2007

Developing/testing a new approach for assessing rapid visual identification of hematological cells using principles of visual cognition: a health science education study

Deborah Elizabeth Fox
Louisiana State University and Agricultural and Mechanical College

DEVELOPING/TESTING A NEW APPROACH FOR ASSESSING RAPID VISUAL IDENTIFICATION OF HEMATOLOGICAL CELLS USING PRINCIPLES OF VISUAL COGNITION: A HEALTH SCIENCE EDUCATION STUDY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Curriculum and Instruction

by
Deborah Elizabeth Fox
B.S., Louisiana State University, 1994
M.A., Louisiana State University, 2001
August 2007

ACKNOWLEDGEMENTS

I, Debbie Fox, would like to sincerely thank the following individuals for his/her role/contributions in the completion of my dissertations research: (a) Brian Fox, my supportive husband (b) Sandy Braden, my wonderful and encouraging mother (c) Dr. James Wandersee, LSU, Major Professor (d) Janssen Burris, Our Lady of the Lake College, Information Technology Department, Programmer for Cell Exam (e) Tammy Montgomery, Pathology Group of Louisiana, medical transcriptionist and (f) Donnell Leblanc, Our Lady of the Lake Regional Medical Center, Hematology Senior Technologist.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The purpose of this study was the development and testing of a novel method for assessment of white blood cell (WBC) identification skills used in the field of Clinical Laboratory Sciences (CLS). A dual format exam was administered to both novices (students) and experts (laboratory professionals). Format 1 was similar to current assessment formats, simply presenting a series of single WBC images for identification. Format 2 applied principles of visual cognition, grouping WBCs for identification by patient and presenting multiple example images from the patient before requesting identification of individual cells. This novel exam format was intended to: (a) provide a contextualized visual background for single cell identifications, (b) mirror the process of WBC identification used in clinical practice, and (c) promote improved performance on difficult/atypical WBC identifications.

The second phase of this study implemented qualitative methods to categorize the general cognitive processing styles used by novices/experts as either analytical or similarity-based. Cognitive processing styles were compared across the 2 levels of expertise as well as across exam formats.

Statistical analyses did suggest that expert performance levels were significantly improved by the novel exam presentation format. Novice performance, however, was not significantly altered by exam format. Evaluation of response times indicated that expert response times were significantly shorter than novice response times in format 2, but not in format 1. In addition, analysis of qualitative data suggested that experts differed significantly from novices in their cognitive verbalizations for format 2, with experts making more statements at a higher cognitive level than did the novices. Format 1 verbalization differences were not found to be significant.

Overall results indicated that the novel exam format invoked experts to implement similarity-based processing, allowing some identifications to be made at the level of the patient case, rather than simply at the feature identification level. Implications of this study include possible alterations to current certification/proficiency exam formats for questions requiring the visual identification of white blood cells. This study also suggests that using patient image sets as instructional stimuli may encourage the development of advanced cognitive processing skills in students.

INTRODUCTION

Joseph Schwab (1973) identifies four parts of the educative process: teacher, learner, subject knowledge, and social milleau. Novak, Mintzes, and Wandersee (2000) add assessment as a fifth and extremely essential element to this process. They explain that "while we believe the primary motivation for learning should be the satisfaction that comes with achieving competence, we need assessment to gauge the degree to which we approach or attain high competence. High-Quality assessment can facilitate high-quality learning" (p.1).

According to the constructivist philosophy (Novak, 1998; Mintzes, Wandersee, & Novak, 1998), assessment is used to evaluate knowledge that is created, not discovered. Methods of meaning creation/negotiation may include the construction of shared meaning between instructor and student as well as the use of certain high-quality forms of knowledge assessment. Gowin (1981) suggests that when meaning negotiation does lead to a high level of understanding, "grasped meaning" has been developed. Identification of this level of understanding is essential and can only be ascertained through the use of well-structured, appropriate, and effective forms of knowledge assessment.

Benjamin Bloom (1956) designates six levels of knowledge: (a) knowledge, (b) comprehension, (c) application, (d) analysis, (e) synthesis, and (f) evaluation. He groups these six knowledge levels into three primary categories: (a) Level I: Recall (which encompasses knowledge and comprehension), (b) Level II: Interpretation (which encompasses application and analysis), and (c) Level III: Problem-Solving (which encompasses synthesis and evaluation). In order to identify true mastery of a subject and the existence of true competency, assessment tools must be developed that ultimately assess knowledge at levels II and III.

Additionally, knowledge assessment tools should assess the development of meaningful learning as defined by David Ausubel (1968). Ausubel defines meaningful learning as a "non-

arbitrary" and "substantive" or "non-verbatim" integration of new information into a person's already existing knowledge framework. He suggests that three things are necessary in order for meaningful learning to occur: (a) Learners must have the necessary prior knowledge, (b) the material itself must be inherently meaningful, and (c) Learners must choose to incorporate the new knowledge into their existing knowledge structure in a non-verbatim, non-arbitrary manner.

Further contributions made by Ausubel in defining the aspects of true learning are embodied by his Cognitive Assimilation Theory. In it, he describes four processes that take place during meaningful learning: (a) subsumption, (b) superordinate learning, (c) progressive differentiation, and (d) integrative reconciliation. During subsumption new, more specific information is incorporated into our existing knowledge structure. Superordinate learning occurs when more general concepts are added to our existing knowledge structure. Progressive differentiation occurs as concepts once grouped together are identified by their differences. Integrative reconciliation takes place when we notice not only the differences between related concepts, but also their similarities.

Acquisition of visual classification skills such as those needed for identification of hematologic blood cells, relies on the use of advanced levels of knowledge and the development of meaningful learning. Although research has not been conducted in the area of visual classification (diagnosis) as it relates to the specific field of hematology, extensive studies have been completed which evaluate cognitive processing and the development of visual expertise in the areas of dermatology, radiology, and pathology.

Cognitive processing models of visual diagnosis have been established in both radiology (mammography) and microscopic pathology (Azevedo & Lajoie, 1998; Crowley, Naus, Stewart, & Friedman, 2003). Both models divide the visual process into three primary components. Azevedo and Lajoie name these three primary components as: data acquisition, data exploration,

and hypothesis generation, while Crowley et al. designate these components as: data examination, data exploration, and data interpretation. Crowley et al. describe the third phase, data interpretation, as hypothesis formation and evaluation thus equating it with the hypothesis generation title chosen by Azevedo and Lajoie.

Both established processing models illustrate that high levels of knowledge (Level II and III in Bloom's taxonomy) are necessary for such visual categorization tasks. Level II knowledge, application and analysis, is clearly used during the data exploration process. For instance, Azevedo and Lajoie state that feature characterization and comparison occur during this phase. Crowley et al. identify many additional instances of knowledge application/analysis occurring during the data exploration phase. Some of these include (a) the association of findings with anatomic location and (b) the determination of finding importance and certainty. Level III knowledge, evaluation and synthesis, is required in the final phase of both cognitive models resulting in the generation of a diagnostic hypothesis. Crowley et al. further explain that after generation of a hypothesis, it may either be confirmed or disconfirmed by the presence and/or absence of supporting findings.

The two representative cognitive models of visual classification/diagnosis also embody the concept of meaningful learning. Progressive differentiation and integrative reconciliation are occurring during the processes of feature comparison and case level comparison as described by Azevedo and Lajoie (1998) and Crowley et al. (2003). Cognitive processing models were developed in both radiology and microscopic pathology for the purpose of informing the development of computerized tutoring systems in each area. Azevedo and Lajoie (1998) explain that the RadTutor system "provides extensive instructional scaffolding during the hypothesis generation phase to ensure that the user has proposed the appropriate hypothesis level" (p. 36).

The necessity of such differing levels of instruction in this computerized system is reflective of the subsumptive and superordinate learning processes that take place during meaningful learning.

Because such visual classification skills rely on the use of advanced levels of knowledge and the development of meaningful learning, assessments which require the visual classification of images inherently function as high quality forms of assessment. This research projects describes the process of visual categorization as it relates to the specific topic of white blood cell identification. The research project explores both assessment format and cognitive processing for two contrasting levels of expertise, the novice and the expert, in order to further describe the differential assessment outcomes/cognitive processes that take place during the development of expertise.

Currently, methods of assessment used in the area of clinical hematology do not precisely mimic the processes used in an actual hematological examination. This researcher explores the use of an alternative image presentation format during hematological assessment. Such an alternative image presentation format may provide options for improving standard assessment methods. In order to further explore the enhancement of hematological instruction and assessment, it is necessary to have a clear understanding of the cognitive processes that take place during the visual categorization of blood cells. Cognitive processing studies have been conducted in dermatology, radiology, and pathology. In such fields, the expert determines the diagnostic category by examining a patient's tissue specimen microscopically (as in the cases of dermatology or pathology) or by examining an X-Ray or skin lesion macroscopically (as in the case of radiology or dermatology). With such cases, experts typically examine a large area of the image and search for any notable abnormalities/lesions before they begin to make a diagnosis. If an abnormality/lesion is found, categorization of that abnormality/lesion then takes place. Visual categorization in clinical hematology has many similarities to the researched fields, but also has

notable differences. This research project only considers the practice of categorizing individual white blood cells. This process is somewhat different than process used in the researched fields because the expert does not have to "search" for specific microscopic fields of interest. Instead, experts move across a specimen slide in a very methodical manner, categorizing the first 100 white blood cells they view. The researcher theorized that the actual cognitive processes used to categorize a single hematological blood cell were very similar to those used in making visual diagnoses in dermatology, radiology, and pathology. Ultimately, this study establishes a specific and defined foundation for cognitive processing in the area of hematology and explores alternate/improved formats for use in hematological assessment.

Research Question

Main Question

How does the assessment design of digital image-based hematological competencies in white blood cell (WBC) identification affect the performance outcomes of experts versus college students and what are the cognitive and visual examination processes used by experts versus college students during WBC morphology identification?

Subquestions

1. What, if any, differential effect does competency test item format and image content have on competency performance outcomes for novice students versus expert professionals in clinical hematology?
2. What interactions, if any, are there between a subjects' response time for an item on a competency assessment and (a) performance outcomes on individual items (b) level of expertise (c) exam format?
3. What are the types of errors revealed during the process of white blood cell identification?

4. What are some explicit cognitive and visual examination processes that are used by students and experts to identify images of white blood cells?

5. How do the cognitive and visual examination processes used in the identification of white blood cell types differ (a) between experts and novices, or (b) within expert and novice groups themselves when image format is altered?

The concepts and methods implemented in this research study are illustrated graphically through the use of a Vee diagram (Novak & Gowin, 1984) as depicted on the following page. The left side of the diagram represents the conceptual or thinking side of the diagram, while the right side of the diagram represents the methodological or doing side of the diagram.

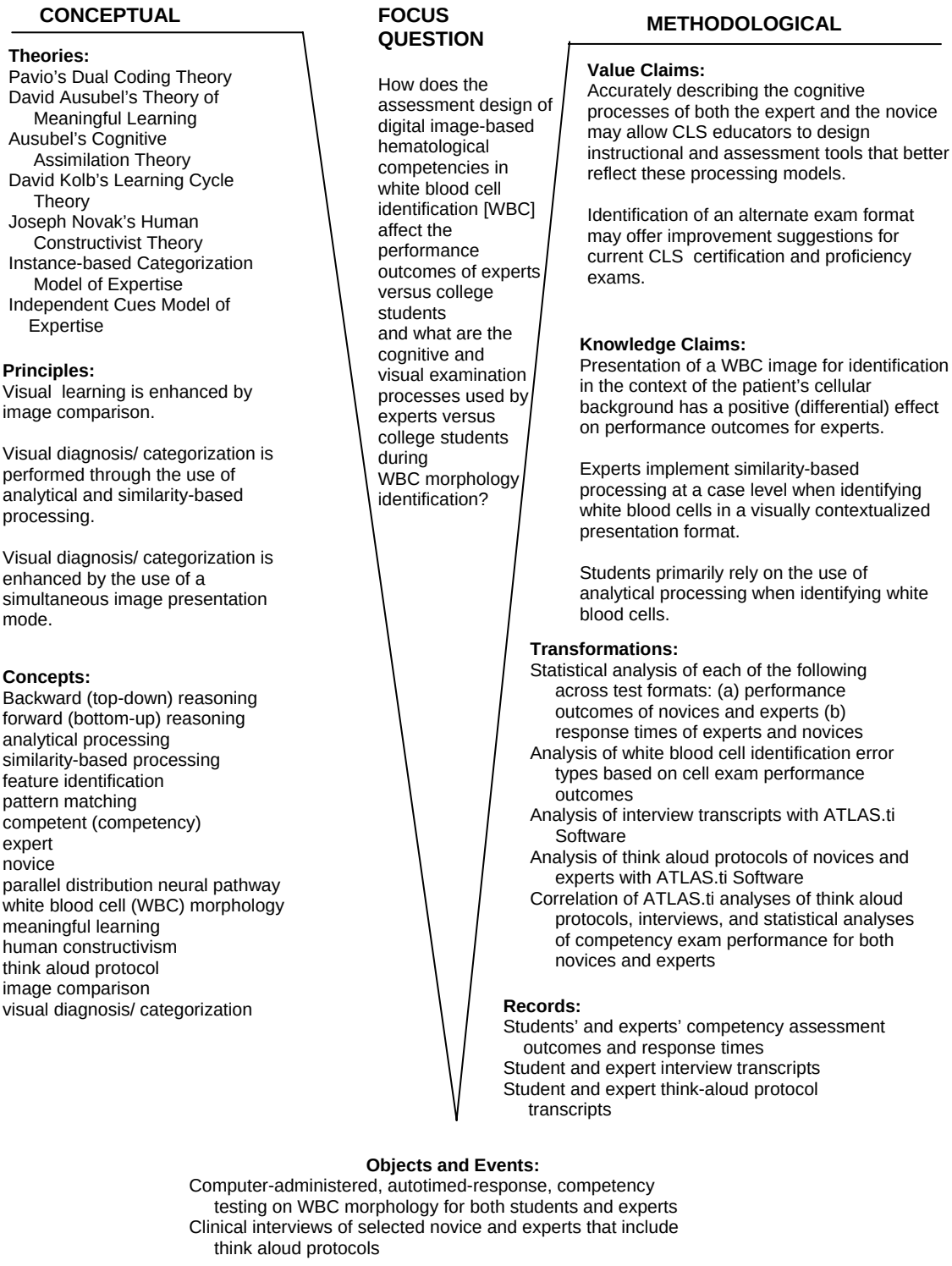


Figure 1 Vee Diagram

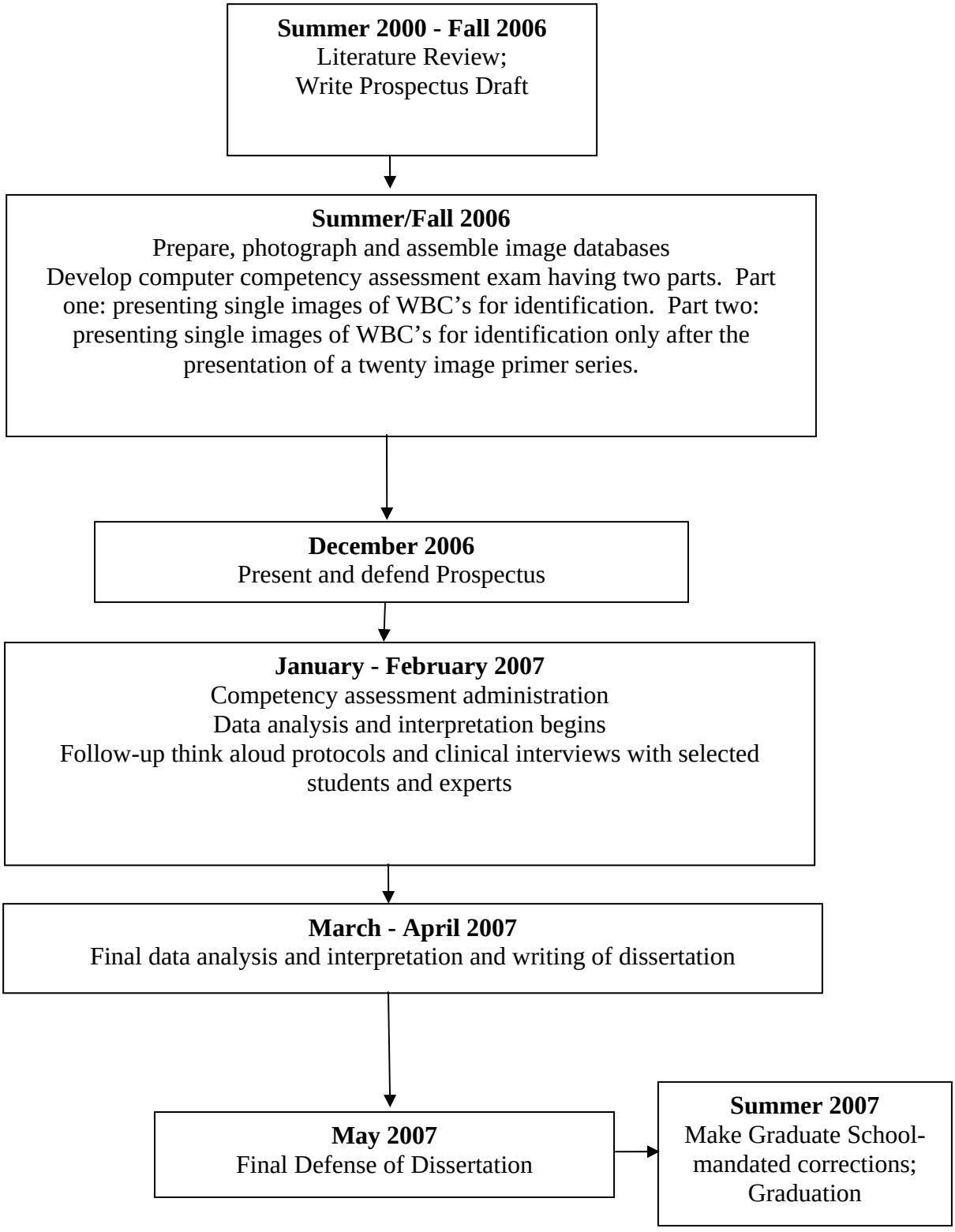


Figure 2 Research Timeline

LITERATURE REVIEW

Health Science Education Context

This educational research study is in the specific area of clinical hematology education. Clinical hematology education is situated in the much larger arena of the health sciences. The arena of the health sciences generally may include a large variety of professions. These professions are: (a) allied health professions, (b) nursing, (c) medicine (physicians), and (d) dentistry. Previously described visual diagnostic research in the areas of dermatology, radiology, and pathology are included as studies in the profession of medicine. This study falls under the venue of the allied health professions, specifically in the area of Clinical Laboratory Sciences (CLS). Allied health professions may include: (a) clinical laboratory sciences, (b) surgical technology, (c) physical therapy, (d) respiratory therapy, (e) emergency health sciences, (f) radiologic technology, and (g) physician assistant.

Researcher Intent

The researcher in this study is a clinical laboratory sciences educator who teaches both beginning and advanced courses in the areas of clinical hematology and clinical chemistry. Her background in basic sciences, clinical laboratory sciences, and curriculum and instruction allows her to consider both the educational and scientific aspects of this study. Her PhD curriculum coursework has taught her that assessment tends to drive instruction and has thus lead to her interest in this study. This research examines current trends in assessment in the area of clinical hematology. Such an examination requires a review of both the current standard in national certification examination of new professionals as well as yearly proficiency testing of existing professionals. The researcher in this study wishes to promote improvement in instruction and assessment in the area of CLS. She hopes that the findings of her study may be helpful in revising the high stakes certification testing currently given in the field by making it and other

examination forms better able to assess the visual skills of future CLS professionals. With this strong intent, the researcher suggests that every improvement in clinical assessment format leads to an increased quality standard in the laboratory field. This further assures that graduating students are entering the profession at a competent level of practice and are able to fulfill their critical role in maintaining quality patient care as a part of today's dynamic and essential health care team.

Scientific Background

Hematology: CBC Analysis

Hematology is the study of formed cellular elements, primarily those present in the blood (Harmening, 2002). The routine laboratory test used for the enumeration white blood cells (WBCs), red blood cells (RBCs), and platelets in the peripheral blood is the Complete Blood Count (CBC). Other hematological values reported as part of a routine CBC are the hematocrit, hemoglobin, and three red blood cell indices: (a) Mean corpuscular hemoglobin (MCH), (b) Mean corpuscular hemoglobin concentration (MCHC), and (c) Mean corpuscular volume (MCV). The CBC also routinely includes a WBC differential count in which white blood cells are enumerated by their cell type and the percentages of each representative cell type are reported.

The WBC differential allows for identification of the five common WBC types (neutrophil, eosinophil, basophil, lymphocyte, and monocyte), as well as any immature, abnormal, or atypical WBC forms (Harmening, 2002). Immature, abnormal, or atypical WBC forms are only found in the peripheral blood when a pathological condition exists. The immature, abnormal, or atypical forms which are enumerated include the blast, promyelocyte, myelocyte, metamyelocyte, plasma cell, and atypical lymphocyte. Today it is standard practice for an automated 5-part differential to be performed on all blood samples for which a CBC is

ordered. A standard 5-part differential is able to enumerate and differentiate the five "normal" WBC types. If the specimen is flagged for review as a result of distributional, morphologic, or instrument failure, a manual differential count will be performed. The first automated hematology analyzer to perform leukocyte differentials was the Techicon Hemalog D which was used at Mayo Clinic from 1975 to 1983. Before automated differential counts became standard practice, a 100- or 200-cell eyecount leukocyte differential was performed on all specimens requiring a CBC (Pierre, 2002). A recent study by the College of American Pathologists (Novis, Walsh, Wilkinson, St. Louis, & Ben-Ezra, 2006) determined that manual scans and/or manual differential counts are performed on approximately 16.2% of all specimens submitted for CBC testing. Specifically, manual scans are performed on approximately 6.5% of specimens submitted for CBC testing, and manual differentials are performed on an average of 9.7% of the specimens. This study found that the most common instrument flag resulting in a manual review is WBC values falling outside of the acceptable reference parameters. Of important note is that in approximately one-third of the cases, study participants discovered new information not available from the automated instrument results/findings upon manual review.

Manual Differential Count

Anticoagulated whole blood, collected in a 3 ml purple-topped vacutainer tube containing Ethylenediaminetetraacetic acid (EDTA) is the type of specimen routinely used for the CBC with differential (Turgeon, 2005). Commonly, the blood is prepared for manual analysis through the preparation of a blood smear on a glass slide using the push-wedge method. The blood smear is then stained using the Wright stain, a Romanowsky-Type stain containing eosin and methylene blue. The methylene blue stains the nucleus and some cytoplasmic structures of the leukocytes, while the eosin stains the other cytoplasmic structures a pinkish color. The cells located in the feathered edge of the smear are examined microscopically using the battlement examination

technique in which the smear is examined in a back-and-forth serpentine manner. The cells are classified by the laboratorian as to their cell type. Proper identification of WBC type is critical in the diagnosis of infection, leukemia, and other disease conditions.

There are, notably, four primary sources of error affecting the outcomes of a manual differential count: (a) observer errors (b) slide distribution errors (c) statistical sampling errors and (d) recording errors (Pierre, 2002). Statistical sampling errors account for the greatest portion of errors. This type of error exists due to the simple fact that the single stained blood smears examined are indeed a random sample of circulating blood leukocytes. Based on the use of 95% confidence limits for a 100-cell count differential, an actual lymphocyte count of 50% may be reliably reported as any percentage between 39% and 61%. Notably, cell types present in the lowest frequency are those associated with the largest amount of statistical sampling error.

Intraobserver variation between expert technologists is responsible for only a very small portion of the total error (Pierre, 2002). Research studies performed between the late 1970's and the mid 1980's determined error rates and types associated with the 100-cell eyecount and a number of 5-part automated differential methods using the National Committee for Clinical Laboratory Standards (NCCLS) 800-cell method as a reference. These studies showed that there was a 0% morphologic false-abnormal rate with manual differential counts. This indicates that expert clinical laboratorians do not misidentify normal cells as abnormal. However, in most cases, these studies did show manual counts to have higher error rates than the reference method or the 5-part automated differentials in the following three areas: (a) morphologic false-normal rate, (b) distributional false-normal rate, and (c) distributional false-abnormal rate. The increase in these three error rates was attributed primarily to the existence of sampling error. Tatsumi and Pierre (2002) explain that the century-old method of manual differential counting "has been generally accepted to be the most reliable standard method for diagnosis of hematologic

disorders with morphologic abnormalities. The method will continue to be used as a reference and diagnostic method as long as the analysis is conducted by experts, even if analysis methods may change year by year in various forms, such as histochemical, immunologic, chromosomal, and genetic" (p. 301). Other sources of error include use of the push-wedge preparation method and reporting formats (Pierre). The push-wedge method concentrates the leukocytes unevenly while differentially distributing specific cell types along the edges, center, and at the end of the slide. Reporting format leads to errors when absolute cell counts are reported instead of relative cell counts. Absolute cell counts are calculated values derived from the total WBC count and the individual cell type percentages. Reporting individual cell types using absolute counts compounds the error rate by introducing error from two separately measured parameters.

In addition to these errors, there are a few other disadvantageous aspects of the eyecount differential to consider. This process is labor intensive, taking from 1.9 to 6 minutes to complete a 100 count differential and requires highly trained technologists (Pierre, 2002). Even with the use of experienced and knowledgeable technologists, the tendency for some degree of interobserver bias among technologists on the criteria for cell identification and as a result, the classification of particular cells within a blood smear will always exist. In fact, a 1994 College of American Pathologist (CAP) report did demonstrate that even the expert technologist cannot reliably differentiate between the segmented neutrophil and the band. CAP's recommendation was that bands not be counted and reported separately from segmented neutrophils on manual leukocyte differentials. It is, however, because of the important role that the manual differential count plays in the diagnosis of leukemias and other disease conditions associated with atypical or abnormal cell morphologies, that the development of continued expertise in this area remains invaluable to the field of Clinical Laboratory Sciences.

Automated Differential Count

Although the leukocyte differential counting began approximately 125 years ago, the first commercial model focused on the production of white cell differentials was the Larc (Corning Glass) in 1974 (Tatsumi & Pierre, 2002). The Larc was based on image processing principles. Peripheral blood smears made with a spun or wedge method were stained using a Romanowsky-type stain. Images of the blood cells were taken using a high-resolution charge coupled device (CCD) camera. Instruments of this type digitized the leukocyte images, performed feature extraction, and matched image features to a cell library using cell identification software. The software then categorized the cells into the five major cell types. The system could analyze 25 – 40 slides per hour resulting in an approximate 1-2 minute differential count. Such instruments did prove to be as accurate and precise as the eyecount leukocyte differential count (ECLDC). One serious weakness of the image processing systems was its inadequacy to grade red blood cell and platelet morphology. The newest automated hematology analyzers produce five-part differential analysis by using a flow cytometer instead of the image processing principle (McKenzie, 2004). The flow cytometer operates on the principle of electrical impendence and optical light scatter and is able to determine cell size, cell complexity, nuclear lobularity, and cytoplasmic granularity. Examples of such instruments include the Cell Dyne 4000 and the Beckman Coulter.

Hematological Expertise

The field of CLS encompasses four main areas of practice: (a) clinical hematology, (b) clinical chemistry, (c) microbiology, and (d) blood banking. The majority of professionals working in the field hold either an Associate or Bachelor's degree. The Associate degree level professional may be referred to as either a Medical Laboratory Technician (MLT) or a Clinical Laboratory Technician (CLT). The Bachelor degree level professional may be referred to as

either a Medical Laboratory Technologist (MT) or a Clinical Laboratory Scientist (CLS). Clinical Laboratory Technician and Clinical Laboratory Scientist are the most up-to-date professional labels, but the alternate terms are still commonly used in job advertisements and other descriptions of the professional field.

Developing the expertise to work as a CLT or CLS requires a specific background in the area of clinical laboratory sciences. The educational training for such professionals most commonly begins with a National Accrediting Agency for Clinical Laboratory Sciences (NACCLS) accredited CLS program. Such programs are available at community, junior, or 4-year colleges/universities. There are also many hospital-based programs that fulfill these academic requirements. After completion of the appropriate degree program, the individual must then pass a national certification examination. In addition to passage of a national examination, all practicing CLS professionals in the state of Louisiana are required to be state-licensed by the Louisiana State Board of Medical examiners (LSBME). Further, medical technologists who have worked for five years in a single concentrated area such as that of hematology are then eligible to sit for a specialty board certification examination. Passage of such an examination earns the individual a specialist certification in the area of hematology.

Current Methods of Assessment in CLS

Because this research project explores assessments to be used in the area of clinical hematology, it is vital to explore the current forms of assessment used in the field of CLS for both graduating students and technologists/technicians currently practicing in the field.

Students completing a NAACLS accredited Medial Laboratory Technician or Medical Technologist program must pass a national certification examination. Two standard examinations are currently offered: (a) The Board of Registry (BOR) Certification given by the American Society of Clinical Pathology (ASCP) and, (b) Certification given by The National

Credentialing Agency (NCA) for Laboratory Professionals. Both examinations are criterion-referenced and computerized in their format. Criterion-referenced examinations have a predetermined minimum score that has been established and is required for passage of the certification examination. The performance of individual examinees is in no way influenced by the performance of examinee peers as is the case with norm-referenced forms of examination.

The ASCP BOR uses computer adaptive testing (CAT) (CAT examination administration and examination results section, para. 1). This testing format is adaptive because each exam given is tailored for the individual examinee based on their question-to-question examination performance. If the examinee answers a question correctly, the next question presented is slightly more difficult than the last. This pattern is continued, until the examinee incorrectly answers a question. The subsequent question presented to the individual is then slightly simpler. Thus, each examination is individual and unique and is appropriately matched to the individual's ability level. Each examination contains 100 questions and has a time limit of 2 hrs 30 min.

The NCA examination includes a total of 180 questions, 30 of which serve as a practice test and are not tallied to compute the final scaled score (NCA Candidate Handbook section, p.8). Students taking this examination have a total of 3 hrs to take the examination. The NCA does not use the CAT method of testing but does offer several different forms of the exam during any one particular examination period. With the use of such computer-administered tests, image-based questions are presented by displaying a single image alongside the appropriate question stem and answer choices.

One form of assessment required for practicing professionals is that which accompanies acquired continuing education hours as is required by most state licensure agencies, as well as the two national certifying agencies. A formal assessment process is involved in acquiring some forms of continuing education units (CEU's) (i.e. self-study courses in the area of CLS). For

example, in Louisiana, since January 1, 1995 every laboratory professional in the state has been required by to complete 12 continuing education credit hrs per year (Allied Health Continuing Education section, para. 3). In 1980, NCA established their recertification policy. All NCA-Certified professionals must undergo recertification every 3 years either by acquiring appropriate and acceptable continuing education credits or by re-examination (Re-Certification section, para. 2). In January 2004, the ASCP BOR established the Certificate Maintenance Program (CMP) requiring all laboratory professionals certified by the ASCP after January 2004 to participate and complete the certification program every 3 years (Certificate Maintanence section, para. 1). Laboratory professionals certified before January 2004 may participate on a voluntary basis.

While both NCA and ASCP BOR require the acquisition of 36 hrs of continuing education credits over a three year period, the specific requirements of each is unique. Both certifying organizations grant continuing education units for a variety of activities including participation in formal, pre-approved sources of continuing education or college/university coursework, presentation of peer-reviewed workshops or lectures, and publication in peer-reviewed journals. The NCA allows the professional to choose the distribution of all 36 hrs of continuing education, whereas the ASCP specifies that of the 36 acquired CMP points one be in safety, and two be in each of the following areas: (a) blood banking, (b) chemistry, (c) hematology, and (d) microbiology. The remaining 25 points can be earned in the area of the professional's choosing.

Another form of assessment in which laboratory professionals may participate is the proficiency surveys administered to all operating laboratories by the College of American Pathologists (CAP). The CAP proficiency survey which contains the "Blood Cell Identification" section is the FH3 Hematology Automated Differentials Survey and is administered 3 times a year. The three survey administrations are denoted as FH3 A-C.

The FH3 survey presents blood cell images in groups of five, each group being preceded with some clinical history about the patient as well as pertinent laboratory findings. Typically, background information includes patient's age, ethnicity, physical symptoms, past/current clinical diagnosis, and past/current laboratory results. Laboratory results typically include total white blood cell count, hemoglobin level, platelet count, and other relevant results. In more than 60% of the cases presented during the last 5 years of CAP survey (2000 FH3 survey C – 2006 FH3 survey B), the 5 photographs were of different white blood cell and/or red blood cell types. In the less than 40% of the cases, a single cell type was repeated more than once within the 5 photograph set. Typically, the image which was repeated had some unique diagnostic significance for the case and was a fairly rare and/or unique cell/image in and of itself. The repeated cells/images were usually not even a cell type found as a standard choice on a differential count reporting format.

Performance of each laboratory facility on blood cell identification items is evaluated based on a refereed model. Laboratories with a good history of performance on previous CAP surveys are asked to serve as referee laboratories in determining the "correct" answer (Carrie Gellings, personal communications, October 13, 2006). Typically, about 20-25 laboratories serve as referees. CAP designates the evaluation criteria for blood cell identification as a 90% referee or participant consensus. If this level of consensus is not reached for a graded item, it is not graded and falls into the "ungraded" or "educational" category instead. Images are presented to participants in one of two formats: (a) Kodachrome photomicrographs (represented by the abbreviation BCK), or (b) printed color photographs (represented by the abbreviation BCP). Performance on each cell/image type is reported separately. Examination of proficiency survey results from 2000C to 2006 B reveal the following:

(a) Neutrophils, basophils, and mature monocytes were most accurately and reliably identified (for >90% of that cell type a 90% consensus level was reached); these cells are all "normally" found in the peripheral blood

(b) Eosinophils, typical lymphocytes, atypical lymphocytes, and blasts were identified with moderate accuracy (for 60-80% of that cell type a 90% consensus level was reached); eosinophils and typical lymphocytes can be found in the normal peripheral blood, but atypical lymphocytes and blasts should usually not be found in the peripheral blood.

(c) Promyelocytes, myelocytes, and metamyelocytes were identified with poor accuracy (for 50% of promyelocytes a 90% consensus level was reached; the 90% consensus level was not reached for any of the myelocytes and metamyelocytes); none of these cells are found in "normal" peripheral blood.

Expertise

In nearly every domain of knowledge or skill certain people exist who have developed exceptional abilities. Such people stand out above the majority in their field and have earned the title and recognition as "experts". Research in the field of expertise knowledge has been of interest in cognitive science and psychology for over fifty years. Interest in the field began with the pioneering research of De Groot on chess expertise in 1946 (Ericsson & Smith, 1991). Ericsson states that: "On the most general level, the study of expertise seeks to understand and account for what distinguishes outstanding individuals in a domain from less outstanding individuals in that domain, as well as from people in general" (p. 1).

