- Innovative use of curricular time and instructional modalities
- Student opportunity, flexibility and choice toward professional growth

A focus on national competencies and course learning outcomes as components of the medical college curriculum was a major step in moving to a competency-based educational approach. This represented a significant paradigm shift to move the institution into harmony with a growing and worldwide emphasis on the use of learning outcomes and competencies for the education of all health care professionals. Each discipline and academic subject contributes its own set of outcomes toward the education of the total physician; therefore, the faculty of each discipline/academic subject need to examine their concepts, generalizations, rules, principles, skills and attitudes to reflect a longer-term concern for development of abilities such as research and learning strategies, critical thinking and communication. In directing attention to how students will ultimately apply the knowledge and abilities they acquire, this approach required our medical educators to look beyond the strict boundaries of disciplinary tradition and demands.

The committee leadership wanted to know if the current faculty and staff perceived the need for change, understood the procedures for change, and were aware of specific decisions that were guiding the college toward a new curriculum. A group of medical educators was charged with conducting formative evaluation of the development process itself. By focusing attention on the process of curriculum change, we hoped to improve the chances that a successful curriculum innovation would be both well accepted and long lasting.

This article presents the first stage of a formative evaluation of one institution's curriculum development process and its impact on faculty, students, and the college. This evaluation was based on the Concerns-Based Adoption Model (CBAM) developed from a set of principles of change described by Hall and Hord[1], which have become part of the established language of change theory. These principles are:

- Change is an ongoing process, not a short-term event, and requires ongoing support and resources and it takes time. It is important to have realistic expectations about the time needed to see significant progress.
- Change occurs in individuals first, then in organizations. Curriculum change cannot succeed unless people are ready and willing to implement it. Individual change is difficult if the organization is not supportive of the change
- People go through change at different rates and in different ways. Individual needs for information and training vary greatly.
- As people implement a new program, their concerns change.
- Change agents must adapt to individuals' changing concerns over time to ensure effective organizational change. Faculty may need opportunities to share their experiences and learn from one another.
- Change agents need to consider the larger system in which a program is being implemented, taking into consideration the impact on other individuals and parts of the program or schedule.

Hall and Loucks[2] developed diagnostic instruments and change strategies from these principles. The model combines an examination of stages of concern and levels of use regarding innovations as a framework for planning staff development interventions. Hall and his colleagues[3] created the Stages of Concern (SoC) questionnaire to measure the model's seven levels of concern that users experience when adopting a new program or practice (Figure 1). In reality, everyone involved with a new program has all of these concerns all the time at varying degrees of intensity.[2-3] By identifying concerns at a specific point in time, change agents can understand the dynamics of individuals' perceptions of change.

Figure 1. Stages of Concern

Stage 0--Awareness. Individuals have little concern or involvement with the innovation.

Stage 1--Informational. Individuals have a general interest in the innovation and would like to know more about the substantive aspects of the innovation.

Stage 2--Personal. Individuals want to learn about the personal ramifications of the innovation. They question how the innovation will affect them and often have intense personal responses to new demands and expectations. Stage 2 is often a hurdle in accomplishing implementation.

Stage 3--Management. As individuals learn the processes and tasks of the innovation, they focus on logistics, schedules, materials, personnel, and equipment that will be needed to make the innovation work. Intense concerns at Stage 3 may last for two or more years of implementation.

Stage 4 --Consequence. Individuals focus on the innovation's impact or relevance for learners. Generally, individuals have already implemented and mastered the management of the innovation.

Stage 5--Collaboration. Individuals cooperate with other users in implementing the innovation to mutually find ways to improve student outcomes.

Stage 6--Refocusing. Individuals consider the benefits of the innovation and think of modifications and alternatives that might work even better.

Use of CBAM has not been reported in the medical education literature on curriculum change; however, the model has been applied successfully in nursing and nursing education research. Lewis and Watson[4] used the SoC questionnaire to measure the concerns of 57 nursing faculty about the use of computer technology. Their pre-post study results suggest that the primary concerns of the faculty were informational and that addressing these concerns through workshops increased interest in the innovation. The study showed that the SoC questionnaire was a useful tool to monitor and manage the process of change. The SoC questionnaire was used by Gwele[5] to measure the concerns of nurse educators (n=93) at four nursing colleges during the implementation of a major legislated curriculum reform. The author concluded that when a staff is required to adopt a major curriculum change the normal progression through the stages of concern is impeded. The study suggested that in these situations it may be important to delay adoption until participants can come to terms with the need to adopt the new curriculum. The SoC instrument was also used in a study to assess the concerns of staff during the installation of a telemedicine system and to assure that concerns were addressed during system implementation. Survey findings were used successfully to modify the implementation and training phases of the program to better meet the needs of the staff.[6]

These studies show the importance of recognizing individual and institutional concerns during the change process. By gathering data based on a strong conceptual framework for documentation of the change process, medical educators can provide the Curriculum Review Committee and faculty with targeted support as they move toward development and implementation of a new curriculum.

For this study, the evaluators established three implementation goals and corresponding evaluation questions (Table 1).

Table 1. Study evaluation tasks and related questions

Evaluation Task	Evaluation Question
Determine the level of faculty preparedness for change early in the change process, and to examine changes in their perception after the process has been in place for several months.	What is the SoC profile for COM faculty and staff before and after implementation of curriculum change?
Determine the specific concerns and level of implementation among various faculty members regarding the Outcomes-Based Curriculum.	What are the specific concerns of faculty and staff during the curriculum change process? How quickly do faculty adopt the key elements?
Use formative data to guide the "next steps" in development and implementation.	What recommendations can be given to the medical school leadership, based on the data gathered, to improve the curriculum change process?

MATERIALS AND METHODS

The study design utilized qualitative approaches for evaluation of the change process. The Concerns-based Adoption Model was selected because it directly focuses on the implementation of curriculum innovations rather than broad organizational change. In addition, CBAM targets the individual or teacher as the critical unit of change. Importantly, the model provided a validated diagnostic instrument. Data were collected through the Stages of Concern (SoC) questionnaire and personal interviews with identified faculty and staff. The evaluation team included four medical education specialists, one physician, and one basic science faculty member. The study proposal was submitted to the university Institutional Review Board (IRB) and received exemption in March 2005.

The stages of concern among medical education faculty and staff were measured through the self-administered Stages of Concern (SoC) questionnaire designed by Hall and his colleagues.[7] The questionnaire, a 35-item rating form, has strong reliability estimates (α= .65 to .86) and internal consistency (α= .64 to .83). The Stages of Concern (SoC) questionnaire was constructed to apply to all educational innovations. Items stay the same with the only change being the insertion of the name of the specific innovation as shown in Figure 2. Each of the seven stages has five items randomly distributed within the questionnaire. "The respondent marks each item on a 0 to 7 Likert scale according to how true it is that the item describes a concern felt by the individual at the present time."[8] The "0" of the scale is recommended for marking items that are completely irrelevant.

session and in the early phases of curriculum revision when most of the five curriculum principles had not yet been incorporated into a new curriculum. To assure anonymity, respondents were requested to mail or FAX their completed questionnaires to an Administrative Assistant assigned to the curriculum committee. The Administrative Assistant, working on a password-protected computer, recorded data on a spreadsheet; no identifying data were retained.

Analysis was conducted using a process described by Hall and associates.[9] Each stage has five items that can provide a raw score between 0 – 35. Raw scores for each individual's questionnaire were tabulated; the group mean was computed for each stage of concern and adjusted to percentiles using a conversion table provided in the scoring manual. The percentile for each stage equals the intensity of concern at that stage. The percentiles are presented as an institutional profile and compared to a typical non-user profile for analysis as recommended by Hall and associates for the introductory phase of a curriculum innovation. A graphic representation or profile of the percentile scores was used for interpretation of the SoC data.

In order to gain a better understanding of the faculty concerns, an interview guide was developed to focus attention on each of the stages of concern while allowing for spontaneity during the interview itself, which provided an opportunity for emergent questions flowing from answers within the interview. In addition to taping all interviews, each interviewer also recorded field notes.

Figure 2. Sample items from SoC Questionnaire for each stage

Stage	0 Irrelevant 1 Not true of me now 2 3 Somewhat true of me now 4 5 6 Very true of me now 7	
Awareness	0. I am completely occupied with other things.	0 1 2 3 4 5 6 7
Informational	1. I have a very limited knowledge about competency-based curriculum.	0 1 2 3 4 5 6 7
Personal	2. I would like to know how my teaching or administration will change in CBE.	0 1 2 3 4 5 6 7
Management	3. I am concerned about not having enough time to organize myself for this approach.	0 1 2 3 4 5 6 7
Consequence	4. I am concerned about students' attitude toward a competency-based curriculum.	0 1 2 3 4 5 6 7
Collaboration	5. I would like to help other faculty in their use of competency-based curriculum.	0 1 2 3 4 5 6 7
Refocusing	6. I now know of some curriculum approaches that might work better.	0 1 2 3 4 5 6 7

Paper copies of the SoC questionnaire were distributed to 60 members of the full-time faculty and relevant staff in May of 2005 along with a brief explanation of the five proposed curriculum elements. The sample was chosen to include all basic science faculty, all clinical chairs, and representative clinical faculty at University Health Care with active roles in teaching at all levels. All staff members with active roles in curriculum implementation were also identified. A follow-up request with the questionnaire attached was sent one month later. A third opportunity to complete the questionnaire was offered to individuals prior to each interview. All three opportunities were given during the summer when classes were not in

Researchers purposely selected 29 faculty and staff because of their potential involvement in the development and implementation of the curriculum. Interviews were completed with 25 faculty and staff; four were excluded due to scheduling conflicts, refusal of interview, or tape malfunction. Interviewees included eight clinicians, thirteen basic scientists, and four administrative/staff personnel. Eight of the interviewees serve on the 22-member Curriculum Review Committee, reflecting a similar distribution of committee members to college faculty and staff. Interview tapes and hand-written notes were transcribed, and names were removed. The transcripts were distributed to the six members of the evaluation team. Interviewers had little previous connection to their interview panel to avoid conflict of interest based on department or work assignments. Each team member coded all interview transcripts individually. Using contextual analysis for coding of responses, several themes emerged for each question that had structural corroboration and referential adequacy.[10] Subsequently, team analysis focused on identifying specific concerns of interviewees, level of understanding, and use of the five key curricular elements. These concerns were then linked to the first four stages of the Concerns-Based Adoption Model – Awareness, Informational, Personal, and Management – because the model focuses on early concerns at the beginning of curriculum development. The use of qualitative reflection by the research team during coding revealed additional linkages and relationships.

Limitations of the Study

The curriculum innovation is still at the conceptual level and does not yet contain critical details that require future faculty work. Although the responses to the questionnaire came from only 53% of the faculty/staff, these are primarily full-time employees and are the most active in the process. Because researchers reported the average of the questionnaire for the institution in the profile, individual differences may have been masked.[11]

RESULTS AND ANALYSES

The demographics of the respondents to the survey showed an overall representation of the college personnel. Several respondents did not respond to all demographic items. When examining their role in the college, 26 (81%) respondents were identified as faculty, and six (19%) respondents were identified as staff. Of those identified as faculty, 12 respondents were basic science faculty and 14 were clinical faculty. There was no category for faculty who are medical education specialists, so these individuals were classified as staff if they did not have direct teaching responsibilities. We did not have a demographic choice for Administrators. Most of the respondents focused their educational work at the MS1-2 level or at the graduate level, although many respondents indicated more than one level of primary teaching responsibility (Table 2). The distribution of years in medical education shows a relatively mature college (Table 2).

All analyses were conducted using the instructions from the SoC Questionnaire Manual. The item responses of the SoC questionnaire were tabulated by stage to determine an individual raw score for each stage, and a group mean for each stage was computed. The group mean was assigned a percentile using the SoC conversion chart from the manual and plotted to create a profile (Table 3).

Table 2. Survey Demographics for Focus of Work and Number of Years in Medical Education

Focus of Work	MS1	MS2	MS3	MS4	Internship	Residency
# Respondents	22	19	9	5	6	11
# Years in Med Ed	1 – 5	6 – 10	11 – 15	16 – 20	21 – 25	25+
# Respondents	3	6	5	6	7	5

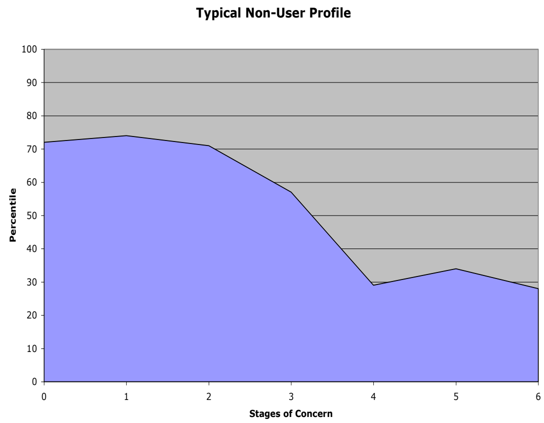
Table 3. Percentiles of Group Means for Each Stage

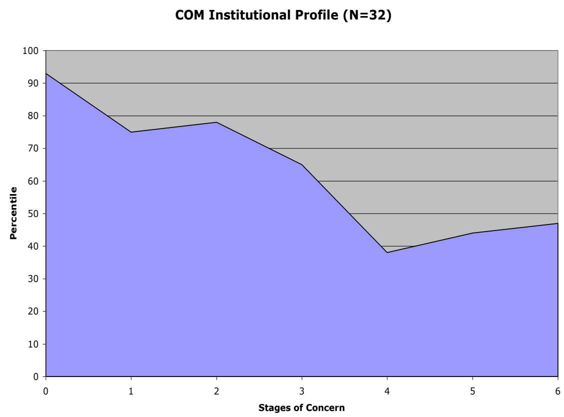
Stage 0 Awareness	Stage 1 Informational	Stage 2 Personal	Stage 3 Management	Stage 4 Consequence	Stage 5 Collaboration	Stage 6 Refocusing
93	75	78	65	38	44	47

Six different interviewers used the protocol, working from varying depths of experience with the curriculum change process, the stages of concern framework, and interviewing skills. One individual refused the interview; two individuals could not be scheduled, and one tape malfunctioned without sufficient backup notes for use. While personal disclosure may be somewhat uncomfortable in the beginning, it is an important and necessary step toward collaboration within the faculty and staff. The researchers maintained confidentiality with regards to the questionnaire results and the corresponding interviews.

The profile of our institution was analyzed holistically in comparison with a typical non-user profile as shown in Figure 3, since the college is at the beginning stage of curriculum development. When looking at the profile of peaks and valleys, the intensity of each stage is not as important as the relationship between stages. The next step was to analyze the individual stages.

Figure 3. Comparison of Typical Non-User Profile and College of Osteopathic Medicine Institutional Profile (N=32)





If you place the COM Institutional Profile on top of the Typical Non-User Profile, you can see that these two profiles are essentially congruent. The institutional profile indicates that our faculty and staff are intensely aware (Stage 0) that the development of a competency-based curriculum is underway, but do not have adequate information (Stage 1) about the process and the specific elements. As such, the intensity of personal (Stage 2) and management (Stage 3) concerns has remained high even after a year of deliberation by the Curriculum Review Committee. Consequence (Stage 4) concerns are much lower and in keeping with the typical profile of non-users; however, collaboration (Stage 5) and refocusing (Stage 6) concerns begin to climb again.

Hall and Hord[12] suggest that any tailing up of the Stage 6, Refocusing Concerns, on a nonuser profile should be taken as a potential warning that there may be resistance to the innovation. An emphasis on refocusing at this early stage can indicate an attitude of disrespect for the innovation. Coupled with the lack of information apparent in the profile, such a tailing up might also indicate that many faculty members have felt left out of the process. Resistance is also indicated by a high Stage 2 Self Concerns percentile. These may be normal characteristics of persons who are uncertain about what will be expected and may indicate self doubts about one's ability to succeed in the new program.[13] In our institution, the high tailing up at Stage 6 Refocusing may reflect the fact that the specifics of the new program have not been developed, only the principles. Hall and Hord interpreted this profile as a positive, rather than distrustful non-user profile.[14] In this evaluation study, the institutional profile was used as a diagnostic tool and established a baseline for further analysis after the curriculum has been fully developed and implementation is under way.

INTERVIEW RESULTS AND ANALYSIS

The follow-up interviews were conducted to gain a sense of what was producing this institutional concerns profile. The interview protocol was organized to reflect the first four stages of concern (awareness, information, personal, and management) since the curriculum had not been fully developed at this time. Within the stages, sometimes overlapping themes emerged that will be used to plan change strategies (Table 4). Respondent quotations from the interviews are often used to support and illustrate the findings.

Table 4. Themes within Stages

Stage	Themes
Stage 0 & 1 Awareness & Informational	Information Flow Driving Force Us versus Them Lack of Attention Curriculum Elements
Stage 2 Personal	Don't Know the Impact Minimal Impact Strong Impact • Curriculum Elements • Resistance • Faculty Development • Personal Time
Stage 3 Management	Resources • Financial • Personnel Faculty & Student Time Assessment & Evaluation Faculty & Staff Development Leadership Issues

Awareness and Information (Stage 0 & 1)

Interviewees were aware that the College of Osteopathic Medicine had a committee charged with curriculum revisions. They also had some awareness of the new national competencies, but did not know about the set of curriculum elements that had been established by the committee to guide the next phase of curriculum redesign.

Information Flow. A majority of interviewees did not have enough information about the process after one year of deliberation by the Curriculum Review Committee. Most of the faculty and staff who understood the process were on the Curriculum Review Committee. Interviewees felt there was no clear structure or place to get information. Most information on the process came through participation in the Curriculum Review Committee or the Academic Affairs [Curriculum] Committee; the process had not filtered into department meetings with the exception of Pharmacology. Faculty and staff members repeatedly described their basic medium for information flow: informal hallway conversations that may have lacked accuracy.

> "People have corridor conversations. We are not trying to second-guess the process."

> "I hear hall talk a little bit, but I would like someone to discuss the overview and where we are going."

Because of this, clinical departments and individual faculty members who are not on campus have more challenges getting the basic information. One clinician acknowledged the information flow from her participation in a systems course and interactions with the systems manager.

> "The curriculum process has been very vague. The process was not spelled out. We need to know the next steps in the process, evaluate it, and discuss it. My faculty members are completely in the dark about the process."

> "Not really enough information, but my systems manager has been very innovative and lets me know what is going on with the process; we are passing ideas back and forth."

Driving Force. Throughout the interviews, it was apparent that the faculty and staff wanted answers to basic questions such as, who makes decisions in curriculum change? What is the driving force for curriculum change? Three basic scientists voiced the concern of others:

> "Usually when something like this is done, it is in response to a concern, so what are the problems? What needs to be fixed?"

> "Are we going to just cut hours? What are the criteria for the changes?"

> "I'm not too happy about that whole thing [new calendar] in terms of what drives the curriculum."

Us versus Them. Another basic scientist expressed a positive perspective describing the Curriculum Review Committee as a group of 'honest, caring, smart people addressing themselves to a problem, which is the same context in which I am working. I trust them to do it.' However, several respondents used the phrase "Us versus Them" with different meanings -- scientist versus clinician, committee member versus non-members, external faculty versus professional educators.

> "The way the school is set up, the curriculum process is driven much more by the basic science faculty and outside consultants that may or may not have active clinical practices or may have some clinical role, but are still more [like] professional educators."

> "There is a false tension between basic science and primary care clinicians. We need the best clinical resource: a specialist. We need to phase down Pathology and Radiology since no one looks at slides and images -- they go to a specialist."

> "Things have been thrown out of the GI system and shouldn't have been. 'They' often don't know the framework of the discipline."

Lack of Attention. Three respondents acknowledged their own lack of attention to the curriculum change process, fearing that it might be their own fault that they didn't have the basic information on the process although they try to stay abreast of the college news through e-mail. While not casting blame, the belief exists that faculty and staff should be informed even if they are not able to attend meetings. A few faculty members who are primarily research-based and a few staff members did not think they needed information on the process. However, staff members recognized the systemic nature of the work and were waiting for decisions in order to do their own planning. One basic scientist deliberately stayed away from the process:

> "My whole goal… was to make sure that I wasn't an active participant. My goal was to avoid getting sucked into another series of ongoing meetings with things that eat up time. I can just let other people do it."

Curricular Elements. The majority of respondents did not have enough information about the five curricular elements that the Curriculum Review Committee had established as guiding principles. They recognized that it was difficult to gauge when they had enough knowledge about innovative practices that had not been fully developed. Others expressed personal concerns in not being able to teach or assess competencies.

> "I don't think we have spent enough time yet in really moving to a competency-based curriculum. There is a lot of work to be done. I think they're going to need concrete illustrations of how the competencies will apply to what they're doing. Pharmacology has a nice example of this."

> "I don't know all the details of what competencies you want, but I'm in full support of them. I don't understand how to teach toward competencies. I don't know how to assess competencies, but I think it's the right way to go."

> "I am aware of the national core competencies, but I don't know how they will be integrated into the curriculum."

However, some respondents expressed knowledge of the curricular elements. Several of the curriculum elements come directly from previous innovations and interests of individual faculty members. One respondent gave the interview team direction in terms of finding articles on curriculum change by other medical schools.

> "Two of my issues—block exams and length of year—have been incorporated into the guiding principles. Everyone can incorporate some of these [elements], which I already do in the courses I direct. I do many things for the second year students that 'they' have no idea I do."

> "The Dean circulates Academic Medicine, and I pull out articles that I feel are germane. I have interests such as grading, lecture time, the value of lecture."

Several respondents from the pharmacology and the neuromuscular medicine department already teach and assess from a competency approach.

> "Understanding competency-based education is different from getting into the classroom. I have integrated several competencies: cost containment, professionalism, medical knowledge, and patient care. I should be more didactic about the competencies with the students. We also use several of the other elements."

In order to move past awareness and informational concerns, one respondent advised that the college leadership for curriculum revision should remember new or returning faculty and staff who may be out of the loop.

> "I came back after the process had already begun, and so I walked into the middle of it. I don't know if there is a website or some resource where I could go to get the history, why we are doing the revision, who is involved, and the goals and objectives of the project. I would like to have information in some format—monthly newsletter with an update or website. I hate to take up someone's time to catch me up."

Personal Concerns (Stage 2)

Personal concerns are an important hurdle to overcome by college leadership and change agents and specific strategies and tactics are often less apparent than for other stages. The interviews confirmed the reason for high concerns regarding the personal impact of curriculum innovation. These concerns lacked focus because of the lack of knowledge and understanding for a large group of respondents.

<u>Impact: Don't Know</u>. Given the lack of information reported by a majority of interviewees, it is not surprising that some faculty members were unable to predict how the curriculum revision process might affect them.

> "I don't know how it will affect me. I don't know how to teach toward competencies, so I don't know how that's going to affect the way I teach."

> "I don't know yet because I don't have any certain information yet in regard to what changes might be recommended to the Dean."

> "I don't have a clue. Obviously I always hope that it's going to be for the better."

<u>Impact: Minimal.</u> Staff members generally felt that the curriculum revision process would have minimal personal impact on them, while the majority of faculty members articulated a range of personal concerns about the process. Among interviewees who felt that curriculum revision would have minimal effect on them personally, the most common reasons were that potential changes did not apply to them or that they were already implementing elements of the proposed curriculum.

> "I don't see myself seeing any fundamental change, because I'm already on the same page."

> "We're already looking at our information from an outcomes perspective and from a competency perspective from matriculation through residency."

<u>Impact: Strong</u>. The remaining respondents shared a mixture of strong personal and general concerns related to the proposed curricular elements and their specific, defined effects. Two major areas of concern centered on reform efforts to integrate the national core competencies and to hold students accountable for course learning outcomes.

> "The integration of the national core competencies, I think, is going to be real tough. . . . I think the two toughies are . . . the integration of the core competencies and the instruction based on learning outcomes."

> "I am fearful of the university interpreting the competencies just from their vantage point and not from the clinical standpoint."

Although approximately one-third of respondents understood that one of the curricular elements was intended to make students more responsible for their own learning, faculty member views on the curricular elements of student choice and student accountability for learning varied widely. Comments in this area reflected the range of faculty experience with lecture alternatives, as well as their opinions about the ability of students to choose learning experiences and to be accountable for their own learning.

> "I'm all in favor of alternative ways to learn. . . . I can't write a prescription for the whole class that's going to work equally well for all [students]. And

they've got to be adults in the matter, to find what works for them."

"I am not an advocate of adult learning theory. I feel these students do not have the skill set to decide they can learn better on their own and decide not to attend class."

"We've essentially eliminated the lectures. There are people that don't even come to lectures and do quite well."

"I think it is important to have options open to the students to organize time and use it most effectively and efficiently. I think we have to accommodate different learning styles."

