

3-28-2013

The Impact of a Classroom Performance System on Learning Gains in a Biology Course for Science Majors

Nilo E. Marin

Florida International University, nmari002@fiu.edu

Follow this and additional works at: <http://digitalcommons.fiu.edu/etd>



Part of the [Higher Education Administration Commons](#)

Recommended Citation

Marin, Nilo E., "The Impact of a Classroom Performance System on Learning Gains in a Biology Course for Science Majors" (2013).
FIU Electronic Theses and Dissertations. Paper 869.
<http://digitalcommons.fiu.edu/etd/869>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

THE IMPACT OF A CLASSROOM PERFORMANCE SYSTEM ON LEARNING
GAINS IN A BIOLOGY COURSE FOR SCIENCE MAJORS

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF EDUCATION

in

HIGHER EDUCATION

by

Nilo Eric Marin

2013

To: Dean Delia C. Garcia
College of Education

This dissertation, written by Nilo Eric Marin, and entitled The Impact of a Classroom Performance System on Learning Gains in a Biology Course for Science Majors, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

George E. O'Brien

Glenda Musoba

Benjamin Baez

Roger G. Gonzalez, Major Professor

Date of Defense: March 28, 2013

The dissertation of Nilo Eric Marin is approved.

Dean Delia C. Garcia
College of Education

Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2013

© Copyright 2013 by Nilo Eric Marin

All rights reserved.

DEDICATION

This dissertation is dedicated to the memory of my father, Leonilo A. Marin, for his unshakable faith in the goodness of people, his indomitable “can-do” spirit, and his enduring legacy of confidence. Herein lies the essence of the book you once predicted I should write.

ACKNOWLEDGMENTS

The completion of every creative piece of work requires the energy, expertise, and guiding hands of many. To all who guided, coached, prodded, and encouraged me to the completion of this study, I am deeply grateful. I extend a special expression of gratitude to my committee chair, Dr. Roger Geértz Gonzalez, and my committee members, Dr. Benjamin Baez, Dr. Glenda Musoba, and Dr. George O'Brien, for their guidance, expert advice, and encouragement. To my friend and colleague, Dr. Steven Obenauf, much appreciation for making it possible to conduct the study, and to my colleagues, Dr. Emmanuel Vrotsos, Professor Gwen Hauer, and Professor Gertudes Velasquez for their patience, cooperation, assistance, and otherwise invaluable role in completing the study. To Dr. Greg Reeder, a heartfelt thanks for providing the training and the clickers for the study. Lastly, to my beloved wife and children, I express my gratitude for their love, support, encouragement, and patience throughout the process.

Collectively and individually, thank you for believing and making it possible!

ABSTRACT OF THE DISSERTATION
THE IMPACT OF A CLASSROOM PERFORMANCE SYSTEM ON LEARNING
GAINS IN A BIOLOGY COURSE FOR SCIENCE MAJORS

by

Nilo Eric Marin

Florida International University, 2013

Miami, Florida

Professor Roger Geértz Gonzalez, Major Professor

This study was conducted to determine if the use of the technology known as Classroom Performance System (CPS), specifically referred to as “Clickers”, improves the learning gains of students enrolled in a biology course for science majors. CPS is one of a group of developing technologies adapted for providing feedback in the classroom using a learner-centered approach. It supports and facilitates discussion among students and between them and teachers, and provides for participation by passive students. Advocates, influenced by constructivist theories, claim increased academic achievement. In science teaching, the results have been mixed, but there is some evidence of improvements in conceptual understanding.

The study employed a pretest-posttest, non-equivalent groups experimental design. The sample consisted of 226 participants in six sections of a college biology course at a large community college in South Florida with two instructors trained in the use of clickers. Each instructor randomly selected their sections into CPS (treatment) and non-CPS (control) groups. All participants filled out a survey that included demographic data at the beginning of the semester. The treatment group used clicker questions

throughout, with discussions as necessary, whereas the control groups answered the same questions as quizzes, similarly engaging in discussion where necessary. The learning gains were assessed on a pre/post-test basis.

The average learning gains, defined as the actual gain divided by the possible gain, were slightly better in the treatment group than in the control group, but the difference was statistically nonsignificant. An Analysis of Covariance (ANCOVA) statistic with pretest scores as the covariate was conducted to test for significant differences between the treatment and control groups on the posttest. A second ANCOVA was used to determine the significance of differences between the treatment and control groups on the posttest scores, after controlling for sex, GPA, academic status, experience with clickers, and instructional style.

The results indicated a small increase in learning gains but these were not statistically significant. The data did not support an increase in learning based on the use of the CPS technology. This study adds to the body of research that questions whether CPS technology merits classroom adaptation.