Definition

Expertise can be defined from a variety of perspectives. Expertise is often associated with age. Age in turn is related to the amount and type of experience in the field that an individual has accumulated (Hoffman, Shadbolt, Burton, & Klein, 1995). Multiple studies in the

area of expertise development including those in chess expertise by Chase and Simon, sports and the arts expertise by Hayes and Bloom, and international performance expertise by Ericsson and Crutcher suggest that about ten years of concentrated experience is necessary for international recognition in a field (Ericsson & Smith, 1991). In fact, many international "performers" enter their field before the age of 6 years old because it does take their entire span of development from early childhood into early adulthood in order to attain a level of expertise. An expert may also be defined by the extent of their memory as well as the organization of their memory. In addition, Hoffman explains that experts usually excel in professional criteria such as graduate degrees, publication record, membership in professional societies, training experience, and licensing. The definition of an expert as derived from a table of "guild" terminology may serve to sum up some of these key characteristics. The "guild" defines an expert as:

> the distinguished or brilliant journeyman, highly regarded by peers, whose judgments are uncommonly accurate and reliable, whose performance shows consummate skill and economy of effort, and who can deal effectively with rare or "tough" cases. Furthermore, an expert is one who has special skills or knowledge derived from extensive experience with subdomains (Hoffman et al., 1995, p. 132).

Ultimately, it is society that decides the basis upon which an expert will be judged (Hart, 1986). Our society chooses its experts based upon personal experiences of aid and exchange of information with such valued people. Expertise is also judged by the development of one's reputation as verbalized by others. In the end, an expert is chosen on the basis of what they can do with the special knowledge they have acquired. In order to be in this position of high esteem and value, Hart further explains that the expert must in some way serve to benefit society as a whole. We rely on experts to fulfill various roles in our society and to act as: (a) a provider of information, (b) a problem-solver, and (c) an explainer.

Acquisition/Development of Expertise

Although it is generally agreed upon that an expert exhibits outstanding performances in their particular domain of knowledge, the exact route whereby experts acquire their unique characteristics is in debate. These exceptional performances may be accounted for either by inherited characteristics, acquired characteristics, or a combination of both (Ericsson & Smith, 1991). Theories dictating that the capabilities for outstanding performance are primarily inherited suggest they may be due to general abilities such as intelligence and personality or to specific abilities such as music ability, artistic ability, or body build for athletes. Howard Gardner's (1983) theory of multiple intelligences suggests the existence of seven distinct and separate forms of intelligence: (a) linguistic, (b) logico-mathematical, (c) spatial, (d) musical, (e) bodily, (f) social, and (g) personal. One of the criteria which define these intelligences is that they must be apparent in select populations. Thus, it could be suggested that possession of a single very strongly developed intelligence could lead to uniqueness and particular expertise (Bruer, 1993).

Some studies have tried to establish that a general intelligence is associated with outstanding performances (Hart, 1986). Galton's 1869 study (as cited in Hart, 1986) examined the familial and genetic origins of eminent individuals in a wide variety of fields who were socially recognized. Galton theorized that the achievement of these individuals was due to both inherited intellectual ability as well as personal motivation. Those eminent individuals selected for the study seemed to come from a small number of families having common ancestors and thus eminence seemed to be genetically determined. Ericsson (2003) explains that "according to Galton, the relevant heritable capacities set the upper bound for the attainable level in physical and mental activities" (p. 96).

Some studies have tried to determine other relatively stable individual characteristics possessed by individuals of outstanding achievement (Hart, 1986). The best example is that of Cattell's 1963 research (as cited in Hart, 1986). He compared the personality profiles of top researchers in the fields of biology, psychology, and physics, with that of teachers and administrators in the field as well as that of the general population. The personalities of those categorized as top researchers showed some amazingly consistent traits. Their profiles found them to be self-sufficient, dominant, emotionally unstable, introverted, and reflective. These profiles exhibited not only unique abilities, but also a level of personal motivation associated specifically with their field of interest. Because the motivation was limited to the researcher's field of expertise, it is suggested that this aspect of the personality profile may be acquired. Despite the underlying implications of both Galton's and Cattell's studies, more recent research has been largely unsuccessful in identifying strong and replicable relations between general inherited characteristics and superior performance.

Because inherited characteristics alone seem unable to account for the superior performance of experts, the primary focus of the remainder of this discussion centers on the characteristics and problem-solving abilities of experts which seem to be acquired over time. Such knowledge and skill may be either acquired through general learning and experience or domain-specific training and practice. In order to determine which characteristics, abilities, and cognitive processes are unique to the hematological expert, one focus of this research was novice/expert comparisons. Novice/expert comparisons allow for the development of expertise models in domain specific fields. Possibly such models may be applied in the improvement of subject-specific teaching methods.

Levels of Expertise

Patel and Groen (1991) classified expertise into five different levels: (a) the beginner, (b) the novice, (c) the intermediate, (d) the subexpert, and (e) the expert. The beginner is someone with everyday, lay domain knowledge of a subject. As learners begin to gain prerequisite knowledge about a subject, they become classified as a novice. In the field of Clinical Laboratory Sciences, an example of a learner at this stage of expertise is a CLS student who is enrolled in the didactic or non-clinical portions of the program curriculum. A learner may be classified at the intermediate level of expertise when they are between the beginner and subexpert classification stages. An example of this is a new CLS professional who has just recently completed their CLS program curriculum. The subexpert has generic knowledge only and does not possess specialized domain knowledge. An example of this is a CLS generalist who performs hematology differential counts only on a limited basis during their daily task routine. The expert has developed very specialized knowledge related to the domain. A CLS technologist who works only in the area of hematology and has a specialist certification in the area or extended years of experience in the area may be classified as an expert.

Other systems of categorizing the development of human expertise have been developed. Dreyfus and Dreyfus (1986) identified five stages in the development of expertise: (a) novice, (b) advanced beginner, (c) competency, (d) proficiency, and (e) expert. In this model, the individual's performance undergoes a major transformation from the novice perspective of context-free and rule-dependent behaviors to the expert perspective of non-reflective, intuitively driven behaviors.

Human Memory

Traditionally, human memory has been divided into short-term memory (STM) and long-term memory (LTM) (Etelapello, 1998). Short-term memory allows for immediate free recall of

items from temporary storage, while LTM allows for retrieval of items only by retrieval cues from a more durable storage area. In George Miller's original 1956 work entitled "The Magical Number Seven, Plus or Minus Two" he establishes that STM, or as he called it at that time the "human channel capacity", has a limited capacity to store only about seven chunks of information at a time (Miller, 1994). These chunks of information that can be stored in STM, correlate directly with existing patterns of information currently held in LTM (Ericsson & Kintsch, 1995).

Expert Memory

An important explanation for the unique memory abilities of the expert involves the development of a long-term memory that is capable of extremely fast retrieval and encoding time (Etelapello, 1998). This unique form of long-term memory is developed only in the domain-specific context of the expert's field. After studies conducted in the late 1980's, Ericsson and Chase developed their skilled-memory theory to account for the exceptional use of an expert's long term memory (Ericsson & Smith, 1991). They suggested that during storage of domain-specific knowledge, experts develop stable retrieval cues for this information. After extensive practice with such cues, the expert's speed of retrieval and encoding when using long-term memory begins to approach the speed of using short-term memory.

Researchers have done various studies to support the fact that this unique memory ability is indeed domain-specific (Ericsson & Smith, 1991). The clearest example of such research was the classic chess research done by Chase and Simon in 1973 in which the superior memory of an expert for chess pieces on a board was tested. If chess experts were presented with a chess position for 5 seconds, their recall of that position would far exceed the recall of a novice. If, on the other hand, both expert and novice were asked to view random chess pieces in meaningless board positions for 5 seconds, the expert would recall no more of the positions than would the

novice. This illustrates that the expert's unique ability to remember large volumes of information is based solely on the recognition of specific chunks of information particular to their domain.

Long-Term-Working-Memory

Theories about human memory were expanded upon in 1995 when Ericsson and Kintsch suggested a third category of human memory besides STM and LTM called long-term-working-memory (LT-WM) (Etelapelto, 1998). The strongest evidence for this is research related to the planning of expert chess moves. The most demanding part of chess for working memory is selection of the next move. During this time, Ericsson and Kintsch suggest that the expert is using the LT-WM to store a long sequence of moves to follow. Some studies have even shown that the more advanced the chess skills of a player, the more elaborate planning that is possible.

LT-WM is termed expertise working memory (ExpWM) by Horn and Masunaga (2006). They recognize that ExpWM functions in the short term but explain that it differs from STM in four very distinct ways. First, the volume of information that can be stored in ExpWM is much greater. Studies have shown, in fact, that for chess experts the limit may be close to forty, as that is the number of possible chess move sequences they are able to mentally consider at a single point in time. In fact, experts show up to a 10-fold increase in performance on tasks in which their acquired memory skills allow their LTM to function in this special short-term capacity (Ericsson & Kintsch, 1995).

Secondly, multi-tasking within their domain of expertise is second nature to an expert even in the face of disruption or distraction (Horn & Masunaga). Ericsson and Kintsch note the ability to continue multiple tasks even after an interruption. They state that information held in the ExpWM (or LT-WM as they termed it) will remain in LTM during the interruption period and can be accessed again simply through reactivation of the appropriate retrieval cues in STM.

The expert cannot, however, multitask efficiently in areas outside their domain of expertise (Horn & Masunaga). For example, chess experts can easily play more than one game at a time without losing track of the moves within any one game. If, however, that same chess expert is posed with the seemingly simple task of determining the meaning of a sentence while also remembering the last word of the sentence, retention of the last word is quite difficult.

A third way in which ExpWM differs from STM is that the order of recall for information is quite flexible. Typically, recall of information is much simpler if it is recalled in the same order in which it was presented. For experts, however, the order of presentation does not seem to matter. The typical memory limit for items to be recalled in the reverse order from that of their presentation is four plus or minus one. Chess experts have been shown to recall sequences of game moves in the reverse order almost as easily as they recall the same moves in the forward sequence. Norman, Brooks, and Allen (1989) also found this to be true in the field of medical diagnosis. When laboratory test results were presented to experts in both an organized, routine manner and a scrambled pattern, the experts recall was unaffected by the presentation format.

Fourth, information held in ExpWM is being held long term but can be recalled more quickly than information held in short term, even when recall is requested unexpectedly (Horn & Masunaga). For example, chess experts can unexpectedly recall the moves in a chess match much more easily and accurately than someone can unexpectedly recall the digits of a phone number just dialed for the first time.

Metacognition

Also very primary to the superior abilities of an expert is the use of metacognition. "Metacognition is the ability to think about thinking, to be consciously aware of oneself as a problem solver, and to monitor and control one's mental processing" (Bruer, 1993, p. 61). Using metacognition allows for self-monitoring by experts. They are able to make more precise

evaluations about their own problem-solving skills than can the novice. Experts seem more aware of when they make errors. In 1978, Simon and Simon noted that physics experts are more likely to double check themselves for mistakes (Etelapelto, 1998). The work of Glaser and Chi in 1988 showed that physics experts were better able to accurately judge the difficulty of problems and therefore more appropriately divide their time. The 1987 study of expert chess players illustrates the self-awareness of experts. Chess experts were more accurate than the novice in predicting how many times they would need to see a chess board before being able to reproduce it. (Etelapelto, 1998).

Some suggest that metacognition is a general skill that can even be used to improve novice performance across domains. Bruer suggests the existence of the intelligent novice who differs from the typical novice based solely on their ability to utilize these metacognitive skills. He claims this technique may even allow the novice to learn a new domain more quickly (Bruer, 1993).

Cognitive Processing

Much of the available evidence seems to indicate that experts solve problems using much different techniques than does the novice. Various terminologies can be used to describe these different problem solving methods, some of which include: (a) deep vs. surface processing, (b) backward vs. forward reasoning, and (c) weak vs. strong methods. But, as a general rule the problem-solving techniques of the expert tend to follow more abstract and complex lines of thinking than does that of the novice.

Although abstraction is commonly associated with expertise and is often seen as beneficial in superior reasoning, the exact level of abstraction must be appropriate for the particular domain. Colleen Zeitz (1997) explains that "a moderately abstract conceptual representation (MACR) is formed through the encoding of the current situation in relation to the

categories that are functional in the domain" (p. 44). Thus, part of becoming an expert is defining the appropriate MACR for processing in the specific domain of interest. Abstraction plays such an important role in expert knowledge because it allows for their complex organization of knowledge. Abstraction allows experts to more easily see patterns in data and information based on their broad range of prior experiences.

A first example of the use of abstraction is seen in the abilities of experts to be deep processors, in contrast to the superficial focus of the novice. A classic example of abstraction can be illustrated by the research done by Chi, Glaser, and Rees in 1982 (Bruer, 1993). The novice and expert physicist were both asked to sort textbook problems into categories based on solution methods. The novice categorized the problems by the objects and features directly mentioned in the problem situation. One such category used for classification was the inclined plane problem. The expert, on the other hand, grouped the problems according to the physical principle or law involved such as Newton's second law of motion regardless of the surface features involved. Similar results were found in the 1983 study by Weiser and Scherz in regards to computer programming (Etalapelto, 1998). When asked to categorize programming problems, the expert utilized solution algorithms, while the novice sorted by area of application. Both examples "indicate that both novices and experts have conceptual categories, but that the experts' categories are semantically or principle-based, whereas the categories of the novices are syntactically or surface-feature oriented" (Etelapelto, p. 39).

Expert Models for Visual DiagnosticTasks

Identification of white blood cells when performing a hematological differential count is a type of visual classification problem solving. Psychological theory and medical decision making research have shown two different processes to be vital in such classification processes: (a) analytical processing and (b) similarity-based processing (Kulatunga-Moruzi, Brooks, &

Norman, 2001). Analytical processing has also been referred to as the "Independent Cues" interpretation while similarity-based processing has been called "Instance-Based Categorization" or "pattern matching" in previous expertise studies and psychological literature (Norman, Rosenthal, Brooks, Allen, & Muzzin, 1989; Norman, Brooks, Allen, & Rosenthal, 1990).

With the use of analytical processing, one makes use of specific clinical features in determining a differential diagnosis or performing a categorization (Kulatunga-Moruzi et al., 2001; Norman et al., 1990). Learners build expertise by acquiring knowledge about those features most useful in such differential determinations. With practice, learners acquire the ability to appropriately weight various features and thus determine the proper differential diagnosis. Analytical processing or "Independent Cues" is a forward reasoning model which suggests that expertise involves the mastery of a complex set of rules.

With similarity-based processing, "learned rules" function as a schema to initially define a category, but such rules are quickly replaced by individual instances or prior examples/cases (Kulatunga-Moruzi et al., 2001). This processing strategy relies on a backward reasoning approach in which a diagnostic/categorization hypothesis is formed first, based on the similarity between previously encountered examples held in memory and the current case. Expertise is developed as the learner builds a vast repertoire of prior examples. Processing using this strategy proceeds in a holistic fashion thus allowing for the unconscious detection of patterns. Dreyfus and Dreyfus (1986) refer to this type of reasoning as "holistic similarity recognition". They explain that segmentation of a pattern into its feature parts does not occur and that rules are not needed.

Both processes have been found to play critical roles in even the most foundational clinical reasoning models. The exact roles of each process in clinical reasoning and the development of visual expertise have been studied in a variety of areas including dermatology,

radiology, and pathology. Much of the early research in these areas present the two processes as competing, but some recent research suggests that the processes may act in a complementary manner during the development of expertise.

Analytical Processing

One early study by Patel and Groen (1986) in the area of visual diagnosis suggested that a forward reasononing model was primary in expert medical reasoning. Patel and Groen used propositional analysis in order to isolate causal networks from the protocols of seven cardiologist specialists during their evaluation of an acute bacterial endocarditis case. They found that in all instances in which the cardiologist expert yielded an accurate diagnosis they relied on an entirely forward reasoning or "bottom-up" pattern. Those experts, on the other hand, who yielded inaccurate diagnoses were found to use a "top-down", backward reasoning pattern during at least some portion of their case analysis.

Other studies outside the area of medical diagnosis have also supported the use of forward reasoning by experts. Larkin and Chabay (1989) examined the problem solving process used by both novice physics students and physics experts (Bruer, 1993). When asked to solve for a specific variable, experts were able to reason in a forward manner from the givens to the goal, based on their qualitative understanding of physical principles. Novices, on the other hand, searched their memory for physics laws or equations containing the variable in question. After choosing an appropriate law, the student started with the desired unknown and tried to work backward to the information given in the problem. Larkin and Chabay concluded that the novice lacked scientific reasoning knowledge and expert schemas. Because the novice lacks the crucial intermediate piece of domain-specific knowledge, he cannot effectively use the forward reasoning process when solving scientific problems as the expert can.

In contrast to the Patel/Groen and Larkin/Chabay studies, the findings of a 1983 radiologic study lead Kundel and Nodine to hypothesize that the use of a "top down" or backward cognitive processing model is primary with experts. The findings suggest that picture perception begins with a general global analysis and that an accurate visual concept for radiologic abnormalities may only be formed if previous examples of the abnormality have been encountered.

Similarity-Based Processing

Evidence for similarity-based processing in clinical reasoning and diagnosis exists in many different areas (Norman & Brooks, 1997). These areas include: (a) evidence of the dichotomous relationship between level of expertise and case performance, (b) evidence of reasoning by both experts and novices that is based on the impact of similarity to prior examples and the specific impact of prior instances on expertise, (c) evidence that experts cannot predict the error of other experts or novices, (d) evidence that individual features are re-analyzed during problem-solving/diagnostic processing, and (e) evidence of rapid, "automatic" or instantaneous processing by clinical experts.

Level of Expertise Versus Case Performance. Evidence of the dichotomous relationship between level of expertise and case performance supports the similarity-based model of processing. Theoretically, if one were to accept the opposing analytical processing model as the governing model for expertise development, one would expect case performance to follow a predictable pattern that would correspond with the individual's level of expertise. Using the analytical processing model, Norman et al. (1989) predicted that expert performance on easy/typical cases should improve much more rapidly than performance on atypical/difficult cases since typical cases exhibit most of the classic features which define a category and atypical cases are more likely to "break the rules". Specifically, Norman et al. predicted that error rates

on typical dermatologic slides should interact with expertise (experts should make relatively fewer errors on typical slides than do those with less expertise). Norman et al.'s study on dermatological expertise was unable to support this prediction. The ratio of errors on typical (easy) slides to total errors remained constant for all levels of expertise (about 40%), despite the fact that the total number of errors did decline as the participants expertise levels increased. Thus, if expert pattern recognition or similarity-based processing is indeed free of detailed feature analysis, it follows that there should be no expected relationship between lesion typicality and expected ease of improvement in diagnostic skill.

Unpredictability of Errors. As shown by the 1989 Norman et al. study, the types of errors exhibited by experts and novices in the field of medical diagnosis is often times unpredictable. Although it is not surprising that it is difficult to predict the errors of a novice due to the inconsistency in their knowledge base, it seems reasonable that a domain expert could reliably predict the types of errors that would be made by another domain expert. This assumption was not supported by the 1989 Norman et al. study. In fact, it was found that when expert dermatologists were asked to predict the errors of fellow experts, 21-64% of the time the erroneous diagnoses suggested were never even mentioned by fellow experts. In only 8-38% of the cases did the first-choice suggested erroneous diagnoses match with the errors of fellow experts. Although the analytical processing model would suggest that performance can be predicted based on typicality, the similarity-based cognition model suggests a strong dependency on prior experience and specific instances. If the similarity-based cognition model is indeed the governing model for developed expertise, the individuality that exists between the sets of prior experiences for each expert would explain the unpredictability of the error types.

Impact of Prior Examples. The debate over the role of prior examples in the learning process is essential in identifying the cognitive process taking place during the development of

expertise. In an "Independent Cues" model, prior examples serve strictly as a means by which to develop and learn how to appropriately apply weighted feature rules. This model assumes that the individual examples themselves have no lasting or profound effect. In contrast to this theory, the "Instance-based" approach supposes that specific prior examples are of particular importance, especially in areas of medicine that rely heavily on the visual domain. During the diagnostic categorization tasks that take place in such arenas, the identification of features may, at times, be somewhat ambiguous. This is especially the case in the area of clinical hematology in regards to the identification of immature and abnormal white blood cells. WBC are characterized based on a variety of features including cell size and shape, cytoplasmic and nuclear color, nuclear shape, nuclear-cytoplasmic ration, chromatin texture, and presence/absence of cytoplasmic granules and nucleoli. The co-existence and configuration of these features is often times very critical. Viewing a variety of example instances has been shown to be critical in the development of visual diagnosis (Brooks, Norman, & Allen, 1991).

The impact of prior examples on the accuracy of dermatologic diagnosis was studied by Norman et al. (1990), Brooks et al. (1991) and Allen, Norman, and Brooks (1992). These studies were able to demonstrate the importance of exposure to multiple, diverse, examples during the development of expertise. The vital role that such prior examples play should not be underestimated. Norman et al. (1990) found that subjects were able to rate the plausibility of differential diagnoses for various test phase dermatological slides much more accurately if they had previously seen a similar slide having the identical diagnosis during the instructional phase. If, however, the slide used during the instructional phase for that diagnosis category was dissimilar to the slide used in the test phase, diagnostic performance declined. Test slides for diagnostic categories studied in the learning phase but dissimilar from the prior examples were identified at approximately the same rate as test slides for diagnostic categories never presented

during the learning phase. This suggests that prior examples have no generalized effect unless they are similar to test phase items. An extension of this study (Brooks et al., 1991) further showed that the use of similarity-based processing does not interact differentially with the level of expertise or with the particular diagnostic strategy used. All levels of experts asked to use two dichotomous diagnostic strategies continued to be affected more strongly by prior instructional slides that were similar rather than dissimilar to the test slides. Allen et al. (1992) tested the effect that prior examples introduced during the learning phase had on the diagnosis of "chameleon" items during the testing phase. "Chameleon" items are items that have some ambiguous features causing their diagnosis to fall into two potentially plausible categories. If the photographic images used during the instructional phase were biasing toward the "correct" diagnosis, then the subject identified the "chameleon" correctly during the test phase. If, however, the photographic images presented during the instructional phase were biasing toward the incorrect diagnosis, the rate of accuracy in diagnosis of the "chameleon" was decreased by approximately 50%. This phenomenon was shown to persist even when the test phase took place an entire week after the practice phase.

Feature Reinterpretation. Use of a non-analytical processing model suggests that features are not merely detected during initial examination of the case itself, but also result from the clinician's individual interpretation of the problem. The clinician's interpretation can be influenced differentially by various factors including clinical histories and comprehensive feature lists. Hatala, Norman, and Brooks (1999) found that presentation of clinical histories in conjunction with ECG influences the diagnostic accuracy of ECG interpretation. Each ECG case used was a bit ambiguous and could be diagnosed with either the "correct" diagnosis or a plausible "alternative" diagnosis. Clinical histories stating the patient's age, gender, referring physician, and referral diagnosis accompanied each ECG. Some histories were consistent with

the "correct" diagnosis and others with the "alternative". The study found that the nature of the clinical histories biased both the diagnostic accuracy of the case as well as the ECG features listed to support the chosen diagnosis. Thus many of the features identified in these cases were chosen in the light of the biasing diagnosis and were not based on the presentation of the ECG itself. The findings of this study mirrored those of any early study using chest radiographs (Norman, Brooks, Coblentz, & Babcook, 1992) in which prior clinical histories for difficult/ ambiguous cases of bronchiolitis also greatly affected both the case diagnosis as well as the feature list generated.

Kulatunga-Moruzi, Brooks, and Norman (2004) state that "clinical features are not self-evident givens but rather are extracted and interpreted in light of the diagnoses being entertained" (p. 563). Kulatunga-Moruzi et al. performed a study in which they discovered that there is "additional information in the perceptual manifestations of the feature that is critical in accounting for diagnosis and categorization" (p. 570). For their study, they generated comprehensive lists of features for photographs associated with dermatological or internal medicine cases. They found that when an expert was presented with the comprehensive list of features before being shown the actual photograph, it decreased the expert's diagnostic accuracy. Diagnostic accuracy levels were much better when experts saw only the photograph. Apparently, consideration of all present clinical features before actually being able to form a diagnostic hypothesis was detrimental to visual categorization accuracy in this situation. Since the case descriptions did contain features both relevant and irrelevant to the correct diagnosis, it is hypothesized that the clinician may have committed to a plausible, yet inaccurate diagnosis upon initial examination of the verbal information. Apparently, the verbal description drew attention to features that would not have been attended to during the normal perceptual processing of the visual stimuli. Acknowledgement of such features increased the difficulty of

the visual categorization task. Even with presentation of the photograph and the apparent visually contradictory evidence, the clinician was not able to overcome the initial erroneous hypothesis. Although it is true that clinical information may be given in both the informational and perceptual forms, a significant conclusion that was drawn from this study was that the perceptual form is the one that provides the discriminatory and specific information needed.

Other studies, such as that of Norman, Brooks, Colle, and Hatala (2000) also support the idea of feature re-interpretation after the formation of an initial diagnostic hypothesis. The advantage of using backward, diagnosis-driven reasoning by the novice was discovered. Use of this reasoning process led to more accurate and specific searches for features and relevant data when composing a feature list for a proposed diagnosis. Like the experts in the Kulatunga-Moruzi et al. (2004) study, novice participants found that once irrelevant features were identified using purely forward reasoning, they were difficult to dismiss when considering a final diagnosis and often contributed to the selection of an incorrect diagnosis.

Automatic/Instantaneous Processing. The "Instance-Based" (similarity-based) model of expertise can be even further supported through the examination of studies that support the automaticity of expertise through analysis of response times and processing timelines. Early studies in the area of dermatology such as Norman et al. (1989), examined the response time of participants. Rapid response times were associated with correct expert responses and slow response times were associated with incorrect expert responses. This response time difference indicated that two different cognitive processes may be taking place. An automatic "pattern-recognition" process may account for the rapid response times of correct identifications. This process allows the dermatologic lesion to be considered as a whole, not on a feature-by-feature basis. An analytical process involving feature-by-feature analysis, may account for the slower

response times of incorrect identifications. The analytical cognitive process may only be applied by the expert when the "instance-based", "pattern-recognition" process has failed.

Also in support of instantaneous expert automaticity, and therefore the departure from feature analysis, are the findings of Crowley et al. (2003) in the area of pathology. Through the use of a combined videotape and verbal protocol analysis, Crowley et al. was able to examine the process timeline that takes place during microscopic pathology diagnosis. Instantaneous processing by experts was indicated by the fact that experts very rapidly identified the anatomic location and began hypothesis formation in comparison to both the intermediate and the novice. In regards to hypothesis formation, experts verbalized the specific hypotheses eventually accepted as the final diagnosis much earlier in the examination process than did the intermediate or expert. In addition, their first statement of hypothesis and their initial statement of final hypothesis occurred in a very short time span. The rapid speed associated with focal lesion detection, hypothesis formation, and anatomical site identification may indicate that rapid instance-based classification can account for the expert speed and performance exhibited in the area of microscopic pathology. These findings substantiate those from Norman et al. (1989) dermatologic study performed more than a decade earlier.

The Continuum of Processing Model

More recent studies support the fact that there may be an interconnective role between these two processing models and even a probable interaction with level of expertise. Two separate studies (Regehr, Cline, Norman, & Brooks, 1994; Kulatunga-Moruzi et al., 2001) in dermatologic diagnostics support this conclusion. Both studies examined the effect that altering instruction during the test phase had on the processing mode used by the study participants. There was an important difference between the participant groups chosen for the two studies. Regehr et al. used medical residents, while Kulatunga-Moruzi et al. used medical students. Both

studies focused on the use of observable outcomes to evaluate the relative use of the two processing models rather than the evaluation of verbal reports. The difference in diagnostic accuracy between typical and atypical cases was used as a measure of analytical processing. The difference in diagnostic accuracy between similar and dissimilar cases was used as a measure of similarity-based processing. The 1994 study, using residents, noted a large similarity effect for both participants given test phase instructions intended to foster analytical processing as well as those participants given test phase instructions intended to foster similarity-based processing. The large similarity effect on those given analytical based instructions was unexpected. The apparent steadfastness of similarity-based processing, even with the deliberate intervention suggests the critical nature of this process in the development of visual expertise. The study also noted a strong typicality effect, but only for those given analytical based instructions. Researchers theorize that as clinicians increase their level of expertise, they shift from the use of analytical processing to the use of a more similarity-based approach. Since the Regehr study participants were residents moving toward an advanced level of expertise, it was assumed that they currently made use of similarity based processing. Their performance could only be differentially improved by giving analytical based test phase instruction. Kulatunga-Moruzi et al. findings supported this theory, by noting a large typicality effect for both instructional formats. The similarity effect in this study was larger for the group given instructions to promote similarity-based processing. It is theorized that because these participants were more novice than in the Regehr study, their primary mode of daily processing was analytical. Thus, only similarity based test instructions could have a differential effect on performance.

Crowley and Medvedeva (2006) further refine the classification problem solving model in the field of microscopic pathology. This refined model depicts a five-step process and spans all 3 expertise levels: expert, intermediate, and novice. This is a developmental model depicting

transitional skill acquisition. It is quite representative of the continuum of processing model suggested above. The five steps in the refined process are identified as: (a) search and detection, (b) feature identification, (c) feature refinement, (d) hypothesis triggering, and (e) hypothesis evaluation. The acquisition of skill progresses on a very clear continuum from novice to expert. The novice is described as having very weak skills in each of the five areas. They possess very limited abilities to evaluate the hypotheses they form because they lack the developed knowledge needed for backward reasoning. The expert has very accurate skills in all five areas and is able to develop an extremely focused set of hypotheses which can then be analyzed using backward reasoning.

Digital Imaging

The popularity of digital cameras has increased since their initial introduction in the early 1990s (Riley, Ben-Ezra, Massey, & Cousar, 2002). Since their introduction, digital cameras have become more technologically sophisticated and economical. With such advances, the digital camera can be more easily and effectively used in areas such as hematology and hematopathology. Commonly imaged specimens include physical lesions, gross lymph node specimens, stained tissue sections of lymph nodes and spleen, and Wright-Stained peripheral blood smears. Lee (2005) states that "of the different disciplines in pathology, hematology has among the most demanding requirements in terms of high image resolution" (p. 151).

Types of Digital Images

Digital images of microscopy can take multiple forms ranging from a high-resolution digital image taken by a digital still camera to a high-resolution, "real-time" image as captured by a digital video camera. One of the more advanced types of digital images that can be created is that of the digital slide.

Virtual microscopy, also called digital microscopy, digital pathology, and virtual pathology, allows examination of digital slide images while simulating the functional features of a real microscope (Lee, 2005). Virtual slide viewing software allows users to pan, zoom, and focus while viewing a digital slide. The panning feature allows virtual microscopy to simulate the field selection capability of a real microscope. Skill in microscopy is developed by allowing the user to select which microscopic fields should be reviewed when analyzing essential diagnostic features of the slide as a whole. The focusing and zoom features allow the user to closely examine particular details of the slide (Hutchinson, Brereton, & Burthem, 2005). A virtual slide is a large panoramic digital image prepared from a single, purposively selected area of the original glass slide. Such large digital images can be created either through the use of a virtual slide scanner or by the stitching/merging of multiple images of sequential, adjacent high power fields. In order to support the focus function, multiple images of the same slide must be taken in varying focal planes thus forming multiple image layers or z-stacks. Such z-stacks allow the user to focus up and down through the various planes of the image. During evaluation of a hematological slide, examination of a single cell layer (focal plane) is usually adequate. Exceptions occur with the examination of bone marrow samples and in the identification of red blood inclusions. Such specimens often require the examination of multiple focal planes in order to achieve adequate cellular detail (Lee, 2005).

Applications

Initially, digital images were used largely and primarily in medical education, but applications for these images are rapidly increasing. Applications in the area of hematology include education, proficiency testing/training, telemicroscopy, and value-added, image-enhanced specimen reporting.

Education. In relation to clinical hematology, the current area of most widespread use is that of education. Digital image atlases are readily available both on the internet and on CD-ROM. Examples of internet sources include the Atlas of Hematology of Nagayo University School of Medicine (http://pathy.med.nagoya-u.ac.jp/atlas/doc/atlas.html), the Bloodline Image Atlas (http://image.bloodline.net/), the American Society of Hematology Slide Bank (http://ashimagebank.org), the WebPath Resource Collection of the University of Utah (http://www-medlib.med.utah.edu/WebPath/HEMEHTML/HEMEIDX.html), and Hemo-Surf-An Interactive Hematology Atlas of the University of Bern (http://www.aum.iawf.unibe.ch/HemoSurf/english.htm) (Riley et al., 2002). Examples of CD-ROM image collections and tutorials are those offered through University of Minnesota's Hematography collection (www.umn.edu/hema) and CACMLE (http://www.cacmle.org/).

Proficiency Testing/Training. Current proficiency testing of hematological morphology is accomplished primarily through the use of 35mm Kodachrome slide images by the College of American Pathologists (Lee, 2005). Some proficiency testing accomplished through the use of the original Kodachrome slide itself and some accomplished through distribution of printed copies of the digitized image. Glass slides, however, are used in some other countries (i.e. the United Kingdom National External Quality Assurance Scheme and the Royal College of Pathologists of Australia). Both forms of media have apparent disadvantages. Kodachrome slides display only a limited field of view, while glass slides can never be reproduced in duplicate exactly and can only be produced from a single sample of peripheral blood in limited quantities.

The use of digital "virtual slides" in the role of quality assessment surveys has been piloted through the UK NEQAS (H) (United Kingdom National External Quality Assessment Scheme for General Hematology) (Burthem et al., 2005). In this pilot study, four different cases

previously provided on glass slides to survey participants were converted to the virtual slide format. One-Third of the UK NEQAS registrants participated in this pilot by assessing the digital slides in the same manner as they would a typical glass slide by listing the five morphological features of the digital slide which they determined to be most important. The results of the study showed very high agreement levels between the previous survey results and the pilot study results. Ideally, future widespread application of virtual microscopy in the area of proficiency testing will eliminate some of the current problems previously discussed.

Virtual Microscopy has also been tested in the area of cytopathology proficiency testing for the Papanicolaou (Pap) test. (Marchevsky et al., 2003). Participants in traditional cytopathologic proficiency testing as administered by the College of American Pathologists are routinely mailed a set of five glass slides of cervical/vaginal material, four times every year. This study used 10 conventional Pap cases and compared the performance of three cytotechnologists and two cytopathologists on both the glass slide and virtual slide. The glass slide was presented to the subjects for examination approximately one year after participant evaluation of virtual slides. Participants were asked to determine the most accurate diagnostic code for each case. All study participants interpreted the diagnostic code correctly for the glass slide. Both cytopathologists interpreted the virtual slide correctly as well. Two of the cytotechnologists earned an 80% on the virtual slide portion while the other cytotechnologist earned a 70%. This study suggests that the use of virtual microscopy is not adequate in the area of cytopathology for proficiency testing requiring diagnostic coding of Pap smears. Previous literature in the area of cytology (Vooijs et al., 1998) suggests the importance of instituting training programs which use new technologies. If technologies such as the virtual slide and digital imaging are incorporated during training, it will assure that all participants feel

comfortable with such new technologies. The use of such tutorials may be critical to ensure the success of virtual proficiency testing in the future in areas of cytopathology as well as others.