Both clinical and basic science faculty noted resistance to change and faculty feelings of territorialism as potential areas of difficulty in revising the curriculum. Respondents cited both a general resistance to change and faculty members' concerns about adequate curricular time for their subject matter as obstacles to be overcome in achieving proposed curricular changes.

"I think culturally you still have a lot of people on this campus who are in the mindset of that 50-minute block of time being how students get the curriculum."

"It's all about 'this is the way I've always done it and this is my time and I'm not giving it up' and we have a lot of work to do there."

"I think the big challenge would be . . . that some of us would have to give up asking our pet [test] questions that reflect what <u>we</u> know and try to develop teaching and learning questions that reflect what students <u>need</u> to know."

Interviewees identified a need for faculty development targeted to assist faculty in applying the planned curricular elements to their actual teaching situations. Activities that model the desired teaching methods and provide opportunities for faculty to develop teaching skills and evaluation methods that have immediate application were most requested.

"I need someone to . . . give me an exact example. I want someone to send me to a resource or show me how I can go from start to finish and fit into that competency picture. I also don't know how to assess competence. And I would like specific examples as opposed to general concepts . . . . I would not want to have to attend weeks upon weeks of lectures to learn that; I'd want to keep it to a minimum."

Another recurring theme was a concern with personal time, both as a limited commodity and as a necessary ingredient for successful curriculum change. Both clinicians and basic scientists expressed concerns about multiple demands placed upon faculty by the proposed curricular changes. One basic scientist summarized it this way:

"When you start talking about small groups, you start talking about clinical relevance, you're talking about a lot more preparation and more faculty members. You need enough faculty members to sit in small groups and have a decent ratio to facilitate attentive students. [Clinical faculty] face similar challenges in terms of having more intimate relationships with students."

One fourth of respondents expressed concerns about time, ranging from whether curriculum changes could be accomplished in the time allotted to how time would be allocated among courses and whether expectations of faculty were reasonable and achievable. Faculty also pointed out that different faculty subgroups – clinicians, basic scientists, full-time, part-time and volunteer faculty – face various challenges in undertaking curriculum revision.

"I feel the issue that has not been addressed is: How much time is allotted to topic areas? This is driving everything else and needs to be addressed."

"It is always hard to get clinicians to change how they approach education. It is a long arduous task to revise lectures, based on the new WebCT™ paradigm. It is going to take a long time to get these fine-tuned."

"To expect [volunteer faculty] to spend that amount of time away from practice is unrealistic."

"The university needs to understand that with faculty who are trying to maintain clinical practices, there is only so much that we can do, and it doesn't mean that we can or should be ignored in the process in making fundamental changes."

"It will be easy for the university to push training by the clinicians out of the 2nd year into the 3rd year and not give us the resources to bring it to fruition. Yet there won't be enough time in the curriculum to teach these students what they need to know—more educational requirements, less time."

Management (Stage 3)

Interviewers asked each person to think of the management implications of changing from current practice to incorporation of five new curriculum principles. Interviewees were very forthcoming in discussions about management issues, and analysis revealed a set of concerns. These findings showed clear thematic issues of resources, faculty and student time, assessment and evaluation, and faculty development, as well as leadership and accountability.

Resources. One-third of all respondents, among both clinical and basic science faculty, were concerned about the financial needs of the college as related to supporting a new curriculum and delivery system, particularly what was defined as an unfunded mandate. They were concerned that the lack of a strong overall national economy would affect the college and its clinical offices directly. One respondent cited a need for grant writing support at the university level and a reliable business office. Another respondent identified the added costs for small group facilitators when moving away from lecture teaching as the new curriculum recommends. One respondent saw a different side of the issue when a perception of limited resources becomes a belief that undermines change, even when not based on accurate information.

Personnel needs were the second and a closely-related resource concern. Basic science instructors asked for technical staff to execute the new curriculum scheduling software and to support faculty who use the WebCT™ course management system. Since WebCT™ is a tool for accomplishing several of the design elements, the faculty want someone to support them, not just train faculty members to do the work, believing that organizing the course website and uploading documents is not a productive use of faculty time. One respondent suggested that faculty should deal with content, sequencing and instruction, not the mini-details of technology use in curriculum delivery.

The Curriculum Office is an important operational office with responsibility for schedules, contracts and student assessment. Their operations received both kudos and caveats that related to resources.

> "The curriculum office is basically overwhelmed. We haven't replaced people who left. It seems that they are having a hard time keeping up with things the way we do it now, much less making a huge transition. They need more personnel and better software."

> "It all comes down to not only money, but somebody who's competent to run the new technology."

Faculty and Student Time. Basic science faculty members were intensely concerned about time, both for faculty and students, as a management element. Respondents noted several key concerns: curriculum change and assignments have been incremental; there is no systematic structure for setting faculty load; and faculty time is often allocated to several degree programs. System course managers are particularly confused by the load obligations. In addition to these larger time issues, one respondent explained that instructors are simply trying to meet daily demands.

One of the first and most visible steps in the curriculum revision process was to shorten the academic year. Because of management constraints, the new calendar was posted in advance of the development of new curricular options including electives and remediation. Although mentioned in the management section of the interview, some responses reflect informational and personal concerns as well.

> "The schedule change is about all that people know about the process. The pedagogy, competencies, [learning] outcomes – they don't know about that."

> "Will faculty involved in remediation or electives need to cut back on their research/scholarship or service to the university/profession?"

> "Will the competency approach and the new calendar bring out space issues?"

> "If we reduce classroom time, everybody needs to hold the line. In the past, there have been issues of extra classes, review sessions, and such that became a ghost curriculum. There is also competition for time between disciplines."

Respondents pointed out that the shorter academic year meets the perceived student demand for dual degree options. Faculty respondents were concerned that the shorter academic year would negatively impact students, but also acknowledged that students will have new responsibilities to come to class prepared.

> "The student time would be my primary concern because I don't know how to estimate well. We can't have unrealistic expectations in the first two years. The rationale to determine content makes sense; however, it must be manageable for the students."

> "If you are doing a case, students will need to read the material before entering class and come prepared."

> "When the student is responsible for the learning outcomes, the information is there even if there is an instructor illness or emergency or snow day. It has given them better options for the use of their time."

> "I get concerned about teaching the appropriate information at the appropriate time. We have to teach for their stage of education."

Assessment and Evaluation. Competency-based education calls for new approaches to student assessment and program evaluation. Concerns arose when the Curriculum Review Committee recommended having block exams, and faculty struggled with the impact on students. Some basic science and clinical professors noted the first step in changing learner assessment is based on a continuum of competency levels from medical school through residency. There were again concerns about resources.

> "Do we have the resources to increase the amount of curriculum that's available on the Internet, handle all

> of our remediation, change the schedule for the summer, and change the tests and measures that we have?"

In order to manage the process, respondents made several suggestions related to management of the assessment and evaluation process.

> "We need to get documentation of grades, both preclinical and clinical, in a timely fashion for the Dean's letter to be completed."

> "The college needs a new computer software system that also reports curricular units to accrediting bodies."

> "Improved evaluation and more timely feedback can lead to evidence-based education."

Faculty Development. The challenge of curriculum change was seen as an opportunity for systemic faculty/staff development. One respondent explained that adopting the new perspective of core competencies requires that all course managers 'see the whole picture.' Faculty respondents also requested more specific training in exam development and test item construction. Both basic science and clinical professors acknowledged the problem of 'refuseniks' and recommended giving support to those who are really interested in making curricular change. In addition to training workshops, one respondent recommended that faculty move beyond the traditional department meetings in order to both learn and develop the competency curriculum.

> "It would be nice to get the first year course directors together periodically and talk about things."

> "We need cross department meetings of all 1st year teachers and all 2nd year course managers."

External clinical lecturers present a challenge for faculty development that is geared toward the implementation of competency-based curriculum.

> "House people are making a list of what students should know, but lecturers are coming in and sometimes their lectures match and sometimes they don't. At the same time you are implementing WebCT™, you've got instructors coming in from the outside that don't know the system."

Leadership and Accountability. The interview process focused on an understanding of the curriculum revision process and the proposed curricular elements. The question of management brought to light concerns about leadership. Individuals complained that they received mixed messages from the leadership and that there was poor coordination of the process. Respondents recommended that the leadership should recruit, support, and capture people interest in the curriculum revision process.

Respondents reflected on past problems with the lack of authority by the college curriculum committee because it does not function as an oversight committee. They recommended a new unit to manage the seamless integration of all five curricular elements, managing both the policies and the logistics. The person in charge of this unit would have leadership and management of the overall look of the curriculum and the delivery as well. However, there were differences in identifying the criteria for this unit leader.

> "The person needs to know where everything is in the curriculum and how it builds to the competencies – all in one place. This person needs to have credibility with the people in the trenches trying to implement this curriculum."

> "I see a new curriculum office…headed up by a PhD or some doctorate level person. Somebody clinical could come who doesn't want to practice any more but may want to get into education. Or maybe it could be a basic scientist who could organize lots of stuff in the curriculum. It can't be someone who is too theoretical."

DISCUSSION AND RECOMMENDATIONS

The SoC questionnaire and the personal interviews were a two-pronged approach to formative evaluation. The questionnaire showed a profile very similar to the typical non-user; therefore, our institutional profile gave us an appropriate baseline for future evaluation. The interviews cast the spotlight on themes that more fully explain the SoC Questionnaire data. The analysis is also significant in helping the leadership determine who is ready for change in a mature institution. From a concerns approach, curriculum leadership must first design strategies to address lower-level concerns such as informational and personal concerns, in order to allow faculty and staff to focus on higher-level concerns of management, consequence, and collaboration.

Interview findings of awareness and information needs were highly consistent with the findings from the SoC profile, which told us that almost all of the responding faculty had an awareness of the reform effort, but few felt that they actually knew what was going on with the curriculum reform. The faculty and staff continue to ask why we are changing if our graduates have done well. There has been no apparent conceptual bridge formulated connecting new professional roles in the national medical system to competency-based medial education. Therefore, the leadership needs to stress the driving forces for curriculum change and establish urgency for the process. The larger picture of competency-based education, including the set of curriculum design elements, was lost when the first visible sign of change was the shorter school calendar without concurrent offerings of summer electives, research opportunities, and remediation. Even those on the Curriculum Review Committee did not recognize the five curriculum design elements as a road map that was generated from hours of discussion in meetings. It is clear that there was no trickle-down effect. Information about both process and decisions needs to be shared in multiple formats and venues and repeated throughout the multi-year process.

The SoC questionnaire showed intense personal concerns. Our analysis of interviews showed a dominant theme of “us versus them”. While this mentality may seem unavoidable because of distinct points of view, everyone should be focused on the successful medical student and resident. So, we ask, how do we break down barriers between groups that fluctuate by role -- between clinical and basic science faculty, between curriculum committee members and other faculty/staff, between administration and faculty, between faculty and medical education specialists? We have changed the COM Faculty Assembly to an evening time slot to encourage participation by clinical faculty. Each clinical department has worked with the MS2 System Course Manager to analyze the course disease index using the criteria of commonality and criticality. The first year Biochemistry course was redesigned to use team-based learning and case-based curriculum with clinical faculty contributing cases and test items.

During the interviews we learned that faculty were unsure how the curriculum design elements would impact them personally, in part because at that time the design elements had not resulted in specific curriculum changes. People who had changed their course to focus on student learning outcomes seemed to think that they were already finished with curriculum change, although their instruction and course assessments may not have changed to learner-centered approaches. A deeper analysis of the interview data indicates faculty insecurity about being successful given the high level of student choice recommended by the curriculum design elements. Faculty members will have less control over the student learning experiences, but expect that the faculty will still be held accountable for student failure. They appear to have little faith that students will actually be held accountable for their own failures. The interviewees gave very specific ideas for needed faculty/staff development related to the curriculum design elements, moving beyond workshops toward demonstration and mentoring. Personal concerns were frequently intertwined and expressed in terms of institutional culture as well as faculty prerogatives. When interviews were analyzed as a whole rather than by individual question, it became clear that initial policy making was critical to the process. The work of curriculum revision requires extension beyond the major committee to a variety of working committees that brings larger numbers of faculty and staff into the process.

The results of the SoC showed a higher concern for personal issues than for management issues; however, the interviewees spent far more time discussing institutional and curriculum management concerns than truly personal concerns. This may be a result of the way that the SoC items are written, a function of the sample, or a function of the faculty member unwillingness to talk about themselves versus their desire to discuss institutional problems that may have been on their minds for some time. Collaboration, as shown by the SoC Questionnaire, was not a strong concern at this stage of the process.

Respondents viewed both additional faculty and support staff as necessary resources to successfully carry out the planned curriculum changes; they also believed that high quality technical support services would be essential in implementing new technologies and sustaining them once adopted. The respondents were clear that the need for resources, both financial and otherwise (personnel, leadership), requires setting priorities at both university and college levels. Resistance to curriculum change often comes from a lack of confidence in the institution being able to meet these resource demands. People recognize that time is also a critical resource that must be managed effectively. Current and future demands on faculty load have not been factored into the plan for curriculum redesign. Demonstration of competency requires extensive and expensive learner assessment and holistic program evaluation, which creates stress within the curriculum system and requires new technology for reporting and feedback. Under the theme of management, the issue of curriculum change has also challenged the current curriculum governance structure. With current changes in institution and college administration, the curriculum process and governance has become even more critical.

Our institution is not alone in struggling with curriculum reform. Other researchers have studied the process of curriculum change. Bernier and coworkers[5] describe experiences of two medical schools over a period of three to four years. Each school changed from a traditional, lecture-based, decentralized, department-based curriculum to a centrally governed, integrated, student-centered curriculum evoking greater collegial relationships among students, faculty and administration. Three lessons were cited as being key to the success of curricular change. First, a broad commitment to change was led by each dean and central administration and reinforced by national education policy and philosophy. Second, widespread inclusion of faculty and students in the process enhanced the willingness of the faculty to cede authority to a new curriculum committee before the design of the curriculum was complete. In addition, the establishment of a centralized governance system early in the process resulted in a strong, empowered curriculum committee with implementation support from the Office of Medical Education.

Our analysis was confirmed when we attended the 2005 annual conference of the American Association of

Colleges of Osteopathic Medicine (AACOM). Dane and Kreuger[16] reported on the curricular change process. Peter Dane (OU/COM) described the tripartite leadership that comes from the Dean, a centralized curriculum authority, and a team of champions who work beyond the basics. He stressed the need for clarity of purpose, describing the philosophy and rationale for their movement toward a new curriculum that was distributed widely in the form of a white paper. He reminded curriculum change agents to honor the present and the past while nurturing seasoned educators. He advised special care of students during the transition year. Paul Kreuger (UMDNJ/SOM) identified four keys to curricular change success that relate well to the stages of concern: a clear mission, strong leadership, adequate resources, faculty and faculty development.

Although ours is a single site, formative evaluation study, many medical educators may see similarities to their own experiences with curriculum revision. Furthermore, the concerns approach and the SoC instrument as well as the findings offer practical information to leaders whose institutions and academic programs are soon to begin curricular change. From our study and the efforts of others, we have learned the importance of setting up a curriculum revision process based on both individual and institutional concerns that have potential for influencing the development and implementation of a new curriculum. The process needs to be systematic with timelines and responsibilities. Everyone needs to be kept informed of the process and the outcomes of deliberations with information coming in multiple formats including a website, hallway bulletin boards, and written materials, as well as key descriptive and research articles. Care should be taken to release information in a logical fashion so that faculty and staff not directly involved in the development stages can understand the reasons for current plans and have an opportunity to respond if they are unhappy with plans at any given stage. Additional resources will be needed to be successful, including: qualified assistance with effective use of WebCT™, grant-writing support, faculty coaching, and implementation of effective evaluation approaches. Understanding the stages of concern can result in more targeted strategies, more relevant workshops, and directed planning to implement the new curriculum plan thereby creating successful, institutionalized change.

REFERENCES

1. Hall, G.E., and Hord, S.M. Implementing change: Patterns, principles, and potholes. *Boston. Allyn and Bacon*. 2001.
2. Hall, G.E., and Loucks, S.F. Teacher concerns as a basis for facilitating staff development. *Teachers College Record.* 1978; 80(l): 36-53.
3. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation. A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979.
4. Lewis, D., and Watson, J.E. Nursing faculty concerns regarding the adoption of computer technology. *Computers in Nursing.* 1997; 15(2):71-6.
5. Gwele, N.S. Concerns of nurse educators regarding the implementation of a major curriculum reform. *Journal of Advanced Nursing.* 1996; 24:607-14.
6. Arwer, J.M., Harris, K., and Dusold, J.M. Application of the concerns-based adoption model to the installation of telemedicine in a rural Missouri nursing home. *Journal for Nurses in Staff Development.* 2004; 20:42-9.
7. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation: A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979.
8. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation: A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979, p. 23.
9. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation: A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979, p.24,26.
10. Eisner E. The enlightened eye: Qualitative inquiry and enhancement of educational practice. *Upper Saddle River, NJ: Merrill/Prentice Hall*. 1998, p. 110.
11. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation: A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979, p. 78.
12. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation: A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979, p.39-40.
13. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation: A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979, p. 72.
14. Hall, G.E., George, A.A., and Rutherford, W.L. Measuring stages of concern about the innovation: A manual for use of the SoC questionnaire. *University of Texas at Austin*, 1979. p. 73.
15. Bernier, G.M., Adler, A., Kanter, S., and Meyer, W.J. On changing curricula: Lessons learned at two dissimilar medical schools. *Academic Medicine*. 2000; 75:595-601.
16. Dane, P. and Kreuger, P. Leading Curriculum Change: Looking Forward, Looking Back. A presentation at the annual meeting of the American Association of Colleges of Osteopathic Medicine. Bethesda MD, June 2005.

An Analysis of First-Year Medical Student Comments in a Peer Evaluation of Professionalism

Scott Cottrell, Ed.D., Julie Balch Samora, MPH, MPA, and James Shumway, Ph.D.

Department of Community Medicine
West Virginia University
PO Box 9007
Morgantown, WV 26506

Phone: (+)1-304-293-0410 Fax: (+)1-304-293-1814 Email: scottrell@hsc.wvu.edu

ABSTRACT

Medical educators are identifying how learning opportunities help further student professionalism, which requires explicit methods of assessment. This paper reports how student comments about peer professionalism may contribute to our understanding and assessment of medical student professionalism. An inductive qualitative analysis identified themes across student comments. Using comment characteristic (positive versus negative) as an independent variable and the mean score of the nine scale questions of the professionalism assessment as the dependent variable, an independent-sample test was calculated. Participants included one-hundred and eleven first-year medical students in a required problem-based learning course. A total of 12 comment themes were distinguished. Out of the total male (292) and female (212) comments, 82% and 88% were positive comments, respectively. There was a significant difference in the mean professionalism score between the students who did and did not receive negative comments, $t = 2.93$, df = 109, $p = .002$. Overall, this study provides an initial framework from which future studies may draw. This investigation will assist educators to interpret the rich and complicated data of student comments, furthering how we figure qualitative information into the overall evaluation of student professionalism.

INTRODUCTION

Medical educators are charged with furthering medical student professionalism skills. Competency-based curriculum literature has provided direction for teaching and assessing student professionalism. For example, Herbert Swick [1] developed a comprehensive definition of professionalism. Frameworks, such as CanMEDS [2], the University of Dundee Three-Circle Model [3], the ACGME Competencies [4], and the AAMC Medical Student Objectives Project [5], have also described the knowledge, skills and attitudes that characterize professionalism. Strides with competency-based curricula have taken the ill-defined construct of professionalism and displayed it as a collection of measurable variables, which has informed educators on assessment development.

Several medical schools are initiating learning opportunities to promote medical student professionalism. Swick, *et al*.[6] conducted a survey that investigated whether medical schools are developing and assessing formal instruction related to professionalism. The results suggest that most medical schools are offering some form of instruction (didactic or experiential) related to professionalism. Tulane University, for example, has developed the Program for Professional Values and Ethics in Medical Education.[7] This program encourages students, residents, and faculty to examine the characteristics of professionalism and discuss how modeling professional attributes can impact health care practices.

Educators also need to identify how learning opportunities help further student professionalism, which requires explicit methods of assessment. Assessment methods to gauge professionalism characteristics are less developed than other constructs, such as medical knowledge.[8] The purpose of this paper is to investigate how student comments on peer evaluations may contribute to our understanding and assessment of medical student professionalism. Peer evaluations may promise to be a valuable source of information. Arnold [9] explained in her review of the literature that students have frequent, close contact with their peers, which lends opportunities to observe behaviors outside the presence of faculty and staff. Student reflections, then, may provide unique insights into peer professionalism.

As educators collect assessment data on student professionalism, questions surface about how to interpret student comments about professionalism of peers. Student understanding of professionalism can be informed by competency frameworks and scale questions on evaluations, but written comments may indicate how students attend to and focus on particular characteristics of professionalism. Exploring how student comments are delineated into themes will further the theoretical development of the complex educational phenomenon of student professionalism. We hypothesized that an examination of student comments will reveal several distinct themes of professionalism behaviors. We also hypothesized that student comments will be aligned with scale questions. That is, students who receive negative comments will have lower scores on scale questions than students who only receive positive comments. This investigation aims to help educators interpret the rich and complicated data of student comments, furthering how we figure qualitative information into the overall evaluation of student professionalism.

MATERIALS AND METHODS

The Professionalism Peer Assessment

In the present study, first-year medical student comments about professionalism of peers were analyzed. The comments were collected from a peer professionalism assessment that was implemented in a problem-based learning course. An important consideration of an assessment instrument is the definition and specificity of the construct.[10] Professionalism was defined as student behaviors that were characterized across nine domains of professionalism, which is the foundation of our medical school's code of professionalism. Using existing literature and frameworks (e.g., Swick's professionalism characteristics, AAMC Medical Student Objectives Project, CanMEDS, the University of Dundee Three-Circle Model, and the ACGME Competencies), a professionalism code was developed by our medical school curriculum committee. The nine domains include:

Honesty and Integrity
Accountability
Responsibility
Respectful and Nonjudgmental Behavior
Compassion and Empathy
Maturity
Skillful Communication
Confidentiality and Privacy in all patient affairs
Self-directed learning and Appraisal skills

Using these domains, the Associate Dean for Student Services and two faculty members in the Office of Medical Education developed 9 broad questions to address each professionalism domain using a 7-point scale. The peer professionalism assessment also includes a general comment question, which allows the evaluators to reflect on their peers' professionalism.

Instrument Implementation

Multidisciplinary faculty designed a problem-based learning (PBL) learning experience to augment an interdisciplinary basic-science course, Human Function. It is a year-long course that combines the disciplines of biochemistry, human genetics, and human physiology. Each of the 14 PBL groups includes 8 students and one facilitator. The course is divided into two, fifteen-week semesters. After the first fifteen-weeks, students are placed into a new PBL learning group with a different facilitator and different students. All students, then, have the advantage of two facilitators each year, and working with different peers in each semester.

Each fifteen-week component includes five cases. Each case confronts students with a complex problem. A packet of information explains a patient’s chief complaints, a psychosocial history, physical symptoms, and particular lab results. Students are asked to share and explore hypotheses of the patient conditions. Students also identify key learning issues, or questions about the material. Students research the learning issues and present information to the group members for the following sessions.

Given the expectations of PBL, it was an appropriate course to implement the professionalism evaluation. As students cooperatively share information, PBL aims to develop several skills that are characteristic of the 9 domains of student behaviors outlined in our code of professionalism. Students, for example, apply interpersonal skills as information is presented and critiqued. Students are also accountable to come to PBL and contribute information to address the learning issues.

During the academic year of 2004-2005, 111 students were given the peer professionalism assessment beginning the last week of each semester. Students were asked to complete peer evaluations for each student in their PBL group. Evaluator names remained anonymous and students were not penalized for failing to complete the peer evaluations.

Research Questions

The analyses of the student comments on a peer professionalism evaluation were addressed by two research questions:
1. What are the underlying themes of student comments about the professionalism of peers?

As Harris [11] noted, qualitative analyses investigate data in the form of words rather than numbers. A qualitative approach evaluates the raw, descriptive data of student comments, and organizes them into themes, which can reveal a conceptual framework. Using the principles of qualitative data analysis, the research process was organized into the following steps: First, student comments were collected from the 2004-2005 peer professionalism evaluations. Student comments were also distinguished between male and female subjects.

Second, the comments were entered into Nudist 6 software for analysis.[12] The principal author conducted the initial interpretation of the data. The author searched for themes, which were used to code the data. This strategy is an inductive method for identifying codes. Third, all authors reviewed the comments and the identified themes. During three, two hour meetings, the authors audited the initial interpretation of the data set. The authors then made modifications to the themes until consensus was achieved. Finally, the data were coded and organized into tables.

2. Is there a significant difference in the mean professionalism scores between students who did and did not receive a negative comment?