TABLE OF CONTENTS

| CHAPTER | | PAGE |
|---------|---|------|
| I | INTRODUCTION | 1 |
| | Need for the Study | 2 |
| | Statement of the Problem..... | 4 |
| | Research Questions..... | 4 |
| | Hypothesis..... | 5 |
| | Audience Response Systems..... | 5 |
| | Theoretical Basis of Audience Response Systems | 7 |
| | Definition of Terms..... | 9 |
| | Delimitations of the Study | 10 |
| | Assumptions..... | 10 |
| | Summary | 11 |
| II | LITERATURE REVIEW | 12 |
| | Audience Response Systems..... | 13 |
| | History of audience response systems | 15 |
| | Classroom performance system | 19 |
| | Technology and audience response systems (ARS) in the 2-year college classroom..... | 21 |
| | Focus of ARS technology | 30 |
| | The Challenges of Integrating ARS Technology into the 2-Year College Classroom | 33 |
| | Cost assessment | 33 |
| | Pedagogical assessment | 34 |
| | Intellectual property rights..... | 34 |
| | Incentives | 35 |
| | Theoretical Foundations..... | 35 |
| | Constructivism | 36 |
| | John Dewey..... | 37 |
| | Jean Piaget | 37 |
| | Lev Vygotsky..... | 38 |
| | ARS Outcomes Assessment..... | 39 |
| III | METHODOLOGY | 57 |
| | Experimental Design..... | 57 |
| | Data Collection | 58 |
| | Subjects | 58 |
| | Procedure | 59 |
| | Instrumentation | 60 |
| | Quantitative Data Analysis | 62 |

| | | |
|----|---|-----|
| IV | RESULTS | 65 |
| | Description of the Sample..... | 65 |
| | Research Question | 66 |
| | Learning Gains..... | 66 |
| | Independent-samples <i>t</i> -test..... | 67 |
| | Analysis of Covariance (ANCOVA) | 67 |
| | Ancillary Analyses..... | 72 |
| | Summary | 74 |
| V | DISCUSSION | 76 |
| | Discussion of Findings..... | 76 |
| | Strengths of the Study | 78 |
| | Fidelity of treatment..... | 79 |
| | Lack of differential attrition..... | 79 |
| | Lack of interaction of selection and maturation | 79 |
| | Power analysis | 79 |
| | Limitations of the Study..... | 80 |
| | Selection bias | 80 |
| | Treatment strength | 80 |
| | Implications of the Study | 81 |
| | Directions for Future Research | 83 |
| | Summary | 84 |
| | REFERENCES | 86 |
| | APPENDICES | 94 |
| | VITA..... | 104 |

LIST OF TABLES

| TABLE | PAGE |
|---|------|
| 1. Distribution of Study Participants by Sex..... | 66 |
| 2. Normalized Gains | 67 |
| 3. Summary of Analysis of Covariance with Pretest as the Covariate | 69 |
| 4. Mean Pretest Scores and Unadjusted and Adjusted Mean Posttest Scores | 70 |
| 5. Summary of Analysis of Covariance with Other Variables as Covariates | 71 |
| 6. Summary of Analysis of Covariance with Ethnicity as the Covariate..... | 72 |
| 7. Summary of Analysis of Covariance with Instructional Delivery Mode as the Covariate..... | 74 |

CHAPTER I

INTRODUCTION

The community college is among one of the major paths to educational achievement in the population. It provides a diverse number of services that benefit the first time in college student; the student returning for a career change, or after an extended absence; the worker whose skill needs to be upgraded; the practitioner that needs continuing education credits; the university student who dual enrolls to save money or returns to start again. The demand for these services is increasing and community support is strong. The 2-year/community college is a major provider of students to 4-year institutions, a leading provider of workforce for the community, and large business entities are partnering with community colleges, providing funds and expertise, to capitalize on this service (American Association of Community Colleges, 2009).

Over the past ten years, alarming trends in community college student attrition and achievement are blamed increasingly on the incongruence between the goal of student learning and the way higher education institutions function. The concern over sub-standard achievement and student achievement coupled with the complex educational environment the community colleges operate under has led to a shift in the learning paradigm, from faculty-centered to a more learner-centered approach (Barr & Tagg, 1995). This focus on learning includes maximizing available technology. Faculty are encouraged to move away from PowerPoint presentations, abandon overhead and slide projectors, and use more encompassing technology that will increase learning in more profound ways (VanWagoner et al., 2005). This framework and the various models and programs that support it address student diversity in learning and provide

encouraging results in academic achievement (Cushman, 1995; Daniels, 1993; Flowers, et al., 2000; Hanson, 2006; Howard, 2002; Outcalt, 2003; Rendón, 1994; VanWagoner, et al, 2005).