Telemicroscopy. Digital imaging also has a dramatic effect on patient care. Telemicroscopy or telepathology is the "sharing of microscopic images via a telecommunication device for remote primary diagnosis, expert consultation, and consensus diagnosis, case conferencing, quality assurance, or education (Riley, Ben-Ezra, Massey, Slyter, & Romagnoli, 2004). Telepathology can be either static or dynamic. If static, only selected digital images are communicated between individuals. Static telepathology is often used as an application for education, quality assurance, and expert consultation. It is infrequently used in remote diagnosis since an accurate diagnosis can not typically be derived by viewing only a few isolated images of a specimen. Dynamic telepathology, on the other hand, involves the continual interactive transmission of images between two parties. Dynamic telepathology allows the off-site consultant to remotely control the microscope at the transmitting site. It has lead to the determination of remote diagnoses in cases which used images of frozen sections and surgical pathology specimens.

Value-Added, Image-Enhanced Specimen Reporting. Value-Added, image-enhanced specimen reporting produces written reports which may be supplemented with the addition of digital images as well as patient and specimen demographics, specimen gross/microscopic diagnosis, final diagnosis, additional laboratory results/studies, and references/relevant Web addresses. Someday, this application may allow for the inclusion of an entire digitized slide in a patient report. This would allow physicians/pathologists/medical technologists to easily compare images from various parts of a single slide. If desired, they could also compare images from two sequential slides on the same patient. The ability to manipulate, adjust, and annotate such digital images will ultimately revolutionize the current medical field. Professionals in the fields of

pathology and clinical laboratory sciences, could, eventually be able to function microscope-free (Riley et al., 2002).

Basic Principle

The digital camera that was used in this research project in order to capture high-resolution digital microscopic images was the SPOT digital camera by Diagnostic Instruments' Insight. SPOT computer software (Diagnostic Instruments, 2002), and an IBM personal computer, Pentium III processor with at least eight GB hard drive and 64 MB of RAM was used to edit the digital images when necessary. This type of digital camera has a CCD (charge-coupled device) image sensor. Silicon photodiodes which cover the image sensor collect photons of light and generate an electrical charge. The cumulative voltage signals are collected and ultimately converted to discrete binary numbers through a process called photoelectric conversion. Red, green, and blue color filters are used in conjunction with the silicon photodiodes in order to make the photodiodes capable of detecting color (Riley et al. 2002). The SPOT camera uses "three shot" digital technology. A rotating filter allows the recording of separate images with all three of the color filters resulting in a vivid, high-resolution microscopic image with high color fidelity.

Advantages

It is clear that the use of digital imaging technology offers many advantages when used in educational settings, proficiency testing, or for the enhancement of patient care (Hutchinson et al., 2005). Most importantly, the resolution achieved with high-quality digital imaging equipment now approaches that of high quality film. The high color quality of digital images does not degrade over time, as does that of glass slides. Additionally, photo-editing software allows for even further adjustment of image color, contrast, etc. in order to achieve optimal

image balance. Such software is also useful for annotating and labeling features of digital images.

Secondly, digital images are economical to produce as well as easy to store and share with others via e-mail, CD ROM, etc (Hutchinson et al., 2005). This allows identical images to be viewed simultaneously at two different physical locations or at two different points in time. It also allows for easy inclusion of digital images in professional presentations and conference seminars.

Third, the panoramic display of images through the use of a digital (or virtual) slide has many additional advantages over that of the single digital image (Hutchinson et al., 2005). The larger the panoramic view shown on the virtual slide the more limited is the effect of selection-bias as it exists with singly selected digital image series. Very large digital slides make almost every feature of the original glass slide available to the viewer. Additionally, virtual slides having z stacks may be coupled with appropriate software to allow for multiple plane focusing and panning (as described earlier), replicating the use of an actual microscope.

Limitations

Despite the many advantages of using the digital image, there are some disadvantages as well (Hutchinson et al., 2005). Creation of the optimal image using a digital camera does take some amount of technical expertise in using the camera software and hardware. Although the creation of virtual slides provides a valuable resource, the slides are somewhat time-consuming to construct and result in the formation of extremely large image files that may require a specialist form for viewing. Additionally, the ability to focus z stacks through virtual microscopy requires dedicated viewing software.

Visual Images

Picture Perception

The perceptual process involved in the analysis of images is quite complex. While some contend that the perceptual process is strictly stimulus-driven and is governed directly by the image observed being "recorded" by the viewer's optics, many others from the "constructivist" viewpoint disagree (Levie, 1987). They argue the importance of the individual's role in constructing meaning from the observed image. They suggest that image meaning is constructed based on each individual's personal experiences and preconceptions. Solso (2003) supports the constructivist viewpoint and explains that perception is motivated by hypothesis testing. Schemas which are individually unique provide the context in which images are interpreted through a top down model of information processing. Solso defines schema as "part of one's mental framework for representing knowledge: specifically, we use the term here for how one might represent an array of interrelated concepts in a meaningful organization" (p. 223). Such schemas are applied to the interpretation of concepts and images in a variety of fields ranging from art to science and are fundamental in the development of individual representations.

Levie (1987) describes the process as being composed of three primary components: (a) attention and scanning, (b) interpreting significant figures and cues, and (c) perceiving global meaning. Attention and scanning occur by combining the processes of foveal fixations and eye pattern movements. Typically, foveal fixations occur for a period of approximately 300 ms per individual fixation. The specific location of each foveal fixation distinctly affects how individuals interpret an image and how it is encoded into memory. Saccade is the term which refers to the extremely rapid eye movements that separates individual foveal fixations.

The study of eye pattern movement has been of great interest particularly in the development of visual expertise in the area of radiologic diagnoses. Nodine, Kundel, Lauver, and

Toto (1996) examine the development of visual search strategies as a function of expertise. The two processes that occur during a visual search are that of the global overview and the focal feature analysis. During the global overview pattern analysis takes place. Subsequently, during the focal analysis feature integration occurs. Study participants varied in expertise, ranging from laypersons lacking both training and experience to mammographers and mammography technicians possessing both training and experience. They were asked to search images of chest radiographs in order to identify the presence of abnormal breast masses/nodules. The later group was the only group that was able to detect the breast lesions with accuracy. They were also the only group that accurately identified mass-free images.

The results of this study can be explained by carefully examining the concept of selectivity as it relates to the process of a visual search. Solso (2003) contends that although human beings are capable of taking in information presented in a variety of different sensational formats, the brain will only focus its attention on particular selected items from such vast sensational information. "Perception is very selective. We attend to only a few of the sights, sounds, and smells available……in our environment" (Fleming & Levie, 1978, p. 7). Zull (2002) further explains that the expert is able to discern which features of the image are important and which are not; to the novice all features of the image have equal significance. Nodine et al. (1996) concluded that the value of radiologic experience is not in the development of search patterns, but in the interpretation of the visual targets once they are fixated. Participants who lacked experience detected candidate targets, but could not accurately discriminate between true and false targets.

Eye movement patterns are affected by a great many factors including viewer expectation, purpose of observation, expertise level, age, and cultural background (Levie, 1987). Francis Dwyer (1978) performed research which evaluated the use of a progressive set of images

of the human heart ranging from a simple line drawing to an actual photograph of the organ. His research showed that the novice (i.e. student) can be very easily overwhelmed by exceedingly realistic/ highly complex images. In such instances, the novice tends to develop a "scanning syndrome". This syndrome is a result of the severe problems the novice has in attending to and interacting with the relevant details of the image. As Dwyer defines it, this syndrome involves the "constant surveillance of the entire perceptual field while not focusing or interacting with any specific stimuli" (p. 6). Dwyer also suggests that if the image is too complex, the novice may contend with each stimulus individually after first surveying the field. Alternatively, they may process the perceived stimuli by creating categories or groups. Undesirably, the learner may simply disconnect themselves from the learning process until the image has left his/her perceptual field.

Cognitive Processing

Concrete experiences result from a person's physical interaction with the world around them. These interactions serve as the primary source of sensory input collected by the back cortex (sensory and postsensory cortex) of the brain (Zull, 2002). The majority of concrete experiences are visual. Images are the easiest form of input for the human brain to remember, thus making them the most effective form of input as well. Zull suggests that there is no limit to the number of pictures that can be stored in our brain. Images that are created from our day to day concrete experiences contain data from all of the senses, and are referred to by Zull as "sense-luscious". He suggests that, if possible, all ideas should be converted into the form of an image.

Visual information is processed through a mechanism known as parallel distributed processing (PDP) (Solso, 2003). This processing mechanism is able to explain why humans are able to recognize and classify visual images so quickly. The many neurons in the visual cortex

of the brain are acting in parallel, rather than in series. Specific features of the image are processed in parallel and then the features are reassembled to form a memory of the original image (Zull, 2002). Many areas of the visual cortex are activated upon the viewing of an image. This simultaneous (or parallel) activation of many different neurons in the brain explains why the examination of an image takes only minutes. In fact, it has been established that the recognition-response time for objects which are highly familiar takes approximately 600-800 milliseconds or less than one second.

Magnetic resonance imaging (MRI) and positron emission tomography (PET) have been used to study how the specific parts of the visual cortex operate and where they are located. Two pathways of information processing have been identified in the visual cortex (Zull, 2002; Solso, 2003). The "where" pathway, which is located in the dorsal part of the visual cortex, depicts spatial information about an object including features of depth, direction, and location. The "what" pathway, which is located in the ventral part of the visual cortex, depicts object classification information including image form, color, and specific information allowing facial recognition. The specific part of the brain that is associated with the processing of color is called the "human V4". It is located more medially than most of the other image processing areas. It is located toward the center of each hemisphere of the brain.

Dual Coding Theory

Alan Paivio's (1991) dual-coding theory (DCT) of information explains the importance of visual images within the structure of human knowledge. The DCT plays a critical role in the study of visual learning because it was the first systematic, objective approach to the study of imagery and its functions. This multiple coding theory describes the mind's representational system and suggests the existence of two separate subsystems: (a) a verbal coding system for storage of verbal/linguistic information and (b) an imaginal coding system for encoding of

nonverbal information (pictures, sounds, smells, and touch). Highly imaginal (concrete) information is encoded in both verbal and imaginal systems. Although separate, these two subsystems work in an interconnected fashion allowing activity in one subsystem to activate activity in the other and allowing for transformations of memories between the two subsystems to occur. The additivity hypothesis proposed by Pavio, suggest that for concepts which are dually coded, the image and verbal codes have additive effects. This hypothesis provides a theoretical explanation for both the superiority of concrete words over abstract words during free recall exercises as well as the superior memory humans have for pictures. The picture superiority effect is justified by Pavio through the fact that dual coding occurs most automatically and easily for pictures. When a picture is viewed, most individuals represent the picture verbally by naming the picture/object as a part of the learning process. In fact, this dual representation system which is used to encode named pictures seems to work equally as well with the recall of imagined words. Use of the image or pictorial code has been found to almost double recall over the use of verbal code only. Verbal processing and visual imagery even have different constraints of organization. Verbal processing must occur sequentially, whereas visual imagery can be processed in the order of individual choice since the entire image itself is available simultaneously.

Mayer and Sims (1994) researched extensions of the dual-coding theory as it related to multimedia learning. Multimedia learning occurs when information is presented visually by animation and verbally by narration. The mind then builds representational connections, building internal representations of the externally presented material in working memory. After the individual builds separate mental representations of both the verbal system and the visual system, referential connections (structural relations) between the two representations are formed. Student performance, as it results from this learning experience, can be evaluated based on the

retention and transfer of conceptual information. Mayer and Sims propose that in order for meaningful learning to occur, leading to the solving of transfer problems, the existence of all three connections (visual representational, verbal representational, and referential connections) are required. Their experiments showed that students are able to build these referential connections and generate creative solutions to problem sets more easily when verbal and visual information are presented simultaneously (contiguously) rather than successively. It was also shown that low-experience learners, having only a small amount of domain specific knowledge, perform significantly better when verbal and visual explanations are presented concurrently. Domain specific knowledge has been shown to help high-experience learners to somewhat compensate for lack of coordination in synchronous instruction. Thus, they do not encounter a differential level of performance in contiguous verses successive modes of instruction. A third aspect examined was the interaction of student spatial ability with the contiguity effect. The contiguity effect was found to be strong for high-spatial ability students but not for low-spatial ability students. A student's increased spatial ability was able to enhance the coordination of verbal and visual instruction. In conclusion, the students most likely to benefit the most from synchronized multimedia instruction are low-experience, high-spatial ability students.

Color

Human beings are able to perceive color in the portion of the electromagnetic spectrum designated as the visible spectrum (Solso, 2003). The visible spectrum extends from a wavelength of 380 nm (deep violet) to a wavelength of 780 nm (red). The human eye is extremely sensitive to color. Tufte (1990) states that, amazingly, a trained colorist can discern between approximately 1,000,000 different colors when asked to differentiate between paired colors in a laboratory setting. The average viewer can discern approximately 20,000 colors.

Vision takes place through use of the structures in our eyes called rods and cones (Solso, 2003). Rods are most useful in their role to detect the black/gray/white stimuli. They are most sensitive to light of wavelength 500 nm. Cones, on the other hand, are most sensitive to light at a wavelength of 550 nm, corresponding to the visible color of yellow-green. It is necessary for colors having wavelengths at the far ends of the visible spectrum to be more intense in order to be easily detected by human vision.

Humans have a trichromatic color vision system that utilizes three different types of cones (Solso, 2003). The three different types of cones differ based on the type of photosensitive pigment they possess. The wavelengths of maximal absorption efficiency for these three cone types are: (a) violet (419 nm), (b) green (531 nm), and (c) yellow-green (559 nm).

Also fundamental when using color in images is color's function. Early studies in the 1950's and 60's focused on the use of color in instructional material presented strictly for informational purposes (Dwyer, 1978). The content materials examined in these studies were not designed to meet any particular educational goal or objective. These studies did not show a clear advantage for using color in the enhancement of instructional materials. Classic studies performed by Dwyer in the 1970s did, however, convincingly illustrate that color did indeed have a significantly positive effect on instructional effectiveness. In fact, the analysis of over one hundred contrasts between color images and identical black and white images showed the significant effectiveness of the color image in every situation.

Color has many functions toward improving instructional effectiveness. These functions include: (a) directing attention, (b) increasing motivation, (c) eliciting emotional response, (d) cueing or coding, and (e) information design (Dwyer, 1978; Goldsmith, 1987; Tufte, 1990). Studies have indicated that color is helpful in drawing the viewer's attention to particular

properties of an object. Color may also aid the viewer in detection of interrelationships or in making fine discriminations (Peeck, 1987). Color has been found to be especially useful in making images more attractive and motivating. This is most helpful for students while reading textual information, particularly if they are low-ability students. Dwyer mirrors the benefits of color by stating "Color not only makes illustrations attractive and emotionally appealing, but it can make them instructionally more effective in facilitating student achievement of specific kinds of learning objectives" (p. 139).

In his 1990 text *Envisioning Information*, Tufte focuses his attention on the use of color in information design. He contends that color can serve in four major capacities in information design. It can be used to: (a) label, (b) measure/quantitate, (c) represent/imitate reality, and (d) enliven/decorate/beautify. He explains that the color in an image must be translated by the viewer into quantitative data and that for each viewer a slightly different perception will ensue. Care should be taken when using color, however, because color adds greatly to the complexity of an image. Increasing the complexity of the transmitted information may overwhelm the viewer's processing ability. It is particularly important that care is taken when viewing time is limited (Peeck, 1987). When applying color to an image the primary constraint is that of human visual memory not the actual ability to discern between color variations (Tufte, 1990). In fact, it has been found that if more than 20 to 30 colors are used to encode abstract information, the use of color may actually have an inhibitory effect on the learning process. It is apparent that applying color to an image is a very complex process that requires purposeful and careful selections. Studies have shown that arbitrary use of color or use of a poorly planned design for color application can easily detract from the learner's instructional gains (Goldsmith, 1987).

Realism

Images and pictures can be portrayed in a variety of forms, but in the study conducted by this researcher the stimuli source used was that of digital photographs. Four major photographic styles are described by Wandersee (2000). These styles include realism, expressionism, formalism, and instrumentalism. Realism allows for nature to be represented in its true to life form. Expressionism is representative of the photographer's own personal experiences. Formalism derives its substance from the form of the photograph instead of from the particular object/topic being photographed. Instrumentalism is used when communicating moral, social, or economic messages. It is noted that the most common type of photograph used in educational materials for biologic sciences is that of the realistic format. The common use of this realistic format has an apparent logic, since the central focus in biology issues is the true nature of living organisms.

The role of realism during instruction was examined closely by Dwyer (1978). He defines realism by explaining that an image which is completely authentic would be so exact in its quality that it would be indistinguishable from the object itself. In a series of detailed studies performed by Dwyer, he introduced four different image types each depicting a human heart (Peeck, 1987). The images were presented in conjunction with a standard 2,000 word text about the heart's structure and function. The four image types included: (a) a simple line drawing, (b) a detailed, shaded drawing, (c) a photograph of an anatomical heart model, and (d) a photograph of the actual organ.

From Dwyer's (1978) studies he found that there is a curvilinear relationship between the amount of realism in a picture/illustration and the amount of measured learning. With the most realistic and complex image types, the amount of stimuli may overwhelm a student, inhibiting the novice's ability to identify stimuli of central interest. Images that are two highly realistic can

actually reduce student learning. At the other extreme, line drawings may not provide enough needed stimuli. This suggests that the best type of picture for use in a learning setting may be a hybrid (photograph/drawing combination).

Several other factors should be considered when using realistic images in an educational setting. First, the amount of image viewing time available affects the effectiveness of using realism in instruction (Dwyer, 1978). With externally paced conditions, which allow only a limited viewing time, pictures having lesser amounts of realistic detail are desired. If the instructional pace is set by the student themselves, providing greater amounts of realistic detail is desired. Under these conditions, the student is able to examine the additional detail provided. Another factor in a student's ability to learn effectively from realistic pictures is the student's prior knowledge and ability to intake the detail of illustrations. A final factor to consider when deciding the degree of realism appropriate for use is the educational objective to be achieved.

Image Comparison: Presentation Modes

Until the late 1960's most of the images presented through media such as film, television, and slides were displayed in a sequential fashion (Perrin, 1969). One of the early theories about the important role of simultaneous images and image comparison in instruction was described by Perrin in his publication, *A Theory of Multiple-Image Communication*. Perrin suggests that when images are presented sequentially, they function in a similar way to verbal language. Several consecutive images/pieces of information must be brought together in order for meaning to be established. Simultaneity, on the other hand, allows the images to interact upon each other at a single moment in time, thus facilitating image comparison. He also notes that the use of simultaneity results in a subsequent increase in information density. Certain visual information may be learned much more effectively if presented in a simultaneous format. Perrin states: "The theory of multiple image suggests that for making contrasts and comparisons, and for learning

relationships, simultaneous images reduce the task of memory (a dimension of visual task) and enable the viewer to make immediate comparisons" (p. 376). Research evidence indicates that use of large images and multiple images further enhances the advantages by using simultaneity. Millard (1964) as quoted in Perrin (1969) underscores the useful role of simultaneity by stating: "Dichotomies, alternative, differenced, likenesses, and many other forms of comparison can likewise be efficiently handled by this method" (p. 369).

Image comparison can be instituted through two different mechanisms: simultaneous (parallel in space) or sequential (parallel in time) presentation modes. Tufte discusses the advantages and disadvantages of both forms of presentation in his 1997 text *Visual Explanations.* Tufte explains that simultaneous image presentation can be described as parallel in space because they are presented in close proximity to each other and appear in a single visual field. Tufte explains that "spatial parallelism takes advantage of our notable capacity to compare and reason about multiple images that appear simultaneously within our eyespan. We are able to canvas, sort, identify, reconnoiter, select, contrast, review ways of seeing, all quickened and sharpened by the direct spatial adjacency of parallel elements" (p. 80). He refers to sequential presentation modes as parallel in time, since the viewing of images occurs segmented by time. Because the images are viewed in two separate presentation fields, the viewer must remember the first image and compare to it the second image. This is often found to be a challenge. One form of this sequential presentation mode has been used by British architect, Humphry Repton. The first image is drawn on a flap of paper and the second image is drawn on the paper to which the flap is attached, directly beneath the flap. Repton flips the flap in order to compare before and after images of architectural reconstruction projects. Tufte explains that rapid "flap flipping" creates an almost simultaneous presentation mode. Tufte feels that this mode of viewing enhances the observed differences between the presented images. It eliminates the necessary

back-and-forth eye movement required to compare images presented adjacently. Even with this advantage however, Tufte still feels that simultaneous comparisons are the most effective.

Visual comparison made via the use of parallel images allows the viewer to note like components of the images including similarities in content, position, or image orientation. Images are most effective when they are presented in a physical manner most appropriate for the instructional purpose (Zull, 2002). Zull explains that the physical arrangement of images stimulates particular neural networks. For example, if features of two images are to be compared, then the images should be presented side by side instead of in series. When images are presented in this simultaneous fashion, the neuronal network for comparison is stimulated. The structure of the brain's neuronal networks reflects its function in many other ways, one of which is illustrated with the use of metaphors. The concepts compared when using metaphors are represented in the brain by neuronal networks that are very similar in their physical structures.

A classic text which examines the cognition of images is *The Psychology of Illustration*. In it, Fleming and Levie (1978) summarize a series of research-based principles about the use of illustrations in instruction. Two of the principles apply directly to the use of images for the purpose of comparison. The first principle states: "Learning to associate or relate two or more objects/events (stimuli and/or responses) is facilitated where they occur or are encountered in contiguity, that is, close together in time or space" (p. 142). The second principle explains that these objects/events "will tend to be perceived as somehow related. Comparisons will be facilitated, both similarities and differences becoming more apparent" (p. 144).

Use of image comparison during the learning process is necessary to develop certain aspects of visual expertise, specifically that of feature recognition. Kim and Astion (2000) examined the modes of image comparison chosen by one hundred and fifty-four second-year

medical students at the University of Washington when working with the user-controlled compare and contrast feature of the Urinalysis Tutor (UAT). All medical students were required to use the UAT as part of their Urinary System Course. Three image-viewing modes were available in the Urinalysis Tutor: (a) single image viewing, (b) paired viewing, and (c) anchored viewing. With anchored viewing, "a single image in one panel is an anchor against which multiple image comparisons were made using the second panel" (Kim et al., p. 349). The study found that the most chosen viewing mode was anchored (41%) followed by single viewing (22%) and paired viewing (11%). Students who used the anchored-viewing mode attained the highest post-test scores, although mean scores were not significantly different from those of students who only used the single image viewing mode. T-Test analysis showed, however, that those students who used the *Compare and Contrast* feature in the crystal section of the Tutor, regardless of the viewing mode, performed significantly better on post-test analysis than those who did not use the feature ($p<0.015$). Although the users did perform better than non-users in the cell and cast sections of the posttest, a significant difference was not noted ($p>0.015$). Researchers felt that the crystal section may have been the only one that led to significant improvement because it contained the widest variety of images.

Images and Assessment

Unfortunately, a large majority of the tests given in the realm of education are verbal in nature and multiple-choice in format. Although multiple choice questions can be constructed to reflect knowledge at the interpretation and problem-solving levels, this is a very difficult task to accomplish. It is also very difficult to adequately test practical skills learned in the science laboratory especially those involving the acquisition of visual identification skills. Sadler (2000) suggests that one of the best ways to increase the difficulty of multiple choice questions is to

include distractors that are derived from common student misconceptions. Such misconceptions may be identified during a clinical interview or through a literature search.

Creation of an image-based test may be a much more appropriate and reasonable way to assess higher levels of laboratory learning and scientific understanding. Image-Based tests are most appropriately chosen when the educational objectives are themselves visual in nature. In such cases, Francis Dwyer (1978) has shown that superior student performance results with the use of visual test forms (drawing/identification tests) instead of non-visual test forms. The image type used in this research study was that of the digital photomicrograph. Photographs have a great many characteristics that make them desirable for use in high-quality and challenging forms of assessment (Wandersee, 2000). Photographs are most useful in the field of science because they act as a form of memory storage. The detailed visual quality of photographs stores details of an object which the human memory could not otherwise specifically store. Images provide a visual peg to which one can anchor other concepts, principles, and theories of relevance in the long-term memory (Pavio, 1991). It is important, however, to assure the use of novel images in assessment. Using images never previously analyzed by the learner assures the elimination of a rote learning effect. If images are used heavily in the learning phase, Wandersee (2000) cautions the learner about the danger of overgeneralization. The learner must keep in mind that they are viewing only a small subset of possible images and that any additional viewed images will continue to have some unique quality. Thus, choosing the appropriate images for learning and assessment can often be very challenging.

METHODS

Institutional Review Board

A request for exemption status for this research project was approved by the LSU Institutional Review Board on October 25, 2006. A copy of this IRB exemption is provided in Appendix A. The consent form signed by the participants in this study is provided in Appendix B. The Certificate of Completion of Human Subjects Protection Training as provided by the National Institute of Health (NIH) is provided in Appendix C.

Sample Population

The 36 participants in this study were composed of two different subpopulations: (a) 14 Clinical Laboratory Science students from Our Lady of the Lake College (OLOLC), novices in the area of clinical hematology, and (b) 22 state-licensed and nationally certified medical technologists or medical technicians who have been practicing professionals for a period of 5 years or more with a concentration in the area of hematology.

OLOLC is a small, private Catholic college with a total student population of approximately 2,000 students. The College's primary educational focus is in the area of health care careers. Table 1 represents a four semester average (Spring 2005, Fall 2005, Spring 2006, Fall 2006) of the College statistics describing the student population of OLOLC based on gender, ethnicity, religion preference, age, and marital status.

OLOLC is divided into three primary undergraduate schools: (a) the School of Health Sciences, (b) the School of Nursing, and (c) the School of Arts and Sciences. The CLS program is included in the School of Health Sciences. The School of Health Sciences typically accounts for about 9.1% of the student population at OLOLC.

Table 1[a]

Population Characteristics of Our Lady of the Lake College Students

Population Characteristic	Percent Population[b]
Gender	
Female	77.2
Male	22.8
Ethnicity	
Black	15.2
American Indian	1.1
White	77.7
Asian	3.5
Hispanic	1.4
Religion	
Catholic	45.6
Protestant (Christian/ Non-Catholic)	5.6
Jewish	0
Other	30.9
Unknown	18.0

[a](Twelfth Class Day Consensus Reports. Retrieved October 18, 2006, from http://www.ololcollege.edu/12thclassday_files/12th_Class_day_files.htm)

[b]Average percentages for the two-year period described have been rounded to the nearest one-tenth percent. Because the values in this table do represent averages for multiple semesters, their totals may not be exactly 100%.

Table 1 Continued

Population Characteristic	Percent Population[b]
Age	
1-17.99	0
18-19.99	1.6
20-21.99	14.2
22-24.99	29.6
25-29.99	25.7
30-34.99	12.9
35-39.99	6.4
40-49.99	7.4
50-59.99	2.1
60.99.99	0
Marital Status	
Divorced	7.8
Married	24.6
Single	65.5
Unknown	6.7
Widowed	0.4

[a](Twelfth Class Day Consensus Reports. Retrieved October 18, 2006, from http://www.ololcollege.edu/12thclassday_files/12th_Class_day_files.htm)

[b]Average percentages for the two-year period described have been rounded to the nearest one-tenth percent. Because the values in this table do represent averages for multiple semesters, their totals may not be exactly 100%.

The Clinical Laboratory Science students who participated in this study were those enrolled in the CLS Bachelor Degree and Clinical Laboratory Technician (CLT) Associate Degree programs at Our Lady of the Lake College in Baton Rouge, Louisiana. Students accepted into the Bachelor Degree Program have completed 2 years of general arts and sciences courses including at least 16 semester hours of biological sciences, 16 hours of chemistry, two semesters of English, college algebra, statistics, and computer science, along with required social science, humanities, and additional elective courses. Students accepted into the Associate Degree Program have completed 1 year of general arts and sciences courses including 10 semester hours of biological sciences, 8 semester hours of chemistry, two semesters of English, and college algebra, along with the required social science, humanities, and additional elective courses.

The Clinical Laboratory Sciences program curriculum is identical for both degree programs for the first 2 semesters or twenty hours of courses. The first 2 semesters of the program encompass basic coursework in all of the major areas of Clinical Laboratory Sciences including clinical microbiology, immunohematology, clinical chemistry, and clinical hematology. The clinical hematology course instructs students in the basics of red blood cells, white blood cells, and platelets including all major pathological hematologic conditions. This course includes an introduction to the performance of manual differential counts. The program curriculums for the Bachelor degree and Associate degree students diverge after the first two semesters of coursework, with the Associate degree students beginning clinical rotations in a hospital setting and the Bachelor degree students continuing with advanced level didactic courses in the area of CLS for one more semester.

Students who participated in this research study had all completed the twenty hours of basic CLS coursework. Quantitative and qualitative study procedures were completed with both

experts and novices during the first month of the semester during which Associate degree students were enrolled in clinical rotations and Bachelor degree students were enrolled in advanced didactic coursework. Study procedures were implemented before any form of novel, additional instruction/laboratory practice was provided for these students about white cell morphology. A brief, standard review was provided to all students in order to ensure that they remembered the basic terminologies and descriptions for all of the major white blood cell types.

In summary, the student population that participated in this research study included 3 Associate degree students and 11 Bachelor degree students. It may be important to note that one of the Bachelor degree students had transferred from another CLS program within the state. Students with such altered course sequences and background/experiences do tend to provide a degree of heterogeneity to the student population at OLOLC. Students in the study population ranged in age from 24 to 50 years of age and included 12 females and 2 males.

The second group of participants was medical technologists and medical lab technicians currently working at Baton Rouge area hospitals. These hospital sites included: (a) Our Lady of the Lake Regional Medical Center, (b) Woman's Hospital, (c) Oschner Medical Center, (d) Baton Rouge General Hospital, and (e) Neuromedical Center Hospital. The 22 laboratory professionals had an average of 15.2 years of practicing experience, with years of experience ranging from 5 years to 33 years. Five of the expert participants held a specialist certification in the area of hematology. The population included 17 females and 5 males.

Pilot Study

Before preparation of images for the cell exam, a pilot test was performed. Three expert professionals with 5 or more years of experience in the specialized area of hematology participated in the pilot study. The pilot study served two purposes which included: (a) determination of time estimates for expert identification of approximately fifty cells, and (b)

verification of the existence of a rate of disagreement amongst experts on certain single cell image identifications. Photographs were taken of fifty microscopic fields from a blood smear of a leukemic patient. These fifty fields contained a total of fifty eight white blood cell images. Participating experts were asked to identify each cell and rate cell identifications as easy, average, or difficult. The table located in Appendix D displays the cell identifications and the difficulty ratings given by each of the three experts who participated in the pilot study. In this pilot study, the three experts reached a consensus agreement for 58.6% of the cells. For 29.3% of the images, 2 out of the 3 experts agreed. For the other 12.1% of the cell identifications, there was no consensus among any of the three experts. The pilot study also found varying levels of agreement amongst the experts in their difficulty ratings. Expert participants reported that it took them an average of fifteen to twenty minutes in order to identify the fifty eight cells. This pilot study indicated that a reasonable level of disagreement amongst expert technologists about the identification of certain white blood cell images does exist. This laid the foundation for the current research study since, in part, it intended to determine if altering the image presentation format could indeed increase expert consensus rates/expert accuracy.

Competency Exam Description

The competency exam was assembled using digital images taken by the researcher with a SPOT® Diagnostic Corporation microscopic camera. Visual Basic® programming software was used to write the examination program itself.

Images for the competency examination were prepared in the following described manner. Twenty Wright-Stained blood smears, each prepared from a different patient, were photographed using the SPOT® Diagnostic Corporation Microscopic Camera. These blood smears were collected from patients who had a wide variety of pathologic conditions including: (a) sickle cell anemia, (b) acute leukemias, (c) multiple myeloma, and (d) infectious

mononucleosis. Twenty-Five white blood cell photographs were taken from each individual blood smear. Five of these images were designated as "test" images, resulting in a total of one-hundred "test" images. Adobe Photoshop® Elements Version 2.0 software was used to color edit, crop, photomerge, and further prepare the images for use in the examination.

Competency exam identifications were separated into four modules in order to assure that fatigue from image-viewing was not a factor in exam performance for any of the participants. No participants took more than thirty minutes to complete a module. Participants completed each of the four modules at a separate sitting.

Modules 1 and 2 of the exam were set-up using what will be referred to as exam format 1 in the rest of this study's discussion. Format 1 of the exam presented the series of one-hundred WBC "test" images sequentially to the participant in a randomly organized manner. A table of random numbers was used to systematically randomize organization of the images within the first two modules. Modules 3 and 4 of the exam were set-up using what will be referred to as exam format 2 in the rest of this study's discussion. Format 2 of the exam presented the same one-hundred WBC "test" images as were presented in format 1 for identification. In format 2, however, these images were grouped by patient and presented in the context of a differential count or patient case. The five "test" images for each patient were presented to the participant only after allowing the participant to view the additional twenty WBC images taken from that patient's blood smear. The purpose of presenting these twenty images before presenting the "test" images was to simulate the scanning of several microscopic fields from the patient's blood smear. This was intended to provide a better context, or frame of reference, for the participant, allowing for more accurate evaluations of the "test" images. Format 2 images were rotated 90^{o} from the image orientation used in format 1. This was done in order to decrease the possibility

of a priming effect that viewing the images previously in format 1 may have had on the participants' performance in format 2.

Exam Directions

Format 1 Directions

"In this module, you will view and identify 50 white blood cell (WBC) images photographed from a variety of Wright-Stained peripheral blood smears. The available responses will be: (1) neutrophil, (2) lymphocyte (normal), (3) monocyte, (4) eosinophil, (5) basophil, (6) metamyelocyte, (7) myelocyte, (8) promyelocyte, (9) blast, (10) lymphocyte (atypical), and (11) plasma cell. Use the mouse to select the cell type that you believe best identifies the cell image presented. Once an answer is selected, you will not be able to change your answer or view the images again. Please perform the task at a rate at which you feel confident.

There is one optional break point in the module – after item 25. Please work on the module continuously, stopping at the optional break point only if necessary. Now, please click on the button below to view the example slides. This will familiarize you with the module format. You may view the example as many times as you wish before starting the exam."