An independent-sample test was calculated with comment characteristic as the independent variable and the mean scale score as the dependent variable. The professionalism assessment included 9 scale questions. The dependent variable (scale score) was the mean score of the nine scale questions included in the professionalism assessment. Using a 7 point scale, mean scale scores could range between 1 and 7. Students who received at least one negative comment (n= 28) and students who did not receive a negative comment (n= 83) were distinguished into two groups: positive and negative comment characteristics. We hypothesized that students who received a negative comment would have a lower mean professionalism score than students who only received positive comments.

RESULTS

Students were relatively compliant. A total of 111 students (63 males and 48 females) were included in this study. Across the 14 PBL groups in each semester, 1,554 peer assessments were distributed. Approximately 93% (1,445) of the peer assessments were completed. Across the 1,445 completed peer evaluations, 504 comments were identified. A total of 12 themes were distinguished (see Table I). Table II identifies the frequencies and percentages of comments for themes across male and female students. Not all students received comments for each theme. The majority of comments were positive. The total for male (n = 238) and female (n = 187) positive comments reported that 84% of the overall comments were positive.

The most commonly identified themes for both male and female students include participation/contributions (n = 105), a global assessment of professional skills (n = 101), collaboration (n = 59) and respectful/nonjudgmental (n= 52) (see Table II). Both males and females received several negative comments relating to participation/ contributions, collaboration, respectful/nonjudgmental, and attendance/timeliness. The most common negative comments for males addressed attendance and timeliness, whereas the female most common negative theme was participation/contributions.

There was a significant difference in mean professionalism scores between students who did and did not receive a negative comment (t = 2.93, df = 109, p = .002). Overall, mean professionalism scores were higher for students who did not receive at least one negative comment (M= 6.23, SD = .29) than students who did receive a negative comment (M = 5.95, SD = .56). The effect size for comment characteristic was adequate (d = .67), suggesting that there is a practical difference in the mean professionalism score between students who did and did not receive a negative comment.

DISCUSSION

The majority of students completed the peer professionalism evaluations. This finding is consistent with previous research, which found that students are willing to complete anonymous, formative evaluations of the professionalism of their peers.[13] The majority of comments reveal that students target a variety of appropriate themes. For example, students attended to their peer participation, attendance, and ability to respect other opinions, which are targeted and important components of professionalism. Overall, the student comments addressed a range of themes, indicating that they can contribute valuable information to the assessment of students' professionalism.

An overall reflection of the themes that captured several negative comments suggests that students are attuned to whether their peers are meeting basic expectations. Students are willing to point out whether their peers are not being punctual, contributing to the group, respecting others, and engaging in discussion. Bryan and coworkers [14] examined a peer assessment of professionalism in a gross anatomy course included in the Mayo medical school curriculum. Consistent with the findings of this study, the authors report that the majority of negative comments address the areas of inter-professional respect and accountability.

This study revealed that students who received a negative comment had an overall lower mean professionalism score than students who did not receive a negative comment. This result is consistent with the research by Rudy and coworkers[15], which reported that positive comments about professionalism in a peer evaluation were associated with higher scale scores.

There are some limitations of this study. For example, the anonymity of the evaluators limited the analyses. Identifying the evaluator gender, for example, will answer whether male and female evaluators are more likely to address particular professionalism themes. In addition, this study used nine scale questions and one general comment question. The scale questions may influence students to reflect on particular themes. Several of the themes are consistent with professionalism frameworks.

Table 1. Definitions of Themes and Sample Positive and Negative Comments

Theme	Definition *Comments that address:*	Sample Comments Positive	Negative
Participation/ Contributions	a student's ability to actively contribute to the discussion and share information to help further peers' understanding of the material.	(She) was willing to volunteer and aid in everyone's understanding of the material.	Sometimes, she should jump in and take some credit for the research she's done.
Global Assessment of Professional Skills	a student's global, overall demonstration of professionalism.	(He) is very professional and competent. He will make a great doctor.	Poor professionalism. I would not recommend him as a future physician to my family or friends.
Collaboration	a peer's willingness to work with the student in a team setting.	I really like working with her. I hope to work with him in the future.	I felt he could be more interactive with the group.
Respectful/ Non-judgmental	a student's ability to respect peers' contributions.	It was clear she respected the opinions of others.	(He) did not seem to care about the feelings of others and often would not see the opinions of others.
Prepared/ Organized	a student's ability to come to class ready to discuss and contribute material.	(She) came to class prepared every day.	I don't think (he) researched the learning issues before class.
Inquisitiveness	a student's ability to ask questions and engage in learning.	He often asked insightful questions that showed his genuine interest in learning.	I don't think he took the learning issues very seriously. (He) hardly ever identified one (learning issue).
Fund of Knowledge	a student's knowledge of the material, ability to address issues that may have otherwise been missed.	She is very knowledgeable about the week's topics.	None Identified
Communication Skills	a student's ability to listen, and present ideas clearly and succinctly.	She always listened in group… (She) presented the material very well.	He sometimes made unrelated comments that distracted the group.
Attendance/ Timeliness	a student's ability to arrive on time and be present for the class meetings.	She was always present and available.	(He) attended very few sessions.
Sense of Humor	a student's ability to exhibit a sense of humor.	She has a great sense of humor.	Her jokes seemed inappropriate at times.
Leadership	a student's ability to take the lead and keep the group on task.	She had the most leader-like position in the group.	None Identified
Appearance	a student's appearance in class meetings.	The best dressed medical student I have seen.	None identified

For example, several themes were related to Swick's professionalism characteristics. Student comments that address participation/contributions, prepared/organized, and attendance/timeliness are all similar to Swick's professionalism characteristics of accountability and responsibility characteristics. Also, the communication, collaboration, respectful/nonjudgmental and self-directed learning/inquisitive themes are aligned. Some unique themes identified in student comments, include fund of knowledge, sense of humor, and appearance. Differently phrased scale items and comment items on peer evaluations may reveal comment themes that were not

Table 2. Frequency and Percentage of Themes across Student Comments about the Professionalism of Male and Female Peers.

Theme	Male				Female				Total
	Positive		Negative		Positive		Negative		
	n	(%)	n	(%)	n	(%)	n	(%)	n
Participation/Contributions	47	(45)	**12**	**(11)**	38	(36)	**8**	**(8)**	105
Global Assessment of Professional Skills	50	(50)	5	(5)	44	(44)	2	(2)	101
Collaboration	26	(44)	**8**	**(14)**	21	(36)	**4**	**(7)**	59
Respectful Non-judgmental	27	(52)	**6**	**(12)**	16	(31)	**3**	**(6)**	52
Prepared/organized	19	(49)	2	(5)	17	(44)	1	(3)	39
Fund of Knowledge	20	(56)	0	(0)	16	(44)	0	(0)	36
Inquisitiveness	18	(55)	4	(12)	10	(30)	1	(3)	33
Communication Skills	19	(61)	2	(6)	9	(29)	1	(3)	31
Attendance/Timeliness	6	(20)	**15**	**(50)**	5	(16)	**4**	**(13)**	30
Sense of Humor	3	(38)	0	(0)	4	(50)	1	(13)	8
Leadership	2	(29)	0	(0)	5	(71)	0	(0)	7
Appearance	1	(33)	0	(0)	2	(66)	0	(0)	3
Total	238	(47)	54	(11)	187	(37)	25	(5)	504

Please note: The four themes with relatively high frequencies of negative comments are in bold.

present in this study. Future research in this area could help answer these important questions.

Future research may also explore how student assessment of their peer professionalism may change over time. The student comments in this study were collected for two semesters of a PBL course. While no discernable differences between the frequencies and themes of comments were ascertained between semesters, future studies can address whether the comments become more sophisticated as students progress through the curriculum. The exercise of evaluating peers early in the curriculum can be characterized as a training tool, which may be repeated to sharpen student appraisal skills. Implementing multiple opportunities to assess peer professionalism may further student ability to offer more precise information about professional behaviors. Future studies, for example, may help qualify whether student comments about their peer professionalism improve between the first-year and the fourth-year of the medical school curriculum.

Further, future studies can examine how student assessment of their peer professionalism may be different in clinical and didactic learning situations. Louise Arnold [9] noted that professionalism assessments must consider the learning context. This study focused on a problem-based learning course. Future studies may examine whether the themes presented in Table 1 reflect themes identified in other learning situations, such as clinical clerkships. In addition, educators may have an effect on the learning context. Educators have diverse characteristics and different teaching styles. Future studies may examine how educators may impact student ability to assess their peer professionalism.

Overall, this study provides an initial framework from which future studies may draw. The themes will allow researchers to deductively examine what professionalism characteristics students may address, particularly for problem-based learning courses. This study also suggests that student comments are valuable sources of information. Students are willing to reflect on the behavior of their peers across a variety of professionalism attributes. Still, the literature recommends that educators should implement several assessments to gauge student professional behavior. We agree with other researchers who argue that student observations about their peers professionalism is only one piece of the assessment puzzle.[16] Educators should exercise caution when using peer comments for evaluation and grading purposes. Students are only beginning to develop their appraisal skills, which may not be as sophisticated as more experienced evaluators.[17] Additional assessments, coupled with student comments, are required to create a more complete and detailed picture of whether students are acquiring and demonstrating professionalism competencies.

REFERENCES

1. Swick, H.M. Toward a normative definition of medical professionalism. *Academic Medicine*. 2000; 75(6): 612-616.
2. CanMEDS. Extract from the CanMEDS 2000 Project Societal Needs Working Group Report. *Medical Teacher*. 2000; 22(6): 549-554.
3. Harden, R.M., Crosby, J.R., Davis, M.H., and Friedman, M. From competency to meta-competency: A model for the specification of learning outcomes. *Medical Teacher.* 1999; 21(6): 546-52.
4. Accreditation Council for Graduate Medical Education (ACGME). *Accreditation Council for Graduate Medical Education Outcomes Project.* Available at http://www.acgme.org/outcome/. 2005; Accessed December 15, 2005.
5. Association of American Medical Colleges (AAMC). *Report 1: Learning Objectives for Medical Student Education. Guidelines for Medical Schools.* Washington, DC: Medical School Objectives Project, AAMC, 1998.
6. Swick, H.M., Szenas, P., Danoff, D., and Whitcomb, M.E. Teaching professionalism in undergraduate medical education. *Journal of the American Medical Association* 1999; 282: 830-832.
7. Lazarus, C.J., Chauvin, S.W., Rodenhauser, P., and Whitlock, R. The program for professional values and ethics in medical education. *Teaching and Learning in Medicine*. 2000; 12(4): 208-211.
8. Dannefer, E.F., Henson, L.C., Bierer, S.B., Grady-Weliky, T.A., Meldrum, S., Nofziger, A.C., Barclay, C., and Epstein, R.M. Peer assessment of professional competence. *Medical Education.* 2005; 39: 713-722.
9. Arnold, L. Assessing professional behavior: yesterday, today, and tomorrow. *Academic Medicine*. 2002; 77(6): 502-515.
10. Devellis, R.F. *Scale development*: *Theory and applications*. London: Sage, 2003.
11. Harris, I. Qualitative Methods. In Geoff Norman (ed.) *International Handbook of Research in Medical Education*, pp. 45-85. Hingham, MA**:** Kluwer, 2002.
12. Nudist 6 software. QSR*:* Victoria: Scolari, 2000.
13. Arnold, L., Shue, C.K., Kritt, B., Ginsburg, S., and Stern, D.T. Medical students' views on peer assessment in professionalism. *Journal of General Internal Medicine*. 2005; 20(9): 819-824.
14. Bryan, R.E., Krych, A.J., Carmichael, S.W., Viggiano, T.R., and Pawlina, W. Assessing professionalism in early medical education: Experience with peer evaluation and self-evaluation in the gross anatomy course. *Annals of the Academy of Medicine*. 2005; 34(8): 486-491.
15. Rudy, D.W., Fejfar, M.C., Griffith, C.H., and Wilson, J.F. Self- and peer assessment in a first-year communication and interviewing course. *Evaluation and the Health Professions.* 2004; 24: 436-445.
16. Lynch, D.C., Surdyk, P.M., and Eiser, A.R. Assessing professionalism: a review of the literature. *Medical Teacher*. 2004; 26(4): 366-373.
17. Ress, C., and Shephard, M. Students' and assessors' attitudes towards students' self-assessment of their personal and professional behaviors. *Medical Education*, 2005; 39: 30-39.

Gender Differences in Academic Qualifications and Medical School Performance of Osteopathic Medical Students

Donna Dixon, Ph.D.

Department of Biomedical Sciences
New York College of Osteopathic Medicine
Northern Boulevard
Old Westbury, NY 11568-8000
U.S.A.

Phone: (+)1-516-686-3769 Fax: (+)1-516-686-3832 Email: ddixon@nyit.edu

ABSTRACT

This study posed the following questions: (1) do women and men differ in their overall medical school performance? (2) are there significant differences in the preadmission academic qualifications of female and male medical students? (3) are gender differences in preadmission qualifications a factor in medical school performance? The study included 705 students in four successive classes at the New York College of Osteopathic Medicine (NYCOM). There was no gender difference in undergraduate GPAs, while the mean total MCAT scores of men in all classes were higher than those for women. There was also no significant gender difference between women and men in their cumulative GPAs for the first two years of medical school. Men had higher mean scores on the Comprehensive Osteopathic Medical Licensing Examination (COMLEX) Level 1, given after the first two years of medical school, but when total MCATs were controlled, there was no gender difference in COMLEX Level 1 performance. MCATs were shown to be correlated with COMLEX Level 1 performance. Clinical performance was determined by scores on clinical subject examinations and the clinically-based COMLEX Level 2 examination. No significant gender differences were seen in these two clinical performance measures. Women outperformed men in evaluations of clinical clerkship performance.

INTRODUCTION

In the last three decades there has been a large increase in the number of women entering the medical profession. In 2004, 50 percent of the entering class of medical students in the United States and 45 percent of the graduating class were women.[1] Despite the increase in numbers of women graduating from medical school, their achievement in academic medicine is less than that of men by the conventional standards of academic rank, salary, and publications.[2-3] Reasons suggested for the gender gap in academic medicine have included sex differences in career goals, the time commitment incompatible with the raising of children, bias, and a lack of institutional support.[2] There have also been recent controversial assertions that women may have less innate science ability.[4]

A limited number of studies have directly compared medical students with respect to their medical school performance. Men were reported to score better on the National Board of Medical Examiners (NBME) examination.[5] Women performed better than men on clinical clerkships and also on clinical performance examinations.[6-7] Osteopathic women medical students outperformed men on clinical clerkships in obstetrics and gynecology.[8]

A search of both MEDLINE and the Educational Resources Information Center (ERIC) did not retrieve any published reports comparing the overall performance of women and men at an osteopathic medical school. An important measure of osteopathic medical school performance is performance on the Comprehensive Osteopathic Medical Licensing Examinations (COMLEX-USA). The COMLEX-USA Level 1 examination is given after the first two years of medical school, and students are tested on basic science knowledge relevant to medical problems. The COMLEX-USA Level 2-CE (cognitive evaluation) is typically taken in the fourth year and tests knowledge of clinical concepts and principles involved in medical problem-solving.

The present study was designed to answer the following questions: (1) Do women and men differ in their overall medical school performance, including the COMLEX-USA examinations? (2) Are there significant differences in the preadmission academic qualifications of female and male medical students? (3) Are gender differences in preadmission qualifications a factor in medical school performance?

MATERIALS AND METHODS

The study involved 705 medical students from four successive graduating classes (2002-2005) at the New York College of Osteopathic Medicine (NYCOM). These were traditional students who had completed their course work in four years and had passed the COMLEX-USA Level 1 and Level 2 examinations. The size of each class (2002-2005) was 168, 180, 182, and 175, respectively. The percentage of women in each class was 48%, 46%, 49%, and 50%, respectively. Performance measures included the medical school GPA, clinical subject examinations, clinical evaluations, and the Level 1 and Level 2-CE COMLEX examinations. GPAs and clinical performance data were obtained from institutional databases, and COMLEX-USA examination scores were those reported by the National Board of Osteopathic Medical Examiners to the institution. The average GPA for the first two years of medical school was the cumulative GPA. Clinical subject examinations were given after the clinical clerkships of osteopathic manipulative medicine, family medicine, medicine, obstetrics and gynecology, pediatrics, psychiatry, and surgery. This measure of clinical performance was the mean of the individual subject examinations (clinical examination). The mean of the clerkship evaluations, based on a five point scale, was the clinical evaluation score (clinical evaluation). COMLEX examination scores and MCAT scores were expressed as differences in the mean scores of the genders in each class to preserve the anonymity of privileged school data.

Preadmission data was obtained from the American Association of Colleges of Osteopathic Medicine Application Service. Preadmission academic variables were the cumulative undergraduate GPAs and the Medical College Admission Test (MCAT) scores. Individual subscores were verbal reasoning (verbal MCAT), physical sciences (physical MCAT) and biological sciences (biological MCAT). The total MCAT score was the sum of the three individual subscores.

Medical school performance was measured by comparing the means of the performance measures using the independent-samples t-test. An analysis of covariance using the total MCAT scores as a covariate was performed. Statistical tests were calculated with SPSS statistical software, version 14.0.

RESULTS

Table 1 shows the cumulative GPAs, clinical evaluation scores, and clinical examination scores of the women and men in each class. There was no difference between women and men in their cumulative GPAs or performance on the clinical subject examinations. Women had higher clinical evaluation scores than men (significant in two classes, $p \leq 0.02$). Table 2 shows the gender comparison of COMLEX-USA examination scores. The data is expressed as the difference in the mean scores for women and men. Men in all classes had higher mean COMLEX Level 1 scores than women, and the differences were significant for three of the four classes ($p \leq 0.02$). There were no significant gender differences in performance on COMLEX Level 2-CE. The class-based analysis showed that there was a consistent pattern of results across the four classes. There was no discernible gender difference in the

Table 1. Gender Comparison of Medical School Performance.

Mean (SD)

		2002		2003		2004		2005	
			*p**		*p*		*p*		*p*
Cumulative GPA	Women	77.22(4.59)	0.08	79.44(3.96)	0.24	78.81(5.45)	0.09	79.51(4.46)	0.96
	Men	78.51(4.83)		80.19(4.49)		77.44(5.23)		79.47(4.79)	
Clinical Evaluation	Women	4.33(0.33)	0.55	4.39(0.49)	0.02	4.32(0.31)	0.27	4.36(0.27)	0.01
	Men	4.29(0.55)		4.24(0.43)		4.28(0.30)		4.26(0.27)	
Clinical Examination	Women	593.84(74.71)	0.18	610.96(83.97)	0.43	609.24(75.01)	0.34	613.30(67.56)	0.76
	Men	610.28(84.13)		620.66(79.98)		598.54(74.38)		609.96(75.76)	

*Significance of gender difference.

Table 2. Gender Comparison of COMLEX-USA Examination Scores

Difference Between the Mean Scores of Women and Men

	2002		2003		2004		2005	
		p		*p*		*p*		*p*
COMLEX-USA Level 1	27.67*	0.01	23.04*	0.02	8.53	0.40	20.27*	0.04
COMLEX-USA Level 2	20.59	0.06	9.51	0.41	9.38	0.35	2.63	0.80

Men's scores are higher than women's, $p \leq 0.02$

performance means in individual classes although the percentage of women increased from the class of 2000 to that of 2005.

The cumulative undergraduate GPAs of the women and men were not significantly different (data not shown). Analysis of the MCAT scores of the students showed gender differences. Table 3 shows the gender comparison of MCAT scores as the difference between the mean scores of women and men. The total MCAT scores of men in all classes were higher than those of women (difference, 0.72-2.32). This difference was significant in three classes ($p \leq 0.00$). Men had higher mean physical MCAT subscores than women in each class ($p \leq 0.00$). Biological MCAT subscores were also higher for male students and were significant in three classes ($p \leq 0.04$). Women had higher verbal MCAT subscores in two classes, and men had significantly higher verbal MCATs in one class. To examine the effect of MCAT scores on medical school performance, an analysis of covariance was done using the total MCAT scores as a covariate. After controlling for the total MCATs, the COMLEX Level 1 scores of women and men were not significantly different (Table 4). The only gender differences now seen were the higher clinical evaluation scores for women (significant in two of four classes, $p \leq 0.03$).

Discussion

The performance of female and male students at NYCOM was equal on the medical school performance measures of cumulative medical school GPA, clinical subject examinations, and COMLEX Level 2-CE examination scores. There was no gender difference in COMLEX Level 1 examination scores when controlling for total MCATs. Women had higher mean clinical clerkship evaluations than men in each class.

Table 3. Gender Comparison of MCAT Scores.

Difference Between the Mean Scores of Men and Women

	2002		2003		2004		2005	
		p		*p*		*p*		*p*
Verbal MCATs	0.59*	0.04	0.24	0.38	0.32†	0.22	0.34†	0.28
Physical MCATs	0.84	0.00	1.29	0.00	0.73	0.00	0.91	0.00
Biological MCATs	0.49	0.04	0.79	0.00	0.31	0.08	0.88	0.00
Total MCATs	1.92	0.00	2.32	0.00	0.72	0.11	1.40	0.00

*Men's mean MCAT scores are higher them women's except where indicated.
†Women's mean MCAT scores are higher.

Table 4. Analysis of Covariance of Medical School Performance Measures*

	2002		2003		2004		2005	
	F-ratio	*p*	F-ratio	*p*	F-ratio	*p*	F-ratio	*p*
Cumulative GPA	0.89	0.35	0.05	0.82	3.44	0.06	0.04	0.84
Clinical Evaluation	0.51	0.47	7.55	0.01	1.20	0.28	4.80	0.03
Clinical Exam Score	0.13	0.72	0.82	0.37	1.30	0.26	0.35	0.56
COMLEX-USA Level 1	3.44	0.07	0.59	0.44	0.35	0.55	3.24	0.07
COMLEX-USA Level 2	1.46	0.23	0.52	0.47	1.44	0.23	0.41	0.52

*Total MCATs as covariate.

While women and men had equivalent undergraduate GPAs, the male students in each of the four classes had higher mean physical and biological science MCATs and total MCAT scores than the women. Physical and biological MCAT scores were previously found to be significant predictor variables for COMLEX level 1 performance at this institution.[9] Total MCATs and physical MCATs were reported as predictors for COMLEX level 1 at other osteopathic schools.[10-11] Since the COMLEX Level 1 exam is given after the first two years of medical school and focuses on basic science knowledge, the higher total mean COMLEX Level 1 scores for men could be partly due to their higher total MCAT scores. When the total MCATs were controlled, there was no significant difference between women and men with respect to scores on COMLEX Level 1. This finding shows that higher performance in the physical and biological parts of the MCATs is correlated with COMLEX Level 1 performance, and the larger percentage of male students with higher MCATs in each class was the cause of the mean gender difference seen. MCAT scores apparently had little effect on performance on the clinically-based COMLEX Level 2-CE examination.

The performance of women in the clinical years was found to be equal to that of men, and higher than men in the case of clinical clerkship evaluations. The ratings for the clerkship evaluations are somewhat subjective. The rating sheets for attitude / professionalism and clinical skills include items such as cooperation, patient communication, and receptivity to feedback. These are areas in which women may excel. Women have been reported to have greater abilities to actively listen and create better relationships with patients than men.[7] It can be postulated that the women's higher clinical evaluation grades found in this study may reflect their better abilities in the areas of cooperation, patient communication, interviewing, and counseling.

This is one of only a few studies that have reported gender comparisons of both preclinical and clinical performance of medical students at an osteopathic or allopathic medical school. The results show that women graduating from this medical school in recent years are the academic equals of their male classmates. It is hoped that in the future, some of these women will choose careers in academic medicine and help to redress the gender inequalities.

REFERENCES

1. Magrane, D., and Jolly, P. The changing representation of men and women in academic medicine. *Association of American Medical Colleges. Analysis In Brief*. July 2005. Available at http://www.aamc.org/data/aib/start.htm Accessed May 22, 2007.
2. Hamel, M.B., Ingelfinger, J.R., Phimister, E., and Solomon, C.G. Women in academic medicine - progress and challenges. *New England Journal of Medicine.* 2006; 355(3):311-312.
3. Wright, A.L., Schwindt, L.A., Bassford, T.L., Reyna, V.F., Shisslak, C.M., St. Germain, P.A., and Reed, K.L. Gender differences in academic advancement: patterns, causes, and potential solutions in one U.S. college of medicine. *Academic Medicine*. 2003; 78:500-508.
4. Handelsman, J., Cantor, N., Carnes, M., Denton, D., Fine, E., Grosz, B., Hinshaw, V., Marrett, C., Rosser, S., Shalala, D., and Sheridan, J. More women in science. *Science*. 2005; 309:1190-1191.
5. Ramsbottom-Lucier, M., Johnson, M.M.S., and Elam, C.L. Age and gender differences in students' preadmission qualifications and medical school performance. *Academic Medicin*e. 1995; 70:236-239.
6. Haist, S.A., Wilson, J.F., Elam, C.L., Blue, A.V., and Fosson, S.E. The effect of gender and age on medical school performance: an important interaction. *Advances in Health Sciences Education*. 2000; 5:197-205.
7. Haist, S.A., Witzke, D.B., Quinlivan, S., Murphy-Spencer, A., and Wilson, J.F. Clinical skills as demonstrated by a comprehensive clinical performance examination: who performs better – men or women? *Advances in Health Sciences Education*. 2003; 8:189-199.
8. Krueger, P.M. Do women students outperform men in obstetrics and gynecology? *Academic Medicine*. 1998; 73:101-102.
9. Dixon, D. Relationship between variables of preadmission, medical school performance, and COMLEX-USA Levels 1 and 2 performance. *Journal of the American Osteopathic Association*. 2004; 104:332-336.
10. Meoli, F.G., Wallace, W.S., Kaiser-Smith, J., and Shen, L. Relationship of osteopathic medical licensure examinations with undergraduate admission measures and predictive value of identifying future performance in osteopathic principles and practice/osteopathic manipulative medicine courses and rotations. *Journal of the American Osteopathic Association.* 2002; 102:615-620.
11. Cope, M.K., Baker, H.H., Fisk, R., Gorby, J.N., and Foster, R.W. Prediction of student performance on the Comprehensive Osteopathic Medical Licensing Examination Level 1 based on admission data and course performance. *Journal of the American Osteopathic Association.* 2001; 101:84-90.

Design, Implementation, and Evaluation of the Transition from Paper and Pencil to Computer Assessment in the Medical Microscopic Anatomy Curriculum

Cynthia J. M. Kane, Ph.D.[1], E. Robert Burns, Ph.D. [1], Patricia S. O'Sullivan[2], Ph.D., Thomas J. Hart, M.S. [3], James R. Thomas, B.M.E.[3], and Ivy A. Pearsall, B.S.[3]

[1]Department of Neurobiology and Developmental Sciences
[2]Office of Educational Development
[3]Office of Academic Computing
University of Arkansas for Medical Sciences 4301 W. Markham Street
Little Rock, AR 72205
USA

Phone: (+)1-501-686-7022 Fax: (+)1-501-526-6756 Email: kanecynthiaj@uams.edu

ABSTRACT

Significant forces are converging to reshape the basic science medical curriculum including advances in educational theory, advances in computer technology, and increased understanding of human disease processes. The explosion of information in the medical field and financial pressure on academic medical centers is stimulating significant change in the way that the basic science curriculum is delivered. This study was undertaken to incorporate these advances into the medical microscopic anatomy curriculum by transition from paper and pencil to computerized assessments of student learning. The goals included: modernization of the mode of student assessment on high stakes examinations, maintenance of high academic standards, maintenance of student performance, acquisition of student experience in computerized testing, and reduction in student and faculty time required for assessment. Design, implementation and evaluation of the transition was documented and analyzed. Freshmen medical students were given either written or computer assessments. Student performance was compared on identical items. There was no performance difference in the overall course or for most of the experimental items. The written format provided an advantage on 3% of the items that were likely cued by proximal items. Student evaluation of the transition was positive. They felt better prepared for future computerized examinations. Student satisfaction with instant scoring of the examination was rated quite high. There was recovery of significant student and faculty time in the assessment process. There was student suspicion of the accuracy of item scoring by the computer. The concern was investigated and student satisfaction was obtained with addition of a minor modification to the programming. Overall, the transition to computerized assessments was successful and productive for both students and faculty.

INTRODUCTION

A dramatic increase in the depth and scope of biomedical knowledge has led to significant changes in the medical curriculum in the basic sciences[1-3]. These changes provide educators a timely opportunity to incorporate advances in computer technology[4,5], educational theory[6], and understanding of human disease[7] in a new era of excellence in medical education.

We applied these advances in a focused manner to successfully transition the high stakes examinations in a freshman basic science course in the medical curriculum from a paper and pencil (i.e., written) format to a computerized format. Computerized assessment of student learning was designed, implemented and evaluated in the context of the medical microscopic anatomy course at the University of Arkansas for Medical Sciences. The course content encompassed histology, cell biology and embryology. The purpose of this study was to determine if students would perform equivalently on identical items presented in either written or computerized format in the course examinations, document student and faculty attitudes about the change to computerization of the major examinations, and identify key elements to a successful transition in this educational context.

Assessment of student learning is essential to the educational process to determine if learning goals are being achieved. Student assessment is also central to curricular development and evaluation of curricular changes. Assessment can be performed by several modalities, but in the basic biomedical sciences it has been traditionally accomplished by evaluation of student performance on paper and pencil written examinations. However, students are required to complete the U.S. Medical Licensing Examination (USMLE) on the computer and, thus, students require educational preparation for the computerized assessment modality[8,9]. Further, computerized assessment can overcome several of the disadvantages of written examinations including: errors in answer transcription by students, criteria for item scoring, and the time and delay inherent to grading of answer sheets; limited incorporation of high resolution, realistic images; and, in particular, the commitment of time and effort on the part of faculty, and the constraints imposed on student and faculty schedules[10]. As a result, computerized assessment holds promise for improvements in student learning, and also student-focused curricular development, as well as productive investment of student and faculty time, and student preparation for the USMLE[8,9].

Although it is generally recognized that equivalent paper and computer examinations can be prepared[8,11,12], it remains essential for the change from paper to computer to be evaluated scientifically for positive or negative effects, in specific contexts, with specific test materials and specific student populations[13]. However, to our knowledge, there is no report with statistical comparison of a successful transition of high stakes assessments from traditional written examinations to computerized examinations in basic science courses in the medical curriculum. Because significant faculty and institutional investment is required to implement the change to computerized assessments in medical education[14], the results reported here may serve as a resource for others pursuing a similar path.

MATERIALS AND METHODS

Design. This study compared student performance on assessments in the medical microscopic anatomy course accomplished by written examinations in Year 1 (Y1) to computerized examinations in Year 2 (Y2). This study was reviewed and classified as exempt by the Institutional Review Board. Participants were the 154 freshmen students taking medical microscopic anatomy (composed of histology, cell biology, and embryology) in Y1 and the 149 freshmen students taking the class in Y2. The students had no prior or concurrent exposure to computerized examinations in the medical curriculum. The students were registered concurrently for two additional courses, gross anatomy and introduction to clinical medicine.

Intervention. The intervention in this study was the administration of 101 items, first, in paper and pencil format (Y1) and, then, in a computerized format (Y2). The content of each item was identical each year and the items were safeguarded in a secure database. The items were one-best-answer, multiple choice questions containing 1 correct and 4 incorrect responses in each item. The items did not contain visual images because of the inherent difficulty and expense to incorporate images into a paper and pencil written examination format. The 101 items were distributed throughout all of the major course examinations, 7 in Y1 and 5 in Y2. The examinations contained a total of 678 items in Y1 and 248 items in Y2. The items were pre-tested in the year prior to this study to determine individual and group item statistics. The pre-test was performed using 151 comparable medical students in the same educational assessment context used in the intervention. Item inclusion in this study was based on difficulty at the medium level (average 75% of the students passing the item) for the group, item reliability of at least 0.85 for the group, and individual item discrimination index ≥ 0.25.

Implementation. For paper and pencil examinations in Y1, the students were given a printed examination booklet and a scanable answer sheet on which answers were marked by filling in printed circles with a pencil. For written examinations, all 154 students sat for each assessment as a single group in a teaching laboratory. The students were allowed 2 hours to complete the examination, all examination booklets were collected, answer sheets were scanned by computer, and scores were posted within 72 hours. After the scores were posted, students were given the opportunity to briefly review the items in a proctored environment and submit a formal written appeal regarding the scoring of an item. Reviews were conducted in the teaching laboratory under the proctoring of the course faculty. The appeal process revealed that no scoring change was necessary for the items included in this study.

Prior to the first scored computerized examination in Y2, students were given the opportunity to participate in a technical training session, which included a simple, very short practice examination, to acquaint them with the computerized assessment design. Approximately 96% of the students participated in the computer training and completed the practice examination.

For computerized examinations, the students were divided into two groups of 72 students each because of a limited number of available computers. The two groups sat for the examination starting at either 8 am or 10 am, based on alphabetical order of last name, and the group start time was alternated for each subsequent assessment. The students sat at individual personal computers in a computerized testing laboratory and logged into the examination with individual coded passwords which changed for each examination session. Examinations were constructed using Authorware® testing software

(MacroMedia). As a security feature, items were presented to each student in a random order to prevent simultaneous presentation of items on multiple computer monitors. Students could move back and forth in the examination to review and change the selected response to any item. The students were allowed 2 hours to complete the examination and, after submitting their responses for grading, they could elect to receive their score immediately on-screen.

A software program was developed to capture student keystrokes on each computer. This was necessary to document student response selection in order to alleviate student concern that the computer did not accurately record their selected responses. In addition, the software was re-programmed to slow down the rate at which a student could scroll through the items and review their responses. This was necessary to allow time for the software to display an item and record the unchanged or changed response before displaying the next item. After the assessment session, students were given the opportunity to briefly review the items and submit a formal written appeal regarding the scoring of an item. The review was conducted on computers in a proctored environment. Faculty review of the student appeals determined there was no change necessary for the scoring of any item included in the study.

Analysis. The class mean Medical College Admissions Test (MCAT) score and mean undergraduate college grade point average (GPA) were compared for each class to determine the equivalence of the student groups. Also, the class mean percentage score and GPA in the medical microscopic anatomy course were compared for each class to evaluate overall course performance. The average difficulty and average discrimination of the items were compared between the two formats. These three statistical comparisons were done using independent sample t-tests. The level of significance was set at p=0.05 for each test.

The proportion of students selecting the correct response for each item was analyzed to compare student performance on the two formats and to determine if one format was more advantageous for any specific item. The proportion of students answering correctly on each format was compared using a test for two independent proportions which is distributed as a *z* statistic. To correct for inflated Type I error, the significance was set at p=0.01 for each comparison.

Student evaluation was obtained at the end of the course by voluntary student completion of a standardized course evaluation which incorporated a Likert scale as well as *ad lib* comments. Questions regarding the computerized examinations were added to the course evaluation to document the student experience. Descriptive statistics of student ratings were calculated.

RESULTS

The aptitude of the two student populations was evaluated (Table 1). There was no statistical difference in the mean undergraduate college GPA. Students in the Y2 group had significantly lower MCAT scores. This was of initial concern because this group would be the first freshman class to encounter the computerized assessment format. While the two groups are statistically different on the MCAT performance, the numerical difference between the two groups is less than a point and the two groups are 0.24 standard deviation units apart. The potential meaning of the significance of the relatively small difference is not obvious. Further, the difference in entering MCAT score did not correspond to subsequent differences in performance in the course. There was no significant difference in student performance in the medical microscopic anatomy course with respect to the mean percent score or the mean course GPA (Table 2).

Table 1. Entering Class Aptitude

	College GPA*	MCAT score
Year 1	3.53 ± 0.34	26.63 ± 3.95
Year 2	3.59 ± 0.31	25.71 ± 3.69
	p > 0.05	p< 0.05

*4-pt scale.

Table 2. Overall Class Performance in the Course

	Course GPA*	Course Score (%)
Written Assessment	2.93 ± 0.87	85.56 ± 6.7
Computerized Assessment	2.97 ± 0.94	86.31 ± 6.5
	p > 0.05	p > 0.05

*4-pt scale.

The average difficulty of the items as a group was similar (p>0.05) between the two years, 0.85 in Y1 and 0.82 in Y2. The average discrimination of the items as a group was also similar (p>0.05), 0.29 in Y1 and 0.28 in Y2.
Student success in selection of the correct response was compared for the same set of 101 items. Students performed similarly on 97% of the items. However, three test items showed a statistically significant difference in class performance in the computerized format compared to the written format (Table 3). In these three cases, overall class performance was lower when the item was presented in the computerized format. Review of each of the three

Table 3. Assessment Items with Difference in Student Performance

	Written Examination Format Year 1				Computerized Examination Format Year 2			
Item	Correct Response (%)			Discrimination Index	Correct Response (%)			Discrimination Index
	Class Overall	Top ¼	Bottom ¼		Class Overall	Top ¼	Bottom ¼	
Q1	55	84	35	0.35	39	59	35	0.27
Q2	79	97	57	0.35	55	73	38	0.38
Q3	87	97	70	0.31	31	44	25	0.13

items suggested that the change in student performance may have been due to context cueing by conceptually related, near proximity items when the item was presented in the written format. Students were allowed full access to all questions and answers in the computerized examinations until they designated their examination as complete and received their score. This allowed the possibility of cueing as students reviewed the items. However, with random presentation of items in the computerized assessment, items were not presented in conceptually related clusters and as a result cueing was less likely.

As detailed in Table 3, the first item (Q1) with differential performance assessed a difficult concept that was cued for students in the top ¼ of the class by 10 proximal items on related concepts. Students in the bottom ¼ of the class did not benefit from the cued items; 35% selected the correct response in both formats.

Q2 and Q3 (Table 3) were two items on a related concept that were presented in immediate sequence on the written examination in Y1. They were neighbored by three additional items on related concepts. The cue was readily discernible by most students, including those in the top ¼ and bottom ¼ of the class.

To capture feedback from the students about their experience with the computerized examinations, five evaluation items were added to the standard student evaluation of the course in Y2. Course evaluations were completed voluntarily by 122 students which was 82% participation in the evaluation process. All items related to the computerized examinations were ranked high by the students with a mean of 4.31/5.0 (Table 4). The overall course evaluation, which included 15 additional items not specific to the computerized examinations, was 4.10/5.0.

Comments about the computerized assessment experience were reported by students on the course evaluation. Descriptive comments about the course were provided by 58 students with 13 students providing specific comments about the computerized assessments. Two of the student comments contained apprehensive statements about the assessments and both related to distrust of the technology. These students wrote, “I don’t think the computer exams were always fair. There were times when I went back and could swear I never put that answer, but there was nothing I could really do about it.” and “I enjoyed taking the computerized exams, but I still feel that some of my answers might have been changed by the computer.” However, overall the students felt confident about the technology, “The computerized exams were a big topic this semester. I think that most of the issues with them were handled appropriately.”

Table 4. Student Evaluation of Computerized Assessment Modality

EVALUATION ITEM	MEAN* ± SD
Attitude about taking the exams on the computer	4.32 ± 1.00
Computerized exams were well constructed, user friendly and accessible	4.06 ± 1.03
Interest in obtaining exam score immediately	4.57 ± 0.83
Usefulness of computerized exams for USMLE preparation	4.33 ± 0.86
Satisfaction with the computerized exam format	4.28 ± 0.83

* Scored on a Likert scale: 1 to 5, poor to excellent. Items completed by 121-122 students.

The positive student evaluation comments focused on ease of reading the questions, preparation for the computerized format of the USMLE, and the ability to obtain instant scoring. “I loved the computerized exams. It was very easy to use and was actually a more relaxing way to take an exam.” “The computer tests were fantastic.” “I loved the computer exams, please keep them.” “I thought the computerized tests were a great idea. I loved knowing my grade immediately after submitting my final answers.” “I loved the computer exams and the instant grading.” “I think the computerized test is the way to go.” “I like having the exams on the computer.” “As for computer testing, I thought it was great and hope that more of our classes use the same approach." "The computer exams are well designed and thought out." "The computerized exams were good in that they are preparing us for the USMLE. I am kind of old fashioned and like the written tests, but since the USMLE is on a computer I want to be prepared and so I want to take more computer tests."

DISCUSSION

Overall, this controlled study documented statistically that computerized assessment did not provide either an advantage or a disadvantage to a population of freshmen medical students taking microscopic anatomy. Student performance was evaluated in a paper and pencil written format versus a computerized format. Items were selected for inclusion in the study based on individual item statistics and group item statistics following pre-testing in a comparable population in the same academic and situational context. While the group performing the computerized tests could be argued to have had a slightly lower aptitude, as measured by MCAT score, that difference did not affect student performance on the assessment items or in the course overall. The two experimental groups were statistically comparable based on college GPA, course GPA, and performance on assessment items.

Student performance was similar on 97% of the identical items regardless which format was used. However, it is noteworthy that student performance was different on 3 of the 101 items. On these three items, advantage may have been conveyed in the written format because of potential cueing of correct responses by neighboring items with related subject content. This aspect of assessment design is often overlooked. The potential for cueing was based on our review of the written examinations in Y1. Cueing is present to some extent in all examinations regardless of whether the format is paper and pencil or computer. Students were allowed full access to all questions and answers in the computerized examinations until they designated their examination as complete and received their score. This allowed the possibility of cueing as students reviewed the items. To advantage, cueing may have been reduced in the computerized format because items were presented in a random order.

In both years, students were permitted 2 hr to complete each examination. There were 678 questions in 7 written examinations in Y1. There were 248 questions in 5 computerized examinations in Y2. Thus, students had more time per question in the computerized format. This design was employed intentionally in the first year of the computerized examinations to offset student tension about the computer interface. We observed that students did not require the additional time with most students completing each session one hour. The difference in time may have provided an advantage to the students in Y2, but no difference in performance was detected.

Student evaluation of the course provided insightful feedback on the transition. By the end of the course, students were almost unanimously pleased with the computerized examinations. Strong positive features from the student perspective included the use of computer technology for examinations, better preparation of test taking skills for the computerized testing format of the USMLE, and instant examination scoring.

The computer literacy of freshman medical students is very high, both at our institution and elsewhere[15,16]. In the year prior to this study, ≥ 85% of the entering medical class at our institution reported their level of computer expertise as 'comfortable' or 'expert', with no one reporting 'no experience'[17,18]. Further, 82% of these students owned a computer, 78% of those who did not planned to purchase one, and 96% intended to use the Internet from home to access their medical coursework[17,18].

Nonetheless, some students were clearly more comfortable with computer technology than others. While some students reported the computerized testing environment to be relaxing, others reported anxiety regarding the technological aspect of the experience. Our current finding that students strongly preferred computerized examinations over written examinations is supported by previous studies. Although other investigators[13] observe that differences in student prior experience with computers does not affect performance on computerized assessments, the present study incorporated an opportunity for no-risk practice in the computerized assessment system, including completion of a practice examination.

Students indicated that the transition to the computerized format was useful to prepare them for other computerized examinations including the USMLE. Proficiency in computerized testing reduced anxiety about the assessment format of future examinations.

Although student comments were generally favorable, students were particularly pleased with the ability to receive a score immediately at the end of their session.

Previous reports suggested that students were displeased by computerized assessments that did not permit them to review and change their selected responses prior to submission for scoring[13]. Students indicated that the ability to review and change their responses was essential prior to scoring. This was true even though the ability to review and change responses did not appear to alter student performance[13]. Instant scoring is one of the potential benefits of computerized examinations compared to written examinations, and is easily incorporated into computerized assessment design. This feature was incorporated into the present study and was rated highly in the course evaluation.

During the first computerized assessment session, one student voiced concern that the selected responses were not accurately recorded by the scoring program. This led to similar concern on the part of other students. Validation of the concern was not obtained. Extensive assessment modeling and software testing by faculty and staff revealed that it might be possible to force the program to malfunction by very rapidly paging through screen views. The software was reprogrammed with a split-second delay between on-screen display of items. The programmed temporal delay was not perceptible during the assessment process. To confirm that the modifications were successful and that each student response was accurately recorded, a supplemental computer software program was written to capture keystrokes. No hardware or software error was documented based on keystroke analysis. The analytical problem-solving response to the student concerns was well received by students as indicated in the course evaluation. However, a couple of students continued to express distrust of the technology. This concern represented the only negative feedback from the students regarding this major transition in high stakes assessments.

A major goal of the transition to computerized assessment was happily realized in the form of recovered time for both students and faculty. Increased faculty and staff investment was required to establish the item database, evaluate item statistics, design the assessment protocol, implement the computerized format for the first time, and respond to student and faculty feedback. However, in subsequent years, the faculty recovered significantly time and effort necessary for preparation, administration, proctoring, and scoring of assessments. Faculty reported satisfaction that high performance standards were maintained.

Students also recovered time and effort with the computerized assessments. Previously, lecture and laboratory content were assessed independently with the use of paper and pencil written examinations of lecture content (2 hr period) and separate microscope-based examinations of laboratory content (2 hr period). Integration of the content reduced the time needed for assessment of both lecture and laboratory content from a total of 4 to 2 hr. In particular, this reduced the burden on students who were scheduled for multiple course examinations in a single day. This represented a real change in strategy that likely allowed students to perform better on the multiplicity of high stakes assessments.

Another major goal of the transition was the opportunity for students to complete self-scheduled assessments. This was realized in the subsequent year when students began to perform self-scheduled assessments in the Learning Resource Center during a one week period. This change was positively received by students. Ultimately, successful computerized assessment with maintenance of high standards for student learning may allow extension of this medical curriculum into a distance education format. Online distance-delivery of the basic science medical curriculum will provide access for a larger, more diverse student population and, perhaps, positively influence diversity among practicing physicians.

The transition from paper and pencil to computerized assessment was successful for both students and faculty. The assessment modality was modified with minimal impact on student performance on individual items or in the course overall. Importantly, computerized assessment provided student practice in preparation for the subsequent USMLE. In addition, both faculty and students recovered significant time and effort. The modality has been further extended since the completion of this study to allow students to perform the assessments by computer on a self-scheduled basis. As a result of this study, the computerized assessment modality was successfully extended to two additional freshman medical courses the following year.

ACKNOWLEDGEMENT

The authors thank Jean Chaunsumlit for entry of the item database and Elizabeth Hicks for topical literature search. CJMK acknowledges the UAMS Teaching Scholars program and Dr. Gwen V. Childs for provision of support, structure, guidance, and encouragement for this research.

REFERENCES

1. Regehr, G. Trends in medical education research. *Academic Medicine*, 2004; 79: 939-947.
2. Heidger, P.M., Jr., Dee,F.,Consoer,D.,Leaven,T., Duncan,J., and Kreiter, C., Integrated approach to teaching and testing in histology with real and virtual imaging. *The Anatomical Record, Part B: The New Anatomist*, 2002; 269: 107-112.
3. Blake, C.A., Lavoie, H.A., and Millette, C.F. Teaching medical histology at the University of South Carolina School of Medicine: Transition to virtual slides and virtual microscopes. *Anatomical Record: Part B, The New Anatomist*, 2003; 275: 196-206.
4. Hastings, N.B., and Tracey, M.W. Does media affect learning: where are we now? *TechTrends*, 2005; 49: 28-30.
5. Vozenilek, J., Huff, J.S., Reznek, M. and Gordon, J.A. See one, do one, teach one: advanced technology in medical education. *Academic Emergency Medicine*, 2004; 11: 1149-54.
6. Sutherland, R., Armstrong, V., Barnes,S., Brawn, R., Breeze, N., Gall, M. Matthewman, S., Olivero, F., Taylor, A., Triggs, P., Wishart, J., and John, P. Transforming teaching and learning: embedding ICT Into everyday classroom practices. *Journal of Computer Assisted Learning*, 2004; 20: 413-425.
7. Sokol, A.J. and Molzen, C.J. The changing standard of care in medicine. E-health, medical errors, and technology add new obstacles. *Journal of Legal Medicine*, 2002; 23: 449-90.
8. Ogilvie, R.W., Trusk, T.C., and Blue, A.V. Students' attitudes towards computer testing in a basic science course. *Medical Education*, 1999; 33: 828-31.

9. Guagnano, M.T., Merlitti, D., Manigrasso, M.R., Pace-Palitti, V., and Sensi, S. New medical licensing examination using computer-based case simulations and standardized patients. *Academic Medicine*, 2002; 77: 87-90.
10. Papp, K.K. and Aron, D.C. Reflections on academic duties of medical school faculty. *Medical Teacher*, 2000; 22: 406-411.
11. Lieberman, S.A., Frye, A.W., Litwins, S.D., Rasmusson, K.A. and Boulet, J.R. Introduction of patient video clips into computer-based testing: effects on item statistics and reliability estimates. *Academic Medicine*, 2003; 78: S48-51.
12. Preckel, F. and Thiemann, H. Online- versus paper-pencil-version of a high potential intelligence test. *Swiss Journal of Psychology*, 2003; 62: 131-138.
13. Luecht, R.M., Hadadi, A., Swanson, D.B., and Case, S.M.. A comparative study of a comprehensive basic sciences test using paper-and-pencil and computerized formats. *Academic Medicine*, 1998.;73: S51-3.
14. Cantillon, P., Irish, B., and Sales, D. Using computers for assessment in medicine. *British Medical Journal*, 2004; 329: 606-609.
15. Seago, B.L., Schlesinger, J.B., and Hampton, C.L, Using a decade of data on medical student computer literacy for strategic planning. *Journal of Medical Library Association*, 2002; 90: 202-209.
16. Jastrow, H. and Hollinderbaumer, A. On the use and value of new media and how medical students assess their effectiveness in learning anatomy. *The Anatomical Record, Part B: The New Anatomist,* 2004; 280B: 20-29.
17. Heard, J., Allen, R., and Schellhase, D. *A comparison of computer skills of freshmen medical students and resident physicians*. In *Eighth International Conference on Medical Education and Assessment*. 1998. Philadelphia.
18. Hart, J. and Allen, R. *Results of the LRC freshmen computer survey 1995-1999*. 1999.