Format 2 Directions

"In this module, you will view white blood cell (WBC) images photographed from Wright-Stained peripheral blood smears from 10 different patients. Images in this module will be presented in ten groups, each group representing a single patient. For each patient, you will first view five composite images. Each composite image will contain four white blood cells representative of the types of cells seen in that patient's blood smear. An automated timing mechanism will advance through these five composite images at a rate of 12 seconds per image. After viewing these composite images, you will be asked to identify five WBC images

photographed from the same patient blood smear. The cells for identification will be presented in the same manner and with the same answer responses as those images previously presented in modules 1 and 2. As before, the available answer responses will be: (1) neutrophil, (2) lymphocyte (normal), (3) monocyte, (4) eosinophil, (5) basophil, (6) metamyelocyte, (7) myelocyte, (8) promyelocyte, (9) blast, (10) lymphocyte (atypical), and (11) plasma cell. The mouse will be used to select a response. Once an answer is selected, you will not be able to change your answer or view the images again. Please perform the task at a rate at which you feel confident. After the completion of each patient, you will be prompted to begin examination of the next group of patient slides. The process will be repeated until all 10 patients have been completed. Please work on the module continuously, stopping at the prompt between patients only if necessary. Now, please click on the button below to view an example patient. This will familiarize you with the module format. You may view the example as many times as you wish before starting the exam. "

Competency Exam Administration

Competency exams were administered to all novices and experts over the same 1 month time period. All participants completed format 1 (modules 1 and 2) on a single day. Then, approximately 1 week later, participants returned to the exam and completed format 2 modules. The 1 week wait period was intended to further ensure that the viewing of images in format 1 did not have a priming effect on outcome performances for format 2.

All study participants were verbally informed that his/her response times were being recorded during the examination. The researcher explained to participants that they simply needed to proceed through the examination at a comfortable, steady pace. Participants were also aware that there was no total time limit for completion of exam modules.

Modules 1 and 2 (format 1) of the exams were administered to novices in a college computer lab between 9 am and 10 am on a Thursday. Modules 3 and 4 (format 2) of the exam were administered to the novices in the same college computer lab between 9 am and 10 am on Thursday, exactly 1 week later. A short break was taken between the completion of modules 1 and 2, as well as between the completion of modules 3 and 4.

Experts performed the exam on their home computers. There was no one suitable time when all 22 experts could be gathered together in a central location for exam administration. Experts were asked to find a time in their home environment during which they could work for an uninterrupted time period of approximately fifteen minutes on each module. Experts were instructed to perform modules 1 and 2 (format 1) of the exam with only a short break between the two parts. They were then instructed to return to the exam approximately one week later in order to complete modules 3 and 4 (format 2). Experts were asked to perform the exam modules at a time when they felt well-rested. Two handouts were given to all experts as a guide for completion of the examination procedures. A copy of the "Expert Participant Directions" summary sheet is included in Appendix E. A copy of "Directions for Exam Set-Up and E-Mailing Results" is included in Appendix F.

Variable module orders were assigned to each participant. All participants completed format 1 modules before completing format 2 modules, but not all participants completed the within-format modules in exactly the same order. Half of the participants for each level of expertise completed the modules in numerical order, 1-4, and the other half of the participants completed the modules out of sequence, module 2, then 1, followed by module 4, then 3.

Determination of Exam Answers

The evaluation criteria used for the exam were patterned after the College of American Pathologists' standard for blood cell identification proficiency testing. Correct answers were

determined by 90% expert consensus. Expert consensus was achieved on forty-four of the one hundred cell identifications. All other items were reviewed by a hematology supervisor and two pathologists specializing in hematology. Images for review were prepared in Powerpoint® using format 2 of the exam. This allowed the reviewing experts to have full control over the images, being able to view them at any desired pace and also to move both forwards and backwards through the presentation. These reviewing experts were also provided with the original patient differential counts and had full access to patient history and diagnosis information. The original slides were also made available to these experts for review. Cell identifications for which convergence occurred between the expert majority, the hematology supervisor, the original patient differential count and at least one pathologist were considered to be confirmed identifications. There were 4 of the 100 cells for which convergence of identification did not occur. These 4 cells were eliminated from the exam and were not considered in the evaluation process.

Exam Composition

The final exam was composed of ninety-six identified cell images consisting of (a) 13 mature myeloid cells (neutrophils, eosinophils, and basophils), (b) 16 immature myeloid cells (metamyelocytes, myelocytes, and promyelocytes), (c) 15 blasts, (d) 16 monocytes, and (e) 36 lymphoid cells. It should be noted that the original intent of the researcher was to somewhat equally represent each of the different cell types on the exam. The original cell groupings intended by the researcher divided the lymphoid cell population into three separate smaller groupings: (a) atypical lymphocytes intended to have a population of approximately 16 cells, (b) typical lymphocytes intended to have a population of approximately 14 cells, and (c) plasma cells intended to have a population of approximately 6. Note that plasma cells were represented on the exam in very low numbers because of the researcher's limited access to this type of

clinical specimens. Upon final determination of correct answers for the exam, it was determined that there were many overlapping cell identifications accepted for these three particular cell types. For instance, there were several cells for which both atypical lymphocyte and typical lymphocyte were judged as acceptable answers. Pathologists also felt that usage of these three terms may tend to vary slightly between expert technologists and different hospitals/clinical sites, with some applying more stringent definitions for the atypical lymphocyte and plasma cell categories than others. For this reason, the cells were ultimately grouped together for evaluation purposes, instead of considering these as separate categories as originally intended.

Quantitative Method: Statistical Analysis

Quantitative Statistical Analysis was performed using t-tests, MANOVA, and ANOVA (Hinkle, Wiersman, & Jurs, 1998). Statistical analysis was performed using SPSS Version 14.0. For all tests, statistical significance was set at $p = 0.05$.

Effect of Module Order on Total Scores and Average Response Times

The order effect of the exam modules on total participant scores and average participant response times was evaluated by using a repeated measure MANOVA for which the within-subject variable was exam format (2 levels) and the between subject factors were expertise (2 levels) and module order (2 levels). The dependent variables were total participant exam score and average participant response time.

Effect of Competency Test Format and Image Content on Performance Outcomes

The primary focus of this study was the differential impact of competency exam format and image content on the performance outcomes of both novice and expert groups. The effect of exam format and item content was evaluated from two different perspectives: (1) in regards to the differential effect on proficiency-type testing outcomes for practicing technicians and

technologists, and (2) in regards to the differential effect on competency-type testing outcomes for Clinical Laboratory Science students.

Evaluation of Impacts on Proficiency-Type Testing. The criteria for evaluating blood cell identifications as defined by CAP were used as the standard for evaluating impacts on proficiency-type testing. Cell identifications for which expert agreement reached a 90% level or better for both of the two formats were considered to be unambiguous cell identifications, because these items would have been graded as correct for all participants on a proficiency-type test. Statistical analysis was only performed on cell identifications for which expert agreement did not reach a level of 90% or greater for both of the two formats. According to CAP guidelines, such items would have been categorized as "ungraded" or "educational" items in a proficiency test outcome report. The effect of exam format and image content on identification of these more difficult/ambiguous cells was the primary interest of this study. Analysis of these items was performed using a one-way repeated measures MANOVA for which the within-subject variable was exam format (2 levels) and the dependent variables were 4 categorical exam scores (immature myeloid, blast, monocyte, and lymphoid).

Evaluation of Impacts on Competency-Type Testing. The academic grading scale used for the CLS program at OLOL College was used as the standard for evaluating the effects on student competency-type testing. The grading scale used was as follows: (1) 94-100 = A (2) 87-93 = B (3) 80-86 = C. The lowest C (80%) is defined as the minimal acceptable level for technical competency. This grading scale is more stringent than the alternate 10-point scale commonly used at other colleges and universities. Therefore, this grading scale serves as a conservative standard for the determination of minimal acceptable student competency levels. Cell identifications for which novices reached an 80% level in both exam formats were omitted from statistical analyses based on their unambiguous identifications. Cell identifications for

which novices reached an 80% competency level or better for both of the two formats were not evaluated using statistical analysis. Items for which the students did not reach an 80% competency level on either one or both of the exam formats were analyzed using a one-way repeated measures MANOVA for which the within-subject variable was exam format (2 levels) and the dependent variables were 5 categorical exam scores (mature myeloid, immature myeloid, blast, monocyte, and lymphoid).

Correlation Coefficient for Item Performance and Item Response Times

Overall interactions between exam item performance outcomes and exam item response times were evaluated using Pearson correlation coefficients. Correlation coefficients were calculated for both exam formats at both the novice and expert levels. These analyses were used to determine the general strength of the relationship between the two variables.

Interaction Between Performance Outcomes and Response Times for Experts

A one-way repeated measured MANOVA was used to evaluate the specific effect of item performance level on categorical item response times for experts. The within-subject variable was the item performance level. Categorical response times on exam items for which experts reached the 90% consensus level were compared to categorical response times on exam items for which experts did not reach the 90% consensus level. Categorical response times functioned as the dependent variable and included the immature myeloid, blast, monocyte, and lymphoid categories.

Interaction Between Performance Outcomes and Response Times for Novices

A paired sample t-test was used to evaluate the specific effect of item performance level on item response times for novices. Response times on exam items for which novices reached an overall 80% competency level were compared to response times on exam items for which novices did not reach the overall 80% competency level. Categorical response times could not

be evaluated for the novice group because of the extremely small and homogenous group of exam items for which an 80% competency level was reached.

Effect of Expertise on Response Times

An independent-sample t-test was used to evaluate the effect of expertise on response times. The dependent variable was the average response time for each of the ninety-six exam items. Expert and novice group means were compared for these ninety-six items for both exam formats.

Effect of Exam Format on Response Times

A 2-way repeated measures ANOVA was used to evaluate the effect that the exam format had on response times for the exam. The within-subject variable was exam format and the between-subject variable was level of expertise. The dependent variable was the experts'/novices' average response time for individual exam items. The ANOVA was used to evaluate the main effect of the format variable as well as the interaction effects between the format and expertise variables.

White Blood Cell Identification Error Types

White blood cell identification error types were determined by categorizing all incorrect exam responses for both novice and expert participants on each format 1 exam item. Defined error types were described using percent frequencies for each of the error categories generated.

Qualitative Methods

The second phase of this mixed methods study applied qualitative methods. Data collection methods included the use of think aloud protocols (Ericsson & Simon, 1993) and interviews. The intent of these data collection processes was exploratory. The researcher was interested in describing/contrasting the cognitive processes used by the novice and the expert.

Participants

Five participants at each level of expertise were selected by the researcher for this part of the research. Participants were chosen based on two main factors: (a) their willingness and availability to participate in this second phase of research, and (b) the participants' characteristics including the relationship between the individual's outcomes in the quantitative research phase and the group mean outcomes. Certain participants drew the researcher's attention based on their concurrence with or deviation from the group outcomes, as evaluated in the quantitative phase of this research.

Overall, experts were chosen based on their total years of experience, their certification as a specialist in hematology, and their performance on the cell exam exercise. The experts interviewed included the four experts with the most experience who held specialist certifications in the area of hematology. All four of these experts scored above the format 2 expert group mean of 86.9 on the cell exam exercise. One of these experts was male and the other 3 were female. The fifth expert interviewed was a slightly younger and less experienced female technologist who scored well above the format 2 expert group mean. A summary of the years of experience for each of the experts is displayed in Table 2. Students were chosen whose scores on format 2 of the cell exercise spanned the range of scores represented by the students. Of the fourteen novices, 5 scored below the novices' group exam mean of 52.9 and 9 scored above the group mean. One of the novices selected for case study participation was male and the other 4 were female. Exam performance scores for case study participants are shown in Table 3.

Multiple Case Study Design

The design for the qualitative phase of this study could be categorized as a multiple case study design. The researcher's analysis focused solely on the performance of cross-case

Table 2

Years of Experience for Expert Case Study Participants

	Years Experience		
Expert Identification	Hematology	Generalist	Other[a]
Expert 1	19	5	-
Expert 2	13	1	-
Expert 3	26	2	5
Expert 4	19	10	-
Expert 5	7	-	-

[a]Other experience for Expert 3 included 3.5 years in chemistry and 1.5 years in clinical lab education

Table 3

Exam Performance for Novice Case Study Participants

Novice Identification	Format 2 Exam Performance Description
Novice 1	Slightly above group mean (54.2)
Novice 2	About 20 points above group mean (72.9)
Novice 3	Slightly above group mean (58.3)
Novice 4	Below group mean (35.4)
Novice 5	About 10 points above group mean (63.5)

analyses. Yin (2003) explains that it is acceptable for the entire report format in a multiple case study to consist solely of cross-case analysis. He further states that individual case summaries are not necessary. The specific design of this case study method would best be characterized as instrumental (Stake, 1995). The researcher is interested in describing/identifying the various

features/components of cognitive processing for both the novice and expert as it relates to the field of hematology. The researcher is not interested in the individuality of each case itself, but instead in how each case represents the cognitive processes typical for their particular level of expertise.

Yin (2003) explains that multiple case study (MCS) design allows the researcher to look for patterns in the results across multiple cases and thus more firmly establishes any derived theoretical propositions. When designing a multiple case study, Yin encourages the use of both literal and theoretical replications. Literal replications occur when you choose a case which you would predict to have the same findings as the original case. Thus, for example, the experts would be classified as literal replications of each other. Theoretical replication occurs when you chose cases that will result in different findings from your original case. Thus, the five novice case studies would provide theoretical replication if the expert was designated as the original case.

Internal validity was established in a variety of ways. First, Yin (2003) contends that the use of the multiple case study method is in and of itself a form of triangulation because it allows for the use of cross-case analysis. Method triangulation was also used. Using two forms of qualitative data collection along with a quantitative form of data collection served as a form of triangulation (Merriam, 1988). Quantitative data such as response times and performance outcomes can be used to further evaluate the types of cognitive processing which are occurring. Such quantitative measures have been shown to correlate with proposed cognitive processing styles in several previous studies involving visual diagnosis.

Another very common form of increasing internal validity is the use of member checking as it is called by Stake (1995) and Tashakkori and Teddlie (1998). With member checking, the informants are asked to review the data collected during interview or observation for

"palatability" as Stake calls it. Yin (2003) also agrees with this form of validation as he suggests review by peers, participants, and informants. Yin explains that the reviewers do not necessarily have to agree with your interpretations or conclusions, but should confirm the accuracy of the facts collected. He states that, at times, this process may even allow some informants to provide the researcher with additional information not thought of during the original interview session. Stake suggests that member checking occurs primarily after all of the data has been collected. Merriam (1988) also suggests the use of this process, but calls it peer examination.

Knowledge Elicitation

In order for research in the field of expert/novice comparisons to be useful to the typical educator, the most appropriate type of knowledge elicitation procedure must be chosen. In order to select the appropriate knowledge elicitation format, one should first understand the structure of expertise approach research. There are three main steps in the research process. First, one must select a representative task that can be performed in a stable laboratory environment that will elicit the desired superior performance. Second, the superior performance is examined using a detailed analysis. Finally, the researcher must attempt to account for the acquisition of the unique characteristics and cognitive processes which have been observed in the expert (Ericsson & Smith, 1991). Educationally, the third and final step is the most important step because it could lead to improvement of teaching applications in the domain-specific task.

Early experts divided knowledge elicitation (KE) methods into two general types: indirect (knowledge obtained from texts and reports) or direct (knowledge obtained by actually observing the expert and probing their reasoning method) (Hoffman et al., 1995). Hoffman et al. (1995) further divided KE into three different categories: (a) analysis of familiar expert tasks such as task analysis and think aloud problem-solving, (b) interviews (structured or

unstructured), and (c) contrived techniques such as rating and sorting tasks. The type of KE chosen must be appropriate for the type of information the researcher wishes to obtain.

In using any KE method there are some potential problems that must be considered by any researcher. The major concern is the formation of the appropriate expert model from the information obtained during KE. The first roadblock that may hinder this task is the tacit nature of expert knowledge. Due to the complex knowledge base of experts, which involves memory "chunking", many expert cognitive activities are quite automatic. This may lead to extreme trouble in the communication of detailed thought processes. Experts may even tend to distort their cognition patterns giving either textbook explanations or oversimplified explanations to the non-expert interviewer (Cooke, 1994). KE is the key to accurate knowledge acquisition and the basis for a suitable model of expert knowledge in domain-specific fields.

Think-Aloud Protocol

The use of think-aloud protocols is a standard technique in cognitive science for the elicitation of verbal reports from study participants, allowing for analysis of thought sequences required for problem-solving, evaluation, or decision making processes (Ericsson & Simon, 1993). Especially important in the use of such protocols is the avoidance of instructions which elicit explanations or descriptions from the participant. Instead, think-aloud protocols attempt to elicit direct verbalizations of cognitive processes, such verbalizations are known as level 1 verbalizations. The researcher desires to have as few intermediate processes occur during the transfer of thoughts stored in STM (short term memory) into verbalizations. If the internal representation of the information in STM is not encoded verbally it must be recoded into a verbal code before verbalization. This is known as a level 2 verbalization. The collection of either level 1 or level 2 verbalizations is adequate as a source of verbal reports for think aloud protocols. With such verbalizations, the organization of the information stored in STM is intact.

Level 3 verbalizations are not desired for collection during a think-aloud protocol because they require the use of additional information not already in STM and thus alter what is originally stored in STM. Such verbalizations often require the filtering of information, the generation of information, or some type of intermediate processing of the information stored in STM in order to provide the verbal information requested by the researcher.

Ericsson offers some suggestions in giving think-aloud instructions. He suggest that the main part of the instructions be very concise and direct and ask that the participant verbalize their thoughts or "inner speech" during the problem-solving process. The researcher may suggest that this procedure may be somewhat familiar to the participant already, if used when thinking through a problem alone.

Some complementary instructions are discussed by Ericsson. One suggested complementary instruction is that for completeness. Other complementary instructions which may be used but are cautioned against by Ericsson are those requesting explanations or those requesting specific content for inclusion in the vocalization. Such requests may tend to induce level 3 verbalizations.

Two sets of think aloud protocols were performed in this study. The first involved the presentation of forty images from format 1 of the exam to the participant. The majority of the images selected for this verbalization task were those found to be more difficult by the expert group. A few "easy" items were also included. Images were included to represent all of the possible WBC types. The second set of images shown to the participants was 5 patient cases selected from format 2 of the exam. The auto timing mechanism present in the actual exam was removed for the think aloud exercise so as to allow participants to move freely through the slide presentation. For both think aloud image sets, participants were asked to verbalize his/her inner speech to the best of his/her ability as he/she went through the WBC identification process for

each image. In this part of the analysis the researcher was most interested in seeing what aspects of the participants' cognitive processing accounted for the majority of their thought processes and subsequently the majority of their verbal description time. The researcher was interested in comparing and contrasting the apparent differences in cognitive processing focus as it applied to the novice and expert levels of thinking.

Interviews

Interviewing was conducted primarily through the use of the standardized open-ended interview format (Patton, 2002). The interview questions were designed utilizing the Patton system which has six main categories of questions: (a) Experience/Behavior, (b) Opinion/Value, (c) Feeling/Emotion, (d) Knowledge, (e) Sensory, and (f) Background. The intent of the interview questions was to further confirm details of the cognitive processing methods used by the novices and experts while performing WBC identifications.

Interview Questions

1. Please describe for me, in as much detail as possible, the general thought process you use when identifying white blood cells while performing a differential count.
2. Through your years of experience, how have you seen your skill as a cell morphologist evolve? (This question was addressed to experts only.)
3. What type of experiences/activities have you found to be the most critical in developing/improving your morphology skills? Specifically, *why* have you found such experiences/activities so critical?
4. In your opinion, what specific cell types are the most difficult to distinguish from each other? Why?

5. When you find yourself trying to differentiate/discriminate between various morphologic cell types while trying to identify a difficult cell, what special thought processes/methods do you use?

Protocol and Interview Analysis: Coding Scheme, ATLAS.ti

All think aloud protocols and interviews were recorded using an Olympus digital recorder and transcribed verbatim for analysis by a medical transcriptionist. Think aloud and interview data was analyzed, coded, and categorized using standard qualitative data analysis methods. Transcripts for think aloud protocols were unitized and categorized based on the constant comparative method (Glaser & Strauss, 1967). ATLAS.ti Version 5.0 (2006) was used to perform this qualitative data analysis process.

Analysis of think aloud data involved the application of many aspects of the verbal data analyses technique as described by Chi (1997). Although the method of data collection used for this study was that suggested by Ericsson and Simon (1993), the researcher was not interested in performing a classic protocol analysis. Chi explains that in protocol analysis the goal is often to identify a match between the participants' verbalizations and elements/operators that are defined a priori. The participants' problem-solving methods are usually matched to an existing problem-solving strategy or model. In this case, the researcher was not interested in building a strategic problem solving model, but was instead interested in examining the types and levels of knowledge used by participants during the white blood cell identification process. This goal ties in much more closely to the goals of Chi's method of verbal analysis. The results of the quantified qualitative codings were also analyzed using inferential statistics. Chi states that "validation is obtained by …applying statistical tests of the quantified qualitative codings to see if the results support a hypothesis" (p. 5). In this study, the researcher hoped to explore the relationships between expertise and cognitive processing, intending to correlate the results of this

study with those previously conducted in the field of visual diagnosis. Answers to interview questions were summarized for each level of expertise using a questions and answer format as suggested by Yin (2003).

Methods Summary

In summary, a mixed research design method was used for this study, the first phase of the research being quantitative with the second phase of research being qualitative. The first phase of the study involved administration of a clinical hematology competency assessment to both novices and experts. Two different formats of the exam were administered. Each study participant completed both forms of the assessment. Statistical analyses were performed in order to determine if exam format had a significant effect on expert/novice performance outcomes or expert/novice response times. Error types were also classified through examination of the quantitative exam performances.

A qualitative case study phase followed the quantitative analyses of exam results. Based on the findings of the quantitative research phase, representative novices and experts were selected for the qualitative phase. Novice and experts were interviewed and probed using think aloud protocols with WBC images selected from each exam format. Qualitative codings revealed during the examination of novice and expert think aloud protocols were categorized and evaluated for both levels of expertise. Patterns in cognitive processing for both levels of expertise were compared and contrasted. Ultimately, outcomes from both study phases were correlated in order to identify any significant implications for warranted changes in the areas of CLS assessment or instruction.

RESULTS AND DISCUSSION

Summary Exam Results

Summary exam results for both novice and expert participants are displayed in Table 4 and Table 5 respectively. Total percent consensus (percent correct) and mean response times are given both the novice and expert populations for all ninety-six cell exam images. Mean response times are reported in seconds.

Table 4

Total Percent Correct and Mean Response Times for Novice Cell Identifications

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
1. Me	64.3	57.1	6.1	3.7
2. N	85.7	92.9	4.3	2.8
3. Bl	28.6	35.7	11.9	6.8
4. N	50	71.4	9.7	3.9
5. N	100	85.7	2	1.7
6. M	28.6	14.3	10.4	4.6
7. At / L / Pl	50	42.9	6.9	3.2
8. At / L	28.6	42.9	6.1	5
9. N	92.9	85.7	1.9	1.6
10. M	21.4	21.4	7.9	5.2
11. L	50	57.1	4.4	3.7
12. Bl	64.3	42.9	6.6	7.3

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 4 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
13. My / Pr	64.3	64.3	9.6	8.8
14. At / L	85.7	85.7	3.2	3.5
15. Bl	42.9	50	7.7	4.6
16. My	50	35.7	8.2	6.9
17. At / L	42.9	42.9	8	6.8
18. M	78.6	71.4	6.9	4.3
19. Bl	71.4	64.3	4.9	6.3
20. At / L	42.9	42.9	6.9	4.6
21. M	85.7	64.3	6.2	5
22. Bl	28.6	35.7	9.7	3.5
23. At / L	28.6	28.6	6.3	6.4
24. E	64.3	50	3.5	3.8
25. Me / My	78.6	42.9	5.5	4.4
26. N	85.7	92.9	3.1	2.7
27. My / Pr	64.3	78.6	5.8	7.6
28. N	92.9	85.7	2.1	2
29. Me / My	57.1	71.4	5.5	5.2

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 4 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
30. M	14.3	28.6	7.6	4.6
31. At / L	78.6	78.6	4.7	5.2
32. My	50	35.7	6.9	8.7
33. At / L / Pl	42.9	21.4	11.5	5.8
34. L	35.7	64.3	5.5	4.3
35. At / L	57.1	57.1	5.4	4.6
36. At / L / Pl	57.1	57.1	5.7	8.1
37. M	35.7	28.6	8.9	6.8
38. At / L	28.6	28.6	5.7	6.4
39. Pl	28.6	50	6.1	6.4
40. E	64.3	64.3	3.1	2.7
41. M	71.4	50	2.9	4
42. M	64.3	57.1	3.1	3.7
43. Me	78.6	50	3.3	5.2
44. At / L	14.3	21.4	6.5	5.3
45. At / L / Pl	50	28.6	7.5	5.8
46. My	28.6	42.9	4.7	6.6

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 4 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
47. At / L	14.3	7.1	6.5	5.7
48. M	71.4	78.6	2.4	3
49. Me	57.1	71.4	5.7	3.6
50. Me / My	35.7	42.9	11.2	4.2
51. At / L	71.4	85.7	7	4.5
52. M	78.6	64.3	3.8	2.4
53. At / L	21.4	14.3	7	7.1
54. M	21.4	28.6	11.7	4.6
55. M	64.3	64.3	7.3	3.4
56. B	71.4	92.9	5	4.4
57. Bl	71.4	71.4	7.4	3.8
58. Bl	50	35.7	6.2	4.6
59. L / At	42.9	57.1	7.2	5.4
60. M	28.6	28.6	5.7	8.9
61. N	92.9	92.9	3.5	1.7
62. At / L	78.6	85.7	5.1	6.3
63. At / L	21.4	14.3	5.6	3.8
64. M	14.3	14.3	4.6	5.8

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 4 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
65. Bl	78.6	71.4	7.3	4.2
66. N	92.9	85.7	2.6	3.4
67. Me	57.1	64.3	8.3	3.9
68. At / L	71.4	78.6	6.4	5.1
69. Bl	71.4	71.4	5.8	4.4
70. At / L	57.1	28.6	8.8	7.3
71. Bl	50	42.9	5.3	4
72. Bl	28.6	21.4	9.4	6.2
73. At / L	42.9	50	8.8	4.3
74. Bl	7.1	14.3	6.4	6
75. At / L	85.7	92.9	4.5	2.3
76. At / L	42.9	42.9	7.2	6.4
77. Bl	42.9	57.1	8.8	2.6
78. At / L	64.3	64.3	5.8	5.2
79. M	78.6	64.3	3.6	2.3
80. At / Pl	50	78.6	3.4	3.8

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 4 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
81. At / L	28.6	42.9	8.4	3.9
82. L	57.1	64.3	5.2	3.9
83. At / L	28.6	42.9	5.9	5.5
84. At / L	78.6	71.4	3.7	2.8
85. E	57.1	57.1	3.5	1.8
86. L	35.7	57.1	6.7	4.2
87. My	64.3	42.9	9.1	5.1
88. My / Pr	50	57.1	7	4.7
89. At / L	57.1	35.7	4	5.7
90. N	100	85.7	1.7	2.3
91. My	42.9	28.6	7.3	5.1
92. Bl	21.4	21.4	5.9	4.4
93. Me	42.9	28.6	5	8.8
94. L	57.1	64.3	4.6	4.4
95. Bl	7.1	7.1	6.3	4.9
96. M	42.9	71.4	10.6	4.5

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 5

Total Percent Correct and Mean Response Times for Expert Cell Identifications

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
1. Me	86.4	95.5	10	2
2. N	95.5	100	4.6	2
3. Bl	40.9	68.2	19.6	6.1
4. N	68.2	68.2	7.6	4.9
5. N	100	100	1.5	1.4
6. M	100	86.4	3.2	3.4
7. At / L / Pl	95.5	95.5	10	4.3
8. At / L	100	95.5	3	2.7
9. N	95.5	100	2	1.3
10. M	22.7	50	11.6	7.5
11. L	100	100	1.8	1.8
12. Bl	95.5	100	4.8	4
13. My / Pr	81.8	77.3	10.1	7
14. At / L	100	100	3.3	2.4
15. Bl	81.8	95.5	5.4	3.3
16. My	81.8	81.7	4.8	4.6

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 5 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
17. At / L	90.9	90.9	6.2	2.8
18. M	81.8	81.8	7.7	3.1
19. Bl	81.8	90.9	4.7	3.6
20. At / L	100	100	3.6	2.8
21. M	63.6	86.4	9.5	4.5
22. Bl	13.6	45.5	12.3	2.8
23. At / L	45.5	72.7	9.3	4
24. E	100	100	2.3	2
25. Me / My	90.9	90.9	6.3	3.7
26. N	90.9	90.9	3.1	3.6
27. My / Pr	68.2	72.7	7.5	5.5
28. N	100	95.5	1.7	1.6
29. Me / My	50	81.8	9.4	3.7
30. M	50	63.6	7.3	6.2
31. At / L	95.5	100	3.7	3.8
32. My	77.3	68.2	4.9	6.9

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 5 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
33. At / L / Pl	63.6	77.3	9.9	8.3
34. L	86.4	86.4	2.7	1.9
35. At / L	95.5	90.9	6.5	2.9
36. At / L / Pl	100	95.5	8	8.1
37. M	18.2	54.5	7.3	4.1
38. At / L	72.7	86.4	6	4.8
39. Pl	81.8	86.4	5.2	4.7
40. E	100	100	2	2.5
41. M	95.5	100	3	1.6
42. M	81.8	90.9	3.1	2.5
43. Me	86.4	100	4	4.7
44. At / L	68.2	68.2	6.9	4
45. At / L / Pl	86.4	90.9	6.6	5.4
46. My	36.4	59.1	8.4	4.3
47. At / L	90.9	90.9	3.2	4.4
48. M	100	95.5	1.9	2.2

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 5 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
49. Me	100	86.4	3.4	4
50. Me / My	77.3	68.2	9.6	4.7
51. At / L	95.5	100	4.4	4.3
52. M	95.5	100	3.6	1.4
53. At / L	81.8	81.8	10.7	5
54. M	77.3	72.7	10.6	7
55. M	100	95.5	2.5	2.8
56. B	100	100	3	2.6
57. Bl	95.5	100	4.2	2.1
58. Bl	68.2	68.2	7.4	2.5
59. L / At	95.5	90.9	3.4	4.3
60. M	68.2	86.4	7.2	5.4
61. N	100	100	2.3	1.5
62. At / L	100	100	3.7	2.8
63. At / L	90.9	100	3.7	2.6
64. M	86.4	90.9	4.8	4.1

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 5 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
65. Bl	95.5	86.4	4.6	3.9
66. N	100	100	3.8	2.5
67. Me	86.4	86.4	4.7	3.1
68. At / L	100	100	3.7	3.3
69. Bl	90.9	100	3.8	2.1
70. At / L	95.5	100	3.6	4.2
71. Bl	77.3	68.2	5.6	5.8
72. Bl	31.8	63.6	8.2	4.9
73. At / L	100	100	3.5	2.8
74. Bl	50	77.3	9.6	8
75. At / L	100	100	4.4	2.8
76. At / L	72.7	86.4	6.6	3.6
77. Bl	77.3	95.5	7.3	3.2
78. At / L	100	90.9	3.5	2.8
79. M	100	95.5	1.8	1.4
80. At / Pl	81.8	95.5	6.6	4.5

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Table 5 Continued

	Total Percent Correct		Mean Response Times	
Image number / Cell Identification	Format 1	Format 2	Format 1	Format 2
81. At / L	95.5	100	5	2.2
82. L	95.5	86.4	2.3	2.1
83. At / L	95.5	100	3.2	2.5
84. At / L	100	100	1.9	1.6
85. E	95.5	100	2.1	1.6
86. L	81.8	86.4	2.7	2.5
87. My	72.7	72.7	5.3	4
88. My / Pr	72.7	68.2	7.6	4
89. At / L	86.4	81.8	6	3.9
90. N	100	100	1.5	1.3
91. My	45.5	81.8	5.2	3.7
92. Bl	36.4	31.8	9.9	4.1
93. Me	90.9	72.7	4.3	5.4
94. L	81.8	100	2.5	1.9
95. Bl	18.2	59.1	14.9	4.5
96. M	95.5	95.5	4.5	3.2

Note. N = neutrophil, E = eosinophil, B = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell

Effect of Module Order on Total Scores and Average Reaction Times

The effect of module order on total participant scores and average response times was evaluated using a repeated measure MANOVA. The mean and standard deviation results for the total scores and average response times for both novice and expert participants are displayed in Table 6. Mean response times are reported in seconds. The "In Sequence" module order refers to those participants who completed the modules in sequential order. The "Out of Sequence" module order refers to those participants who completed the modules in the following order: 2, 1, 4, 3.

Table 6

Means and Standard Deviations for Total Participant Scores and Average Participant Reaction Times by Module Order (In Sequence Versus Out of Sequence)

	Novice		Expert	
Exam Outcomes	In Sequence	Out of Sequence	In Sequence	Out of Sequence
		Means		
Format 1 Score	54.31	52.54	82.86	80.96
Format 2 Score	52.96	52.84	87.23	86.55
Response Time 1	6.64	5.66	5.61	5.40
Response Time 2	4.83	4.67	3.63	3.66
		Standard Deviations		
Format 1 Score	15.14	16.55	6.22	6.52
Format 2 Score	15.40	13.03	6.86	6.42
Response Time 1	4.77	0.98	1.93	1.80
Response Time 2	2.97	0.93	1.17	1.01

MANOVA results showing the main effect for module order and related interaction effects for order are displayed in Table 7. The main effect for module order on total participant scores and average response times was not statistically significant ($p > 0.1$). The statistically non-significant effects for expertise x module order interaction, format x module order interaction, and format x expertise x module order interaction further show that the module order sequence assigned for each participant did not significantly alter the overall exam results. Module order effects for each dependent measure were also statistically non-significant. The average exam score resulted in an $F(1) = 0.103$ and a $p = 0.751$, while the average response time resulted in an $F(1) = 0.222$ and a $p = 0.641$. The lack of statistically significant effects for order suggests that exam scores and response times were fairly uniform regardless of the order in which the exam modules were completed by each participant.

Table 7

Multivariate Analysis of Variance for Module Order

Source	F value	P value	Eta^2
	Between Subjects		
Module order	0.128	0.880	0.008
Expertise x module order	0.067	0.935	0.004
	Within subjects		
Format x Module order	0.654	0.527	0.040
Format x Expertise x Module order	0.208	0.813	0.013

Question 1: What, If Any, Differential Effect Do Competency Test Item Format and Image Content Have on Competency Performance Outcomes for Novice Students Versus Expert Professionals in Clinical Hematology?

The effect of exam format and image content were evaluated from the prospective of proficiency-type testing for experts and from the prospective of competency-type testing for

students. The effects from these two different perspectives were evaluated separately for each of the two levels of expertise using two separate one-way repeated measures MANOVA.

Expert Exam Performance Results

A 90% expert consensus level was achieved for forty-four of the ninety-six WBC images. Experts did not reach the 90% consensus level for the other fifty two exam items. Table 8 summarizes the categorical exam breakdown based on these two levels of performance. Information is displayed for each of the five major WBC categories represented on the exam.

Table 8

Categorical Breakdown of Experts' Exam Performance

Exam Breakdown	Mature myeloid	Immature myeloid	Blast	Monocyte	Lymphoid
		> 90% Consensus			
No. of Exam Items	12	1	3	6	22
% of WBC Category Represented	92.3%	6.3%	20%	37.5%	61.1%
		< 90% Consensus			
No. of Exam Items	1	15	12	10	14
% of WBC Category Represented	7.7%	93.8%	80%	62.5%	38.9%

Evaluation of Impacts on Proficiency-Type Testing. A one-way repeated measures MANOVA was used to evaluate the effect of exam format on expert performance for the fifty two cells for which experts did not reach a 90% agreement level. Subcategory scores were calculated for each of four major WBC subtypes (immature myeloid, blast, monocyte, and lymphoid). All of the mature myeloid cells except for one were identified at a ninety percent agreement level. The single remaining mature myeloid cell was identified at a 68.2% agreement

level by experts on both exam formats. It was apparent that format did not have a differential effect on this one item or on the other twelve unambiguous mature myeloid cells. For this reason the mature myeloid category was not further evaluated in this analysis. Of important note is that the mature myeloid and lymphoid cells were identified most frequently at the >90% consensus level. This suggests that these cell types possess the least ambiguous features and may be easily recognized through a feature-based or pattern matching type cognitive process. The results of this analysis are displayed in Table 9.

Table 9

Means and Standard Deviations for Expert Categorical WBC Exam Scores on Items with Average Performance < 90%

	Mean		SD		MANOVA	
WBC Category	Format 1	Format 2	Format 1	Format 2	F value	p value
Immature Myeloid	74.24	78.18	14.59	16.86	1.890	0.184
Blast	56.06	70.83	21.38	28.50	11.531	0.003
Monocyte	65.00	76.36	18.96	24.98	7.591	0.012
Lymphoid	77.60	84.75	17.13	16.26	12.833	0.002

The main effect for format on expert performance was statistically significant, $F(4,18) = 10.065$ and $p = 0.000$. Experts performed significantly better on format 2, in which cell identifications were made in the context of a patient differential background. The strength of the relationship, as indexed by eta-squared, was 0.691. Univariate analysis for the various subcategory scores showed that format had a statistically significant effect ($p < 0.05$) on expert performance in the blast, monocyte, and lymphoid subcategories. There was no statistically significant effect ($p > 0.1$) for format on expert performance in the immature myeloid subcategory.

Novice Exam Performance Results

Students achieved an overall 80% competency level on ten of the ninety-six WBC images. Students did not reach the 80% competency level on the other eighty six items. Table 10 summarizes the breakdown for these two levels of performance. Information is displayed for each of the five major WBC categories represented on the exam

Table 10

Categorical Breakdown of Novices' Exam Performance

Exam Breakdown	Mature myeloid	Immature myeloid	Blast	Monocyte	Lymphoid
		>80% Competency			
No. of Exam Items	8				2
% of WBC Category Represented	61.5%				5.6%
		< 80% Competency			
No. of Exam Items	5	16	15	16	34
% of WBC Category Represented	38.5%	100%	100%	100%	94.4%

It should be noted that the only two cell types for which a level of 80% competency was reached were the mature myeloid cells and the lymphocytes. These were the same two WBC subcategories that appeared to be the most unambiguous for experts as well.

Evaluation of Impacts on Competency-Type Testing. A one-way repeated measures MANOVA was used to evaluate the effect of exam format on novice performance for the eighty six cells for which students did not reach the 80% competency level. Subcategory scores were calculated for each of the five major WBC subtypes (mature myeloid, immature myeloid, blast, monocyte, and lymphoid). The results of this analysis are displayed in Table 11.

The main effect for format on novice performance was not statistically significant, $F(5,9) = 1.258$, $p = 0.360$. The strength of the relationship, as indexed by eta-squared, was 0.411. This indicates that novice exam scores on the two exam formats were not significantly different. Univariate analysis for the various subcategory scores further supported the main effect findings by indicating that format did not have a statistically significant effect ($p > 0.1$) on any of the five subcategory scores evaluated.

Table 11

Means and Standard Deviations for Novice Categorical WBC Exam Scores on Items with Average Performance < 80%

	Mean		SD		MANOVA	
WBC Category	Format 1	Format 2	Format 1	Format 2	F value	P value
Mature Myeloid	61.43	67.14	37.18	32.92	2.167	0.165
Immature Myeloid	55.39	50.91	25.80	22.70	0.828	0.379
Blast	44.29	42.86	24.74	28.72	0.129	0.726
Monocyte	50.02	46.90	22.33	27.27	0.737	0.406
Lymphoid	45.79	48.54	18.56	14.12	0.264	0.616

Discussion of Cell Exam Format

The impact of exam format on WBC identification was the central interest and focus in this study. The results of this study clearly show that presentation of cells in the context of a patient differential background has a positive effect on expert performance outcomes. The patient presentation/case study format does not, however, have a significant effect on novice performance outcomes. These results can be justified by the fact that novices are very "rule-dependent" and generally use information in a very context-free manner (Dreyfus and Dreyfus, 1986). In fact, in the field of nursing the Dreyfus model has been applied. Benner (1984)

explains that "following rules legislates *against* successful performance because the rules cannot tell them (the novice) the most relevant tasks to perform in an actual situation" (p. 21). When identifying hematological cells, novices do not possess the context-dependent hematological experience necessary to break away from reliance on rules and feature lists. Consequently, they cannot make proper use of the differential context provided to them in format 2 of the exam. Format 2 may have also had beneficial effects for the expert by allowing them to view the various cell types found on the patient blood smear in a simultaneous presentation format prior to the identification of the "test" images. Simultaneous presentation modes encourage the institution of the image comparison processes (Perrin, 1969; Tufte, 1997; Zull, 2002).

Both CAP proficiency surveys and format 2 of the researcher's exam provide a contextualized presentation setting for WBC images, although the contextualization is provided through different mechanisms. The presentation of CAP proficiency surveys is somewhat contextualized in that participants are given patient clinical histories, demographics, clinical diagnosis, and/or past/current laboratory findings. The CAP proficiency surveys provide the contextualized setting through the use of written text. Format 2 of the researcher's cell exam, on the other hand, is contextualized in its visual presentation. In general, the CAP surveys were not found to provide significant visual contextualization, since only five images were presented per case and most commonly these five images did not include duplicates of a single cell type.

Results of the researcher's study did correlate with the summary findings presented for the CAP 2000-2006 proficiency surveys. CAP proficiency surveys showed that the WBC categorical area with the weakest overall performance was that of the immature myeloid cells. A 90% consensus level was reached for only 50% of the promyelocytes and for none of the metamyelocytes/myelocytes presented over the 5 year period. This correlates with the findings of the present study in which it was determined that the immature myeloid population was the only

difficult/atypical cell population for which performance outcomes did not differentially improve in format 2 of the exam.

Question 2: What Interactions, If Any, Are There Between a Subjects' Response Time for an Item on a Competency Assessment and (a) Performance Outcomes on Individual Items (b) Level of Expertise (c) Exam Format?

Correlation Coefficients for Item Performance and Item Response Times

The general strength of the relationship between exam item performance outcomes and exam item response times was determined by evaluating Pearson correlation coefficients for both exam formats. The results of this analysis are shown in Table 12. All analyses confirmed the inverse relationship between performance outcomes and response times by returning negative coefficients of correlation in every case. All correlation coefficients were < - 0.5 and so suggest a moderate to strong inverse relationship between the two variables. The correlation coefficients for novices in both formats remained fairly constant, returning values between -0.5 and -0.55. The correlation coefficient for experts in format 1 suggests a much stronger relationship between performance outcomes and response times than does the correlation coefficient for format 2.

Table 12

Pearson's Correlation Coefficient Between Exam Item Performance Outcomes and Exam Item Response Times

Format	Correlation Coefficient
Format 1	
Novices	-0.546
Experts	-0.740
Format 2	
Novices	-0.506
Experts	-0.536

Interaction Between Performance Outcomes and Response Times for Experts

The second analysis which was performed to evaluate the relationship between performance outcomes and response times was a one-way repeated measure MANOVA performed for the expert group. This analysis was used to evaluate the relationship between item response times and item performance levels in the four WBC subcategories previously evaluated for research question 1 (see Table 13). Item response times are reported in seconds. Mature myeloid cells were not included because of the apparent unambiguous nature of the mature myeloid identifications for the expert group.

Table 13

Expert Means, Standard Deviations, and MANOVA Results for Categorical Response Time Averages: > 90% Consensus Versus < 90% Consensus

	Mean		SD		MANOVA	
WBC Category	> 90% Consensus	< 90% Consensus	> 90% Consensus	< 90% Consensus	F value	p value
			Exam Format 1			
Immature Myeloid	6.27	6.56	5.73	2.34	0.055	0.817
Blast	4.23	8.97	2.42	5.28	25.713	0.000
Monocyte	2.87	7.22	1.37	4.23	23.245	0.000
Lymphoid	4.22	6.20	1.73	2.36	21.997	0.000
			Exam Format 2			
Immature Myeloid	3.70	4.51	3.16	1.54	1.347	0.259
Blast	2.74	4.37	1.92	2.21	27.599	0.000
Monocyte	2.12	4.78	0.66	2.51	28.937	0.000
Lymphoid	3.28	4.03	1.35	1.62	11.327	0.000

The MANOVA showed that for the blast, monocyte, and lymphoid subcategories the response times for those items for which 90% consensus was reached was significantly shorter than the response times for those items for which 90% consensus was not reached ($p < 0.001$). The main effect for the level of performance (> 90% consensus versus < 90% consensus) on response time in format 1 was statistically significant, $F(4,18) = 11.337$, $p = 0.000$. The effect size as represented by eta^2 was 0.716. The main effect for the level of performance on response time in format 2 was also statistically significant, $F(4,18) = 14.014$, $p = 0.000$. The effect size as represented by eta^2 was 0.757.

Interaction Between Performance Outcomes and Response Times for Novices

A paired samples *t* test compared the response time for items with competency levels > 80% with response times for items with competency levels < 80%. The difference between mean response times based on level of performance was not evaluated by WBC sub-category for the novice group. Because there were only ten items for which novice performance fell above the 80% competency level and because these ten items only included mature myeloid cells and lymphocytes, sub-category evaluation was not reasonable. Mean and standard deviations are shown in Table 14. Findings for the novice mirror those previously suggested for the expert. Response times for format 1items with competency levels > 80% were significantly shorter than response times for format 1 items with competency levels < 80%, $t(9) = 5.491$, $p = 0.000$. Evaluation of format 2 response times showed statistical significance as well with $t(9) = 4.097$ and $p = 0.003$.

The significantly shorter response times found for both expert and novice groups on those items for which correct responses were given supports the findings of Norman et al. (1989) in regards to automatic/instantaneous processing. Such response times may suggest the use of two

Table 14

Novice Means, Standard Deviations, and T-Test Results for Response Time Averages: >80% Competency Versus < 80% Competency

Competency Level	Mean	Standard Deviation
	Exam Format 1	
> 80% Competency	2.89	0.999
< 80% Competency	7.96	2.351
	Exam Format 2	
> 80% Competency	2.40	0.688
< 80% Competency	5.22	1.841

different cognitive processes, one used in the determination of correct answers and one used when the initial process breaks down, resulting in incorrect responses. The slower response times for incorrect answers may also simply represent the breakdown of the current cognitive process in use. Response times alone are not enough for a sound evaluation of probably cognitive processing styles.

Effect of Expertise on Response Times

Independent samples *t* test analysis suggested that mean item response times for novices did not differ significantly from mean item response times for experts ($p > 0.1$) in format 1. *T* test analysis did show, however, a significant difference ($p < 0.05$) between mean item response times for novices and experts in format 2. Refer to Table 15 for descriptive statistics and specific *t* test results.

Effect of Exam Format on Response Times

ANOVA results suggest that novices and experts respond significantly faster in format 2 than in format 1. Mean and standard deviations for response times are displayed in Table 15.

Table 15

Means, Standard Deviations, and T-Test Results for Average Response Times: Novice Versus Expert

	Mean		SD		t-test	
Exam Format	Novice	Expert	Novice	Expert	*t* value	p value
Format 1	6.15	5.54	2.33	3.21	1.504	0.134
Format 2	4.76	3.65	1.68	1.63	4.651	0.000

The main effect for format was statistically significant with F (1,190) = 90.810 and p = 0.000. The effect size is expressed by an eta^2 value of 0.323. Expertise did not have a differential effect on format response times, as shown by the interaction effects statistics for format x expertise, F(1,190) = 2.105 and p = 0.148 (eta^2 = 0.011). Exam format 2 resulted in shorter response times for both novices and experts. This may have simply occurred because format 2 allows multiple examples of a single cell type to be examined before an identification of like cells occurs.

Question 3: What Are the Types of Errors Revealed During the Process of White Blood Cell Identification?

Error types were analyzed for both experts and novices by evaluating the performance outcomes from format 1 of the exam. This format was chosen for evaluation since it contained a larger respresentation of the error types made by experts than did format 2. The results of this analysis are shown in Table 16. Five main categories of error types were generated: (a) myeloid maturation stage, (b) cell lineage, (c) subclassification of lymphocytes, (d) subclassification of mature myeloid cells, (e) blasts misidentified as lymphocytes, and (f) lymphocytes misidentified as blasts. The myeloid maturation stage error type included misclassification of a metamyelocyte, myelocyte, or promyelocyte cell as a more mature or less mature myeloid cell

stage. The cell lineage error type included the misidentification of a cell as being from a different cell lineage.

Table 16

Percent Frequencies for WBC Identification Error Types

Error Types	Error Frequencies	
	Experts	Novices
Myeloid Maturation Stage	**97 / 382 (25.4%)**	**123 / 625 (19.7%)**
Cell Identified as More Mature	50 / 382 (13.1%)	65 / 625 (10.4%)
Cell Identified as Less Mature	47 / 382 (12.3%)	58 / 625 (9.3%)
Cell Lineage	**150 / 382 (39.3%)**	**334 / 625 (53.4%)**
Lymphoid Cells Misidentified as Myeloid	30 / 382 (7.9%)	119 / 625 (19.0%)
Lymphoid Cells Misidentified as Monocytes	13 / 382 (3.4%)	69 / 625 (11.0%)
Monocytes Misidentified as Lymphoid Cells	41 / 382 (10.7%)	41 / 624 (6.6%)
Monocytes Misidentified as Myeloid Cells	36 / 382 (9.4%)	64 / 625 (10.2%)
Myeloid Cells Misidentified as Monocytes	5 / 382 (1.3%)	24 / 625 (3.8%)
Myeloid Cells Misidentified as Lymphoid Cells	25 / 382 (6.5%)	17 / 625 (2.7%)
Subclassification of Lymphocytes	**24 / 382 (6.3%)**	**37 / 625 (5.9%)**
Subclassification of Mature Myeloid Cells	**6 / 382 (1.6%)**	**26 / 625 (4.2%)**
Blasts Misidentified as Lymphocytes	**77 / 382 (20.2%)**	**40 / 625 (6.4%)**
Lymphocytes Misidentified as Blasts	**15 / 382 (3.9%)**	**35 / 625 (5.6%)**
Miscellaneous	13 / 382 (3.4%)	30 / 625 (4.8%)

Error types involving the subclassification of lymphocytes included the misclassification of normal lymphocytes, atypical lymphocytes, or plasma cell as an alternate cell type within the lymphoid cell lineage itself. Frequencies for lymphocyte subclassification errors are displayed in Table 17.

Table 17

Percent Frequencies for Error Types Made in the Subclassification of Lymphocytes

Error Types	Error Frequencies	
	Experts	Novices
Normal Lymphocytes Misidentified as Atypical Lymphocytes / Plasma Cells	11 / 382 = 2.9%	8 / 625 = 1.3%
Atypical Lymphocytes / Plasma Cells Misidentified as Normal Lymphocytes	4 / 382 = 1.0%	7 / 625 = 1.1%
Atypical Lymphocytes/ Normal Lymphocytes Misidentified as Plasma Cells	7 / 382 = 1.8%	20 / 625 = 3.2%
Plasma Cells Misidentified as Atypical Lymphocytes	2 / 382 = 0.5%	2 / 625 = 0.3%

Error types involving the subclassification of mature myeloid cells included the misclassification of mature neutrophils as eosinophils or basophils and the misclassification of eosinophils as basophils or basophils as eosinophils Frequencies for mature myeloid subclassification errors are displayed in Table 18.

Table 18

Percent Frequencies for Error Types Made in the Subclassification of Mature Myeloid Cells

Error Types	Error Frequencies	
	Experts	Novices
Neutrophils Misidentified as Eosinophils / Basophils	5 / 382 = 1.3%	7 / 625 = 1.1%
Eosinophil / Basophil Identification Reversal	1 / 382 = 0.3%	19 / 625 = 3.0%

The last category of error types was that of the miscellaneous error types. Miscellaneous error types included those errors that could not be classified into any other major error type category. Table 19 displays the frequencies for the miscellaneous error types.

Table 19

Percent Frequencies for Miscellaneous Error Types

Error Types	Error Frequencies	
	Experts	Novices
Blasts Misidentified as Monocytes	9 / 382 = 2.4%	16 / 625 = 2.6%
Monocytes Misidentified as Blasts	4 / 382 = 1.0%	7 / 625 = 1.1%
Lymphocyte or Blast Misidentified as an Eosinophil / Basophil	0%	7 / 625 = 1.1%

Question 4: What Are Some Explicit Cognitive Visual Examination Processes That Are Used by Students and Experts to Identify Images of White Blood Cells?

Cognitive and visual examination processes used by novices and experts in the identification of WBCs were evaluated using ATLAS.ti. The major categories which resulted from iterative coding of the think aloud protocols for both novices and experts were (a) data description, (b) data analysis, and (c) data interpretation.

Data description included Level I type processing, described in Bloom's taxonomy (Bloom, 1956) as the knowledge and comprehension (or recall) levels of thinking. The main subcategories involve the identification or description of the white blood cell features present or absent in the cells being identified. Feature identification categories generated include: (a) cell color, (b) cell maturity, (c) cell shape, (d) cell size, (e) cytoplasm amount, (f) cytoplasm color, (g) cytoplasm shape/margins, (h) cytoplasm texture, (i) granules absent, (j) granules present, (k) halo present, (l) Nuclear:cytoplasmic (N:C) ratio, (m) nuclear color, (n) nuclear maturity, (o) nuclear location, (p) nuclear shape, (q) nuclear size, (r) nuclear texture, (s) nucleoli absent,

(t) nucleoli present, (u) physical relationship of white blood cells to red blood cells, (v) reactivity, (w) vacuoles absent, (x) vacuoles present, and (y) feature presence uncertainty. The data description category also included descriptions of surrounding red blood cells or platelets. Example quotations from each data description subcategory are given in Table 20.

Table 20

Example Quotes for Data Description Codes Used in White Blood Cell Identification

Categorical Code	Example Quotes
Cell color	Novice: "It's kind of dark."
Cell maturity	Novice: "It's real immature."
	Expert: "It looks somewhat immature."
Cell shape	Novice: "It looks really perfectly round almost."
Cell size	Novice: "I see a huge cell size."
	Expert: "It's very large."
Cytoplasm amount	Novice: "Little cytoplasm."
	Expert: "It doesn't have a lot of cytoplasm."
Cytoplasm color	Novice: "The cytoplasm is pinkish blue, light-colored."
	Expert: "It has a dark blue cytoplasm."
Cytoplasm shape	Novice: "The cytoplasm is more spread out."
	Expert: "It has the cytoplasmic protrusions."
Cytoplasm texture	Expert: "Has a ground glass appearance."
Granules absent	Novice: "There is no granulation."
Granules present	Novice: "It has some light pink granules".
	Expert: "I see some red azurophilic granulation."
Halo present	Novice: " It has a halo on the side of the nucleus."
	Expert: "It has a perinuclear clearing."
N:C ratio	Novice: "The nucleus is taking up almost all of the cytoplasm."
	Novice: "The cytoplasm to nucleus is 2:1."

Table 20 Continued

Categorical Code	Example Quotes
Nuclear color	Novice: "It's really dark purple in the nucleus."
Nuclear location	Novice: "The nucleus is pushed off to the side."
	Expert: "The cell's nucleus is somewhat eccentric."
Nuclear maturity	Expert: "Really immature-looking nucleus."
Nuclear shape	Novice: "Starting to lobe."
	Expert: "It has a peanut shaped nucleus."
Nuclear size	Novice: "A large nucleus."
Nuclear texture	Novice: "There is some density in the nucleus."
	Expert: "Brain-Like convolutions of the nucleus."
Nucleoli absent	Novice: "I don't see any definite nucleoli."
Nucleoli present	Expert: "It has two distinct nucleoli."
Physical relationship to RBCs	Novice: "Encroaching red blood cells."
	Expert: "It's pushed up against the red blood cells."
Reactivity	Expert: "Looks very reactive. The coloration is reactive-looking."
Vacuoles absent	Novice: "No vacuoles."
Vacuoles present	Expert: "The cytoplasm is a little vacuolated."
Feature presence uncertainty	Novice: "Hard to tell if there is granulation."
RBC / platelet examination	Expert: "There are some large platelets present in the smear. Red cell morphology looks normal."

Data analysis and data interpretation both involve Level II type processing, as described by Bloom's taxonomy. Knowledge about WBC morphology must be both applied and analyzed in order to infer the proper cell identification for each white blood cell viewed.

Data analysis included the correlation of individual features and observed cellular patterns with each other and with the hypothesized or selected cell identifications. This requires

differentiation between the various cell features or cellular patterns observed/described and the weighting of their importance in the inference of final cell identifications. Data analysis also involves the comparison of white blood cells with others presented simultaneously or sequentially. These comparisons occur in the form of whole cell comparisons or on a feature by feature level. Data analysis also involves the defense and rationalization for both positive statements of cell identification as well as statements of non-identification. Detailed definitions for sub-categorical codes in the data analysis grouping are given in Table 21.

Table 21

Definitions for Data Analysis Categorical Codes Used in White Blood Cell Identification

Categorical Code	Categorical Definition
Correlation of Features/ID	Correlation of cell features with each other or with possible identifications. Includes (a) searching for the presence of cell features which correlate with the hypothesized identification, (b) notation of features whose presence contradict each other, (c) explanations for non-identifications, and (d) notation of features which contradict the selected identification.
Comparison to typical/normal	Comparison/contrast of WBC to what is typical/normal. Includes the comparison of WBCs to what has been encountered during past hematological experience.
Comparison of size to RBCs	Comparison/contrast of WBC to surrounding RBCs in size.
Comparison to other WBCs	Comparison/contrast of WBC to other adjacent WBCs.
Further analysis desired	Desire for analysis of further information including other similar WBCS and other laboratory results
Diagnosis inferred	Inference of patient disease diagnosis

Example quotations from each of the data analysis subcategories are given in Table 22.

Table 22

Example Quotes for Data Analysis Categorical Codes Used in White Blood Cell Identification

Categorical Code	Example Quotes
Correlations of features / identifications	Novice: Explanation for non-ID: "It's not a myelocyte because I don't see the nucleus pushed off to the side." Novice: Notation of feature that contradicts selected ID: "The cytoplasm is lighter, but I'm still going to name this one a blast." Expert: Notation of feature that contradicts selected ID: "There is vacuolization which is normally seen in monos but this is definitely a lymphocyte." Expert: Explanation for non-ID: "It's not a blast because the chromatin is too clumped." Expert: Notation of contradictory features: "It's granulated but it has an immature nucleus."
Comparison to typical / normal	Novice: "But it doesn't look like a typical plasma cell." Expert: "It's not really a regular-looking monocyte." Expert: "We're getting into some weird cells. They look nasty."
Comparison of size to RBCs	Novice: "Closer in size to the red blood cells."
Comparison to other WBCs	Novice: "This one is more rounded than the others." Expert: "The nucleus doesn't really look as immature as cell B does." Expert: "Definitely looking at all the cells, I can now say I 'm looking at vacuolated cytoplasm."
Further analysis desired	Expert: "There again, I would like to see more cells like this." Expert: "I would definitely want to know what the cell counts were and the patient's history was."
Diagnosis inferred	Expert: "Because of the different cell lines, I'm thinking maybe some kind of chronic leukemia." Expert: "There is maybe an infection going on due to monocytes and the immature neutrophils." Expert: "We could have either acute leukemia or some kind of leukoerythroblastic reaction."

The third main category of qualitative coding for the white blood cell identification scheme is data interpretation. Data interpretation requires the interpretation of individual cell features as well as cellular patterns which are recognized by the hematologist in order to solve individual cell identifications. Data interpretation may invoke the naming of a single cell identification type or the identification of an individual cell by its lineage. Detailed definitions for sub-categorical codes in the data interpretation grouping are given in Table 23. Example quotations from each of the data interpretation subcategories are given in Table 24.

Table 23

Definitions for Data Interpretation Categorical Codes Used in White Blood Cell Identification

Categorical Codes	Categorical Definition
Hypothesis	WBC identification considered, but not selected as the final cell identification.
Specific cell ID	Identification of a cell by stating the specific cell type.
ID by lineage	Identification of a cell by stating the cell's lineage rather than the specific cell identification.
Non- ID	Statement of non-identification.
Lineage reference	Identification of the general cell lineage for a cell that is also specifically identified.
Transitional ID	Identification of a cell as being transitional between two specific cell maturation stages.
Variant ID	Identification of a cell as a cell development stage/variation other those eleven cell types specified for the competency exam.
Unnamed ID	Cell for which no type of final identification was given.

Table 24

Example Quotes for Data Interpretation Categorical Codes Used in White Blood Cell Identification

Categorical Codes	Example Quotes
Hypothesis	Expert: "It could be a monocyte."
Specific cell ID	Novice: "I would probably say it's maybe an atypical lymphocyte.
ID by lineage	Expert: "Definitely lymphoid cells."
Non-ID	Expert: "It's definitely not a blast."
Lineage reference	Expert: "This is in the neutrophilic series."
Transitional ID	Expert: "Cells A, C and D all look like they're myelocytes starting to become metas."
Variant ID	Expert: "I would tend to call this an atypical prolymphocyte."
Unnamed ID	Expert: "I really don't know what it is."

The last main category of qualitative coding for the white blood cell identification scheme is meta-reasoning. The ability to think about the process of cognition itself and the monitoring of one's own thought processes is metacognition (Bruer, 1993). In this instance, meta-reasoning refers to the novices' and experts' abilities to monitor their own certainty about the cell identifications selected. Experts also commonly rate the difficulty of the identifications that they are encountering. Detailed definitions for sub-categorical codes in the meta-reasoning grouping are given in Table 25. Example quotations from each of the meta-reasoning subcategories are given in Table 26.

Table 25

Definitions for Meta-Reasoning Categorical Codes Used in White Blood Cell Identification

Categorical Codes	Categorical Definition
Uncertainty	Includes (a) general statements of uncertainty, (b) specific statements of cell identification uncertainty, and (c) the need for identification confirmation.
Difficulty Evaluation	Statement regarding difficulty level of cell identification.

Table 26

Example Quotes for Meta-Reasoning Categorical Codes Used in White Blood Cell Identification

Categorical Codes	Example Quotes
Uncertainty	Novice: “I’m really not sure about this one.”
Difficulty evaluation	Expert: “This is a hard cell.”

Qualitative Coding Total Counts

The iterative qualitative coding process used to analyze the format 1 and format 2 think-aloud protocols for both novice and experts generated four sets of categorical coding counts. Format 1 think-alouds generated fairly equal volumes of total statements for each level of expertise, 701 total statements for the novice group and 683 total statements for the expert group. Form 2 think alouds, however, generated almost twice as many verbalizations for the novice group as for the expert group. Novices made a total of 1520 statements with an average of 304 statements per novice and experts made a total of 872 statements with an average of 174.4 statements per expert. Format 2 think alouds required participants to view a total of 100 images, while format 1 think alouds only required participants to view 40 images. The fact that experts verbalized a similar number of statements to the novices in format 1, but not in format 2 may

suggest that the cognitive processes of the expert and novice are similar in format 1, but diverge significantly in format 2. It also may indicate that format 2 allows experts to function at a higher level of expertise than format 1 does. Experts functioning at the highest level of expertise are much more likely to perform "contextually based intuitive actions that are difficult or impossible to report verbally" (Ericsson, 2006, p. 12). Total categorical counts for both expert and novice are presented in Appendix G.

Question 5(a): How Do the Cognitive and Visual Examination Processes Used in the Identification of White Blood Cell Types Differ Between Experts and Novices?

Cognitive Processes of Experts and Novices Compared for Identification by Patient Format

Categorical codings for "identification by patient" think aloud protocols were compared for the two levels of expertise. A Mann-Whitney, the non-parametric equivalent of an independent *t*-test, revealed that categorical codings for experts differed significantly from categorical coding for novices. Specifically, analysis revealed that novices verbalized significantly more statements in the data description category than did experts ($p < 0.01$). Experts, however, verbalized significantly more statements in the data analysis and data interpretation categories than did the novices ($p < 0.05$). Mean percent frequencies, standard deviations, and Mann-Whitney results are represented in Table 27.

Because significant differences between experts and novices were found in the data description, data analysis, and data interpretation categories, subcategories in these areas were further analyzed for specific statistical differences. Mann-Whitney analysis revealed that experts and novices did not vary significantly from each other in any of the data description subcategories. A qualitative observation made by the researcher was that verbalizations in the "granules present" category made by experts were much more defined and specific than those made by the novices. Experts specified the types of granules present by using precise terms in their descriptions. These terms included (a) primary, (b) secondary, (c) specific, (d) toxic, and

Table 27

Means, Standard Deviations, and Mann-Whitney results for Overall Categorical Percent Totals in the Cell Identification by Patient Format

	Mean		SD		Mann-Whitney	
Overall Category	Novice	Expert	Novice	Expert	U value	p value
Data Description	53.63	25.86	6.34	7.05	0.000	0.008
Data Analysis	5.29	14.16	2.69	3.12	0.000	0.008
Data Interpretation	39.97	56.14	6.50	7.00	1.000	0.016
Meta-Reasoning	1.10	2.78	0.75	1.68	5.000	0.151

(e) azurophilic. Novices, however, simply noted the presence of granules in general, sometimes including the color of the granules.

Data description categories which resulted in 5% or more of the total verbalizations and included verbalizations from both novices and experts were (a) cell color, (b) cell maturity, (c) cell size, (d) cytoplasm color, (e) granules present, (f) nuclear shape, (g) nuclear texture, (h) nucleoli present, and (i) vacuoles present. Table 28 contains mean percents, standard deviations and Mann-Whitney results for all data description subcategories.

Evaluation of subcategories within the data analysis category revealed that the majority of the novices' verbalizations involved the correlation of cellular features with each other or the correlation of cellular features with suspected identifications. The second most significant type of verbalization for the novice was the comparison of white blood cells to others in the visual field. Percentages of expert verbalizations were more evenly split, however, involving feature/identification correlation, comparisons of the cell for identification to the idea of normal/typical, and comparison of the cell for identification to other white blood cells. Mann-Whitney analysis revealed 3 statistically significant differences between novice and expert

Table 28

Mean Percents, Standard Deviations, and Mann-Whitney results for Qualitative Coding of Cell Identifications by Patient, Data Description Category

	Mean		SD		Mann-Whitney	
Data Description Category	Novice	Expert	Novice	Expert	U value	p value
Cell color	2.13	3.47	4.33	4.85	9.000	0.548
Cell maturity	1.11	14.71	1.06	9.15	4.000	0.095
Cell shape	1.42	0.00	2.74	0.00	5.000	0.151
Cell size	3.00	5.89	2.57	9.91	11.000	0.841
Cytoplasm amount	2.38	1.85	2.03	2.55	8.000	0.421
Cytoplasm color	10.75	5.21	4.92	3.93	4.000	0.095
Cytoplasm shape/margins	1.03	3.65	1.44	3.48	9.000	0.548
Cytoplasm texture	0.08	1.89	0.18	1.93	6.000	0.222
Granules absent	1.35	2.55	2.36	5.71	9.000	0.548
Granules present	4.51	12.07	4.09	5.15	3.000	0.056
Halo present	0.40	0.33	0.68	0.75	11.000	0.841
N:C ratio	3.48	0.43	3.97	0.95	6.000	0.222
Nuclear color	4.11	0.00	6.07	0.00	5.000	0.151
Nuclear maturity	0.00	6.34	0.00	8.41	5.000	0.151
Nuclear location within cell	4.83	1.20	3.55	1.75	3.000	0.056
Nuclear shape	8.62	11.56	4.87	9.54	12.000	1.000
Nuclear size	1.27	0.37	1.20	0.83	6.000	0.222
Nuclear texture	2.61	6.38	5.40	4.89	7.000	0.310

Table 28 Continued

	Mean		SD		Mann-Whitney	
Data Description Category	Novice	Expert	Novice	Expert	U value	p value
Nucleoli absent	0.32	0.00	0.33	0.00	5.000	0.151
Nucleoli present	4.19	5.53	2.80	5.61	12.000	1.000
Physical relationship to RBCs	2.21	0.85	2.48	1.91	8.00	0.421
Reactivity	0.00	2.33	0.00	3.36	5.00	0.151
Vacuoles absent	0.16	0.00	0.22	0.00	7.500	0.310
Vacuoles present	5.30	8.16	4.00	4.37	7.00	0.310
Feature presence Uncertainty	0.56	0.80	0.45	1.09	11.500	0.841
RBCs /platelets	0.00	4.44	0.00	9.94	10.000	0.690

cognitive processing within the data analysis category. Experts were shown to have made a significantly larger percentage of statements in the "comparison to normal/typical" and the "diagnosis inferred" category than did the novice ($p < 0.05$). Novices were shown to have made a significantly larger percentage of statements in the "comparison of size to RBCs" ($p < 0.05$). Means, standard deviations and Mann-Whitney results for the data analysis category are shown in Table 29.

Statements of data interpretation made in the "cell identification by patient" format were somewhat similar for both experts and novices. Although experts did verbalize more lineage references, transitional identifications, and variant identification, only the percentage of lineage references and variant identifications were found to be statistically different from the novices. Table 30 displays the means, standard deviations, and Mann-Whitney results for the subcategories within data interpretation.