Weekly Open-Book Open-Access Computer-Based Quizzes for Formative Assessment in a Medical School General Pathology Course

Margret S. Magid, M.D.[1], Merril K. Schindler, M.A., M.S.L.S.[2]

[1]Department of Pathology
[2]Gustave L. and Janet W. Levy Library
Mount Sinai School of Medicine
1 Gustave L. Levy Place
New York, NY 10029

Phone: (+)1-212 241-7459 Fax: (+)1-212-289-2899 Email: margret.magid@mountsinai.org

ABSTRACT

Frequent testing, enhanced by computer delivery, provides a valuable means of formative assessment through timely review of course material, prompt feedback and image display. We introduced weekly computer-based quizzes in a medical school General Pathology course. Quizzes were released for several days following each of five weekly modules and represented, in total, 10% of the final grade. We hoped to further enhance the value of frequent computer quizzes by introducing two new features. First, we offered independent, open-access scheduling with the ability for the student to take the quizzes independently at any site with Internet access. By doing so we conserved in-class hours for instruction and eliminated the need for a designated testing site and faculty supervision. Secondly, we permitted an open-book format to encourage directed course review and decrease stress. Data were generated on specific questions from the quiz and on the entire quiz by individual student and class. Prompt analysis of results permitted timely remediation of problematic topics and identified students at academic risk. In the three years of study, student feedback has been highly favorable, particularly with regard to reinforcing understanding of topics for study, including both those formally presented in class and those assigned as independent learning, and for the impetus to "keep up" with course material. Performance on the final examination showed a statistically significant improvement after introduction of the quizzes. Weekly quizzes, enhanced by self-scheduled computer delivery and open-book format, are a valuable teaching tool for formative assessment.

INTRODUCTION

Formative assessment, that is, the process by which feedback is utilized to modify ongoing teaching and learning, has been lauded as an important process to enhance learning [1]. Feedback is a critical aspect of formative assessment, in that it provides information on what the student has learned and can dictate both teacher- and student-generated actions to correct deficiencies. Frequent testing has been suggested as one of the ways formative assessment can be offered to students, as it provides both the means by which material can be assessed shortly after delivery and also the possibility of timely feedback to both students and teachers [2,3]. Other reported benefits of frequent evaluation include improvement of student performance on final examinations [4-8] and reduction of examination anxiety [2,3].

Computer-based delivery of frequent quizzes is a further modification recently introduced within the medical school curriculum [9,10]. The computer offers the additional advantages of immediate scoring and feedback, thus increasing the effectiveness of formative assessment, as well as increased flexibility and enhanced linking of images and other media [9,10].

We introduced weekly computer-based quizzes into the General Pathology course at Mount Sinai School of Medicine and currently have three years experience with this initiative. Our various objectives reflected both the previously recognized advantages of frequent testing, as well as specific issues within the General Pathology course.

The General Pathology course is a six-week course given in the spring of the first year to approximately 120 students. The course introduces basic pathological concepts, introduced as weekly modules covering Cell Injury, Inflammation and Repair, Hemodynamic Disorders, Immunopathology, and Neoplasia, within a 22-hour schedule utilizing a mixed lecture/small group/laboratory format. The course concludes with a two-hour, in-class, multiple-choice examination. Due to limitations of time, there had been no additional formal testing within the course. An inordinate eighty-five percent of the student's final grade was determined by performance on the final examination, with instructor evaluation of student laboratory/small group performance contributing the remaining 15%.

There were several disadvantages inherent in the course format. The students had no opportunity prior to the end of the course to formally assess their retention and understanding of course material. The students were also concerned about related practical considerations – while a few sample questions from previous examinations were available to the students for review, details on examination question content and level of difficulty were not obvious. Another consequence of the limited course hours was that only a portion of the General Pathology material that the faculty deemed important could be presented by formal didactic instruction. The students were required to acquire the remaining material through their textbook readings – a task that appeared to the faculty to be indifferently performed by many students. An additional issue for the pathology faculty was that General Pathology runs contemporaneously, at least in part, with several other first-year courses, including Immunology, Pathogenesis and Mechanisms of Host Defense, Epidemiology, and the Art and Science of Medicine. The students were inclined to devote the major part of their studying to these larger courses at the expense of General Pathology.

In sum, specific objectives we hoped to address in weekly quizzes included: 1) Provide formative assessment to medical students on material both presented in class and assigned for independent study; 2) Reduce the importance of the single final examination on the student grades; 3) Provide examples of topics, and their level of detail, considered to be important by the faculty; 4) Encourage weekly review of the course material (particularly in view of the competition presented by the other major courses); and 5) Provide image reinforcement.

In addition to the above, we chose to introduce two elements into our computer quiz program that had not been previously described and that we hoped would make it easier to administer the quizzes and improve the learning experience for the students. First, we offered the students self-scheduling of quizzes and the ability to take them independently at any site with Internet access. By this means we could avoid both use of class time for quizzes and the necessity of arranging for Medical School space and personnel to administer the exam. Secondly, we allowed the students to refer to their books during the quiz, which we believed would facilitate directed learning in a less stressful environment.

MATERIALS AND METHODS

One quiz was designed for each of the five major modules of the course, and was rewritten each year. Each quiz consisted of 10-20 case-based questions that were frequently linked to both gross and microscopic images. Our computer system is WebCT™. Each question was created using an html table, which allowed it to be transferred to the exam with its format intact. This allowed images with legends to be included, which is more difficult when done directly within the system. Answer format was multiple-choice (single-best-answer) or matching (Figure 1). General Pathology, as every course at the Mount Sinai School of Medicine, has its own website within the WebCT™ course management system, available on our Intranet. Access by students to the site is secured by login and password. Each quiz was released on the course website following the in-class formal presentation of the week's material, which occurred on Mondays and Tuesdays of each week. The quizzes could be accessed at the student's convenience from any computer – indeed, one student took the quiz while at home in California! Quizzes were available each week for a limited five-day period, from Tuesday evening through Sunday evening. Although not required, the quizzes were emphasized to the students as an important tool for self-study. Quizzes were to be taken independently, but were open-book. During the open period, students were allowed unlimited access to the questions, until such time that they formally submitted their results. Each Monday morning, the quizzes were scored by computer and released to the students with their grades and the correct answers. On Monday evenings, second-year medical student teaching assistants met with the General Pathology students to review the previous module's material and go over questions they might have about the computer quiz from the previous week. A value of 10% of the final grade, or 2% per quiz, was selected to provide a "safety cushion" toward the final grade, yet minimize the harm to the final grade if a student was unable to complete one of the quizzes due to conflicting academic or personal obligations. Each quiz remained available for review for the remainder of the course, even for those students who had not formally taken the quiz.

Group and individual student compliance and performance were analyzed with the WebCT™ toolset. Data analysis provided the discrimination factor for each question, as well as the frequency that each answer was selected, which permitted analysis of the students' thought processes. The duration of time each student had the quiz site open was also documented. Specific questions relating to the computer quizzes were incorporated into the student course evaluations, required at the end of the every course at Mount Sinai School of Medicine, for analysis of student opinions. Quiz and final examination grades were reviewed to determine if the individual student performance on quizzes within a given year correlated with performance on the final examination. Additionally,

the whole class performance on the final examination (means and standard deviations) was analyzed by a 2-sample t-test comparing the two years prior to the introduction of quizzes to the years afterwards.
The Associate Dean of Undergraduate Medical Education, Mount Sinai School of Medicine, submitted and received Institutional Internal Review Board (IRB) waiver to use any student survey data for publication.

RESULTS

Student compliance in taking the quizzes ranged from 95-99%. Results were generated on specific questions and the entire quiz by individual student and class, with the mean grades for individual quizzes ranging from 86 to 98%. We observed that student compliance in taking the test was highest in the first year (99%), and for the subsequent two years dropped a little to and stayed stable at 95%. We believe that the exceptionally high student compliance in the first year reflects the novelty of the computer quizzes in the curriculum. In subsequent years other courses in the first year of the medical school followed the example of General Pathology and introduced computer quizzes. The mean performance on the five quizzes in the first year (94%) was slightly higher than in the following two years (90 and 89%, respectively), probably for similar reasons.

Analysis regarding performance by the upper and lower 25% of the class allowed calculation of a discrimination factor for each question. The results led to several interpretations: 1) when both the upper and lower 25% of the class scored high, the concept was deemed to have been learned by all; 2) when the upper 25% performed well and the lower 25% poorly, the high discrimination factor indicated a concept that was more difficult and worthy of additional emphasis for the weaker students (Table 1); 3) when both the upper and lower 25% scored poorly, the conclusion was either that the topic was inadequately covered by the course and presented by the textbook in a confusing manner, or that the question was poorly written. If it was determined that a question was poorly written, it was discarded and the students were not penalized for incorrect answers.

In the infrequent event that a student elected not to take a quiz, the score for that quiz was recorded as zero, which was added to the rest of the scores to arrive at a final quiz grade. It was very unusual that a student scored less than 75% on any individual quiz. Review of the student performance in this circumstance indicated that the problematic questions were generally based on assigned readings and not on topics discussed in class. When we looked at the time that the student devoted to the quiz, we frequently found that the student spent less than an hour, indicating in our opinion less of an inability to answer specific questions than a disinclination to spend the time reading the appropriate section in the textbook. As noted earlier, following the quizzes students were provided with correct answers and the assistance of teaching assistants (and faculty, if necessary) for clarification of problematic questions.

The timely generation of test results and analysis permitted prompt remedial action as indicated – either by email to the class or in-person by the Course Director at the next lecture. The Course Director noted that an unexplained failure of a student to take more than one quiz often correlated with academic or personal difficulties that potentially dictated various levels of intervention, ranging from contact by email or personal conference with the Course Director, to involvement of the medical school deans.

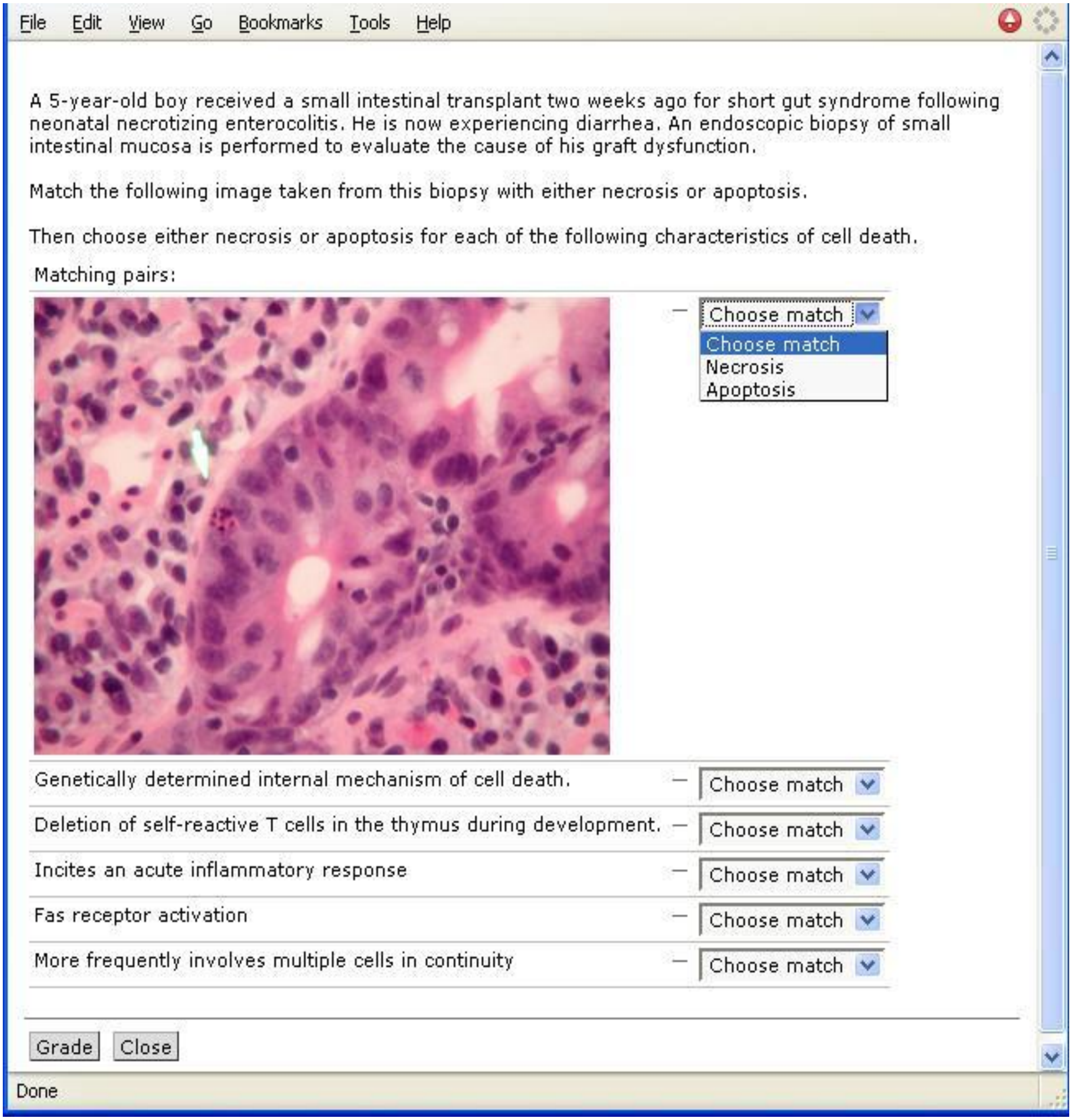
We followed two different methods of evaluating student feedback. In the first two years that we offered computer quizzes, we asked two relatively general questions: "Is it worthwhile to continue the quizzes as a means of teaching and reinforcing selected course material?", and "Is it worthwhile to continue the quizzes as a percentage of your final grade?" Affirmative answers to both questions averaged 90% of the class, and we did not find that there was any significant change from the first year to the second year. In the third year of our program, we posed different and more specific questions to the students and utilized a Likert score in an attempt to generate a more detailed analysis (Table 2). As demonstrated, responses were generally favorable. The strongest level of agreement was seen for the role of quizzes in reinforcing understanding of topics for study, including both those formally presented in class and those intended as independent learning. Surprisingly, the value of quizzes in decreasing anxiety for the final examination, either by giving examples of questions or by providing partial credit toward the final course grade, was deemed the least important justification. In free text many students volunteered that the quizzes forced them to keep up with the course material. The teaching assistants reported that attendance was low at their weekly reviews of the quizzes; most students volunteered that the feedback they had received on incorrect answers was sufficiently clear to make additional review unnecessary.

The mean and standard deviation of the student performance on the final examination for the years prior to the introduction of the weekly quiz program were 86.5 and 5.90, respectively, and for the years after were 89.8 and 5.45. By a 2-sample t-test the improvement in the mean grade was significant at a p value of <0.001. Correlation between quiz grades and performance on the final examination for individual students could not be assessed, since the grade spreads on both the quizzes and final examination were too narrow to permit analysis.

DISCUSSION

The use of frequent quizzes has been recommended for a variety of didactic purposes that relate in various ways to formative assessment. In the literature, the effect of

Figure 1. Question from the Quiz on Cell Injury. This matching question was designed for reinforcement of material presented in the class and the textbook. It includes a microscopic image for identification.



frequent testing on final grades has been generally beneficial [4, 5, 7, 8], although not uniformly so [6, 11-13]. However, other educational objectives of frequent testing have been identified, including providing early identification of individual students with problems or more generally problematic course material requiring clarification [2,3], reducing student anxiety about examinations by encouraging regular studying and providing familiarity with types of questions asked [2], and generally engendering an improvement in student approaches and attitudes toward learning by encouraging a consistent pattern of studying [4,5]. Furthermore, in several studies students commonly favored frequent quizzes, which they claimed help them keep up with work [3,13], and were more likely to favorably rate courses offering frequent testing [4,7,10].

The computer is recognized as a powerful educational tool, and has been utilized in innumerable on-line didactic exercises and tutorials. Within the last ten years computer-based quizzes have been introduced into medical school courses at a few institutions [9,10]. These institutions observe that computer-based quizzes have certain advantages over paper-based testing, including flexibility of delivery that allows closer integration with instructional material, ease of incorporating images and other media into questions, providing immediate feedback, and facilitation of item banking. The programs of computer-based quizzes described in the literature offer mastery type of quizzes at special supervised computer testing centers at specific times that are arranged in advance.

Table 1. Example of Quiz Data Analysis Evaluating Results for All Students Who Took the Quiz on Inflammation & Repair. Specific questions are identified by keywords in the left column. The highlighted question on mechanisms of pleuritis shows a low mean class grade and a high discrimination factor between the upper and lower 25% of the class, which indicates the need for additional review of the topic. Prompt computer generation of data facilitates timely addressing of the problem.

Summary: Inflammation and Repair

[Submissions] [Detail] [Reports] [Summary]

Add to question statistics

Title	N	% Correct Of:			Discrimination	Score	
		Whole Group	Upper 25%	Lower 25%		Mean	SD
chf	112	100	100	100	0.00	100.0%	0.0
repair	112	83	100	68	0.51	91.3%	20.9
myocardium	112	96	100	92	0.26	96.4%	18.6
pleuritis	112	36	100	13	0.66	47.9%	41.9
mediators	112	96	100	94	0.20	98.8%	7.2
complement	112	97	100	100	-0.12	97.3%	16.2
cellular events	112	93	100	86	0.36	98.0%	8.5
					Overall Mean:	89.3%	

Our computer-based tests differed from these programs by having an open-book format and open scheduling. Students were free to access the quiz at any time (within a five-day "window") and from any location with an Internet connection, providing scheduling flexibility to the students, eliminating the necessity for proctored supervision, and conserving in-class hours for instruction. The open-book format was selected to encourage reading and study in a stress-free setting. Specific questions relating to weekly reading assignments encouraged review of the material, which could be done during the course of the open-book examination. Inclusion of such subject matter, which had not been presented formally in course contact sessions, encouraged the self-directed study that is implicit in the current emphasis on lifelong learning. The students were instructed to work independently. Some of the faculty expressed skepticism that the students would refrain from collaborating on answers. While this was a risk, there were several reasons that collaboration was not likely: 1) the quizzes were "low stakes"; 2) Mount Sinai School of Medicine operates on an Honor System; and 3) the students knew that images and questions from the quizzes, in shuffled context, would appear on the final examination. Prompt feedback heightened the value of the quizzes for formative assessment of the students. Individual students had the opportunity to clarify promptly areas of confusion with teaching assistants or with the course faculty. For the instructors, immediate analysis of student answers also permitted timely intervention on occasional topics shown to be problematic to the class. The computer-based format also provided the opportunity to introduce and reinforce images, which is particularly desirable in a Pathology course. While the Course Director anticipated identifying students in academic trouble following poor performance on the quizzes, an unexpected result of the process was identification of individuals with personal difficulties who had repeatedly failed to take the quizzes. Student compliance was high and feedback was favorable. In evaluation surveys, students specifically praised the quizzes for highlighting topics deemed important in the course and providing them with a means for regular self-assessment and motivation to keep up. We were unable to correlate the individual student performance on quizzes with their performance on the final examination, due to generally high grades. However, we noted a statistically significant improvement in the mean class grades on the final examination in the years following introduction of the computer quiz program.

Based on our experience with the computer quizzes, we made several changes in subsequent years to facilitate smooth functioning of the program. We conducted more rigorous pre-quiz review to eliminate ambiguous questions. A set of detailed instructions on how to access the quiz and submit answers was added as an introduction to the first computer quiz. We were able to anticipate and warn students about procedural difficulties, such as the inability to access quizzes if the students had blocked "pop-ups" on their computers. The option of make-up quizzes was eliminated as it was too disruptive to the timetable of the quizzes. (We found that the students accepted this policy if it was explained in advance.)

Table 2. Student Responses to Computer Quizzes. Students were asked to grade their responses to specific questions on a 5-point Likert scale (1 being "strongly disagree", and 5 being "strongly agree").

Question	Score
"The computer-based quizzes emphasized some of the important key concepts required for the course"	4.34
"The computer-based quizzes encouraged me to read about topics not covered in class"	4.19
"The computer-based quizzes helped indicate my level of understanding of the course material that was addressed in the quizzes"	4.02
"The computer-based quizzes decreased my level of anxiety for the final exam by giving me examples of exam questions"	3.97
"The computer-based quizzes decreased my level of anxiety for the final exam by counting for 10% of my final grade"	3.64

CONCLUSIONS

By both student and faculty assessment, weekly computer-based quizzes provided valuable formative assessment in General Pathology. We found that the two new features we introduced into our computer quiz program enhanced its usefulness for both faculty and students. The introduction of independent self-scheduling of quizzes conserved in-class hours for didactics and eliminated the need for faculty supervision and a designated testing site. Open-book format encouraged directed weekly review of course material, including topics to be learned by independent reading assignment, in a stress-free environment. Favorable student evaluations demonstrate that computer-based open-book quizzes provide enriched curriculum support and student satisfaction. Mean class performance on the course final examination showed a statistically significant improvement following introduction of the weekly computer quizzes, providing an objective measure of its beneficial effect on learning.

Acknowledgment: The authors thank John Doucette, PhD., for assistance in statistical analysis.

REFERENCES

1. Rushton, A. Formative assessment: a key to deep learning? *Medical Teacher*. 2005; 27(6): 509–513.
2. Sporer, R. The no-fault quiz. *College Teaching*. 2001; 49(2): 61.
3. Robins, L.S., Fantone, J.C., Oh, M.S., Alexander, G.L., Shlafer, M., and Davis, W.K. The effect of pass/fail grading and weekly quizzes on first-year students' performances and satisfaction. *Academic Medicine*. 1995; 70(4): 327-329.
4. Geist, J.R., and Soehren, S.E. The effect of frequent quizzes on short- and long-term academic performance. *Journal of Dental Education*. 1997; 61(4): 339-345.
5. Duty. R.C. Weekly or biweekly quizzes in organic chemistry: Does it make a difference? *Journal of Chemical Education*. 1982; 59(3): 518-519.
6. Freilich, M.B. Frequent quizzing, the final exam, and learning. *Journal of Chemical Education.* 1989; 66(3): 219-223.
7. Bangert-Drowns, R.L., Kulik, J.A., and Kulik, C-L.C. Effects of frequent classroom testing. *Journal of Educational Research*. 1991; 85(2): 89-99.
8. Negin, G.A. The effects of test frequency in a first-year torts course. *Journal of Legal Education*. 1981; 31(3-5): 673-676.
9. Peterson, M.W., Gordon, J., Elliott, S., and Kreiter, C. Computer-based testing: Initial report of extensive use in a medical school curriculum. *Teaching and Learning in Medicine*. 2004; 16(1): 51-59.
10. Miller, A.P., Haden, P., Schwartz, P.L., and Loten, E.G. Pilot Studies of in-course assessment for a revised medical curriculum: II. Computer-based, individual. *Academic Medicine*. 1997; 72(12): 1113-1115.
11. Olsen, R.E., Weber, L.J., and Dorner, J.L. Quizzes as teaching aids. *Journal of Medical Education*. 1968; 43(8): 941-942.
12. Beaulieu, R.P., and Utecht, K.M. Weekly administered formative tests of a modified format and student performance. *Journal of Instructional Psychology*. 1990; 17(2): 55-58.
13. Haberyan, K.A. Do weekly quizzes improve student performance on general biology exams? *The American Biology Teacher*. 2003; 65(2): 110-114.

Test-taking Behaviors on a Multiple-Choice Exam are Associated with Performance on the Exam and with Learning Style

John A. McNulty, Ph.D.[1], Beth Sonntag, M.Ad.Ed.[2], James M. Sinacore, Ph.D.[3]

[1]Department of Cell Biology, Neurobiology and Anatomy
[2]Teaching and Learning Center
[3]Department of Preventive Medicine and Epidemiology
Loyola University Chicago
Stritch School of Medicine
Maywood, IL 60153

Phone: (+)1-708-216-5161 Fax: (+)1-708-216-3676 Email: jmcnulty@lumc.edu

ABSTRACT

The goal of the study was three-fold: to investigate medical student behaviors (e.g., changing answers) while taking high-stakes multiple-choice exams; to determine if specific behaviors were associated with performance on the exam; and to determine if there are associations of learning style, as measured by Kolb's Learning Style Inventory (LSI), with test-taking activities. We developed high-stakes, on-line exam applications that included server event logs, which provided a time sequence of entries/activities that students made while taking their exam. This new paradigm allows collection of detailed test taking behaviors that can be used to test a variety of hypotheses. Test taking activities were extracted from the event logs for a mid-term anatomy exam given to freshman medical students. Although student exam-taking activities showed considerable variability, one notable finding was that when students changed answers, they were 3 times more likely (on average) to change their answer from incorrect-to-correct than they were to change it from correct-to-incorrect. Correlation of test-taking behaviors with performance on the exam revealed that there were significant negative correlations with the number of times answers were changed, and the number of times answers were changed from correct-to-incorrect. There were also significant associations of learning styles with exam-taking behaviors. The most consistent differences between students who did the best on the exam (Assimilators) and those who did the worst (Accommodators), concerned the frequencies with which answers were changed. Differences between Accommodators and Convergers, who received the next highest average score, involved behaviors related to variables other than changing answers (e.g., time spent reviewing the exam and number of questions marked for review). In summary, the use of objective computer entry logs allowed a better understanding of the associations of test-taking behaviors with academic performances and with learning styles. Based on these findings, learning strategies might be designed to help students cope with courses that rely heavily on multiple-choice exams for assessing student achievement.