Table 29

Means, Standard Deviations, and Mann-Whitney Results for Qualitative Coding of Cell Identification by Patient, Data Analysis

	Mean		SD		Mann-Whitney	
Data Analysis Category	Novice	Expert	Novice	Expert	U value	P value
Correlations of features/IDs	44.28	29.43	12.93	19.83	6.00	0.222
Comparison to typical/normal	7.05	31.48	8.19	13.63	1.00	0.016
Comparison of size to RBCs	17.97	0.00	20.27	0.00	2.500	0.032
Comparison to other WBCs	30.70	24.14	23.53	15.78	11.00	0.841
Further analysis desired	0.00	1.60	0.00	2.20	7.500	0.310
Diagnosis inferred	0.00	13.34	0.00	15.13	2.500	0.032

Table 30

Means, Standard Deviations, and Mann-Whitney Results for Qualitative Coding of Cell Identification by Patient, Data Interpretation

	Mean		SD		Mann-Whitney	
Data Interpretation Category	Novice	Expert	Novice	Expert	U value	p value
Hypothesis	4.60	1.40	3.92	1.51	7.000	0.310
Specific cell ID	88.11	78.56	7.18	9.71	5.000	0.151
ID by lineage	2.26	3.28	3.88	3.56	7.000	0.310
Non-ID	0.92	1.69	1.00	1.30	7.000	0.310
Lineage reference	0.89	3.92	2.00	2.34	2.000	0.032
Transitional ID	0.156	2.45	0.35	2.46	6.000	0.222
Variant ID	1.55	8.24	3.47	5.06	2.000	0.032
Unnamed ID	1.52	0.46	2.16	0.70	10.500	0.690

Only two subcategories were defined within the meta-reasoning category. Although novices did make a larger percentage of statements about uncertainty while experts made a larger percentage of statements about difficulty, neither of these differences proved to be statistically significant. Table 31 displays descriptive statistics and Mann-Whitney results for the meta-analysis category.

Table 31

Means, Standard Deviations, and Mann-Whitney Results for Qualitative Coding of Cell Identification by Patient, Meta-Reasoning

	Mean		SD		Mann-Whitney	
Meta-Reasoning Category	Novice	Expert	Novice	Expert	U value	p value
Uncertainty	68.00	36.67	41.47	37.55	7.5000	0.310
Difficulty	12.00	63.33	17.89	37.55	3.5000	0.056

Cognitive Processes of Experts and Novices Compared for Single Cell Identification Format

Categorical codings for "single cell identification" think aloud protocols were compared for the two levels of expertise. A Mann-Whitney analysis indicated that there was no overall effect for expertise. Verbalizations in only 1 out of the 4 categories showed significant differences between novices and experts. Experts did verbalize significantly more statements in the data analysis category than did novices ($p < 0.05$). Mean percent frequencies, standard deviations, and Mann-Whitney results are represented in Table 32.

Since initial Mann-Whitney analysis did uncover significant differences between the amount of cognitive processing taking place in the data analysis category for the experts and novices, subcategories were further analyzed for differences using the Mann-Whitney test. Novices were found to make a significantly higher percentage of size comparisons to red blood

Table 32

Means, Standard Deviations, and Mann-Whitney for Overall Categorical Percent Totals in the Single Cell Identification Format

	Mean		SD		Mann-Whitney	
Overall Category	Novice	Expert	Novice	Expert	U value	p value
Data Description	63.39	48.53	9.55	13.98	4.000	0.095
Data Analysis	3.97	13.43	3.61	7.46	2.000	0.032
Data Interpretation	31.42	37.43	6.12	9.63	7.000	0.310
Meta-Reasoning	1.31	0.98	2.08	1.12	12.000	1.000

cells and experts were discovered to make a statistically significantly greater number of statements expressing their desire for further analysis (analysis of other laboratory data, diagnosis, patient history, etc.). Quantitative results from this analysis can be found in Table 33.

Table 33

Means, Standard Deviations, and Mann-Whitney Results for Qualitative Coding of Single Cell Identifications

	Mean		SD		Mann-Whitney	
Data Analysis Category	Novice	Expert	Novice	Expert	U value	p value
Correlations of features/IDs	20.00	51.64	29.82	22.79	5.000	0.151
Comparison to typical/normal	11.67	28.20	16.24	14.05	5.000	0.151
Comparison of size to RBCs	51.67	0.00	45.80	0.00	2.500	0.032
Comparison to other WBCs	16.67	3.68	28.87	5.15	10.500	0.690
Further analysis desired	0.00	12.56	0.00	7.85	2.500	0.032
Diagnosis inferred	0.00	3.83	0.00	4.20	5.000	0.151

Question 5b: How Do the Cognitive and Visual Examination Processes Used in the Identification of White Blood Cell Types Differ Within Expert and Novice Groups Themselves When Image Format Is Altered?

Wilcoxon Signed Ranks tests, the non-parametric equivalent to the correlated groups *t*-test, were performed to discover if changing the think-aloud protocol format revealed any shift in processing focus for either the novice or the expert. As previously explained, format 1 required the viewing of 40 single WBC images and format 2 required the viewing of 100 WBC images (20 four-cell composites and 20 single WBC images). Simply because there were more images viewed in format 2 than in format 1, the number of data interpretation statements was expected to and did increase for both levels of expertise. The primary interest of this study was to find out if the mechanism by which the novice and expert reached the level of data interpretation significantly shifted when the image presentation format was changed. For this reason, only the three non-interpretative data processing categories were considered in these analyses.

Cognitive Processes of Novices (Cell Identification by Patient Versus Single Cell Identification)

The Wilcoxon Signed Ranks test revealed that the novices' non-interpretative processing focus did not shift when the think aloud protocol was altered. Mean percentages for statements verbalized in the data description, data analysis, and meta-reasoning categories did not significantly change ($p > 0.1$) as a result of format change. Table 34 displays these mean percentages and Wilcoxon Signed Ranks test results.

Cognitive Processes of Experts (Cell Identification by Patient Versus Single Cell Identification)

Wilcoxon Signed Ranks test revealed that the experts' non-interpretative processing focus did shift significantly when the think aloud protocol was altered. Mean percentages for statements verbalized in the data description were significantly greater ($p < 0.05$) in format 1 than in format 2. Table 35 displays these mean percentages and Wilcoxon Signed Ranks results.

Table 34

Means, Standard Deviations, and Wilcoxon Signed Ranks Test for Overall Novice Categorical Percent Totals, Single Cell Identification (Format 1) Versus Cell Identification by Patient (Format 2)

	Mean		SD		Wilcoxon Signed Ranks	
Overall Category	Format 1	Format 2	Format 1	Format 2	Z value	p value
Data Description	92.05	89.36	5.83	4.33	-0.674	0.500
Data Analysis	6.03	8.75	5.57	4.01	-0.674	0.500
Meta-Reasoning	2.01	1.89	3.20	1.41	-0.135	0.893

Table 35

Means, Standard Deviations, and Wilcoxon Signed Ranks Test for Overall Expert Categorical Percent Totals, Single Cell Identification (Format 1) Versus Cell Identification by Patient (Format 2)

	Mean		SD		Wilcoxon Signed Ranks	
Overall Category	Format 1	Format 2	Format 1	Format 2	Z value	p value
Data Description	76.76	58.54	12.37	11.48	-2.023	0.043
Data Analysis	22.02	32.22	12.76	3.72	-1.214	0.225
Meta-Reasoning	1.63	6.50	2.04	4.15	-1.753	0.080

Case Study Interview Results

Question 1a: Novice Response. Question 1 was "Please describe for me, in as much detail as possible, the general thought process you use when identifying white blood cells while performing a differential count." The two part answer given by all novices included: (a) the use of specific cellular features in the classification of white blood cells, and (b) a description of the need for viewing similar cells. Novice 5 explained "The first thing I try to do is scan through the slide and just get a feel for what's on the slide." Novice 3 supported this same idea by saying "If

I can't see the features that I'm looking for in what I'm looking at – the cytoplasm, the nucleus, the size- I try to look around the field to see if I can see something to support my answer."

Question 1b: Expert Response. The answers given by the experts represented two slightly different viewpoints. Experts 1, 2, and 4 described somewhat different processes than did Experts 3 and 5. Experts 1, 2, and 4 all stated that they first examine the results of the complete blood count (CBC). The automated CBC results are analyzed for the presence of abnormal red blood cell, white blood cell (WBC), and platelet counts. The blood smear would then be scanned for WBC numbers/appearance, RBC morphology, and platelet clumping. WBC features such as nuclear texture and cytoplasmic granularity/texture would then be used to differentiate between the various cell types/lineages. If abnormalities were seen upon manual review of the blood smear, experts indicated that they would then correlate the manual differential counts with any automated instrument flags indicating the presence of suspect abnormal WBC counts or suspect abnormal WBC types. The patient history information (diagnosis) would also be reviewed for patients having abnormal manual cell counts.

Experts 3 and 5 indicated that instead of beginning with an examination of the CBC and available patient data, they would instead begin directly with a review of the manual blood smear itself and begin to do the differential count. As Expert 5 explained, "I don't want to be swayed by the CBC data." Both of these expert technologists preferred instead to focus on the morphology of the cells themselves. Also, they both explained the importance of scanning various fields of the blood smear. Expert 3 indicated that if she had difficulty with the identification of a cell by examining its features, she would "scan the slide first and try to give thought as to what kind of company these cells are keeping." Expert 5 further explained that "If you see one (abnormal cell), you're not going to call it because it could just be there….. If you call it, the doctor is going to be upset about it. If you see two or more, then it's a problem. Then

you have to call them". These two technologists indicated that only on very difficult and in rare cases would they refer to available laboratory results and diagnosis information on the patient. Additionally, they agreed that review of such laboratory results would only take place after performing the initial manual differential count.

Question 2: Expert Response. Question 2 was "Through your years of experience, how have you seen your skill as a cell morphologist evolve?" This question was only asked to the expert participants. Experts expressed that their confidence in their own cell identification skills and their speed of identification had increased over the years. Their knowledge about each individual cell type has also grown immensely over the years. Expert 2 felt that "you become accustomed to what you're looking at and usually you become accustomed to the abnormal things you're looking at". Expert 5 explained that by retaining the information he learns from each new cell/patient he encounters he adds to his knowledge with each new experience. Several experts also explained that they learned how to look at the entire patient picture, instead of just at the WBC cells. Expert 4 stated that "I know now to look at the CBC, to look at the patient's diagnosis, and to get an overall picture of everything that's going on in the blood other than just the white blood cells. When you know the overall picture, it's a lot easier to distinguish what the cells are that you are looking at."

Question 3a: Expert Response. Question 3 was "What type of experiences/activities have you found to be the most critical in developing/improving your morphology skills? Specifically, why have you found such experiences/activities so critical?" Experts listed the following experiences as being the most critical in developing/improving their morphology skills: (a) concentrated work experience in hematology, (b) teaching WBC morphology to students, (c) reviewing leukemic/abnormal blood smears with a pathologist or experienced technologist, (d) performing differential counts on bone marrows, (e) performing advanced

hematological tests such as flow cytometry, (f) holding senior technologists position and having to make final decisions on cell identifications, (g) discussing difficult cells with colleagues, and (h) reviewing cell morphology books that provide detailed explanations for the basis of cell identifications.

The experiences/activities listed by technologists as critical for building morphology skills were important for a variety of reasons. Concentrated work experience in hematology and performing differential counts on bone marrows were found to be critical because of the chance for exposure to a larger number of WBC as well as wider range of WBC varieties. Reviewing/discussing abnormal and leukemic smears with pathologists/other technologists and reviewing cell morphology books were useful because of the exposure to different viewpoints that could be used to mold an individual technologist's ideas. Teaching students and holding senior technologists positions were deemed critical because they required the technologist to build confidence and competence in cell identification skills. Both of these positions also required the technologist to be able to clearly verbalize accurate reasoning for the determination of cell morphologies. Performing advanced hematological testing such as flow cytometry aided the technologist in providing an overall or complete picture in regards to the patient diagnosis.

Question 3b: Novice Response. Novices all agreed that seeing large volumes of images, whether presented digitally or under a microscope, was the activity most critical in building their morphology skills. Novice 4 justified this by saying, "every lymphocyte looks different. They kind of have the same characteristics, but sometimes they don't all look exactly the same, so just looking at all different kinds of cells, different patients, different blood smears.". They all also valued hearing/reading detailed morphology descriptions from a text/during a classroom presentation. One-on-one interaction with the instructor in discussing cellular morphologies viewed under the microscope was also deemed critical.

Question 4a: Expert Response. Question 4 was "In your opinion, what specific cell types are the most difficult to distinguish from each other? Why?" The experts reached consensus on two difficult cell pairings: (a) monocytes and lymphocytes, and (b) atypical lymphocytes and blasts. Cells in each of these pairings were described as having characteristic cell features that were very similar to each other.

Question 4b: Novice Response. Novices did not reach consensus on difficult cell pairings. Difficult cell pairings listed by the novices included (a) monocytes and blasts, (b) lymphocytes and blasts, (c) monocytes and promyelocytes/myelocytes, (d) blast and lymphocyte (e) lymphocyte and plasma cell, (f) blast and promyelocyte, and (g) lymphocyte and monocyte. As with the experts, novices explained that they found differentiation between these particular cell types difficult because of the feature similarities between the paired cell types.

Question 5a: Expert Response. Question 5 was "When you find yourself trying to differentiate/discriminate between various morphologic cell types while trying to identify a difficult cell, what special thought processes/methods do you use?" Experts agreed that the identification of difficult cells requires scanning the slide to find similar cells for comparison. For further confirmation, each expert explained that they would use the techniques previously outlined for interview question 1 (reliance of feature discrimination/differentiation or review of CBC and patient diagnosis).

Question 5b: Novice Response. Novices came to the consensus that the identification of difficult cells requires not only careful examination/discrimination based on the cellular features present, but also a comparison to other white blood cells from the same patient blood smear. Novice 5 states that "I just remember the identifying features…just run through that in my head...I try to identify as many features as I can and overanalyze the cell to try to find something that identifies it as one or the other." Novice 4 even mentioned the importance of the cellular

background of the smear itself in the identification of a single cell. She explained that myeloid cells may be seen in the presence of other immature myeloid cells, but that the mature lymphocytes would be more likely to be seen in a background of mature monocytes.

Discussion of Cognitive Processing Styles

There are several aspects of the quantitative and qualitative results that can be correlated in order to draw appropriate conclusions regarding the cognitive processing styles used by both experts and novices. Correlation between various aspects of the study results leads to several conclusions including: (a) both experts and novices rely, at least partially, on an analytical type processing/feature evaluation protocol when identifying white blood cells, especially in the "single cell identification" format, (b) viewing images in the "cell identification by patient" format institutes a cognitive processing shift on difficult cells for experts but not for novices, (c) experts institute similarity-based processing when viewing images in exam format 2, and (d) experts exhibit greater metacognitive abilities than do novices.

Evidence for Analytical Processing

Quantitative analysis of the exam results showed that for both experts and novices the largest majority of cells identified above the designated consensus/competency levels were mature myeloid cells and lymphocytes (primarily those with unambiguous features) (Table 8 and 10). Experts identified > 90% of the mature myeloid cells accurately and > 60% of the lymphocytes correctly, while novices identified > 60% of the mature myeloid cells, but only about 6% of the lymphocytes with > 80% competency. The fact that experts exhibited their best performance on cell types that were easy/typical and that they performed significantly better on such easy/typical cells than did the novices suggests an analytical rather than a similarity-based model of cognitive processing. These findings do directly contradict the evidence found by Norman et al. (1989) in the area of dermatological expertise. Norman et al. (1989) determined

that the ratio of errors on typical slides remained constant for all levels of expertise, thus supporting the existence of a similarity-based model of expertise.

Expert performance for the mature myeloid and lymphocytic cells did correlate reasonably well with the 2000-2006 CAP proficiency summary results presented previously in the literature review. CAP results showed that for greater than 90% of neutrophils and basophils (two of the mature myeloid subtypes) the 90% consensus level was reached. For eosinophils (the third mature myeloid subtype), however, only 78% of the cells were identified above the 90% consensus level. These CAP results suggest, however, reasonable correlation with the > 90% performance level on mature myeloid cells found in the current study. Performance on the lymphocytic cells in CAP proficiency testing showed that 60-70% was identified at the consensus level, thus correlating with the > 60% performance level determined by the current study.

Statistical analysis of qualitative codings generated for the "single cell identification" format suggests that the cognitive processes for both the expert and the novice rely heavily on feature (data) description and do not differ significantly from each other in this area (Table 32). The only categorical area of coding for which the novice and expert did differ significantly was in data analysis, with the experts making more statements in this category. The only subcategory for which a significantly larger percentage of verbalizations occurred for the expert than for the novice was in their desire to perform further analytical processes (Table 33). This suggests the experts' awareness of their own increased use of analytical processing and the apparent decrease in their ability to perform higher level cognitive functions in the "single cell identification" format.

Evidence for Cognitive Processing Shifts by Experts

Results from the quantitative portion of the study clearly indicate that a cognitive processing shift occurs when images are presented to experts using exam format 2. A cognitive processing shift was not indicated for the novice group. There were three differential quantitative outcomes which indicated the presence of a processing shift for the experts. First, exam format 2 resulted in significant performance increases in the blast, monocyte, and lymphocyte subcategories for experts (Table 9), while no significant performance increases were found for the novices (Table 11). Second, correlation coefficients between exam performance outcomes and response times (Table 12) shifted significantly for the experts between format 1 (r = -0.740) and format 2 (r = -0.536), while the novice correlation coefficient remained fairly constant (r = -0.506 to -0.546). Third, although average expert response times were shorter that average novice response times, they were not significantly different from each other in format 1 (Table 15). Expert response times were, however, significantly shorter than novice response times in format 2.

The differential shortening of expert response times and the coinciding categorical performance increases observed with exam format 2 suggests a potential shift in cognitive processing from an analytical form to a similarity or instance-based cognitive processing style. Previous studies in dermatology and microscopic pathology suggest that a quickening in response times and an improvement in performance may indicate the implementation of a more automatic, pattern-recognition process (Norman et al., 1989; Crowley et al., 2003).

Statistical analysis of qualitative codings also supported the existence of a cognitive processing shift for experts. The analysis of format 2 think-aloud protocols indicated that experts verbalized significantly different types of statements than did novices (Table 27). Experts verbalized a significantly smaller percentage of data description statements and a

statistically larger percentage of data analysis and data interpretation statements. Because statistical analysis indicated no overall effect for expertise on the format 1 verbalizations, but a significant overall effect for expertise on the format 2 verbalizations, it is hypothesized that the expert has shifted from the focused reliance on feature characterization to some higher levels of processing requiring comparison and other pattern-matching processes.

Further statistical analysis comparing format 1 and 2 verbalizations for the experts in the non-interpretive data processing categories were vital as well (Table 35). Experts verbalized a significantly smaller percentage of statements in the data (feature) description category when viewing cells in the "cell identification by patient" format than when viewing cells in the "single cell identification" format. Cognitive processing shifts were not suggested for the novice group since their distribution of verbalizations in the non-interpretive data processing categories did not differ significantly between the two presentation formats.

Evidence for Use of Similarity-Based Processing in Format 2 by Experts

A variety of additional qualitative evidence supports the use of similarity-based processing by experts when viewing format 2 cell identifications. Analysis of the format 2 verbalizations from the data analysis category (Table 29) revealed that experts make a significantly larger percentage of statements comparing the cell for identification to the concept of normal or typical than do novices. This suggests that the experts are relying on prior experiences in order to generate this concept of "normal" and therefore is using "instance-based categorization" in their reasoning. In conjunction with this analysis, it was shown that experts also made a significantly greater percentage of data interpretation statements (Table 30) in the "variant identification" category than did novices. This again implies that experts have a large number of past experiences to guide them, allowing them to recognize white blood cells as variants from the normal.

Further evaluation of cell exam results yielded additional support for the theory of similarity-based processing use. Because accuracy in identification of difficult/atypical cell types was increased after viewing patient differential simulations, it may be theorized that similarity-based processing is being applied at the level of the patient case and not at the level of the individual cell itself. Once the current patient case is matched to a prior instance from the expert's experience, instance-based categorization at the diagnostic level can occur. This allows for the recognition of individual cells within difficult cases by allowing the expert to apply newly available context-dependent knowledge about the particular cell in question. Recognition of such individual cells may have not been possible in the "single cell identification" format because in order for the expert to access the similarity-based information required for proper cell identification a contextualized representation of the cell was required. This theory is consistent with the idea of feature re-interpretation and the importance of the perceptual information derived from visual stimuli (Hatala et al., 1999; Kulantunga-Moruzi et al., 2004). This theory also coincides with the findings reported by Chi (2006) about the context-dependent nature of domain-specific knowledge for an expert. He reports that when expert physicians were presented with contextual cues such as patient symptoms, medical charts, as well as pictures of the patient, the physician was able to diagnose the patient much more accurately.

The existence of such case level pattern matching is further suggested by expert responses to interview question 1 as well as format 2 think-aloud verbalizations. Cross-Case analysis of question 1 revealed that three of the experts routinely evaluate the holistic patient picture (including diagnosis and CBC data) when making case level analyses, while the other two experts specifically evaluate white blood cell morphology when making case level comparisons with past experiences. Occurrences of broad case-level analyses are illustrated through format 2 think-aloud statements made by Expert 2: "We've had ALL (acute

lymphoblastic leukemia) patients look just like this". Expert 3 made format 2 think-aloud statements indicative of a morphological-type case level comparison when he stated, "I've gone to the pathologists with cells like this and they said just to call them plasmacytoid lymphs. Some of these hard cells could be the same thing". Statistical analyses of "cell identification by patient" verbalizations also showed that only experts made statements which suggested pathologic diagnoses while performing the cell identification exercises, again indicating their ability to process at the level of the holistic case.

Metacognition

Correlation of quantitative and qualitative data also illustrated the unique metacognitive abilities which were apparent for experts, but not for the novices. Detailed analysis of subcategories existing within the meta-reasoning category for the "cell identification by patient" format revealed that experts do make a higher percentage of statements (Table 31, $p = 0.056$) about the difficulty level of the cell identification than does the novice. This may represent their strong metacognitive abilities. Additional evidence of this ability to accurately evaluate cell difficulty was revealed through correlation of cell exam error types and cross-case analysis of answers to interview question 4. The three most common error types revealed through quantitative analysis of expert exam performance were (a) incorrect cell maturation stage for myeloid cells (25%) (b) blasts misidentified as atypical lymphocytes (20.2%), and (c) monocytes misidentified ad lymphocytes (10.7%). During the interview, experts did repeatedly name the atypical lymphocyte/blast pairing as well as the lymphocyte/monocyte pairing as being difficult. Although no technologist specifically identified a difficulty in determining myeloid maturation stage, one technologist did recognize the difficulty in identifying the promyelocyte cell.

Novices, on the other hand, did not display the ability to verbalize thoughts about cell difficulty during the think-aloud protocols and did not predict their error types as closely as did

the experts. The three most prominent error types for the novices were (a) incorrect myeloid maturation stage (19.7%), (b) lymphocytes misidentified as myeloid cells (19.0%), and (c) lymphocytes misidentified as monocytes (11.0%). One novice mentioned difficulty in differentiating blasts from the promyelocyte, which corresponds with the error identified as incorrect myeloid maturation stage. Another student listed lymphocytes and monocytes as being difficult to distinguish from each other. No novice recognized their confusion of lymphocyte and myeloid cells. In summary, there was a general lack of consensus amongst the novices on which cells were difficult and, in general, novices were not able to predict the cell types that they actually did have the most trouble in identifying. These conclusions indicate that novices in the area of hematology have not developed the metacognitive skills which are possessed by experts.

CONCLUSION

The visual identification of white blood cells is an essential hematologic skill, one which requires dedicated time commitments and vast clinical experience in order for true expertise to develop. The most important and essential outcome of this study is that of the apparent benefit derived by experts from the presentation of white cells for identification in the context of the patient's cellular background. Implications for current proficiency testing may include the use of a more extensive white blood cell image series for each patient case presented. Such images could be included alongside the written contextualizing information already provided by CAP. An integral part of the design for such image sets would require the presentation of multiple white blood cells in the simultaneous presentation format in order to allow for effective between cell comparisons. Inclusion of such an image series would promote the expert's use of similarity-based processing and developed visual perceptual skills, both which could be applied at the case level. Use of such image sets would also aid in creating a proficiency exam that more closely mirrored the skills required in day-to-day clinical practice.

Implications for current certification/competency testing may also include the use of more extended white blood cell image sets with questions that currently require the identification of a single white blood cell. Such image sets could potentially be included in Powerpoint® style presentation slides. Although this study did suggest that outcome performances for the average student would not be differentially improved by altering the presentation format, it seems reasonable to assume that the performance of a student who has achieved advanced skill levels in the area of hematology may be differentially affected. Another very reasonable application of this suggested format improvement would be its use in the preparation of questions for specialist certification in the area of hematology. It is important to note that consideration of such changes in certification format comes at a historically critical time in the field of Clinical Laboratory

Sciences. At the March 2, 2006 Annual Clinical Laboratory Educator's Conference held in San Antonio, Texas it was announced that the National Credentialing Agency for Laboratory Personnel, Inc. (NCA) and the Board of Registry (BOR) of the American Society for Clinical Pathology (ASCP) have made plans to unite in the formation of a single credentialing agency (Fritsma, Summer 2006). With such plans still underway, now is an ideal time for those in the field of CLS to consider new ideas in the creation of a joint certification exam.

Further areas of application of these study results may be in the area of teaching. Instructors in the area of Clinical Laboratory Sciences may wish to promote the development of the contextualized visual process by instituting aspects of (a) David Ausubel's (1968) theory of meaningful learning, (b) Ellen Langer's (1997) mindful learning theory, and (c) David Kolb's learning cycle as described by Zull (2002). Each of these learning theories supports the use of contextualized learning in building developed connections within and between vast knowledge networks. Building such knowledge networks may allow the novice to more easily visualize the comparisons/contrasts between various cell morphologies as well as recognize the presence of cell morphology patterns and their association with case level meaning. Use of comprehensive patient image sets instead of random compilations of single white blood cell images in classroom teaching may help to further contextualize a student's knowledge, building connections to case level knowledge and promoting the use of similarity-based processing.

Results from the qualitative portion of this study indicate that although experts may institute some form of analytical processing when presented with single white blood cell images, they may rely more heavily on the use of similarity-based processing when white blood cell images are presented in the context of a patient differential. Results suggest that novices rely very heavily on the use of analytical processing or feature assessment of white blood cells for all formats of image presentation. Novices are inept, however, at fully instituting this processing

model because they are only able to apply it for the most unambiguous examples of the most typical cell types. Novices are inefficient at the differential weighting of cellular features which is a crucial part of the analytical processing model (Kulatunga-Moruzi et al., 2001; Norman et al. 1990). Several more in-depth qualitative studies would be required in order to further investigate the true nature of cognitive processing as used by the novice and expert in the field of clinical hematology.

Study Limitations

There are several limitations in the ability to generalize from the results of this study to the larger population of CLS students and experts. The most apparent limitation is the use of a convenience sample, including students from Our Lady of the Lake College and professionals from Baton Rouge area hospitals. Although OLOLC has a CLS curriculum that is similar to other CLS programs nationwide, each program may differ in their content emphasis, methods and modes of delivery, course assignments, and course sequencing. Clinical practice in Baton Rouge area hospital laboratories should, as well, mirror those at other hospital labs nationally. The possibility of geographic variations does exist, however, especially since several of the Baton Rouge area hospitals are under the guidance of the same group of pathologists, and are likely to follow very similar rules of thumb. Findings from this study could be further validated by using a much larger, and more nationally representative population as well as through the development of a much more extensive exam containing a larger number of cellular images.

One existing limitation for the quantitative study was determination of the correct cell identification for each exam item. A very methodical process was used for the identification of all cells and involved the correlation of expert data with several other sources of cell identification including the original patient differential report, evaluation of the original blood smear, and pathologist identifications. Four cells for which consensus identification could not be

achieved were completely removed from exam analysis. Despite the use of these extensive processes, the possibility for error still exists. Another limitation of the cell exam was the possible bias that existed in the selection/availability of the patient blood smears used for creation of the exam. Although blood smears did represent patients having a variety of illnesses, the researcher's access to patient specimens was limited to those blood samples collected over the last two years by technologists working at Our Lady of the Lake Regional Medical Center.

Further limitations of the study exist for the qualitative portion of the study. Since this study only involved the analysis of 5 case studies at each level of expertise and a normal distribution of data could not be assumed, non-parametric inferential statistics were performed on the mean categorical percentages generated. Using non-parametric statistics is not, however, as powerful as using parametric statistics. A future study in which cognitive processing styles for the novice and expert were of central focus and which utilized a larger and more heterogeneous sample population would undoubtedly yield more generalizable results. With only 10 case studies to evaluate, the amount of individual variation between case study participants was very likely to have an affect on the overall results. Also inherent to limitations of the qualitative methods was the non-random selection of case study participants. Although cell exam performance was used as a criterion in participant selection, availability and willingness of the individual was also a primary factor in the selection process.

Other limitations to the qualitative portion of the study encompass issues of reliability and validity. Although the researcher used member checking, triangulation of data between quantitative and qualitative study results, and made multiple passes at analyzing the qualitative data, it may have been possible that another researcher would have coded the data in a slightly different manner.

Recommendations for Future Research

Recommendations for future research include a wide variety of studies. The first of these suggested studies would be a nationally based study to further evaluate and validate the apparent effect that visual presentation context seems to have on expert performance during proficiency-type testing. Research questions that could be considered in such a study would include the following: (a) Are there significantly different outcome effects if experts are allowed to view patient background slides at their own pace instead of using an automated timing device?, and (b) Does exam format effect performance outcomes for clinical laboratory generalists in a significantly different way from clinical laboratory hematology specialists?. Future studies could also further investigate the lack of differential effect that format had on performance outcomes for immature myeloid cells. A more concentrated study with a much wider variety of cell examples could be prepared and tested. Continued research could also more specifically evaluate that effect that exam format has on students who have acquired various levels of hematological understanding. Such a study could be used to determine if exam format differences can provide possible benefits for students who acquire clinical skills that are above average.

Although more than a decade of research exists on the cognitive processing strategies which are used in other areas of visual diagnosis including dermatology, radiology, and microscopic pathology, this is the first study to try to describe the cognitive processing strategies applied specifically in the area of clinical hematology. Further studies targeted specifically for this purpose could be planned. Such studies could use a quantitatively larger number of participants who represent a variety of heterogeneous characteristics, serving to provide a better cross-section of clinical laboratory professional characteristics.

REFERENCES

Allen, S. W., Norman, G., & Brooks, L. R. (1992). Experimental studies of learning dermatologic diagnosis: The impact of examples. *Teaching and Learning in Medicine*, 4(1), 35-44.

Allied Health Continuing Education. (n.d.). Retrieved October 10, 2006, from http://www.lsbme.louisiana.gov/allied_health.htm.

Ausubel, D. (1968). *Educational psychology: A cognitive view*. New York: Rinehart and Winston.

ATLASti Scientific Software Development GmbH (2006). ATLASti (Version 5.0) [Computer software]. Berlin, Germany: ATLASti Scientific Software Development GmbH.

Azevedo, R., & Lajoie, S. P. (1998). The cognitive basis for the design of mammography interpretation tutor. *International Journal of Artificial Intelligence in Education*, 9, 32-44.

Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice*. Menlo Park, California: Addison-Wesley Publishing Company.

Bloom, B.(1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook* I: *Cognitive domain*. New York: David McKay.

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General*, 120, 278-287.

Bruer, J. T. (1993). *Schools for thought: A science of learning in the classroom.* Cambridge: MIT Press.

Burthem, J., Brereton, M., Ardern, J., Hickman, L., Seal, L., Serrant, A., et al. (2005). The use of digital 'virtual slide' in the quality assessment of haematological morphology: Results of a pilot exercise involving UK NEQAS(H) participants. *British Journal of Haematology*, 130, 293-296.

CAT examination administration and examination results. (n.d.) Retrieved October 19, 2006, from http://www.ascp.org/Certification/CertifyingExaminations/ cert_procedures/ administration/default.aspx

Certificate Maintanence. (n.d.). Retrieved October 18, 2006, http://www.ascp.org/ certification/CMP/

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21-30). Cambridge, NY: Cambridge University Press.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6(3), 271-315.

Cooke, Nancy J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, 41, 801-849.

Crowley, R.S., & Medvedeva, O. (2006). An intelligent tutoring system for visual classification problem solving. *Artificial Intelligence in Medicine*, 36(1), 85-117.

Crowley, R.S., Naus, G.J, Stewart, J.S., & Friedman, C.P. (2003). Development of visual diagnostic expertise in pathology: an information-processing study. *Journal of the American Medical Informatics Association*, 10(1), 39-51.

Diagnostic Instruments, Inc. (2002). SPOT (Version 3.5.2) [Computer software]. Sterling Heights, MI: Diagnostic Instruments, Inc.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: The Free Press.

Dwyer, F.M. (1978). *Strategies for improving visual learning*. State College, PA: Learning Services.

Ericsson, K. A. (2006). An introduction to *The Cambridge handbook of expertise and expert performance:* its development, organization, and content. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 587-611). Cambridge, NY: Cambridge University Press.

Ericsson, K.A. (2003). The search for general abilities and basic capacities. Theoretical implications from the modifiability and complexity of mechanisms mediating expert performance. In R.J. Sternberg & E.L. Grigorenko (Eds.), *The psychology of abilities, competencies, and expertise*. (pp. 93-125). Cambridge, United Kingdom: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.

Ericsson, K.A. & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data.* (Rev. ed.). Cambridge.:MIT Press.

Ericsson, K. A., & Smith J. (1991). Empirical study of expertise: Prospects and limits. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: prospects and limits* (pp. 1-38). Cambridge: Cambridge University Press.

Etelapelto, A. (1998). *The development of expertise in information systems design*. Jyvaskyla: University of Jyvaskyla.

Fleming, M., & Levie, W. H. (1978). *Instructional message design: Principles from the behavioral sciences.* Englewood Cliffs, NJ: Educational Technology Publications.

Fritsma, G. A. (Summer 2006). NCA and BOR consider new credentialing agency. *Clinical Laboratory Science,* 19(3), 132-133.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Glaser, B.G, & Strauss, A.L. (1967). *The discovery of grounded theory; strategies for qualitative research.* Chicago: Aldine Publishing Company.

Goldsmith, E. (1987). The analysis of illustration in theory and practice. In H. A. Houghton & D. M. Willows (Eds.), *The psychology of illustration* (pp. 53-85). New York: Springer-Verlag New York, Inc.

Gowin, D.B. (1981). *Educating*. Ithaca, NY: Cornell University Press.

Harmening, D.M. (2002). *Clinical hematology and fundamentals of hemostasis* (4th ed.). Philadelphia: F.A. Davis Company.

Hart, A. (1986). *Knowledge acquisition for expert systems.* New York: McGraw Hill.

Hatala, R., Norman, G. R., & Brooks, L. R. (1999). Impact of a clinical scenario on accuracy of electrocardiogram interpretation. *Journal of General and Internal Medicine*, 14, 126-129.

Hinkle, D.E., Wiersma, W, & Jurs, S.G. (1998). *Applied statistics for the behavioral sciences* (4th ed.). Boston: Houghton Milton Company.

Hoffman, R. R., Shadbolt, N. R., Burton, A.M., & Klein, G. (1995). Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes,* 62(2), 129-158.

Horn, J., & Masunaga, H. (2006). A merging theory of expertise and intelligence. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 587-611). Cambridge, NY: Cambridge University Press.

Hutchinson, C. V., Brereton, M. L., & Burthem, J. (2005). Digital imaging of haematological morphology. *Clinical Laboratory Haematology*, 27, 357-362.

Kim, S., & Astion, M. (2000). Patterns of image comparison using compare and contrast feature in Urinalysis Tutor™. *British Journal of Educational Technology*, 31(4), 349-357.

Kulatunga-Moruzi, C., Brooks, L.R., & Norman, G. R. (2001). Coordination of analytic and similarity-based processing strategies and expertise in dermatological diagnosis. *Teaching and Learning in Medicine*, 13(2), 110-116.

Kulatunga-Moruzi, C., Brooks, L.R., & Norman, G. R. (2004). Using comprehensive feature lists to bias medical diagnosis. *Journal of Experimental Psychology*, 30(3), 563-572.

Kundel, H.L., & Nodine, C.F. (1983). A visual concept shapes image perception. *Radiology*, 146, 363-368.

Langer, E. (1997). *The power of mindful learning*. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc.

Larkin, J. H., & Chabay, R. W. (1989). Research on teaching scientific thinking: Implications for computer-based instruction. In L. B. Resnick & L. E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research*. Association for Supervision and Curriculum Development.

Lee, S. (2005). Virtual microscopy: Applications to hematology education and training. *Hematology*, 10 (Supplement 1), 151-153.

Levie, W. H.. (1987). Research on pictures: A guide to the literature. In H. A. Houghton & D. M. Willows (Eds.), *The psychology of illustration* (pp. 3-50). New York: Springer-Verlag New York, Inc.

Marchevsky, A. M., Wan, Y., Thomas, P., Krishnan, L., Evans-Simon, H., Haber, H. (2003). Virtual microscopy as a tool for proficiency testing in cytopathology. *Archives of Pathology & Laboratory Medicine*, 127, 1320-1324.

Mayer, R.E., & Sims, V.K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 86(3), 389-401.

McKenzie, S. B. (2004). *Clinical laboratory hematology*. Upper Saddle River, NJ: Pearson Education, Inc.

Merriam, S. B. (1988). *Case study research in education: A qualitative approach*. San Francisco: Jossey-Bass Publishers.

Microsoft Visual Studio (2005). Microsoft Visual Basic (Version 2005) [Computer software]. Redmond, Washington: Microsoft Corporation.

Millard, W. L. (1964). Visual teaching aids: Production and use. *The encyclopedia of photography*. New York: The Greystone Press.

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 101(2), 343-352.

Mintzes, J.T., Wandersee, J.H.,& Novak, J.D. (Eds.). (1998). *Teaching science for understanding: A human constructivist view*. San Diego, CA: Academic Press.

NCA Candidate Handbook. (n.d.). Retrieved October 19, 2006, from http://www.nca-info.org/examinations_application-handbook.asp

Nodine, C. F., Kundel, H. L., Lauver, S.C., & Toto, L.C. (1996). Nature of expertise in searching mammograms for breast masses. *Academic Radiology*, 3, 1000-1006.

Norman, G. R., Brooks, L. R. (1997). The non-analytic basis of clinical reasoning. *Advances in Health Science*, 2, 173-184.

Norman, G.R., Brooks, L.R.., & Allen, S. W. (1989). Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology. Learning Memory, and Cognition*, 15(6), 1166-1174.

Norman, G.R., Brooks, L.R.., Allen, S.W., & Rosenthal, D. (1990). Sources of observer variation in dermatologic diagnosis. *Academic Medicine*, 65(9), S19-S20.

Norman, G.R., Brooks, L.R., Coblentz, C.L., & Babcook, C.J. (1992). The correlation of feature identification and category judgments in diagnostic radiology. *Memory and Cognition*, 20(4), 344-355.

Norman, G.R., Brooks, L.R.., Colle, C. L., & Hatala, R. M. (2000). The benefit of diagnostic hypothesis in clinical reasoning: Experimental study of an instructional intervention for forward and backward reasoning. *Cognition and Instruction*, 17(4), 433-448.

Norman, G.R., Rosenthal, D., Brooks, L.R., Allen, S.W., & Muzzin, L.J. (1989). The development of expertise in dermatology. *Archives of Dermatology*, 125, 1063-1068.

Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps™ as facilitative tools in schools and corporations*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Novak, J.D., & Gowin, D.B. (1984). *Learning how to learn*. Cambridge, MA: Cambridge University Press.

Novak, J. D., Mintzes, J.J., & Wandersee, J. H. (2000). Learning, teaching, and assessment: A human constructivist perspective. In J. J. Mintzes, J. H. Wandersee,& J. D. Novak (Eds.), *Assessing science understanding: A human constructivist view* (pp. 1-16). San Diego, CA: Academic Press.

Novis, D. A., Walsh, M., Wilkinson, D., St. Louis, M., & Ben-Ezra, J. (2006). Laboratory productivity and the rate of manual peripheral blood smear review: A college of American pathologists Q-probes study of 95141 complete blood count determinations performed in 263 institutions. *Archives of Pathology & Laboratory Medicine*, 130, 596-601.

Paivio, A. (1991). Dual coding theory: retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255-287.

Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.

Patel, V.L. , & Groen G. J. (1991). The general and specific nature of medical expertise: a critical look. In A. Ericsson and J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 93-125). New York: Cambridge University Press.

Patton, M.Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage Publications, Inc.

Perrin, D. G. (1969). A theory of multiple-image communication. *AV Communication Review*, 17 (4), 368-382.

Peeck, J. (1987). The role of illustrations in processing and remembering illustrated text. In H. A. Houghton & D. M. Willows (Eds.), *The psychology of illustration* (pp. 115-151). New York: Springer-Verlag New York, Inc.

Pierre, R. V. (2002). Peripheral blood film review: The demise of the eyecount leukocyte differential. *Clinics in Laboratory Medicine*, 22(1), 279-297.

Recertification. (n.d.). Retrieved October 18, 2006, http://www.nca-info.org/recertification_application.asp

Regehr, G., Cline, D. J., Norman, G. R., & Brooks, L. R. (1994). Effects of processing strategy on diagnostic skill in dermatology. *Academic Medicine*, 69, S34-36.

Riley, R.S., Ben-Ezra, J.M., Massey, D., & Cousar, J. (2002). The virtual blood film. *Clinics in laboratory medicine*, 22(1), 317-345.

Riley, R.S., Ben-Ezra, J.M., Massey, D., Slyter, R., & Romagnoli, G. (2004). Digital photography: A primer for pathologists. *Journal of Clinical Laboratory Analysis*, 18, 91-128.

Sadler, P. M. (2000). The relevance of multiple-choice testing in assessing science understanding. In J. J. Mintzes, J. H. Wandersee,& J. D. Novak (Eds.), *Assessing science understanding: A human constructivist view* (pp. 251-274). San Diego, CA: Academic Press.

Schwab, J. (1973). The practical 3: Translation into curriculum. *School Review*, 81, 501-522.

Solso, R. L. (2003). *The psychology of art and the evolution of the conscious brain.* Cambridge, MA: The MIT Press.

SPOT RT color digital cameras. (n.d.) Retrieved from August 25, 2006, http://www.nuhsbaum.com/SPOTCAMERA.htm

SPSS Inc. (2003). SPSS (Version 14.0) [Computer software]. Chicago: Prentice Hall.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining the qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.

Tatsumi, N., & Pierre, R. V. (2002). Automated image processing: Past, present, and future of blood cell morphology identification. *Clinics in Laboratory Medicine*, 22(1), 299-315.

Tufte, E.R. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.

Tufte, E.R. (1997). *Visual explanations*. Cheshire, CT: Graphics Press.

Turgeon, M. L. (2005). *Clinical hematology: Theory and procedures* (4th ed.). Philadelphia: Lippincott, Williams, and Wiklins.

Vooijs, G. P., Davey, D. D., Somrak, T. M., Goodell, R. M., Grohs, D. H., & Knesel, E. A. (1998). Computerized training and proficiency testing. *Acta Cytologica*, 42(1), 141-147.

Wandersee, J. H. (2000). Designing an image-based biology test. In J. J. Mintzes, J. H. Wandersee,& J. D. Novak (Eds.), *Assessing science understanding: A human constructivist view* (pp. 129-143). San Diego, CA: Academic Press.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.).Thousand Oaks, CA: Sage Publications.

Zeitz, C. M. (1997). Some concrete advantages of abstraction: How experts' representations facilitate reasoning. In P. J. Feltovich, K. M. Ford, & R. R. Hoffman (Eds.), *Expertise in Context: Human and Machine* (pp. 43-65). Menlo Park, CA: AAAI Press.

Zull, J. E. (2002). *The art of changing the brain: Enriching the practice of teaching by exploring the biology of learning*. Sterling, VA: Stylus Publishing.

APPENDIX A

IRB EXEMPTION

IRB #: E3456 LSU Proposal #:_____ **Revised:** 10/04/2006

LSU INSTITUTIONAL REVIEW BOARD (IRB) for HUMAN RESEARCH SUBJECT PROTECTION

578-8692 FAX 6792
Office:203 B-1 David Boyd Hall

APPLICATION FOR EXEMPTION FROM INSTITUTIONAL OVERSIGHT

Unless they are qualified as meeting the specific criteria for exemption from Institutional Review Board (IRB) oversight, ALL LSU research/projects using living humans as subjects, or samples or data obtained from humans, directly or indirectly, with or without their consent, must be approved or exempted in advance by the LSU IRB. This Form helps the PI determine if a project may be exempted, and is used to request an exemption.

Instructions: Complete this form.

Exemption Applicant: If it appears that your study qualifies for exemption send:

(A) Two copies of this completed form,
(B) a brief project description (adequate to evaluate risks to subjects and to explain your responses to Parts A & B),
(C) copies of all instruments to be used. If this proposal is part of a grant proposal include a copy of the proposal and all recruitment material.
(D) the consent form that you will use in the study. A Waiver of Written Informed Consent is attached and must be completed only if you do not intend to have a signed consent form.
(E) Certificate of Completion of Human Subjects Protection Training for all personnel involved in the project (including students who are involved with testing or handling data) at http://cme.cancer.gov/clinicaltrials/learning/humanparticipant-protections.asp. (Unless already on file with the IRB.)

Study exempted by
Louisiana State University
Institutional Review Board
203 B-1 David Boyd Hall
225-578-8692
[illegible] Chair

to: ONE screening committee member (listed at the end of this form)in the most closely related department/discipline or to IRB office.

If exemption seems likely, submit it. If not, submit regular IRB application. Help is available from Dr. Robert Mathews, 578-8692, irb@lsu.edu or any screening committee member.

Principal Investigator Deborah Elizabeth Fox **Student?** Y Y/N

Ph: (225)-931-8969(cell) or (225)-768-1727 (W) **E-mail** dfox@ololcollege.edu **Dept/Unit** ETPP

If Student, name supervising professor Dr. James Wandersee Ph: 578-2348
Mailing Address 12352 Old Mill Drive, Geismar, La. 70734 Ph (225)-313-3019
Project Title Developing/ testing a new approach for assessing rapid visual identification of hematological cells using principles of visual cognition

Agency expected to fund project N/A
Subject pool (e.g. Psychology Students) Clinical Laboratory Science students attending Our Lady of the Lake College and practicing medical lab technicians and medical technologists.
Circle any **"vulnerable populations" to be used**: (children <18; the mentally impaired, pregnant women, the aged, other). Projects with incarcerated persons cannot be exempted.
I certify my responses are accurate and complete. If the project scope or design is later changed I will resubmit for review. I will obtain written approval from the Authorized Representative of all non-LSU institutions in which the study is conducted.

PI Signature Deborah E. Fox **Date** 10/23/06 (no per signatures)

==========

Screening Committee Action: Exempted X Not Exempted ___ Category/Paragraph _____

Reviewer S. Kim MacGregor Signature S. Kim MacGreg Date 10/25/06

APPENDIX B

PARTICIPANT CONSENT FORM

1. Study Title: Developing/ testing a new approach for assessing rapid visual identification of hematological cells using principles of visual cognition.

2. Performance Site: Our Lady of the Lake College, Baton Rouge, La.

3. Investigator: The investigator in this study, Debbie Fox, is available M-F 9:00 a.m. – 5:00 p.m. at 768-1727.

4. Purpose of the Study: The purpose of this study is to explore/ evaluate the role that assessment techniques/ formats which use digital images for the measurement of competency in the visual identification of white blood cells have on expert and novice performance. The study will also explore and describe the differential cognitive processes that are used by the expert and the novice during the morphological identification of white blood cells.

5. Subject Inclusion: Students enrolled in a CLS degree program at Our Lady of the Lake College who have taken a basic course in hematology. Nationally-certified, state-licensed medical technologists or medical technicians who are currently practicing at Baton Rouge area hospitals.

6. Number of Subjects: 35-40

7. Study Procedures: Computer-administered competency testing on the visual identification of white blood cells will be administered to both students and experts. Computer-administered competency testing is estimated to take approximately 1 1/2 hours per expert and approximately 2 1/2 hours per students. Clinical interviews of selected students and experts that include the use of think aloud protocols will be conducted after completion of competency testing. Interviews will be conducted in 1 to 1 1/2 hour sessions, with no more than two sessions scheduled per selected interviewee. If two sessions are scheduled with a single interviewee, these sessions will be scheduled on separate days of the week and will be scheduled at least two days apart.

8. Benefits: The study may yield interesting and valuable information about assessment in the field of clinical hematology as well as information about the cognitive processing operations/differences between the expert and the novice. Both students and experts will be given full access to all study results which will include both expert and student cell identification statistics. Examination of expert cell identifications by students may serve as a primary learning tool in improving visual identification skills in student participants. Comparison of expert results with each other by experts themselves may serve as a potential source of continuing education as the practicing professionals discuss key and relevant features of controversial cell types.

9. Risks: There are no known risks to participation in this research project.

10. Right to Refuse: Participation in this research study is voluntary. Subjects may choose not to participate or to withdraw from the research study at any time without penalty or loss of any benefit to which they might otherwise be entitled.

11. Privacy: Results of the study may be published, but no names or identifying information will be included in the publication. The subject's identity will remain confidential unless disclosure is required by law.

12. Signatures:

The research study has been discussed with me and all of my questions have been answered. I may direct additional questions regarding study specifics to the investigator. If I have questions about subjects' rights or other concerns, I can contact Robert C. Mathews, Chairman, LSU Institutional Review Board, (225)-578-8692. I agree to participate in the study described above and acknowledge the researchers' obligation to provide me with a copy of this consent form if signed by me.

_____ _____

Subject Signature Date

APPENDIX C

NIH CERTIFICATE

Completion Certificate

This is to certify that

Debbie Fox

has completed the **Human Participants Protection Education for Research Teams** online course, sponsored by the National Institutes of Health (NIH), on 08/17/2006.

This course included the following:

- key historical events and current issues that impact guidelines and legislation on human participant protection in research.
- ethical principles and guidelines that should assist in resolving the ethical issues inherent in the conduct of research with human participants.
- the use of key ethical principles and federal regulations to protect human participants at various stages in the research process.
- a description of guidelines for the protection of special populations in research.
- a definition of informed consent and components necessary for a valid consent.
- a description of the role of the IRB in the research process.
- the roles, responsibilities, and interactions of federal agencies, institutions, and researchers in conducting research with human participants.

APPENDIX D

PILOT STUDY DATA

Table D1

Cell Identifications and Difficulty Ratings from Pilot Study

	Cell Identification			Difficulty		
Image Number	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3
1	N	N	N	E	E	A
2	M	Me	Me	A	A	A
3	Bl	Bl	Bl	E	E	D
4	L	L	L	A	E	A
5	My	N	Me	E	E	A
6	Bl	Bl	Bl	E	E	A
7	L	My	L	E	E	A
8	N	N	N	E	E	E
9	Bl	At	Pr	A	A	A
10	L	My	L	E	A	A
11	N	Me	N	E	A	E
12	N	N	N	E	E	E
13	Bl	At	Pr	E	A	A
14	N	N	N	E	E	E
15	L	L	L	E	E	E

Note. N = neutrophil, Ba = basophil , Me = metamyelocyte, My = myelocyte, Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte, L = normal lymphocyte, Pl = plasma cell, E = easy, A = average, D = difficult

Table D1 Continued

	Cell Identification			Difficulty		
Image Number	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3
16	L	L	L	E	E	E
17	L	Me	L	E	E	E
18	M	Pl	L	D	D	A
19	M	L	L	A	E	A
20	Bl	Bl	Bl	E	A	A
21	M	Bl	My	D	A	D
22	L	Me	L	E	A	E
23	L	L	L	A	E	E
24	Bl	Bl	Bl	E	E	A
25	Pl	At	Pl	D	E	D
26	Bl	Bl	Bl	A	E	A
27	Bl	Bl	Pr	D	E	A
28	L	L	L	A	E	A
29	L	L	L	E	E	E
30	L	L	L	D	D	A

Note. N = neutrophil, Ba = basophil , Me = metamyelocyte, My = myelocyte,

Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte,

L = normal lymphocyte, Pl = plasma cell, E = easy, A = average, D = difficult

Table D1 Continued

	Cell Identification			Difficulty		
Image Number	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3
31	N	N	N	A	E	E
32	Bl	Bl	Bl	E	E	A
33	Bl	Bl	Bl	E	E	D
34	L	L	L	A	E	E
35	N	N	N	E	E	E
36	At	L	My	A	E	A
37	Bl	Bl	Bl	A	E	A
38	L	L	L	E	E	D
39	Bl	Bl	Bl	A	E	A
40	Bl	Bl	Bl	D	E	A
41	N	N	N	E	E	A
42	M	M	L	D	D	D
43	Bl	Bl	Bl	E	E	A
44	Bl	Bl	Bl	A	E	A
45	Bl	Bl	Bl	D	D	A

Note. N = neutrophil, Ba = basophil , Me = metamyelocyte, My = myelocyte,

Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte,

L = normal lymphocyte, Pl = plasma cell, E = easy, A = average, D = difficult

Table D1 Continued

	Cell Identification			Difficulty		
Image Number	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3
46	Bl	Bl	Bl	A	E	A
47	Me	N	N	A	E	A
48	M	L	L	A	E	A
49	L	L	L	D	E	E
50	M	L	L	E	E	E
51	L	L	L	A	E	A
52	Ba	Ba	Ba	E	E	E
53	Bl	Bl	Bl	D	E	A
54	L	N	My	D	D	D
55	M	L	L	A	E	E
56	Bl	Bl	Bl	A	E	E
57	M	N	M	A	D	A
58	M	L	L	D	E	E

Note. N = neutrophil, Ba = basophil , Me = metamyelocyte, My = myelocyte,

Pr = promyelocyte, Bl = blast, M = monocyte, At = atypical lymphocyte,

L = normal lymphocyte, Pl = plasma cell, E = easy, A = average, D = difficult

APPENDIX E

EXPERT PARTICIPANT DIRECTIONS

Technology Requirements

1. Home (personal) computer with internet access and a CD ROM drive.
2. A personal e-mail account.

Exam Code

Exam Code	E25

Please begin the exam with Module 1 and perform the other modules in the order indicated below.

Module Order	Module Number
1st	1
2nd	2
3rd	3
4th	4

- Perform each module when you are **well-rested** and have some **uninterrupted** time available
- Module 1 and 2 are estimated to take approximately 10 minutes each
- Module 3 and 4 are estimated to take approximately 20 minutes each
- Please perform each module at a separate sitting
- After completing Modules 1 and 2, wait approximately **1 week** before completing Modules 3 and 4
- Upon completion of all modules, e-mail the file to dfox@ololcollege.edu
- Completed exam file will be located on your C drive; file name = exam code

Target Completion Date =Monday, February 19th

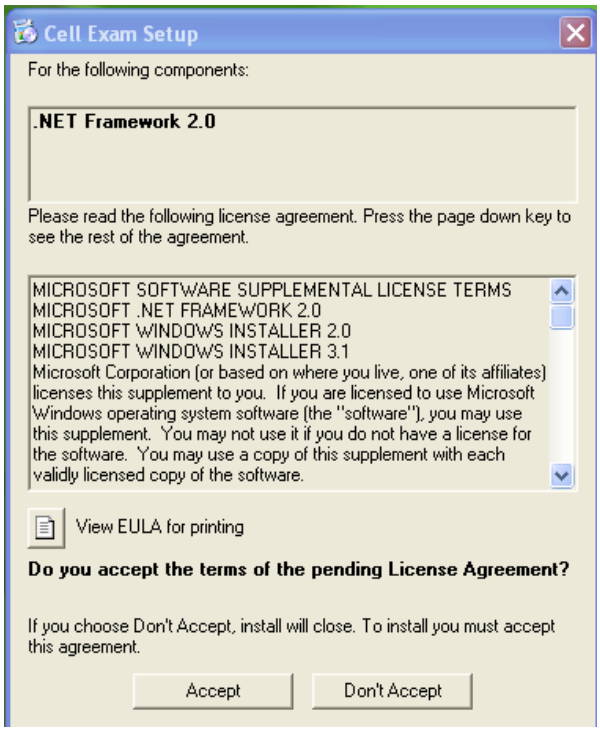
APPENDIX F

DIRECTIONS FOR EXAM SET-UP AND E-MAILING RESULTS

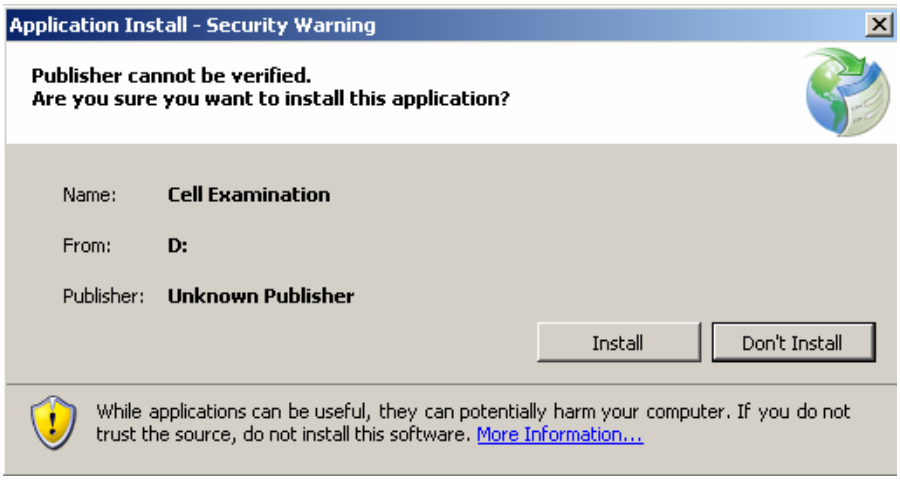
Loading Cell Exam file and beginning exam:

1. Place the CD ROM in the disk drive.

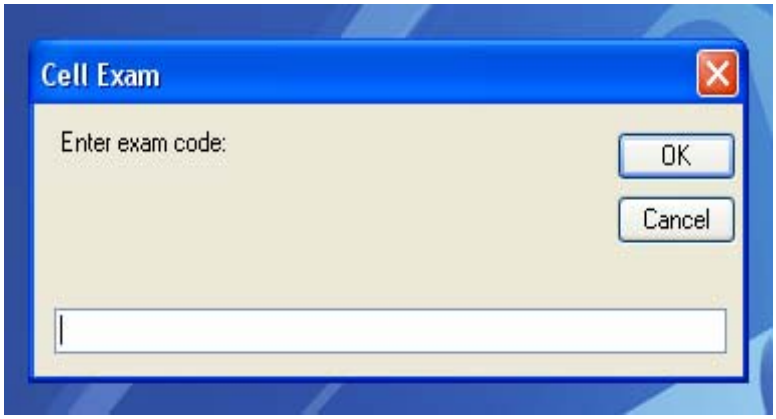
2. If NET Framework is not installed on your computer and you have internet access, your computer will attempt to automatically load the program. When it does, you will have to click "Accept" to accept the licensure agreement.



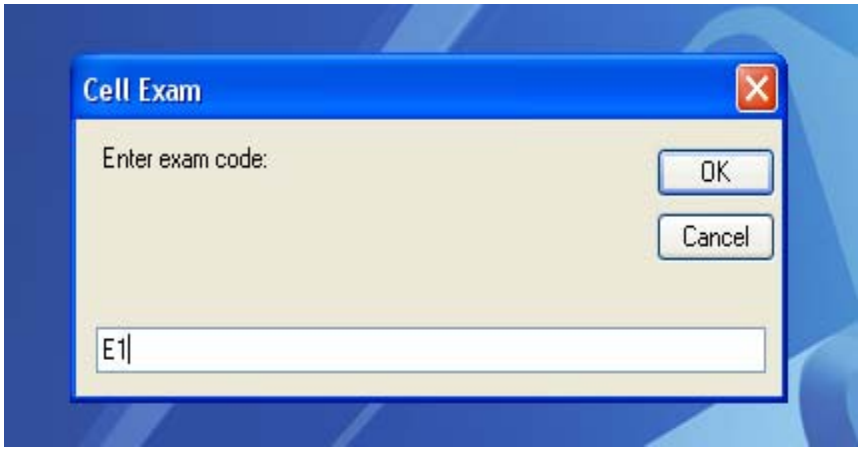
3. The following screen will appear. Click "Install".



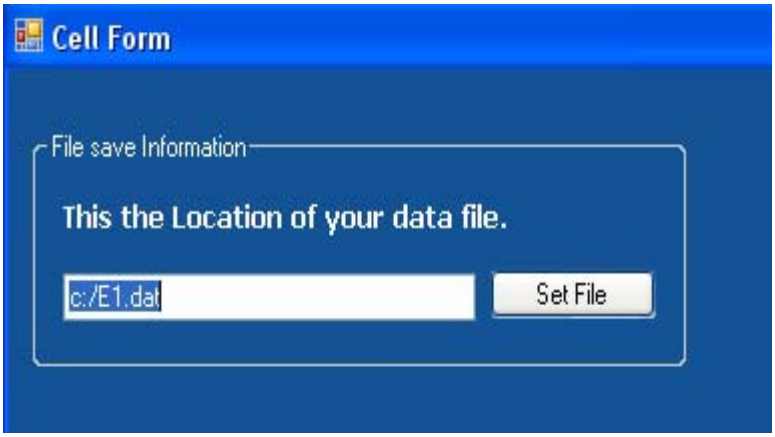
4. Once the cell exam begins, you will be prompted to enter an exam code.



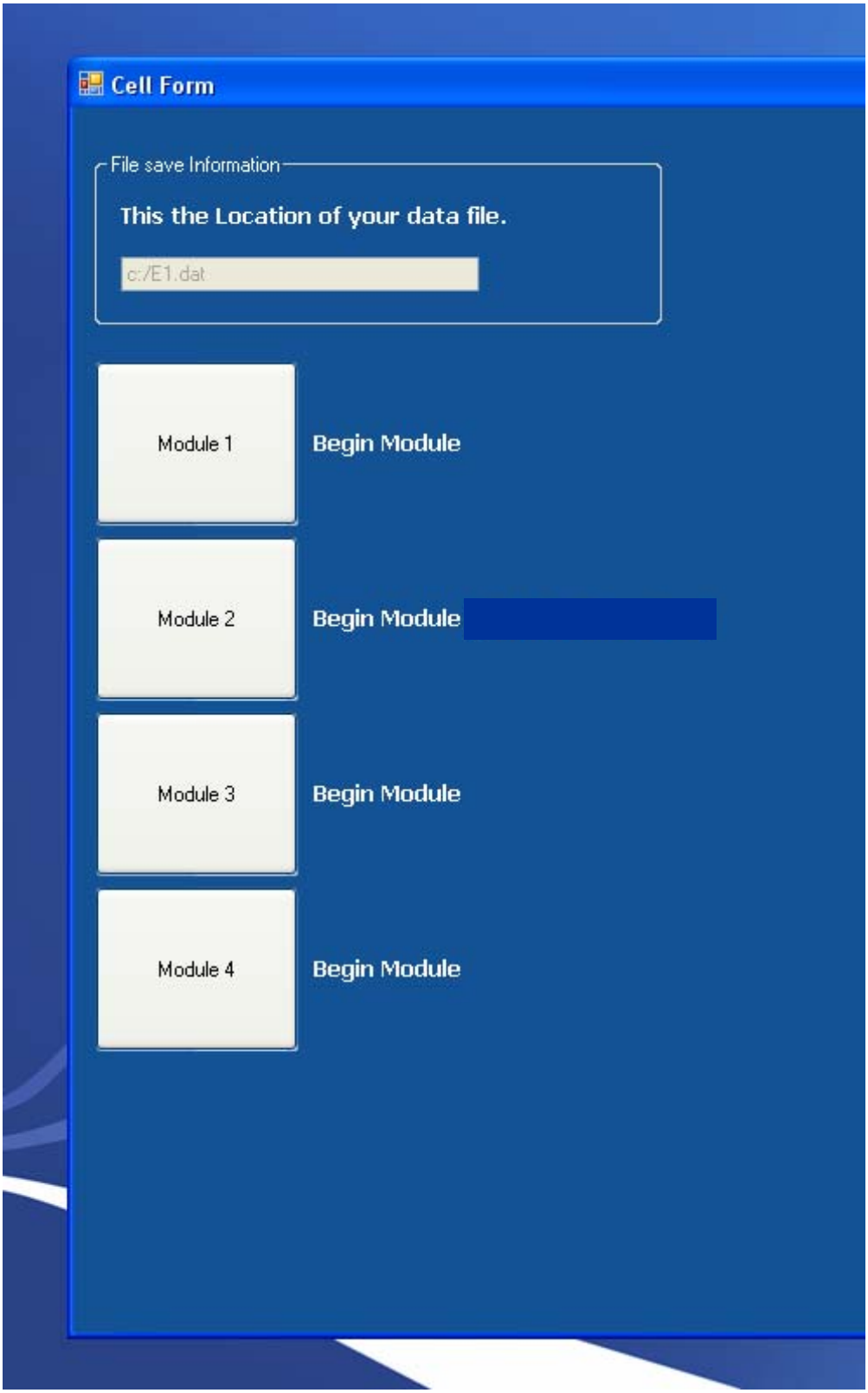
5. Enter the exam code issued to you. (i.e. E1, E2).



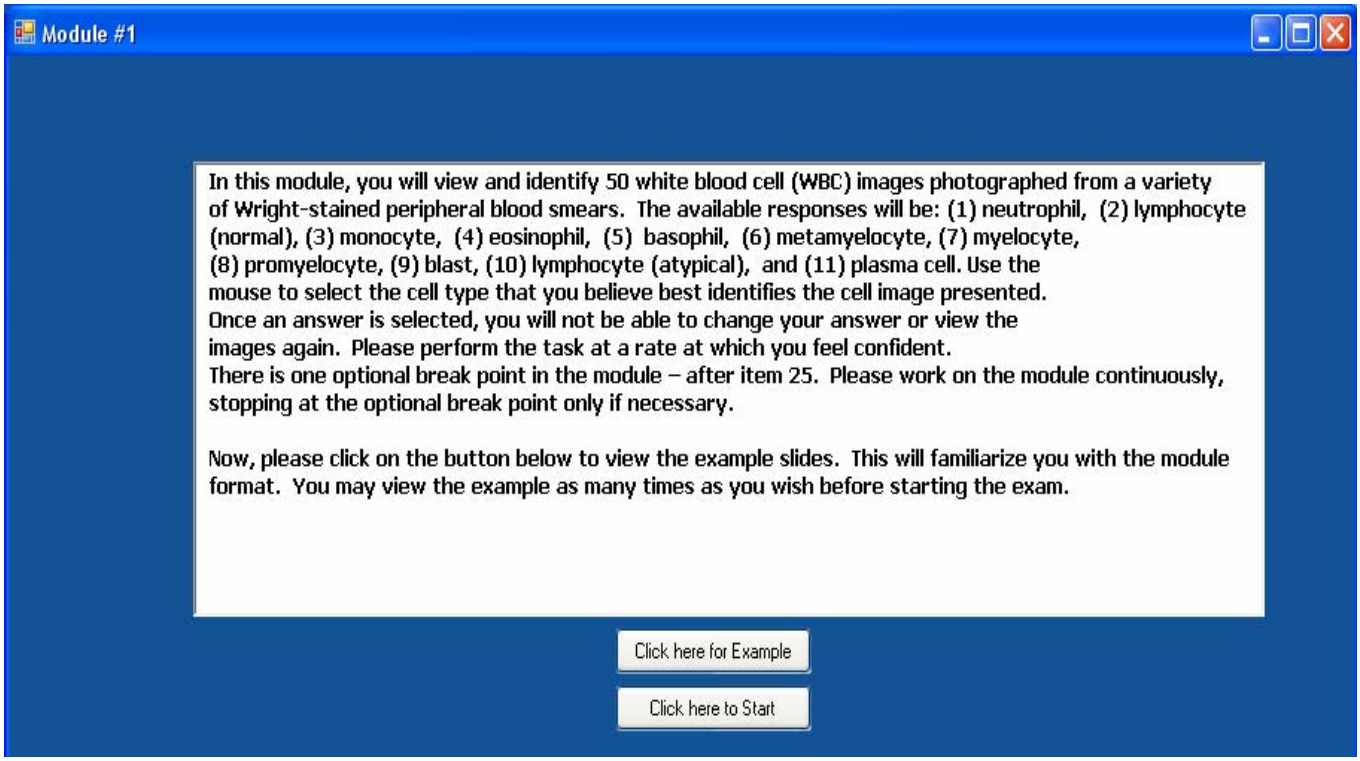
6. Click “OK”.
7. You will now view a screen which displays to location to which the data file will be saved. The data file is set to automatically saved to the C drive of your computer.
8. Click “Set File”.