INTRODUCTION

Computer technology is increasingly crucial to the training of physicians (cf. [1]). One area where computers are rapidly impacting training at all levels is their use in assessments, especially those involving multiple-choice exams. Although student attitudes to computerized testing are generally positive [2], this new testing environment has raised concerns such as effects of computer anxiety and attitudes, sufficiency of resources, and reliability as compared to paper-and-pencil multiple-choice tests [3-7].

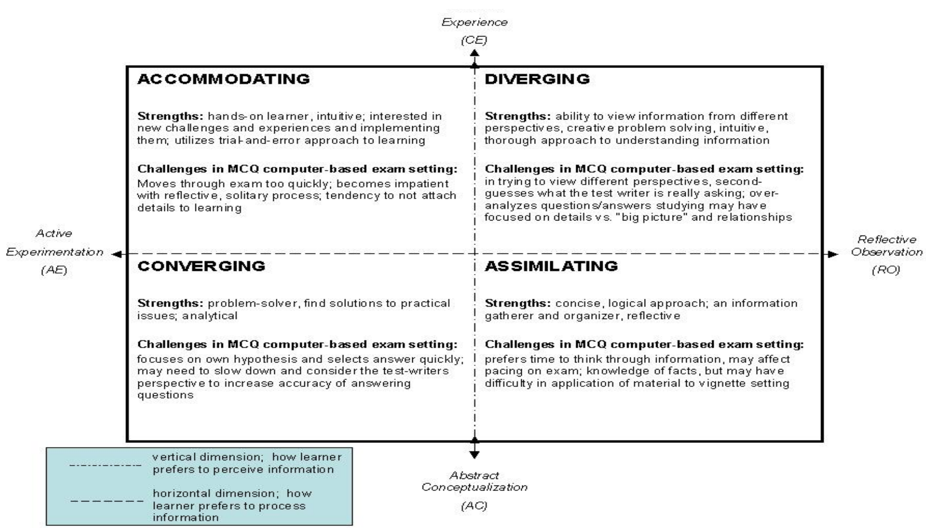
Several studies have demonstrated associations between learning style and academic performance across the professional spectrum of medical training [8-12] as well as in other disciplines [13, 14]. More specifically, Lynch et al. [9] showed that medical students with a learning style preference of abstract conceptualization tended to perform better on the multiple-choice United States Medical Licensing exam (USMLE). This association between learning style and performance on multiple-choice exams requires a better understanding because of the broad implications both to the individual student progress and assessment in the medical curriculum. As computerized exams become more commonplace in the medical curriculum, possible effects of learning style on test-taking behaviors and performance on exams are more easily studied because of the ease with which objective data can be collected unobtrusively.

We undertook the present study as part of an initiative at the Stritch School of Medicine to develop web-based applications for delivery of high-stakes exams throughout the 4-year curriculum. The applications include server event logs of individual student actions while taking exams. These event logs were used to determine the association of test-taking behaviors with performance on the exam and with learning style as measured by Kolb's Learning Style Inventory (LSI) [15].

Kolb's theory of learning style consists of four stages in the cycle of learning: concrete experience (CE), reflective observation (RO), abstract conceptualization (AC) and active experimentation (AE). These four stages represent two dimensions. CE and AC represent the vertical dimension of perceiving information. RO and AE represent the horizontal dimension of processing information. These two dimensions create four modes or styles of learning (Figure 1). The completed Learning Style Inventory (LSI) provides a score for each stage in the learning cycle. By calculating the results of the LSI and applying the scores to the Learning Style Type Grid, a learning style is defined based on each individual's preference for how they perceive (CE-AC=perceiving score) and process information (RO-AE = processing score); thus, designating the quadrant that defines their learning style. The four modes or learning styles result from these dimensions: Diverging (CE/RO); Assimilating (RO/AC); Converging (AC/AE) Accommodating (AE/CE) (Figure 1).

The purpose of our study was three-fold: to investigate

Figure 1. Descriptions of each of the four learning styles based on Kolb's Experiential Learning Model, which reflects learners' preferences in perceiving and processing information. The descriptions include strengths of each of the learning styles, challenges each learning style faces with multiple choice question (MCQ) exams, and suggested strategies learners could apply to further develop their test-taking skills. (Adapted from Kolb DA. Experience as the Source of Learning and Development. Englewood Cliffs, NJ: Prentice Hall, 1984:42, with permission.)



medical student behaviors (e.g., changing answers) while taking high-stakes multiple-choice exams; to determine if specific behaviors were associated with performance on the exam; and to determine if there are associations of learning style, as measured by Kolb's Learning Style Inventory (LSI), with test-taking behaviors.

MATERIALS AND METHODS

Administration of Kolb's Learning Style Inventory: The paper-based version of Kolb's LSI (Version 3) was administered to the first year medical class (n=137) during orientation week. An overview of Kolb's Model of Experiential Learning was presented to the students, followed by their completion and scoring of the inventory. The assessments were collected and calculations were verified and recorded. Additional small group workshops were given to the students, which provided a more in-depth discussion of learning style preferences, strengths and challenges. Administration of the LSI was coordinated through the Teaching and Learning Center at the Stritch School of Medicine (SSOM). The distribution of learning styles was as follows: Convergers (n=64); Accommodators (n=19); Divergers (n=15); and Assimilators (n=39).

On-Line Exam Application: The web-based exam database and applications were developed as part of the Loyola University Medical Education Network (LUMEN) to provide on-line exams. The exam interface (Figure 2) allows students to cross out answers, add notes, and submit answers for later review in the event they are unsure of the answer. The exam applications were constructed utilizing Allaire's ColdFusion (v4.5) for middle-tier application development and Microsoft SQL Server (v7.0) for database services. Secure portal access is accomplished through the "myLUMEN" student portal, which uses Secure Socket Layer (SSL v3.0) with 128-bit digital certificates provided by Versign.

Administration of On-Line Exam: The exam analyzed for this study was the second midterm comprising 75 multiple-choice questions (see Figure 2 for an example of the questions). It was given to students (n=137) in the first year medical anatomy course. Students were assigned to specific computer stations in the learning laboratories and they accessed the exam through their secure student portal. They were given 2 hours to complete the exam.

Data Collection and Analysis: Server event logs for each student were exported into an Excel database. These logs included a timeline (by seconds) for all mouse click and keystroke activities from the time each student opened the exam to the time he/she clicked the "Finalize Exam" button. Each of the variables analyzed in this study (See Table 1) was manually extracted from the event logs. The total time spent taking the exam was the time from release of the exam to the time each student finalized his/her exam. The time spent reviewing the exam was estimated by subtracting the time when the last question was viewed from the total time of the exam. All data were entered into an Excel spreadsheet and the names of students deleted prior to further analyses in order to preserve anonymity. The data were analyzed using analysis of variance, Pearson correlation coefficients (after showing that the data were normally distributed), and the Kendall Coefficient of Concordance [16]. Given that eight tests of significance each needed to be conducted with the analysis of variance and Pearson coefficients, a Bonferroni adjusted p value of .0063 (.05/8) was used. No adjusted p value was used with the Kendall coefficients because only two such coefficients were computed.

The study was approved by the IRB.

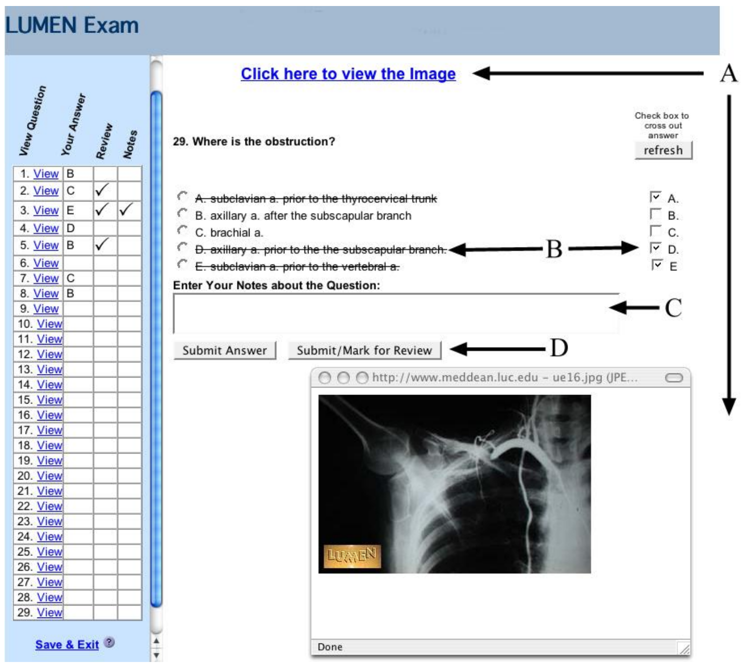
RESULTS

Student behaviors while taking a multiple-choice exam. Table 1 illustrates the considerable variability in the activities of students while taking the exam. For instance, the total number of times students submitted answers for the 75 questions ranged from 76 to 266. The number of questions marked for review ranged between 0 and 67, and students spent an average of about 30 minutes reviewing their exam. The maximum number of answers changed was 21 (out of 75), with 99% (135/137) changing 1 or more answers. When students changed answers, they were 3 times more likely (on average) to change their answer from incorrect-to-correct than they were to change it from correct-to-incorrect.

Are exam-taking behaviors associated with performance on the exam? Pearson's correlations were used to evaluate whether any of the variables measured were associated with performance on the exam. There were no significant correlations with time spent on the exam, time spent reviewing the exam, the number of answers submitted, nor the number of questions marked for review (Table 2). However, there were significant negative correlations with the number of times answers were changed, and the number of times answers were changed from correct-to-incorrect. Interestingly, there was no correlation between grade and the number of times answers were changed from incorrect-to-correct.

Are exam-taking behaviors associated with learning style? We were next interested in determining whether learning styles were associated with differences in exam activities. Results from the analysis of variance did not reveal any association of the Kolb learning styles when the data were sorted according to Kolb's four learning styles with the variables listed in Tables 1-3 (p's ranged from .268 to .842; data not shown). However, a 5% difference in grade between Assimilators (86.1) and Accommodators (81.1) as seen in Table 3 prompted us to consider the possibility that a combination of behaviors accounted for the differences in performance on the exam. Trends in the data were evident when the group averages were ranked (Table 3). The most consistent differences between students who did the best on the exam (Assimilators) and those who did the worst (Accommodators), concerned the

Figure 2. The exam interface showing links for opening multimedia in a separate window that can be moved (A). The students have options to cross out answers (B) and a text box for notes (C). If text is entered a checkmark appears next to the question in the list of questions to the left. Students can submit answers for later review (D), which then marks that question in the list of questions to the left.



frequencies with which answers were changed from incorrect-to-correct and vice-versa (Table 3). Differences between Accommodators and Convergers, who received the next highest average score, involved behaviors related to variables other than changing answers (e.g., time spent reviewing the exam and number of questions marked for review). Associations of learning style with test taking behaviors were tested using the Kendall Coefficient of Concordance test [16] using the rank orders. Learning styles were found to be significantly associated with answer changing (W = .91, χ^2 = 8.20, df = 3, p = .042) as well as with behaviors other than answer changing (W = .68, χ^2 = 10.20, df = 3, p = .017).

DISCUSSION

Our first goal to study student behaviors while taking an on-line exam illustrates an important advantage of server-based applications to collect objective data unobtrusively. A number of basic statistics can be useful for administration of future exams. For instance, knowing the average amount of time that students spend on individual questions can be used to better judge an appropriate time for the whole exam. Additionally, determining the frequency with which individual questions are marked for review can assist course directors and faculty in evaluating the appropriateness of the questions (e.g., is the question

Table 1. Statistics for each of the variables for the whole class.

	Mean (±sem)	Range
Time spent taking exam (min)	104.8 (±1.4)	43 - 120
Time spent reviewing exam (min)	32.1 (±1.5)	1 - 82
Number of submitted answers	120 (±3.5)	76 - 266
Number of questions marked for review	12.9 (±1.1)	0 - 67
Number of times answers were changed	6.0 (±0.3)	0 - 21
Number of answers changed incorrect-to-correct	3.4 (±0.2)	0 - 11
Number of answers changed correct-to-incorrect	1.1 (±0.1)	0 - 5
Number of answers changed incorrect-to-incorrect	1.4 (±0.1)	0 - 10
Grade (%)	84.8 (±0.7)	60-100

Table 2. Correlations of variables with grade (n=137). Significant p values were Bonferroni adjusted.

	Pearson Correlation	P value
Time spent taking exam (min)	.069	
Time spent reviewing exam (min)	.128	
Number of submitted answers	.147	
Number of questions marked for review	-.086	
Number of times answers were changed	-.245	p<0.004
Number of answers changed incorrect-to-correct	-.012	
Number of answers changed correct-to-incorrect	-.238	p<0.005
Number of answers changed incorrect-to-incorrect	-.407	p<0.0005

ambiguous).

Our finding that students who change answers are more likely to change them from incorrect-to-correct is consistent with several other reports [17-20]. This tendency was more recently confirmed in a study using midterm results from an undergraduate psychology course where tabulations of erasure marks showed that 51% of changed answers were from wrong to right outnumbering changes from right to wrong by a factor of 2-to-1 [21]. The authors further showed that 54% of the students improved their grades by changing answers. Even more impressive were the present findings that 85% of the medical students improved their grade by changing answers; 5% made their grade worse; while the remaining 10% had no net change in grade.

A second goal of our study was to determine if specific behaviors were associated with performance on the exam. Although changing answers from wrong to right improved the score for most students in the present study, this activity was not correlated with grade on the exam. In fact, students in the lower quartile changed, on average, a slightly higher number of answers from incorrect to correct compared to students in the upper quartile (data not shown). In contrast, there were significant negative associations between grade and the frequency with which students changed their answers from right to wrong and from wrong to wrong. The overall increase in number of changed answers from those students who received a lower grade most likely reflects a general lack of confidence in their knowledge. Their greater tendency to change wrong answers to other wrong answers is certainly consistent with being unsure of their knowledge while taking the exam.

A third goal of the study was to examine if there were associations of exam-taking behaviors with specific learning styles. The wide variability in the behaviors associated with taking exams was not entirely unexpected in view of similarly large variability in medical student use of computer aided instructional resources [22]. Many of the factors explaining this variability have not been determined, however some possible associations with personality/learning styles have been investigated [13, 23, 24]. Results from the present study suggest that learning styles

Table 3. Rankings of group averages for each of the Kolb's Learning styles (4 = highest average). High and low rankings are bolded to highlight the trends in two separate groupings; those involving changing answers (below double line) and those involving all other behaviors (above double line). Average grades are shown in the table only as a reference for the rankings.

	Con	Acc	Div	Asm
Time spent taking exam (min)	3	**1**	2	4
Time spent reviewing exam (min)	**4**	**1**	3	2
Number of submitted answers	**4**	**1**	2	3
Number of questions marked for review	**4**	**1**	3	2
Number of times answers were changed	**4**	**2**	3	1
Number of answers changed incorrect-to-correct	2	**1**	3	**4**
Least number of answers changed correct-to-incorrect	3	**1**	2	**4**
Least number of answers changed incorrect-to-incorrect	2	**1**	3	**4**
Grade (%, $\pm$standard deviation)	85.5 $\pm$8.7	81.1 $\pm$7.7	83.0 $\pm$8.9	86.1 $\pm$7.8

are associated with a combination of behaviors while taking multiple-choice exams. More specifically, Accommodators, who received the lowest average score on the exam, took the least amount of time for the exam, spent the least amount of time reviewing the exam, spent the least amount of time answering questions, had the fewest number of submitted answers, marked the fewest questions for review, changed the fewest answers from wrong-to-right, and changed the most answers from right-to-wrong.

These findings are consistent with the Accommodating style, which reflects a learning preference for action and implementation (active experimentation - AE); thus, reducing the learner's tendencies toward reflection and review (reflective observation - RO). Accommodators utilize an intuitive, trial-and-error approach to solving problems (concrete experience - CE) and in new learning situations, tend to rely heavily on others for information and details rather than their own analytic ability (abstract conceptualization - AC). Therefore, the stationary, non-interactive environment of the standardized exam setting is in direct conflict with the strengths that optimize and reflect the learning preferences of the Accommodating style.

Our study showed that Assimilators and Convergers had the best scores on the exam, which is consistent with other research showing that students with a preference for abstract conceptualization (AC) tended to perform better on multiple-choice exams [9, 10]. The present results extend these findings by providing insights into specific activities and behaviors that are associated with performance on exams. Acquiring a better understanding of the association of learning styles with academic performances is important for those courses that rely heavily on multiple-choice exams for assessing student achievement. For the assessor, it can contribute to creation of an assessment tool that reflects a balance of diverse questions to challenge all learning styles in content and process. Probably more importantly, for the medical student, it can help them adapt their learning style preferences to optimize their test-taking strategies. Learning strategies might be designed to help students cope with a variety of examination settings. For example, Accommodators preparing for multiple-choice exams could practice organizing and analyzing their own information, identify patterns, build conceptual models, slow down and reflect before action. By utilizing an awareness of learning styles in correlation with exam behaviors and results, students can better prepare for the ongoing multiple-choice exams that the health profession requires. It is important for learners to not only set goals for development per their current learning style (see Figure 1), but to increase their awareness of strengths and challenges of other learning styles.

A noteworthy limitation of this study was sample size. We found that group sizes for the four Kolb learning styles only provided an average of 22% power for achieving statistical significance with analysis of variance at the .05 level given the magnitude of differences that were observed in the data. As the Kolb's LSI is completed by future classes, data will be combined for analysis.

Having only a single examination was another limitation of our study and we are working to develop report applications, which will allow more efficient extraction of data from the event logs. When these are complete, we plan to extend our research to address a number of important questions that arise from this study. First, will intervention influence a student's behavior on subsequent exams? Learning assistance offered to students having academic difficulties could include analyses of their activities during the exam, which may provide helpful insights into their test-taking skills allowing them to modify behaviors that are associated with poor performance as discussed above (see Figure 1). A second important question is whether individual exam behaviors are consistent from exam to exam and from course to course? The data collected for this study were from an exam in gross anatomy, which emphasizes factual knowledge. Would similar behaviors persist in a course requiring more conceptualizations? Based on significant correlations of individual performances from exam to exam in the anatomy course (unpublished observations), we predict that exam-taking behaviors are also consistent on exams within a course. Finally, to what degree do test-taking behaviors change over the curriculum and are these changes associated with changes in learning style that are known to occur with medical training [25-27]?

ACKNOWLEDGEMENTS

We gratefully acknowledge the following individuals who facilitated development and implementation of the applications for the LUMEN on-line exams (alphabetical order): Dr. A. Chandrasekhar, Dr. B. Espiritu, E. Fabro, G. Klitz, A. Hoyt, R. Naheedy, and R. Price.

REFERENCES

1. Seago, B.L., Schlesinger, J.B., and Hampton, C.L. Using a decade of data on medical student computer literacy for strategic planning. *Journal of Medical Library Associations.* 2002; 90(2):202-9.
2. Ogilvie, R.W., Trusk, T.C., and Blue, A.V. Student's attitudes towards computer testing in a basic science course. *Medical Education.* 1999; 33:828-31.
3. Clariana, R., and Wallace, P. Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal Educational Technology.* 2002; 33(5):593-602.
4. Hahne, A.K., Benndorf, R., Frey, P., and Herzig, S. Attitude towards computer-based learning: determinants as revealed by a controlled interventional study. *Medical Education.* 2005; 39:935-43.
5. Lee, G., and Weerakoon, P. The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based

and paper-and-pen multiple-choice tests. *Medical Teaching*. 2001; 23(2):152-7.
6. Kreiter, C., Peterson., M.W., Ferguson, K., and Elliott, S. The effects of testing in shifts on a clinical in-course computerized exam. *Medical Education.* 2003; 37:202-4.
7. Rozell, E.J., and Gardner, W.L.I. Cognitive, motivation, and affective processes associated with computer-related performance: a path analysis. *Comparative Human Behavior.* 2000; 16:199-2000.
8. Markert, R.J. Learning style and medical students' performance on objective examinations. *Perception Motor Skills.* 1986; 62(3):781-2.
9. Lynch, T.G., Woelfl, N.N., Steele, D.J., and Hanssen, C.S. Learning style influences student examination performance. *American Journal Surgery*. 1998; 176(1):62-6.
10. Newland, J.R., and Woelfl, N.N. Learning style inventory and academic performance of students in general pathology. *Bulletin Pathology Education*. 1992; 17:77-81.
11. Chapman, D.M., and Calhoun, J.G. Validation of learning style measures: implications for medical education practice. Medical Education. 2006; 40:576-83.
12. Smits, P.B.A., Verbeek, J.H.A.M., Nauta, M.C.E., Ten Cate, T.J., Metz, J.C.M., and Van Dijk, F.J.H. Factors predictive of successful learning in postgraduate medical education. *Medical Education*. 2004; 38:758-66.
13. Cordell, B.J. A study of learning styles and computer-assisted instruction. *Comparative Education.* 1991; 16(2):175-83.
14. Hamkins, G. Motivation and individual learning styles. Engineering Education. 1974; 64:408-11.
15. Kolb, DA. Experiential Learning: Experience as the Source of Learning and Development. *Englewood Cliffs, N.J.: Prentice-Hall, Inc.* 1984.
16. Siegel, S., and Castellan, N.J. Nonparametric statistics for the behavioral sciences. 2nd ed. *New York: McGraw-Hill.* 1988.
17. Ferguson, K.J., Kreiter, C.D., Peterson, M.W., Rowat, J.A., and Elliott, S.T. Is that your final answer? Relationship of changed answers to overall performance on a computer-based medical school course examination. *Teaching and Learning in Medicine.* 2002; 14(1):20-3.
18. Fischer, M.R., Herrmann, S., and Kopp, V. Answering multiple-choice questions in high-stakes medical examinations. *Medical Education*. 2005; 39:890-4.
19. Range, L.M., Anderson, H.N., and Wesley, A.L. Personality correlates of multiple choice answer-changing patterns. *Psychology Reports.* 1982; 51:523-7.
20. Vispoel, W.P. Reviewing and changing answers on computer-adaptive and self-adaptive vocabulary tests. *Journal of Educational Measures*. 1998; 35:328-45.
21. Kruger, J., Wirtz, D., and Miller, D.T. Counterfactual thinking and the first instinct fallacy. *Journal of Personality and Social Psychology*. 2005; 88(5):725-35.
22. McNulty, J.A., Halama, J., Daudzvardis, M.F., and Espiritu, B. Evaluation of web-based computer-aided instruction in a basic science course. *Academic Medicine.* 2000; 75(1):59-65.
23. McNulty, J.A., Espiritu, B., and Halsey, M. Medical student use of computers is correlated with personality. *Journal of the International Association of Medical Science Educators.* 2002; 12:9-13.
24. McNulty, J.A., Espiritu, B., Halsey, M., and Mendez, M. Personality preference influences medical student use of specific computer-aided instruction (CAI). *BMC Medical Education.* 2006; 6:7.
25. Lindemann, R., Duek, J.L., and Wilkerson, L. A comparison of changes in dental students' and medical students' approaches to learning during professional training. *European Journal of Dental Education*. 2001; 5(4):162-7.
26. Newble, D.I., and Gordon, M.I. The learning style of medical students. *Medical Education*. 1985; 19(1):3-8.
27. Slotnick, H.B. How doctors learn: education and learning across the medical-school-to-practice trajectory. *Academic Medicine.* 2001; 76:1013-26.