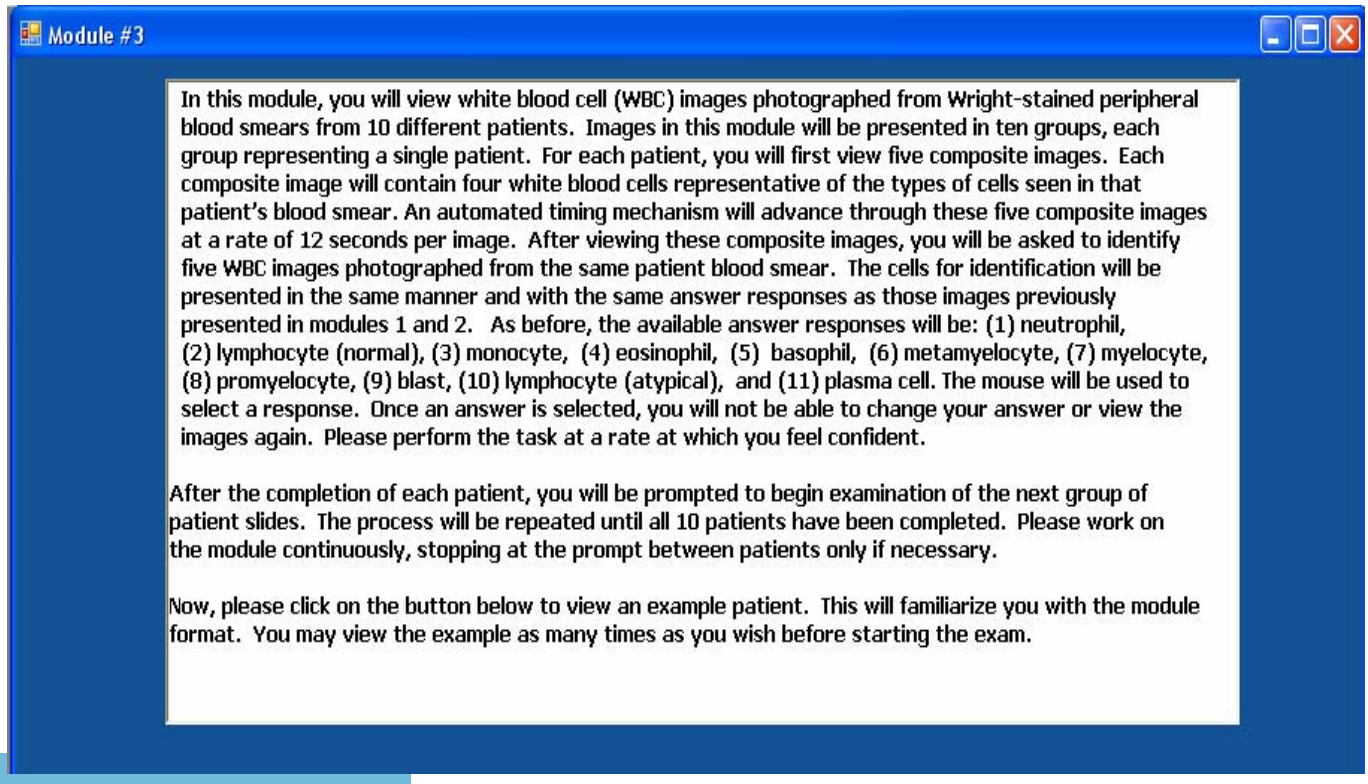


9. Click on Module # which you would like to begin.

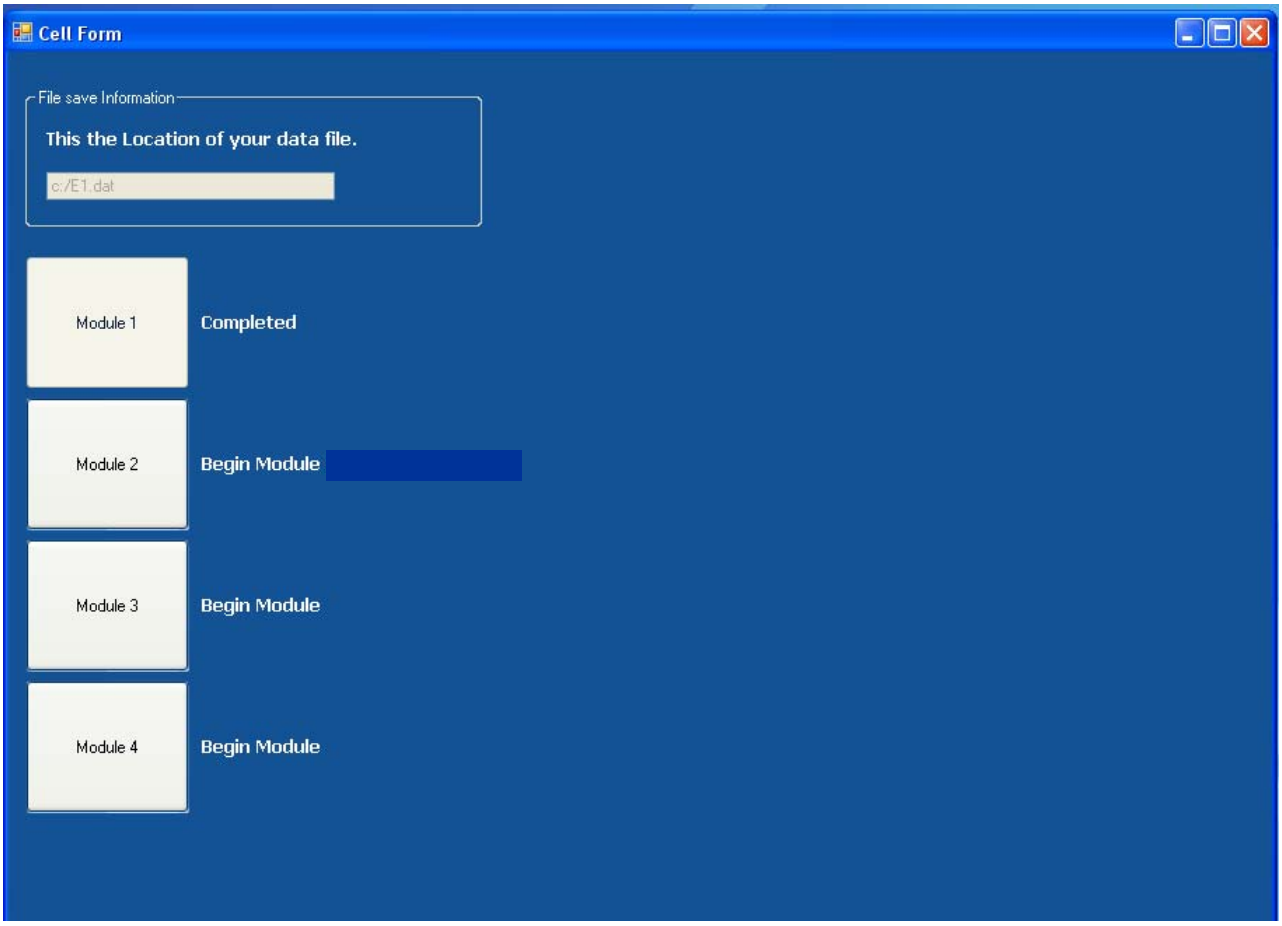


10. Detailed directions and an example will be provided for each Module.





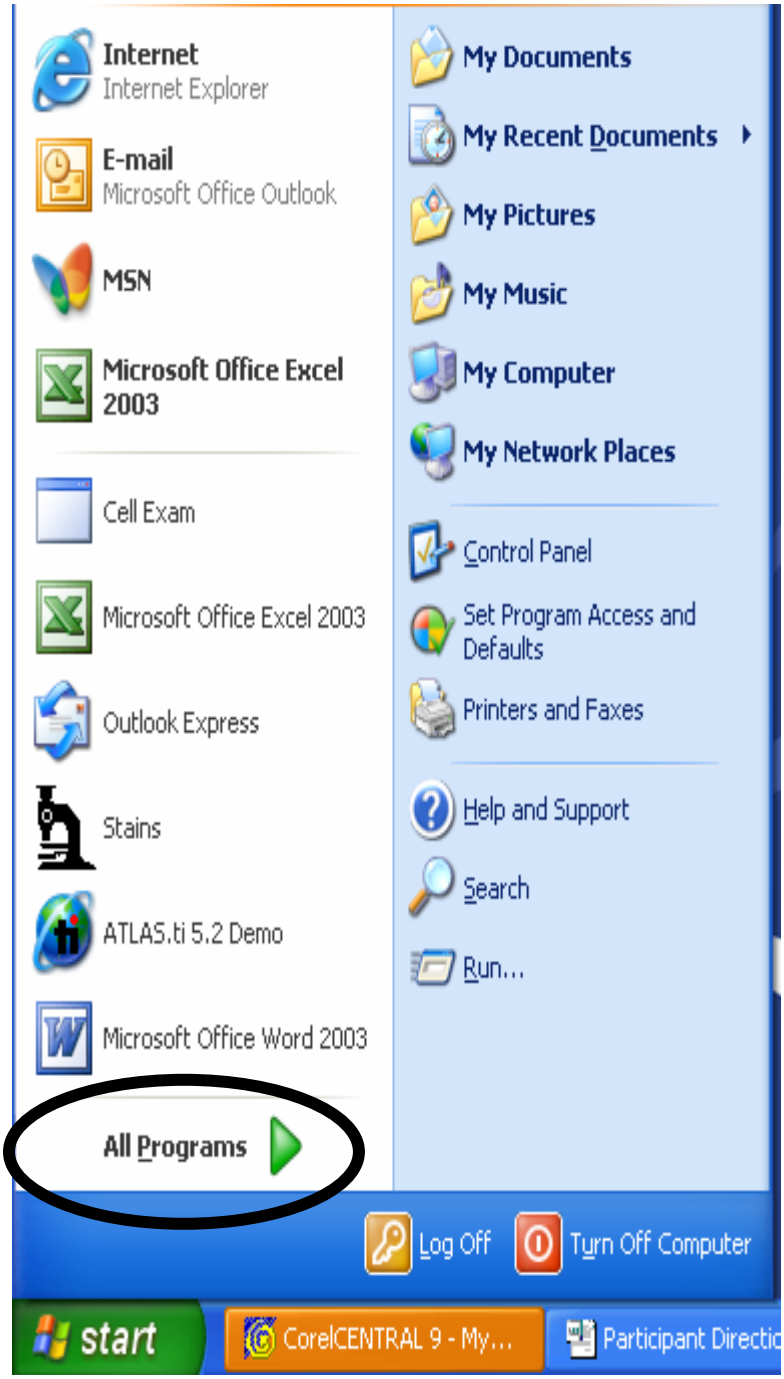
11. After completing a single module, you will be automatically returned to the Main module selection screen.



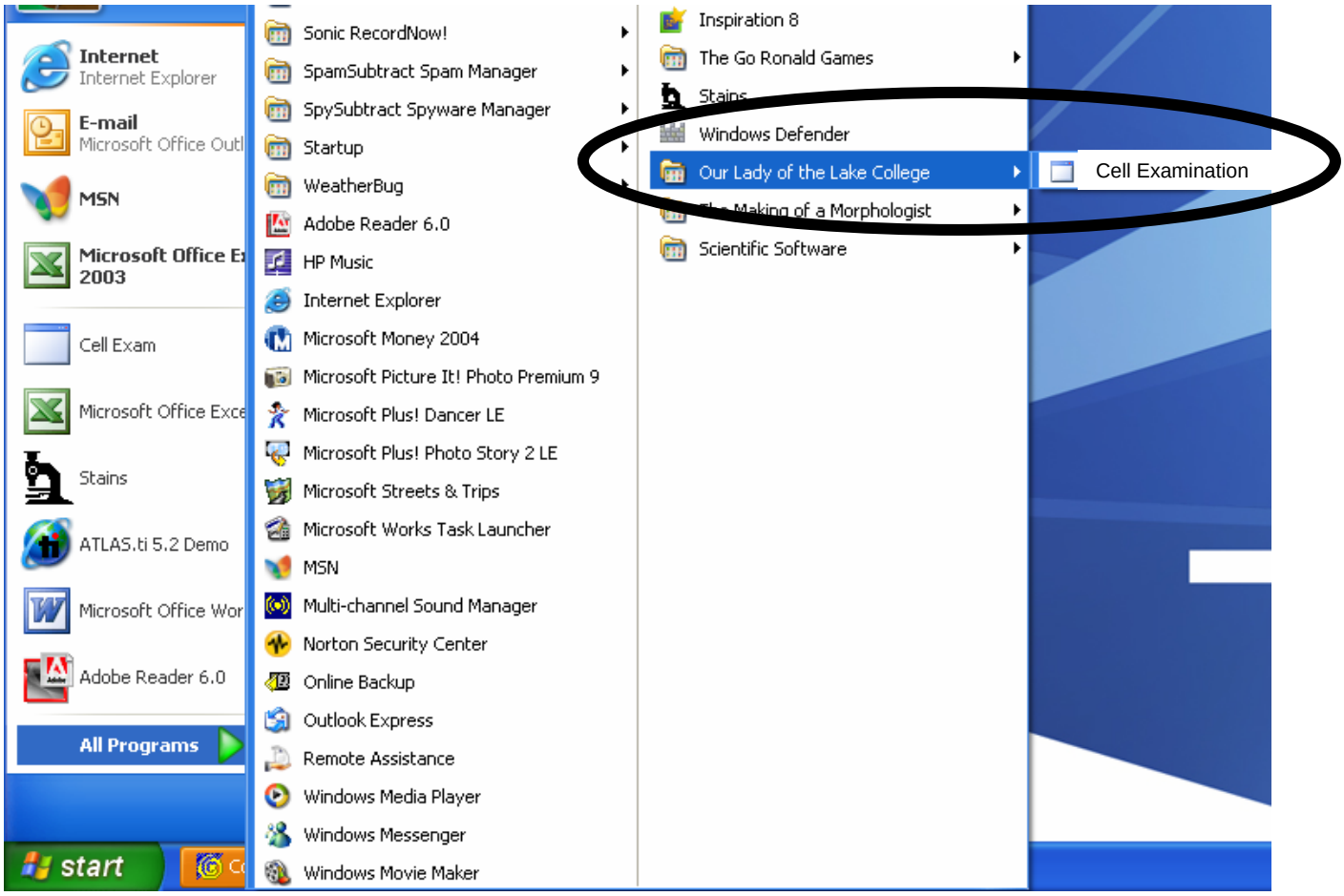
12. Click the red “x” in the upper right corner of the screen to close the exam after finishing the module.

13. When you return to the exam at a later time in order to begin another module, click on the “Start” menu located in the lower left corner of your main computer screen.

14. Place your mouse arrow on “All Programs”.



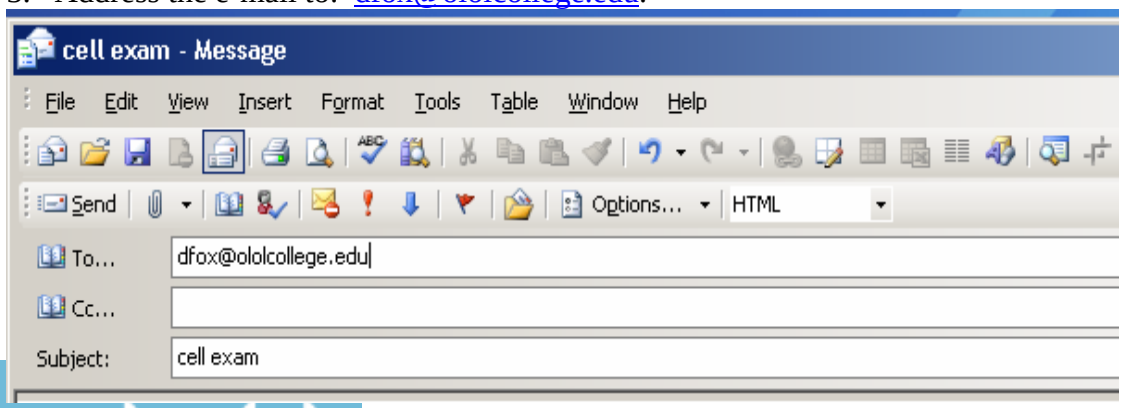
15. Place your mouse arrow on the “Our Lady of the Lake College” folder and click on “Cell Examination”.



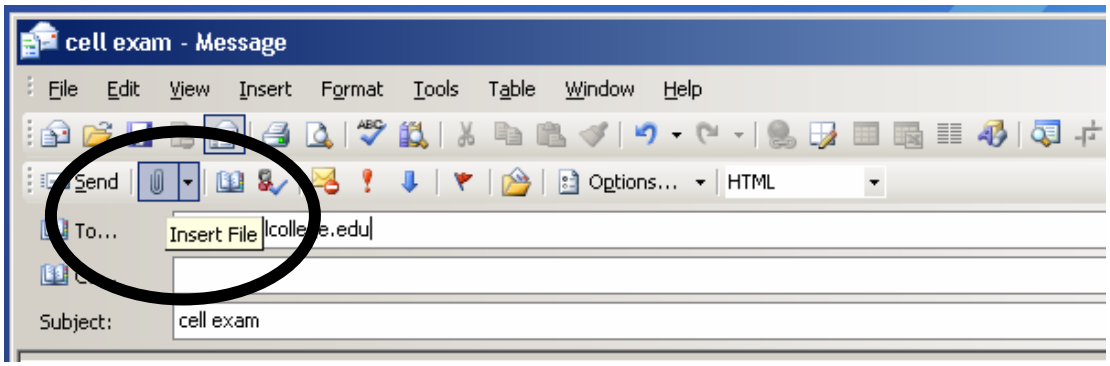
16. You must **type in your assigned exam code each time you re-enter** the exam. (See step 4, above)

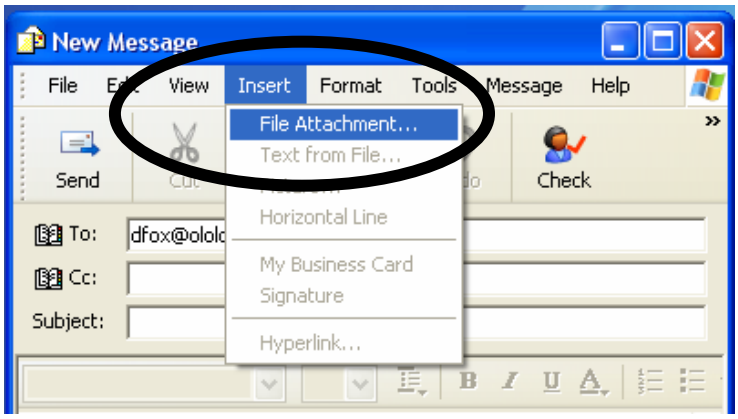
E-mailing exam results:

1. Open your e-mail account.
2. Create a new e-mail.
3. Address the e-mail to: dfox@ololcollege.edu.

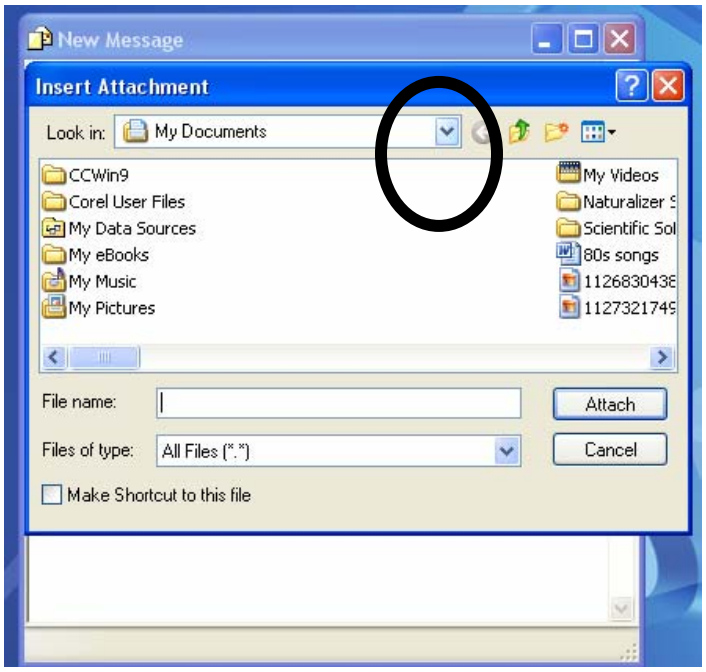


4. Access the menu bar at the top of the screen and click the paper clip icon or “Insert file” button. (Two examples follow, each having a different screen format.)

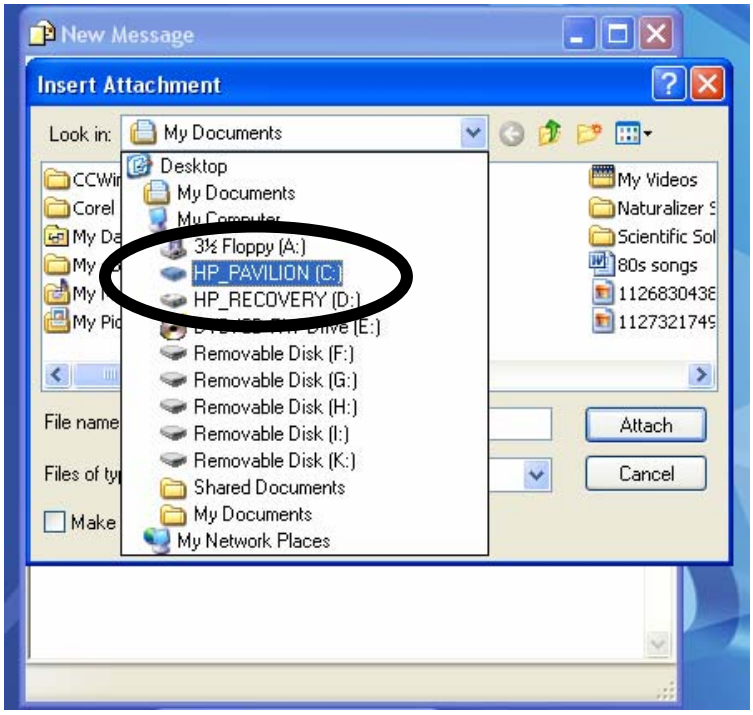




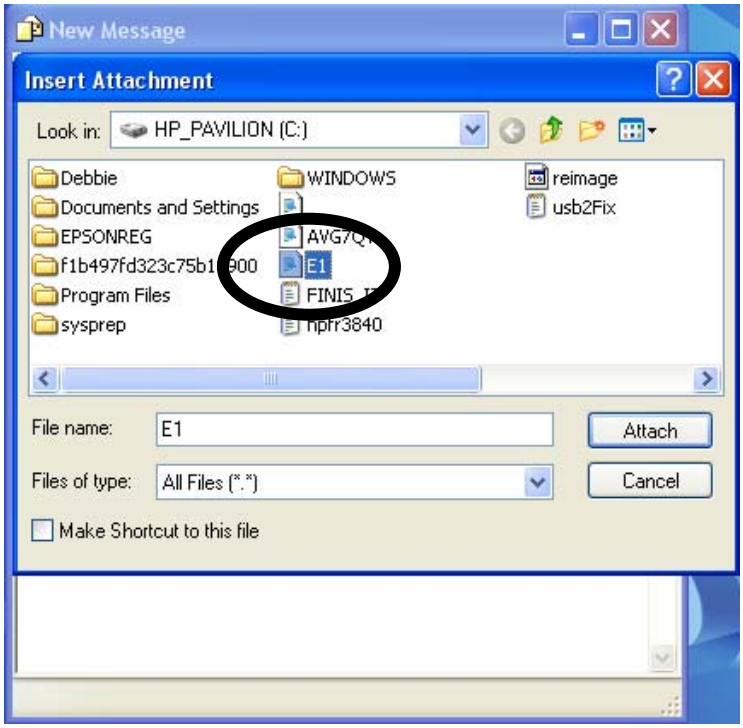
5. Click the down arrow next to the open “Look in:” Browser box.



6. Click on the “C:” drive selection.



7. Double click on the file containing your exam code, (i.e. E1, E2).



8. Click “Send”.

APPENDIX G

RAW COUNTS FOR QUALITATIVE DATA

Table G1

Novice Raw Counts for Subcategories Within Data Examination, Cell Identification by Patient Format

Data Examination Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Cell color	2	0	0	21	0	23
Cell maturity	4	4	0	6	0	14
Cell shape	0	0	1	16	1	18
Cell size	16	0	12	7	3	38
Cytoplasm amount	13	2	2	3	10	30
Cytoplasm color	43	15	19	21	38	136
Cytoplasm shape/ margins	9	2	0	1	1	13
Cytoplasm texture	0	0	0	0	1	1
Granules absent	14	0	1	2	0	17
Granules present	7	3	12	29	6	57
Halo present	0	1	0	0	4	5
N:C ratio	0	15	6	23	0	44
Nuclear color	37	0	0	9	6	52
Nuclear maturity	-	-	-	-	-	-
Nuclear location within cell	13	13	2	26	7	61

Table G1 Continued

Data Examination Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Nuclear shape	21	5	17	28	38	109
Nuclear size	2	0	8	2	4	16
Nuclear texture	31	0	1	0	1	33
Nucleoli absent	0	2	1	0	1	4
Nucleoli present	7	12	20	1	13	53
Physical relationship to RBCs	5	0	0	15	8	28
Reactivity	-	-	-	-	-	-
Vacuoles absent	1	1	0	0	0	2
Vacuoles present	27	8	17	0	15	67
Feature presence Uncertainty	1	1	2	3	0	7
Red blood cells/ Platelets	-	-	-	-	-	-

Table G2

Novice Raw Counts for Subcategories Within Data Analysis, Cell Identification by Patient Format

Data Analysis Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Correlations of features/IDs	4	2	6	14	11	37
Comparison to typical/normal	1	0	0	3	4	8
Comparison of size to RBCs	1	2	4	1	0	8
Comparison to other WBCs	8	1	0	19	5	33
Further analysis desired	-	-	-	-	-	-
Diagnosis inferred	-	-	-	-	-	-

Table G3

Novice Raw Counts for Subcategories Within Data Interpretation, Cell Identification by Patient Format

Data Interpretation Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Hypothesis	9	0	1	9	11	30
Specific cell ID	101	58	126	121	113	519
ID by lineage	3	6	0	0	0	9
Non-ID	0	0	1	2	3	6
Lineage reference	0	3	0	0	0	3
Transitional ID	0	0	0	0	1	1
Variant ID	10	0	0	0	0	10
Unnamed ID	6	0	0	4	0	10

Table G4

Novice Raw Counts for Subcategories Within Meta-Reasoning, Cell Identification by Patient Format

Meta-Reasoning Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Uncertainty general	1	0	1	1	2	5
Uncertainty of ID	3	0	2	2	3	10
Review needed (pathologist/technologist)	-	-	-	-	-	-
Difficulty evaluation	1	0	2	0	0	3

Table G5

Expert Raw Counts for Subcategories Within Data Examination, Cell Identification by Patient Format

Data Examination Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Cell color	1	0	0	7	1	9
Cell maturity	9	9	0	8	5	31
Cell shape	-	-	-	-	-	-
Cell size	1	0	2	14	0	17
Cytoplasm amount	0	0	2	3	0	5
Cytoplasm color	0	4	3	3	1	11
Cytoplasm shape/ margins	3	0	0	3	2	8
Cytoplasm texture	1	0	2	2	0	5
Granules absent	0	0	6	0	0	6
Granules present	10	4	5	3	4	26
Halo present	0	0	0	1	0	1
N:C ratio	0	0	1	0	0	1

Table G5 Continued

Data Examination Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Nuclear color	-	-	-	-	-	-
Nuclear maturity	6	7	0	1	0	14
Nuclear location	0	0	1	0	1	2
Nuclear shape	3	2	7	3	7	22
Nuclear size	1	0	0	0	0	1
Nuclear texture	0	3	6	2	2	13
Nucleoli absent	-	-	-	-	-	-
Nucleoli present	3	0	7	2	1	13
Physical relationship to RBCs	0	0	2	0	0	2
Reactivity	1	3	0	1	0	5
Vacuoles absent	-	-	-	-	-	-
Vacuoles present	2	5	2	7	2	18
Feature uncertainty	1	0	1	0	0	2
Red blood cells / Platelets	12	0	0	0	0	12

Table G6

Expert Raw Counts for Subcategories Within Data Analysis, Cell Identification by Patient Format

Data Analysis Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Correlations of features/IDs	10	6	13	2	4	35
Comparison to typical/normal	8	9	2	9	10	38
Comparison of size to RBCs	0	0	0	0	0	0
Comparison to other WBCs	3	1	7	10	7	28
Further analysis desired	1	0	0	1	0	2
Diagnosis inferred	4	10	0	2	1	17

Table G7

Expert Raw Counts for Subcategories Within Data Interpretation, Cell Identification by Patient Format

Data Interpretation Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Hypothesis	0	2	1	0	3	6
Specific cell ID	104	42	88	66	93	398
ID by lineage	4	6	1	1	2	14
Non-ID	4	1	3	0	1	9
Lineage reference	8	4	2	4	1	19
Transitional ID	4	0	0	5	4	13
Variant ID	15	7	2	12	4	40
Unnamed ID	1	1	0	0	0	2

Table G8

Expert Raw Counts for Subcategories Within Meta-Reasoning, Cell Identification by Patient Format

Meta-Reasoning Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Uncertainty of ID	1	0	0	1	0	2
Review needed (pathologist/technologist)	0	2	0	1	2	5
Difficulty evaluation	3	4	3	0	6	16

Table G9

Novice Raw Counts for Subcategories Within Data Examination, Single Cell Identification Format

Data Examination Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Cell color	0	5	0	0	1	6
Cell maturity	1	4	0	0	0	5
Cell shape	-	-	-	-	-	-
Cell size	7	6	1	9	1	24
Cytoplasm amount	0	0	1	0	0	1
Cytoplasm color	24	1	12	24	17	78
Cytoplasm shape/ margins	2	0	7	0	0	9
Cytoplasm texture	-	-	-	-	-	-
Granules absent	0	0	1	5	0	6
Granules present	14	8	8	13	11	54
Halo present	1	0	3	3	3	10
N:C ratio	0	5	16	11	1	33

Table G9 Continued

Data Examination Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Nuclear color	5	1	0	4	6	16
Nuclear maturity	-	-	-	-	-	-
Nuclear location within cell	7	7	6	5	3	28
Nuclear shape	10	4	8	14	16	52
Nuclear size	0	1	2	14	4	21
Nuclear texture	19	1	0	1	1	22
Nucleoli absent	0	0	1	2	0	3
Nucleoli present	4	1	4	13	6	28
Physical relationship to RBCs	7	10	0	0	3	20
Reactivity	0	0	0	1	0	1
Vacuoles absent	-	-	-	-	-	-
Vacuoles present	11	1	4	8	8	32
Feature presence uncertainty	2	0	0	2	1	5
Red blood cells / Platelets	-	-	-	-	-	-

Table G10

Novice Raw Counts for Subcategories Within Data Analysis, Single Cell Identification Format

Data Analysis Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Correlations of features/IDs	0	4	4	0	0	8
Comparison to typical/normal	0	0	3	0	1	4
Comparison of size to RBCs	2	2	3	2	0	9
Comparison to other WBCs	0	0	2	0	2	4
Further analysis desired	-	-	-	-	-	-
Diagnosis inferred	-	-	-	-	-	-

Table G11

Novice Raw Counts for Subcategories Within Data Interpretation, Single Cell Identification Format

Data Interpretation Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Hypothesis	2	0	1	0	8	11
Specific cell ID	38	40	40	40	40	198
ID by lineage	1	0	0	0	0	1
Non-ID	0	0	0	1	1	2
Lineage reference	0	0	1	0	0	1
Transitional ID	-	-	-	-	-	-
Variant ID	1	0	0	0	0	1
Unnamed ID	-	-	-	-	-	-

Table G12

Novice Raw Counts for Subcategories Within Meta-Reasoning, Single Cell Identification Format

Meta-Reasoning Category	Novice 1	Novice 2	Novice 3	Novice 4	Novice 5	Novice Total
Uncertainty general	-	-	-	-	-	-
Uncertainty of ID	1	1	0	0	6	8
Review needed (pathologist/technologist)	-	-	-	-	-	-
Difficulty evaluation	0	0	0	0	1	1

Table G13

Expert Raw Counts for Subcategories Within Data Examination, Single Cell Identification Format

Data Examination Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Cell color	0	0	0	0	5	5
Cell maturity	0	0	3	2	2	7
Cell shape	0	1	0	1	0	2
Cell size	3	1	3	1	7	15
Cytoplasm amount	1	0	0	0	3	4
Cytoplasm color	25	6	4	0	1	36
Cytoplasm shape	0	2	7	2	0	11
Cytoplasm texture	5	0	1	0	1	7
Granules absent	11	0	1	0	1	13
Granules present	21	9	11	19	5	65
Halo present	2	2	1	2	2	9
N:C ratio	-	-	-	-	-	-

Table G13 Continued

Data Examination Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Nuclear color	2	0	0	0	0	2
Nuclear maturity	1	7	0	9	0	17
Nuclear location	3	0	0	1	0	4
Nuclear shape	16	3	3	10	5	37
Nuclear size	1	0	0	4	1	6
Nuclear texture	23	2	3	6	4	38
Nucleoli absent	2	0	0	0	0	2
Nucleoli present	3	1	0	4	2	10
Physical relationship to RBCs	1	0	0	7	6	14
Reactivity	5	2	1	9	2	19
Vacuoles absent	-	-	-	-	-	-
Vacuoles present	10	10	0	3	1	24
Feature uncertainty	3	1	0	1	0	5
Red blood cells / Platelets	-	-	-	-	-	-

Table G14

Expert Raw Counts for Subcategories Within Data Analysis, Single Cell Identification Format

Data Analysis Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Correlations of features/IDs	11	4	9	9	6	39
Comparison to typical/normal	1	2	5	6	13	27
Comparison of size to RBCs	-	-	-	-	-	-
Comparison to other WBCs	1	0	0	0	3	4
Further analysis desired	0	1	4	2	5	12
Diagnosis inferred	0	0	2	1	1	4

Table G15

Expert Raw Counts for Subcategories Within Data Interpretation, Single Cell Identification Format

Data Interpretation Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Hypothesis	1	3	1	3	3	11
Specific cell ID	38	38	37	33	37	183
ID by lineage	-	-	-	-	-	-
Non-ID	1	1	3	1	0	6
Lineage reference	3	10	1	8	3	25
Transitional ID	0	2	3	3	0	8
Variant ID	1	0	0	4	2	7
Unnamed ID	1	0	0	0	1	2

Table G16

Expert Raw Counts for Subcategories Within Meta-Reasoning, Single Cell Identification Format

Meta-Reasoning Category	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert Total
Uncertainty of ID	-	-	-	-	-	-
Review needed (pathologist/technologist)	-	-	-	-	-	-
Difficulty evaluation	1	0	3	1	1	6

VITA

Deborah Elizabeth Fox was born in Midland, Michigan. She has been a resident of Louisiana since 1981. She graduated from Louisiana State University with a Bachelor of Science degree in microbiology in 1994. After completing a one-year medical technology program at Our Lady of the Lake Regional Medical Center in Baton Rouge, Louisiana, she was certified by the American Society of Clinical Pathologists (ASCP) as a medical technologist in 1995. She worked as a medical technologist at Our Lady of the Lake Regional Medical Center in the area of microbiology for three and one-half years. In August 2001, she earned a Master of Arts in curriculum and instruction from LSU. She worked as a laboratory coordinator/adjunct instructor for the Clinical Laboratory Sciences program and Division of Arts and Sciences at Our Lady of the Lake College (OLOLC) for three and one-half years. She has served as full-time faculty in the OLOLC Clinical Laboratory Sciences program for five years, teaching coursework in hematology, clinical chemistry, phlebotomy, urinalysis/body fluids, and education/research.

Deborah's interest in the field of science education blossomed during her senior year of college when she served as a teaching assistant in general microbiology laboratories taught to freshman and sophomore undergraduates. Her interest has always been in the hands-on laboratory aspects of the science field and has become specialized over the years to encompass a health science focus.

Deborah is married to Brian Fox and has one son, age 3, named Logan. She is a 15 year member of the American Society for Clinical Laboratory Science (ASCLS). She has served as state membership chair for the Louisiana society of this organization for the past two years. In 2006, she presented a seminar entitled "Meaningful Mapping" at the annual Clinical Laboratory Educator's Conference in San Antonio, Texas. She has also been a 10 year member of the American Society for Clinical Pathologists.