Creating a Functional and Adaptable Web-based Atlas for Medical Students: Implementation of Database Technology

Benjamin P. Rosenbaum[1], Sanjay G. Patel[1], H. Wayne Lambert, Ph.D.[2, 3]

[1]Vanderbilt University School of Medicine
Nashville, TN 37232.
[2]Departments of Cell and Developmental Biology and Cancer Biology
Vanderbilt University School of Medicine
Nashville, TN 37232.
[3]University of Louisville School of Medicine
Department of Anatomical Sciences and Neurobiology
Louisville, KY 40292

Phone: (+)1-502-852-2024 Fax: (+)1-502-852-6228 Email: wayne.lambert@louisville.edu

ABSTRACT

At Vanderbilt University, an interactive, Web-based histology atlas was transformed into a dynamic database-driven tool through utilization of a MySQL® database and the PHP scripting language. The advantages of this transformation were substantial. For the faculty, data associated with the histological images were manipulated quickly and easily to affect the entire website without the involvement of a technology expert. Features such as practice exams, keyword searches, slide lists, and site usage analysis took minutes to develop after the conversion to a database-driven atlas. For the medical students, the database-driven interactive atlas could quickly adapt to meet their educational needs. The implementation of a database-driven histology atlas was well-received by students and faculty due to the ease of data entry, the increased functionality, and the unlimited potential to adapt to meet the pedagogical demands of students.

INTRODUCTION

During the summer of 2004, the construction of an interactive, Web-based histology atlas began in the histology course at the Vanderbilt University School of Medicine (http://www.mc.vanderbilt.edu/histology/atlas/ chapters/example; Figure 1). With only two faculty members responsible for teaching 106 medical students, the development of an interactive, Web-based histology atlas addressed the specific needs of the Vanderbilt students. This approach to an educational adjunct was necessary because of the lack of course-specific software packages and the previous effective use of computer-aided instruction (CAI) in fields that emphasize visual learning, such as anatomy, histology, and neuroanatomy.[1-9] In addition, evidence showed that CAI in the microscope laboratory increases student confidence in structure identification, student-to-student interactivity, and the quality of questions students posed to faculty, which ultimately leads to higher student satisfaction in histology courses.[10-12] Furthermore, CAI tools with sound pedagogy have been able to promote effective learning while helping faculty to overcome the shortage of personnel support in laboratory sessions and reductions in contact hours.[1,13-15]

In anticipation of launching the atlas in January 2005, it was determined that it would be strengthened by implementation of a database structure. The development of the database structure would allow the technological components of the atlas to be decoupled from the data management aspects, which would enable faculty members to act as the data or content experts. The faculty would be able to develop the tool through assistance of a technology expert according to the needs of the students and faculty. In general, technology experts can design tools so that the content expert is able to add, edit, and delete the data associated with the tool in any manner desired. If designed properly, faculty can manage the data effectively for the lifetime of the tool without the future involvement of the technological expert. Therefore, the database structure

Figure 1. Example interactive histology atlas Web pages

was implemented to help to maximize the capability, functionality, and adaptability of a CAI tool while saving the faculty time in construction and editing.

Medical educators benefit from Web-based, database-driven search engines, such as Google™ or PubMed, but they have failed to harness database technology to improve CAI applications. Our goal was to implement database technology to improve an interactive, Web-based histology atlas.

MATERIALS AND METHODS

Atlas Functionality

The atlas is organized by course chapter. Each chapter's laboratory manual text is online and linked to histological structures shown on a separate image Web page. Image Web pages display high resolution (150 ppi) digital images of referenced histological specimens and a brief description about relevant structures contained within the field of view. Links are embedded in text descriptions and when these colored text links are "rolled over" with a computer mouse, the structure of interest is highlighted. This highlighted structure of interest is replaced with the original, non-highlighted, photo when the mouse "rolls off" the link. The images are preloaded so that the speed at which the transition occurs gives users the impression that the structure "lights up" when the mouse rolls over the active text and the structure "lights off" when the mouse rolls off the active text. Additionally, the glass slide number referenced in the laboratory manual, description of specimen, animal model, embedding medium, and magnification of the specimen are listed on each image page to provide students with more information about the image.

Data Handling

The atlas utilizes a MySQL® database and PHP as the Web page programming language. The data presented in the static version of the interactive atlas were analyzed to create the table structure necessary to store the information in an intuitive manner. The core of the interactive histology atlas utilized a small number of relational database tables that enabled the creation of several powerful dynamic tools. In such a relational scheme, there was only one table entry for each item of interest (e.g. course chapter or microscope slide). Therefore, the data were consistent and easily edited to affect the entire site.

Data Extracting

Once the structure of the database tables was established, a PHP script was written that extracted all of the data present in the static version of the interactive atlas and placed this information into the database tables automatically. Previously, the online histology atlas consisted of hundreds of individual Web pages with histological images and information specific to each image. This crucial step eliminated the need for someone to manually go through the hundreds of Web pages and images to enter the data associated with each image into the database. Once the data from the static version of the interactive atlas were inserted into database tables, it was then possible to develop the advanced elements of the site.

Data Display

The conversion to using a database directed the need to use a programming language (PHP in this case) to access and to display the data. PHP was used to create small computer programs (scripts) for each type of content page on the website. The ability to utilize one script for each type of content page ensured that the format of the Web page could be customized for the content depending on the user's internet browser and operating system.

Survey

A survey completed by medical students who had taken the histology course was conducted two months after the course ended. The survey consisted of fifteen Likert scaled questions/comments (Strongly Agree = 5, Strongly Disagree = 1) and two free response questions. The questions covered topics related to integration of the interactive Web-based atlas into the Vanderbilt histology curriculum and the technological features available on the interactive atlas (Table 1). The survey focused on a six week time period during which the first thirteen of twenty-four atlas chapters were available.

RESULTS

The Vanderbilt histology atlas was successfully converted into a database-driven interactive atlas in December 2004. The time required to develop computer scripts to extract data from the previous atlas Web pages into a database (MySQL®) took approximately 3 hours. The time spent creating some of the features after converting to a database are listed in Table 2. Some of those features are further described in this Results section.

Microscope Slide List

A microscope slide list allowed students to see all of the images associated with a given microscope slide. Because the Vanderbilt histology course followed the typical method of presenting basic tissue types before introducing organs, the microscope slide list provided a way to review all of the basic tissue types present on one slide. The list also gave students an opportunity to see the other chapters in which a microscope slide was referenced.

Search Function

The search feature provided students the ability to search all of the images contained within the interactive atlas for a structure or tissue of interest. With hundreds of images in the interactive atlas, the search feature allowed students to

Table 1. Interactive atlas student satisfaction and usage survey results

Question	Value
How often did you use the interactive atlas (when it was complete for that section of the course)?	Never 1; Once 3; Once a week 13; Multiple times each week 37; Daily 5
When did you most often use the interactive atlas?	Before lab 17; During lab 19; After lab as a review 27; Before quizzes and exams 49
I found the interactive atlas more helpful than other atlases (Gold Standard Media, Ross, Wheater's, Junquiera, etc)	*Mean: 4.34 Std Dev: 0.71
When the interactive atlas was available, it increased my confidence and ease of identification of structures.	*Mean: 4.63 Std Dev: 0.61
When the interactive atlas was available, my likelihood of attending lab... (1 = increased, 3 = stayed the same, 5 = decreased)	Mean: 2.78 Std Dev: 0.72
I was more productive when the interactive atlas was available.	*Mean: 4.31 Std Dev: 0.97
When the atlas was available, I did not attend lab as much because I could rely on the atlas to provide sufficient information regarding the topic to be covered in the laboratory.	*Mean: 1.93 Std Dev: 0.87
I felt that the interactive atlas was more helpful than other atlases (Gold Standard Media, Ross, Wheater's, Junquiera, etc) because it followed along with the Vanderbilt course.	*Mean: 4.32 Std Dev: 0.78
I hope the interactive atlas is completed for future Vanderbilt classes.	*Mean: 4.85 Std Dev: 0.36
I anticipate referencing the interactive histology atlas during my 2nd year pathology course.	*Mean: 4.12 Std Dev: 0.83
The ability to highlight structures using the mouseover made identification and learning the structures easier than referencing pictures with arrows.	*Mean: 4.72 Std Dev: 0.45
I felt the functions (search function, slide list, self-test, etc.) added to the interactive atlas throughout the semester made the atlas a better educational tool.	*Mean: 4.53 Std Dev: 0.68
I used the self test	Never 1; Once 6; 1-5 times 28; 6-10 times 17; Greater than 10 times 7
I found the self test helpful in reviewing for quizzes and exams.	*Mean: 4.22 Std Dev: 0.81
I felt that the atlas positively affected my learning experience by making structure identification a quicker and less confusing process.	*Mean: 4.66 Std Dev: 0.51

*Likert scale (Strongly Agree = 5, Strongly Disagree = 1)

quickly find additional examples of structures reinforced their learning. The results provided a link to the lab manual content page and the image itself.

Self-Test

After the course started, a student suggested creating a self-test function to review material. Because the strength of CAI tools lies in their ability to have interactive features, a self-test was developed where students selected the number of questions they would attempt and the chapters from which those questions were generated. Each self-test was unique, dynamic, and tested the student knowledge of histological images in the database.

Data Entry

Table 2. Time developing selected atlas features after database was in place

Feature	Time To Develop
Data Extracting	3 hours
Smart Navigation	2 hours
Microscope Slide List	15 minutes
Search Function	15 minutes
Self-Test	2 hours
List of Chapters	30 minutes
Intellectual Property	2 hours

The chance for random errors being included in data entry was drastically reduced through the creation of one common entry and display process. After the data from the static interactive atlas were read and loaded into the database tables, tools were created for the course faculty to add and edit new images and data to the interactive atlas. The options for adding and editing data included the use of a simple Microsoft® Excel spreadsheet, direct editing in the database via the Web-based tool phpMyAdmin (http://www.phpmyadmin.net), or an intuitive text format where the course faculty could enter free flowing text about many different microscope slide images in a Microsoft® Word or standard text document.

Intellectual Property

In order to protect the work of the creators of the interactive atlas, a PHP script was written to add a watermark dynamically to every image used in the site. With hundreds of images being used in the interactive atlas, this saved countless hours of work for the authors. Without this script, the authors would have had to paste their watermark on each individual photo in the atlas.

Furthermore, the interactive atlas was linked to the medical student homepage that requires medical students to login via their standard university login ID and password. Therefore, the interactive atlas was only available to its intended audience and its content was protected. With non-medical students enrolled in the course, an alternative login system was created to bypass the medical student homepage and allow specified users entry into the site.

Site Usage

To better understand how the interactive atlas was used, each student login session was tracked to determine the length of stay, the number of pages viewed, the type of computer used to access the site, and if the self-test was accessed. All data were recorded anonymously with the IP address as the only identifying information. Tracking the IP address provided a method to determine if there were malicious attempts made against the website. Over the course of the semester, 1,146 login sessions were recorded for 106 students. Out of 1,146 login sessions, the self-test was accessed in 276 of these sessions (23%). The duration of each login session lasted an average of 57.7 minutes, and Figure 2 displays the range in the duration of each login session. During each session, the students accessed an average of 81.2 pages, and Figure 3 displays the range in the number of pages viewed for each session. In total, the interactive atlas saw approximately 1,200 hours of usage throughout the six weeks of the course for which it was complete (thirteen of twenty-four course chapters). Further data collected showed a wide variety of computer hardware and software used to access the site: operating system, internet browser, and screen resolution.

Survey Results

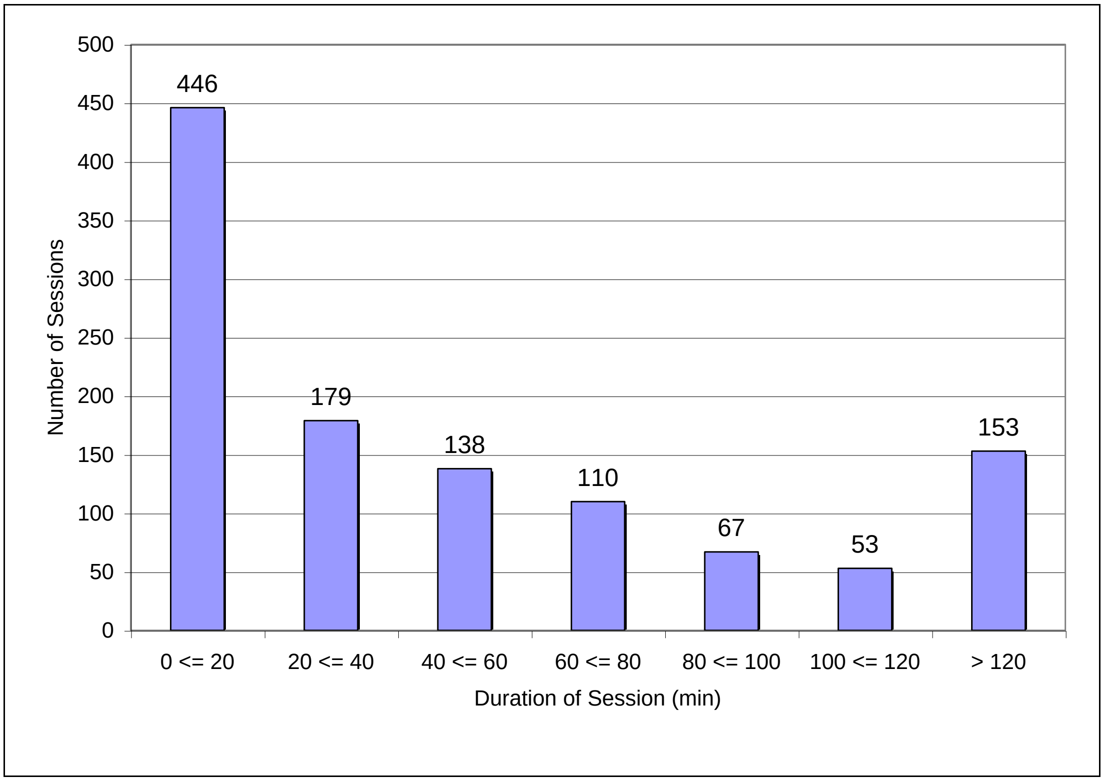
59 of 106 students responded (56%). One question asked "How often did you use the interactive atlas?" and 55 (93%) students responded that they utilized the interactive atlas at least one time per week with 42 (71%) students accessing the atlas multiple times per week or daily. When the question "When did you use the interactive atlas?" was posed, 49 (83%) stated they utilized this Web-based tool to study before quizzes and tests. The survey results also indicated 17 (29%) students accessed the atlas to preview structures before entering the laboratory, 19 (32%) students utilized the tool during lab, and 27 (46%) students reviewed the atlas following the laboratory sessions. The complete survey results are presented in Table 1.

DISCUSSION

Converting the Vanderbilt University interactive histology atlas, a static educational resource consisting of hundreds of Web pages, into a database-driven educational tool allowed easy alteration of the tool and created unlimited potential for expansion, similar to other database applications.[16] The need for flexibility and expansion was evident given the wide variety of hardware and software platforms used to access the site. Because of the adaptability of the database-driven atlas, students were able to view the contents of the atlas with the internet browser and computer with which they felt most comfortable.

A database-driven interactive atlas allowed for the creation of a number of advanced features, such as smart navigation, a search function, and a student self-test, which enabled the students to learn interactively and to review the course material. The implementation of the database also enabled protection of intellectual property and a way for faculty to analyze the Web page usage. With a database present, faculty experienced a drastic reduction in the time spent generating new chapters within the atlas due to the PHP computer scripts that were able to search, find, and place information in appropriate files, specifically designed to ease the data entry process. These features

Figure 2. Histogram of duration of login session



were made possible through the implementation of a database (MySQL®) and a dynamic programming language (PHP).

With the database structure in place, future work can be directed at improving the advanced features already developed and creating more interactive learning tools. The student survey results and use of the atlas (1,200 hours in six weeks) provided evidence that the implementation of the database structure was successful enhancing student learning. The survey results showed that students felt the atlas positively affected their learning experience by making structure identification a quicker and less confusing process (4.66 ± 0.51 Likert response). Moreover, students also benefited from the self-test (4.22 ± 0.81 Likert response), but suggested more time should be spent generating new identifying questions. The challenge in developing educational tools is often devoting time to the most appropriate resources.[17] With the database-structure implemented, construction of additional self-quizzing features can easily be developed in the future.

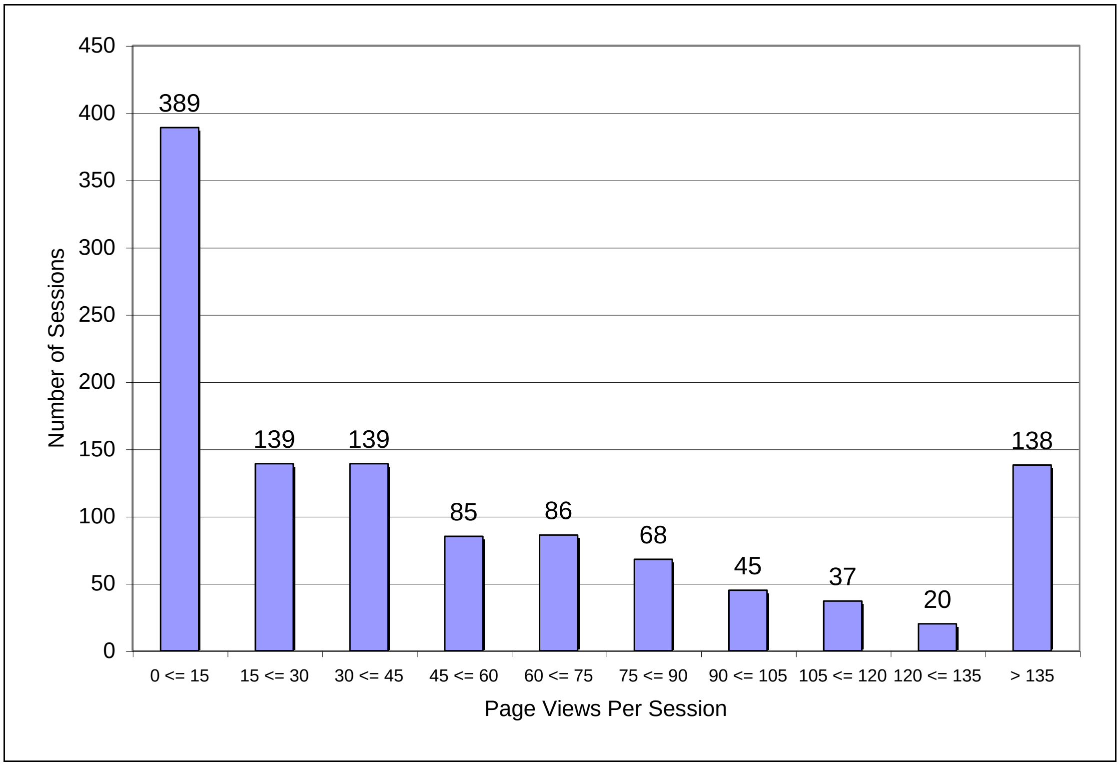
A database-driven educational tool can be altered quickly to meet the demands of students, which is crucial considering the growing technological expectations of students and the necessary considerations of these expectations in creating or modifying new multimedia tools.[18] With a database implemented, changes to the site layout or data presentation were made in seconds by altering one PHP script as opposed to the hours it would have taken previously to edit hundreds of Web pages. Because all of the Web pages referenced the same database, whenever data were edited, added, or deleted, all pages that referenced the data were automatically updated.

Therefore, implementation of the database saved faculty countless hours in atlas management time.

From the site usage analysis, it was evident that the interactive atlas was used in multiple ways. Data showed that the mean site stay was 57.7 minutes, and Figure 2 shows the variability in the duration of time students used the interactive atlas. These data suggest the atlas functioned as a quick reference (e.g. using the search function), but was also utilized during prolonged study sessions, especially before quizzes and exams. Figure 3 further supports this statement by showing great variability in the number of pages viewed per session. The students averaged 81.2 page views per session, but this histogram reveals that students utilized this CAI application for different purposes. These data emphasized the importance of having a tool that was flexible and suitable for users in many capacities.

Even if a tool were not database-driven, it would still benefit from the implementation of a scripting language. By using a scripting language like PHP, any changes in design would be easily accommodated. A simple script could be written to perform a variety of automated tasks: renaming files, extracting data from static pages, reorganizing content on static pages, changing link names, and formatting to name a few of the capabilities. Changing data manually is tedious and often not feasible.

Figure 3. Histogram of page views per login session



The time needed to create an automated script is drastically less than the time that it would take for numerous changes to be made manually.

Designing a tool that utilizes databases and dynamic scripting languages allows the tool to be placed on the Web and accessed from any location. Compared to a stand-alone application installed on the computer of the user, using the Web as the medium for access has many advantages. Given the importance of and widespread availability of the Web in educational settings, students can access the atlas whenever they desire.[4,5,19,20] The atlas was created for display in Hypertext Markup Language (HTML), which was compatible with many computer operating systems and Web browsers. In sharp contrast, multiple versions of a stand-alone application would have to be developed to accommodate users with different operating systems. Placing the atlas on the Web gave students access to the tool wherever they wanted with the platform that they chose to utilize. Moreover, developing the tool on the Web allowed it to be integrated with the Vanderbilt university username and password system to protect the content of the atlas. The Web also enabled the creators to track the usage of the CAI application. Furthermore, utilization of the Web provided a way to make instantaneous changes and to add new features. Finally, the Web-based atlas provided the framework and the unlimited potential to integrate information with another course. A Web-based atlas enables directed instruction coupled with high-resolution images, movies with extensive narration, Web-based animations, and other interactive features to supplement course material.[21,22] For example, the interactive atlas could be incorporated, merged, or used for comparison with an interactive atlas in a pathology course. The power of two related tools would provide students with more resources to enhance their learning, which is an important part in integrated learning.[23] Given readily available internet access, it is the opinion of the authors that all tools, even those distributed with educational books, should be available online and not distributed as stand-alone applications for installation on a user's local machine. The effects of such a design choice severely limit the possibilities for expansion and accessibility.

As the interactive atlas is used in the histology course, more advanced features will be developed according to insight and student feedback. For example, the interactive atlas may dynamically insert the text from the lab manual onto the page showing the image for structure identification. With the previous static design, this task would have taken weeks to accomplish. In addition, a feature could be created to allow course instructors the ability to create a session to preview a laboratory exercise or to review for an exam. With such a feature, course instructors would choose from the existing set of slides, provide additional comments about why the slide is being reviewed, and then place them in the appropriate order for the students to review. This function would provide the instructors the ability to highlight the most important images or concepts upon which they would like students to focus. Also, the incorporation of a medical dictionary into the interactive atlas is possible. With each new medical course, students face the challenge of learning new medical terms. Because the interactive atlas uses a database and PHP, every word on a page that matches an entry in a medical dictionary could be defined. Such a feature could provide the definition when the user moves his or her mouse over the word. Finally, the interactive atlas could be modified as a medium to test students throughout the course in laboratory or kodachrome quizzes using images not previously seen. If such a feature were implemented, data could be collected to compare students' performance on quizzes with and without the use of the atlas. Assessing student learning is an important aspect of any educational tool, and comparison data would enable analysis of the effectiveness of the interactive atlas.[24] In previous studies, students felt the interactive nature of CAI applications led to increased productivity and efficiency, and with the tracking capabilities of a database-driven atlas, more comparison data could be generated for analysis.[25,26]

CONCLUSIONS

At Vanderbilt University School of Medicine, an interactive, Web-based histology atlas was created to place the students' laboratory manual on the Web and to display hundreds of histological images via the use of HTML. The atlas was converted into a database-driven tool using a Web programming language (PHP) and a MySQL® database. The conversion streamlined the development and modification of the atlas, and it enabled the implementation of advanced features.

Our experience shows it is possible and efficient to convert existing CAI applications created without database technology to an underlying database structure. The database structure permitted maintenance and upgrades without continual reliance upon a technology expert. Such a conversion maximizes the capability, functionality, and adaptability of a CAI tool while saving time in construction and editing.

ACKNOWLEDGEMENTS

The authors would like to thank the Vanderbilt University Departments of Cell and Developmental Biology and Cancer Biology for their support of the interactive atlas. Drs. Lynn Matrisian and Cathy Pettepher financially supported the creation of the interactive atlas. The authors would also like to thank Nader Shourbaji for his preliminary work on the atlas and Davin Chark for his vision in its graphical design.

REFERENCES

1. Lei, L.W., Winn, W., Scott, C., and Farr, A. Evaluation of computer-assisted instruction in histology: effect of interaction on learning outcome. *Anatomical Record Part B, New Anatomist*. 2005; 284(1):28-34.
2. Sinav A., and Ambron R. Interactive Web-based programs to teach functional anatomy: the pterygopalatine fossa. *Anatomical Record Part B, New Anatomist*. 2004; 279(1):4-8.
3. Letterie G.S. Medical education as a science: the quality of evidence for computer-assisted instruction. *American Journal of Obstetrics and Gynecology*. 2003; 188(3):849-53.
4. Nieder G.L., and Nagy, F. Analysis of medical students' use of Web-based resources for a gross anatomy and embryology course. *Clinical Anatomy*. 2002; 15(6):409-18.
5. McLean M. Web pages: an effective method of providing CAI resource material in histology. *Medical Teacher*. 2001; 23(3):263-9.
6. Heidger P.M., Jr., Dee, F., Consoer, D., Leaven, T., Duncan, J., and Kreiter, C. Integrated approach to teaching and testing in histology with real and virtual imaging. *Anatomical Record*. 2002; 269(2):107-12.
7. Hayes M., and Billy, A. Web-based modules designed to address learning bottlenecks in introductory anatomy and physiology courses. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*. 2003.
8. McNulty, J.A., Halama, J., and Espiritu, B. Evaluation of computer-aided instruction in the medical gross anatomy curriculum. *Clinical Anatomy*. 2004; 17(1):73-8.
9. Foreman, K.B., Morton, D.A., Musolino, G.M., and Albertine, K.H. Design and utility of a Web-based computer-assisted instructional tool for neuroanatomy self-study and review for physical and occupational therapy graduate students. *Anatomical Record Part B, New Anatomist*. 2005; 285(1):26-31.
10. McMillan, P.J. Exhibits facilitate histology laboratory instruction: student evaluation of learning resources. *Anatomical Record*. 2001; 265(5):222-7.
11. Lehmann, H.P., Freedman, J.A., Massad, J., and Dintzis, R.Z. An ethnographic, controlled study of the use of a computer-based histology atlas during a laboratory course. *Journal of the American Medical Informatics Association*. 1999; 6(1):38-52.
12. Cotter, J.R. Laboratory instruction in histology at the University at Buffalo: recent replacement of microscope exercises with computer applications. *Anatomical Record*. 2001; 265(5):212-21.
13. Elizondo-Omana, R.E., Morales-Gomez, J.A., Guzman, S.L., Hernandez, I.L., Ibarra, R.P., and Vilchez, F.C. Traditional teaching supported by computer-assisted learning for macroscopic anatomy. *Anatomical Record Part B, New Anatomist*. 2004; 278(1):18-22.
14. Sims, R. Interactivity: A forgotten art? *Computational Human Behavior*. 1997; 13:157-80.

15. Chou, C. Interactivity and interactive functions in Web-based learning systems: a technical framework for designers. *British Journal of Educational Technology*. 2003; 34:265-79.
16. Lee, M.Y., Albright, S.A., Alkasab, T., Damassa, D.A., Wang, P.J., and Eaton, E.K. Tufts Health Sciences Database: lessons, issues, and opportunities. *Academic Medicine*. 2003; 78(3):254-64.
17. Turchin, A., and Lehmann, C.U. Active Learning Centre: design and evaluation of an educational World Wide Web site. *Medical Informatics and the Internet in Medicine*. 2000; 25(3):195-206.
18. Huang, C. Designing high-quality interactive multimedia learning modules. *Computerized Medical Imaging and Graphics*. 2005; 29(2-3):223-33.
19. Cook, D.A. The research we still are not doing: an agenda for the study of computer-based learning. *Academic Medicine*. 2005; 80(6):541-8.
20. Michaels, J.E., Allred, K., Bruns, C., Lim, W., Lowrie, D.J., Jr., and Hedgren, W. Virtual laboratory manual for microscopic anatomy. *Anatomical Record Part B, New Anatomist*. 2005; 284(1):17-21.
21. McLean, M. Introducing computer-aided instruction into a traditional histology course: student evaluation of the educational value. *Journal of Audiovisual Media in Medicine*. 2000; 23(4):153-60.
22. Brisbourne, M.A., Chin, S.S., Melnyk, E., and Begg, D.A. Using Web-based animations to teach histology. *Anatomical Record*. 2002; 269(1):11-9.

Medical Student Attitudes Towards and Perception of the Basic Sciences in a Medical College in Western Nepal

P. Ravi Shankar, M.D.[1], Arun K. Dubey, M.D.[2], P. Subish, M.Pharm.[2], Dinesh K. Upadhyay, M.Pharm.[2]

[1]Department of Pharmacology and Medical Education
[2]Department of Pharmacology
Manipal College of Medical Sciences
Pokhara, Nepal

Phone: (+)1-00977-61-440600 Fax: (+)1-00977-61-440260 Email: ravi.dr.shankar@gmail.com

ABSTRACT

The basic sciences in Nepal are taught during the first four semesters of the undergraduate medical course. The objectives of this study were to obtain information on the attitudes towards and perception of the basic sciences among second and fourth semester students and note the association, if any, of the attitude with the respondent personal characteristics. The study was carried out among second and fourth semester medical students at the Manipal College of Medical Sciences, Pokhara, Nepal during March 2005. The student attitudes towards the basic sciences were assessed by noting their degree of agreement with a set of nine statements using a modified Likert-type scale. The median scores for the different statements were compared among the various subgroups of respondents. Sixty-two of the 75 second semester (82.7%) and 73 of the 75 fourth semester (97.3%) students successfully completed the questionnaire. Female students were more in agreement with the statement 'Faculty members excite students' curiosity through the teaching of the basic sciences' compared to male students (p=0.034). The female students had a more positive opinion regarding the 'perceived effectiveness of medical education' compared to males (p=0.003). Significant differences in the median scores of some statements were seen between the second and fourth semester students. Greater emphasis should be placed on the psychological aspects of medical treatment and on teaching about applying basic sciences to clinical medicine. Similar studies in other medical schools would add to this data. The students overall had a positive opinion towards the basic sciences.

INTRODUCTION

At the Manipal College of Medical Sciences (MCOMS), Pokhara, Nepal the basic science subjects (Anatomy, Physiology, Biochemistry, Pharmacology, Pathology and Microbiology) are taught during the first four semesters of the undergraduate medical (MBBS) course. Teaching and learning of Community Medicine starts from the first semester and continues until the seventh semester. The curriculum of Kathmandu University with which the college is affiliated was revised in 2001. The revised curriculum now emphasizes integrated, organ-system based teaching of basic medical science subjects. These subjects are linked with Community Medicine and frequent hospital visits are emphasized to give a clinical orientation to the subjects.[1] The students are introduced to clinical knowledge and skills utilizing a problem solving approach.

Traditional basic science teaching has been as individual disciplines unrelated to other basic science disciplines and the student future medical practice.[2] Though an integrated approach to teaching is followed in our institution, assessment is still done subject wise. Subjects are taught to prepare students for assessments and unfortunately the process of integrating with other basic science and clinical subjects is not always emphasized. It has been observed that basic science knowledge learned in a clinical context is better comprehended and more easily applied by the students.[3,4]

At MCOMS, we use a combination of didactic lectures, problem-stimulated learning (PSL) sessions, student seminars and community-based learning (CBL) for teaching the basic sciences. During the clinical orientation module, history taking and physical examination, both general and system wise are covered.

To measure student attitude towards and perception of the basic sciences, an instrument was developed by West, Mennin, Kaufman, and Galey.[5] The same instrument was used by other authors in later studies[2,6] and has also now been applied to our students. Information on student attitudes towards basic sciences are so far lacking in colleges in Nepal and would also provide additional data for medical schools in other countries. Hence the present study was carried out. The objectives of the study were to:

a) obtain information on the attitudes towards and perception of the basic sciences among second and fourth semester medical students and
b) note the association, if any, of the attitude with personal and demographic characteristics of respondents.

MATERIALS AND METHODS

Participants: The study was carried out among second and fourth semester medical students at the Manipal College of Medical Sciences, Pokhara, Nepal during March 2005. The second and fourth semester class consists of 75 students each and all students were invited to participate. The objectives and the scope of the study were explained and the students were assured that individual demographic and personal information collected would be kept confidential. Verbal consent was taken. However, written informed consent was not obtained. Sixty-two second semester and 73 fourth semester students participated in this study.

Instrument: The demographic and personal characteristics including age, sex, nationality and occupation of parents were noted. Information was collected on the method of financing of medical education and the preferred specialty for post graduation. The respondents were asked whether they would consider basic sciences as a career option. Information on whether the respondents were graduates of a graduate study program and the place of residence was noted. The students were asked to rate themselves academically as excellent, good, average and poor.

The questions designed by West and coworkers[5] were presented in English. The questionnaire used is shown in the Appendix. The students were asked to denote their degree of agreement with the individual statements using a 5-point Likert type scale. Items 1, 2 4 and 5 assess the perceived importance and relevance of basic sciences for clinical medicine while items 6, 7 and 9 assess the perceived effectiveness of medical education. The forms were distributed to the students during the pharmacology practical sessions. The objectives of the study were explained to the students and they were invited to participate. The completed forms were collected at the end of the session.

Analysis: The median scores for the different statements were compared among different subgroups of respondents. The median scores of individual statements and of the two subgroups of statements were compared. The students were divided into subgroups according to nationality, method of financing of medical education, preferred specialty for post graduation, occupation of parents and place of residence. The median scores of the statements and the two subgroups were compared among the second and fourth semester students. For the purpose of analysis, the occupation of parents was grouped into three: students with both parents' doctors, students having one doctor parent and students without doctor parents. The subjects for post graduation were grouped into basic sciences, medical specialties and surgical specialties. The Mann-Whitney U test was used for dichotomous variables and the Kruskal Wallis test for the others. A p value < 0.05 was taken as statistically significant.

RESULTS

Sixty-two of the 75 second semester (82.7%) and 73 of the 75 fourth semester students (97.3%) successfully completed the questionnaire. The demographic characteristics of the responders and non-responders are comparable. The majority of the respondents (88.9%) were 19, 20 and 21 years of age. Seventy-eight respondents (57.8%) were male. Forty-seven respondents (34.8%) were Nepalese while 76 (56.3%) were from India. The demographic and personal characteristics of the respondents are shown in Table 1. The majority of students [105 students (77.8%)] were self-financed and from urban areas [118 students (87.4%)]. The majority of students rated themselves as good or average (not shown). Basic sciences were not preferred subjects for post graduation. However, 35 of the 70 self-financing students (50%) responded that they may consider the basic sciences as a career option.

Only 4 students (3%) had joined the undergraduate medical course after first completing a graduate course of study. A majority of students [91 students (67.4%)] did not have doctor parents and rated themselves as average students [86 students (63.7%)]. The characteristics are shown in Table 1. The median score for the two negative statements about basic sciences (statements 1 and 2 in the Appendix) were showing high disagreement with the statements. The median scores for the other statements (statements 3, 4, 5, 6, 7, 8 and 9) which were positive about basic sciences were showing low agreement with the statements. The students had a favorable opinion regarding the basic science subjects.

The median score and the interquartile range for the nine statements and the two subgroups of statements are shown in Table 2. Table 3 shows the median scores of the statements according to selected demographic characteristics of the respondents. Female students were more in agreement with the statement "Faculty members excite students' curiosity through the teaching of the basic sciences" (statement number 9) compared to male students ($p=0.034$). The scholarship students more strongly disagreed with statement 2 (basic science research being far removed from clinical medicine) compared to the self-financing students. The female students had a more positive opinion regarding the "perceived effectiveness of medical education" (sum of the scores of the statements 6, 7 and 9) compared to males ($p=0.003$).

The median scores of the individual statements and of the two subscales were compared among the second and fourth semester students. The significant differences have been shown in Table 3. The fourth semester students agreed more with statement 5 that applying basic sciences to clinical

Table 1. Demographic and personal characteristics of student respondents.

Age (in years)		**Gender**		**Nationality**		**Financing**		**Subject post graduation**		**Occupation of parents**		**Self-assessment**	
Variable	**No. (%)**	**Variable**	**No. (%)**	**Variable**	**No. (%)**	**Variable**	**No. (%)**	**Variable**	**No. (%)**	**Variable**	**No. (%)**	**Variable**	**No. (%)**
19	23 (17)	Male	78 (57.8)	Nepalese	47 (34.8	Scholarship	30 (22.2)	Basic sciences	8 (5.9)	Both doctors	14 (10.4)	Below average	42 (31.1)
20	63 (46.7)	Female	57 (42.2)	Indian	76 (56.3)	Self-financing	105 (77.8)	Medical	36 (26.7)	One doctor	30 (22,2)	Average	86 (63.7)
21	34 (25.2)			Sri Lankan	11 (8.1)			Surgical	39 (28.9)	None doctor	91 (67.4)	Good	5 (3.7)
Others	15 (11.1)			Others	1 (0.7)			Undecided	52 (38.5)				

Age (in years)		Gender		Nationality		Financing		Subject post graduation		Occupation of parents		Self-assessment	
Variable	No. (%)	Variable	No. (%)	Variable	No. (%)	Variable	No. (%)	Variable	No. (%)	Variable	No. (%)	Variable	No. (%)
19	23 (17)	Male	78 (57.8)	Nepalese	47 (34.8	Scholarship	30 (22.2)	Basic sciences	8 (5.9)	Both doctors	14 (10.4)	Below average	42 (31.1)
20	63 (46.7)	Female	57 (42.2)	Indian	76 (56.3)	Self-financing	105 (77.8)	Medical	36 (26.7)	One doctor	30 (22,2)	Average	86 (63.7)
21	34 (25.2			Sri Lankan	11 (8.1)			Surgical	39 (28.9)	None doctor	91 (67.4)	Good	5 (3.7)
Others	15 (11.1)			Others	1 (0.7)			Undecided	52 (38.5)				

Table 2. Median score and interquartile range for the nine statements and the two subscales.

Characteristic	Median	Interquartile range
Statement 1	4	1
Statement 2	3	2
Statement 3	2	1
Statement 4	2	2
Statement 5	2	1
Statement 6	2	2
Statement 7	2	1
Statement 8	2	1
Statement 9	2	1
Importance for clinical medicine	11	1
Perceived effectiveness of medical education	7	3

medicine is a skill which should be learned early on in medical education (p= 0.025). The second semester students agreed more with statement 6 regarding acquisition of facts in basic sciences compared to the fourth semester (p=0.022). The second semester students agreed more with the statement number 9 that faculty members excite student curiosity through the teaching of basic sciences compared to the fourth semester (p=0.091).

Table 4 shows the mean ± SD scores of the nine statements in our study compared to those of the study by Kaufman and Mann and Custers and Ten Cate. Some of the scores reported in our study were lower and others higher than those reported in previous studies. Lower scores denote a greater degree of agreement with the statements. Our scores for the statement 1 "A physician can effectively treat most medical problems without knowing the details of the biological processes involved" was slightly higher compared to that reported previously. The scores for the statement number 2 (basic science research being far removed from clinical medicine) were lower. Our students showed a far greater degree of agreement with statement 4 that knowledge of biological mechanisms is the most important facet of a good physician. They also tended to agree more with statement 9 "Faculty members excite students' curiosity through teaching basic sciences" compared to previous studies.

DISCUSSION

Rangachari has written that "The basic medical sciences play elegantly the role of villains, subjecting students to volumes of increasingly arcane information".[8] It has been observed that each discipline is convinced that "it represents the most important part of medicine and turns on the spigot of information" forcing students to question the relevance or irrelevance of the information.[8]

MCOMS admits students from Nepal, India and Sri Lanka for the MBBS course. A few students come from other countries. The Nepalese scholarship students are selected through an entrance examination conducted by the Ministry of Education (MoE). The Nepalese self-financing students are selected on the basis of an entrance examination conducted by Kathmandu University (KU) with which the college is affiliated. The students from India and Sri Lanka are selected on the basis of their twelfth standard marks.

Like in previous studies[2,6] the sex ratio of the respondents in our study was not equal (M/F=78/75). In our study the female students were more in agreement with statement 9 (Table 3). There is a hypothesis that males may be more technically oriented than females and may value the basic sciences more. In a previous study, females agreed more with the statement that psychological factors are just as important in the healing process as physical factors.[2] The female respondents had a more positive opinion of the effectiveness of medical education.
We compared our study with those conducted by Custers & Ten Cate at Utrecht[2] and by Kaufman & Mann at Dalhousie[6] which had used a similar questionnaire. These studies had compared student attitudes towards basic sciences under an innovative integrated curriculum. Differences in the curriculum, teaching methods and cultural differences may be confounding variables in the comparison. However, our results are quite similar to those reported by other studies (Table 4), with statement number 4 providing the highest variance.

The scholarship students disagreed more with statement 2 (Table 3). A greater proportion of self-financing students (50% of the total) answered that they may consider the basic sciences as a career option.

There were no significant differences in the statements according to other demographic characteristics. We have only shown the statements where there were significant differences between the subgroups of respondents. Our students had a better opinion regarding the 'perceived importance and relevance of the basic sciences for clinical medicine' (sum of statements 1, 2, 4 and 5) and 'perceived effectiveness of medical education' (sum of statements 6, 7 and 9) compared to that reported previously.[2,6] However, our study was conducted among only the basic science subjects compared to the previous studies which were done among students doing their clerkship.

Table 3. Median scores of the statements according to selected demographic characteristics.

Characteristic	Statement number	Median scores	Statistical tests
Gender Male Female	9	3 (2.25)	Mann-Whitney U test, U = 1762, p =0.034
Financing Scholarship Self-financing	2	4 (1.25) 3 (2)	Mann Whitney U test, U= 1202, p =0.040
Gender Male Female	Sum of statements 6, 7 and 9	8 (2) 7 (2)	Mann Whitney U test U= 1557, p = 0.003
Semester Second Fourth	5	2 (1.25) 1 (1)	Mann Whitney U test, U = 1791, p =0.025
Semester Second Fourth	6	2 (2) 3 (1.5)	Mann Whitney U test, U = 1760, p =0.022
Semester Second Fourth	9	2 (2) 3 (1.5)	Mann Whitney U test, U = 1892.5, p =0.091

There was a positive change towards statement 5 as we moved from the second to the fourth semester (Table 3). This is a welcome sign. However, the second semester was more positive regarding the role of the faculty members in exciting student curiosity through the teaching of basic sciences.

Our students were of the opinion that detailed knowledge of the biological processes involved is necessary to effectively treat patients compared to the results from the previous studies (statement number 1). They also agreed to a greater extent with the statement 9. These are welcome signs for a basic science educator. However, the students also agreed to a greater extent with statement 2 regarding the low usefulness of basic sciences for the practicing doctor. This is a dissenting statement in the overall positive opinion regarding basic sciences and the reasons have to be investigated further. Despite the limitations of the comparison previously mentioned, the comparison with two previous studies[2,6] gave us an idea regarding the perception and attitude towards basic sciences of students in MCOMS compared to those reported from developed countries.

Though this was a preliminary study, certain tentative conclusions can be drawn regarding student perceptions and attitudes towards basic sciences. Though there is no formal instruction on integrative and holistic medicine in our curriculum, students agreed that psychological factors play as important a role as physical factors in healing. Integrative medicine may be considered for inclusion in the curriculum if this finding is confirmed by larger studies. Applying the basic sciences to clinical medicine is a skill which should be learned early on in medical education. We are of the opinion that this skill is not much emphasized in our institution and a greater degree of coordination and integration with the clinical disciplines

Table 4. Comparison of the results of the present study with that of previous studies.

Statements	MCOMS, Pokhara Second semester (n= 62) Mean (SD)*	MCOMS, Pokhara Fourth semester (n= 73) Mean (SD)	Custers & Ten Cate Old current (n=73) Mean (SD)	Custers & Ten Cate New current (n=89) Mean (SD)	Kaufman & Mann Conventional current (n =52) Mean (SD)	Kaufman & Mann Current (n =72) Mean (SD)
1	3.97 (1)	4 (1.14)	3.97 (0.95)	3.58 (1.03)	3.87 (1.14)	3.68 (1.1)
2	2.93 (0.99)	3.08 (1.1)	3.62 (0.98)	3.34 (0.95)	3.83 (0.98)	3.54 (1.01)
3	1.9 (0.93)	1.92 (1.1)	2 (0.94)	2.05 (0.96)	**	**
4	2 (1)	2.04 (1.05)	3.75 (1.01)	3.66 (1.1)	3.27 (1.12)	3 (0.98)
5	2.03 (1.02)	1.73 (1.03)	1.87 (0.86)	1.75 (0.75)	1.94 (1.06)	1.93 (0.78)
6	2.1 (1.1)	2.7 (1.8)	3.34 (1.14)	4.07 (0.84)	2.9 (1.14)	3.28 (1.13)
7	2.58 (1.17)	2.29 (1.06)	2.34 (1.07)	2.11 (0.92)	2.33 (1)	1.96 (1.06)
8	2.14 (1.04)	1.96 (0.9)	2.17 (0.97)	2.05 (0.86)	2.29 (0.87)	1.99 (1.5)
9	2.32 (1.15)	2.66 (1.19)	3.1 (1.01)	2.58 (0.9)	3.31 (0.98)	2.82 (0.83)

* The statements were rated on a five-point scale, ranging from 1 (strongly agree) to 5 (strongly disagree). ** This statement was not included in the questionnaire used at Dalhousie university.

are required. The students were confused about whether they should concentrate on facts or on concepts during the basic sciences course.

Medical education is a dynamic process and modifications in teaching-learning methodology should be periodically considered. Our study suggests a positive attitude towards basic sciences and a longitudinal study will reveal whether the attitude persists during the latter years of study. Other medical schools may carry out a similar study to find out the student perception and attitudes towards basic sciences. The point whether the students feel that basic sciences are of importance in the practice of clinical medicine is especially important. It is important that basic science teaching emphasizes concepts, excites student curiosity and teaches them the skill of applying basic sciences in clinical medicine. Students would be more motivated to learn basic sciences if they feel it is relevant to clinical practice and is important to their future role as a doctor. Student attitude is of primary importance to basic science educators as it may directly affect student interest in, motivation towards and learning of the basic sciences. With the decreasing emphasis on basic sciences already referred to, such studies will be helpful to underscore the relevance and importance of basic sciences in medical education. Longitudinal studies will also be important.

Our study had limitations. The study was carried out among only two semesters of medical students in one medical college in Nepal. The sample size was small. The clinical years of study and doctors doing their clerkship were not included. Longitudinal studies of student attitudes as they progress through medical school were not carried out. Further studies with a larger sample size and in other medical colleges in Nepal are required.

References

1. Kathmandu University. Curriculum for Bachelor of Medicine and Bachelor of Surgery (MBBS) Part one Basic Medical Sciences Third version. *Dhulikhel:2001*.
2. Custers, E.J.F.M., and Ten Cate, T.J.O. Medical students' attitudes towards and perception of the basic sciences: a comparison between students in the old and the new curriculum at the University Medical Center Utrecht, the Netherlands. *Medical Education.* 2002; 36: 1142-1150.
3. Patel, V.L., Evans, D.A., and Kaufman, D.R. Reasoning strategies and the use of biomedical knowledge by medical students. *Medical Education.* 1990; 24: 129-136.
4. Hmelo, C.E. Cognitive consequences of problem-based learning for the early development of medical expertise. *Teaching and Learning in Medicine.* 1998; 10: 92-100.
5. West, M., Mennin, S.P., Kaufman, A., and Galey W. Medical students' attitudes towards basic sciences: influence of a primary care curriculum. *Medical Education.* 1982; 16: 188-191.
6. Kaufman, D.M., and Mann, K.V. Basic sciences in problem-based learning and conventional curricula: students' attitudes. *Medical Education.* 1997; 31: 177-180.
7. Rangachari, P.K. Basic sciences in an integrated medical curriculum: the case of pharmacology. *Advances in Health Sciences Education*. 1997; 2: 163-171.
8. Vogel, W.H. Relevance of 'irrelevant' facts in medical education: the value of basic science teaching for later medical practice. *Academic Medicine.* 1993; 63 Suppl: S 27-28.

Appendix:

Questionnaire used to evaluate student attitudes towards the basic sciences (The questionnaire was developed by West, Kaufman & Galey. We have added specific demographic and personal characteristics of our interest)

Age: **Sex:** **Nationality:**

Scholarship/Self-financing Graduation: yes/ no **Preferred subject for post graduation:** Basic sciences/ Medical specialties/ Surgical specialties **Occupation of parents:** Both doctors/ One doctor/ None doctor **Basic sciences for post graduation:** Yes /no **Place of residence:** Urban/Rural **Self-rating as a student:** Excellent/ Good/Average/Poor

For the following statements denote your degree of agreement with each individual statement using the following scale (1 = strongly agree with the statement, 2 = agree, 3= neutral, 4 = disagree with the statement, 5 = strongly disagree with the statement)

1. A physician can effectively treat most medical patients without knowing the details of the biological processes involved.
2. Most basic science research is so far removed from clinical medicine that its usefulness to the practicing doctor is slight.
3. Psychological factors are just as important as physical factors in the healing process.
4. Of the facets of a good physician, his/her knowledge of biological mechanisms is most important.
5. Applying the basic science of medicine to clinical practice is a skill which should be reinforced early on in medical education.
6. It is first necessary to learn as many facts as possible in the basic sciences and then learn to apply them later on in the clinical years.
7. What students should learn in the basic sciences are the general concepts, in order that they might have a good working knowledge without having to know all the facts.
8. The information and knowledge I have gained to date are fundamental to my future role as a physician.
9. Faculty members excite students' curiosity through the teaching of the basic sciences.