Need for the Study

The purpose of this study was to assess the impact of an Audience Response System (ARS) on student academic performance in 2-year college biology courses for science majors because testing and research should precede technology implementation in the classroom. VanWagoner, Bowman, and Spraggs (2005) stated that: “New technology is adopted with forethought, testing, and planning” (p. 45). This is of particular interest in the sciences because the introductory biology course for science majors is the primary gatekeeper course that influences student persistence into professional studies such as pharmacy, medicine, nursing, dentistry, and a degree in the sciences, such as biology. Although audience response systems have been used in a wide variety of class sizes, formats, and courses, this has been primarily in the university, elementary, and K-12 settings (Caldwell, 2007; Judson & Sawada, 2002). This study, therefore, added to the knowledge base of scholarly research as it relates to the effectiveness of audience response systems in the science curriculum by analyzing its application and impact to a 2-year college science environment.

The critical shortage of science professionals in the nation has led to an increased interest in the science programs. Many of the students attracted to the programs are typical “at-risk” community college students, although a few may have good science backgrounds from high school (Garcia, 2006). The biology course for science majors has few pre-requisites other than a recommendation to complete a basic-college level

chemistry course. Institutional data gathered informally at Broward College indicate that about 40% of students enrolled in biology for science majors classes, and that are taught by full time instructors, do not persist to the next level, either through poor performance or through attrition. The data analyzed were limited to the full time instructors' grades because their data are continuous, more numerous, and are less subjected to bias from lack of tenure and student evaluation, as is perceived for adjunct instructors. Garcia (2006) experienced a 48% attrition rate in the sample size of a similar study on the effects of Supplemental Instruction.

The use of computers as instructional tools has greatly increased since the 1980s. Research on the effectiveness of instructional technology, however, is a mixed body of work (Caldwell, 2007; Jonassen, 1991; Judson & Sawada, 2002). The development of new technologies outpaces the research such that many institutions rush into their use without adequate testing. Audience response systems are used in a variety of educational environments (Caldwell, 2007). Before these technologies, including the audience response system, undergo widespread adoption, research should be conducted to explore the possible impacts of the technology. The information gathered from this study indicates to policymakers and college administrators that the audience response system is not one of the strategies that can be used to help overcome the numerous barriers to success of the non-traditional/at-risk student attracted to the science programs. It addressed the question of testing and researching before implementing the technology. Additionally, it indicated that the ARS technology is not a technique that could improve how students learn. Moreover, it added valuable information to the knowledgebase of

scholarly research as it relates to the effectiveness of technology in the science curriculum.

Statement of the Problem

This research was primarily focused on assessing the impact of an Audience Response System (ARS), specifically the Classroom Performance System (CPS), on student academic performance in a 2-year college biology course for science majors. Some courses were taught using an audience response system and other courses were taught the traditional way. The results of both conditions were compared and analyzed using appropriate statistical methods.

Research Questions

This study addressed the following questions:

First, will students who participate in a Classroom Performance System (“Clicker” technology) format in a biology course for science majors have higher achievement scores, measured as normalized learning gains, than students in the same course that is not formatted for delivery with Classroom Performance System technology?

Second, are the differences in learning gains between the groups influenced by sex, GPA, academic status, prior experience with the technology, and ethnicity?

These questions have become more relevant today due to the rapid advances in information technology, the subsequent explosion of the Internet and the changes in economics of information (Trumbach, 2006).

Hypothesis

Students using a Classroom Performance System (CPS) in a 2-year college biology course will have higher mean learning gains than students who do not use a Classroom Performance System, after controlling for sex, GPA, ethnicity and student academic status.

Audience Response Systems

Technology has gradually increased in use as an instructional resource in education. The use of technology is considered an instructional enhancement in the classroom and many researchers have shown that there is a relationship between learning and instructional technology (Clark, 1983; Kearsley, 1996): either it is effective for learners or makes no difference for learners (Cuban, 1986). Gauging learning through effective feedback in the classroom has been traditionally done through written feedback, verbal feedback, and peer-group discussion; and although adequate, no one method has proven superior over another (Merrill, 1987; Mason & Bruning, 2001). Technology advancements and computer-based feedback methods, however, have dramatically enhanced the capabilities of formative assessment through questioning and feedback (Roschelle, Penuel, & Abrahamson, 2004). Roschelle et al. (2004) proposed the diagram reproduced in Figure 1 to show “how pedagogy and technology realize multiple desirable outcomes” (p. 1).

One such advancement is the ARS which allows the collection and analysis of large feedback data sets. According to Woods and Chui (2003), ARS allow instructors to instantly assess whether learning is taking place. As a technology that uses wireless transmitting hardware, computers, and database software, ARS has evolved into a system

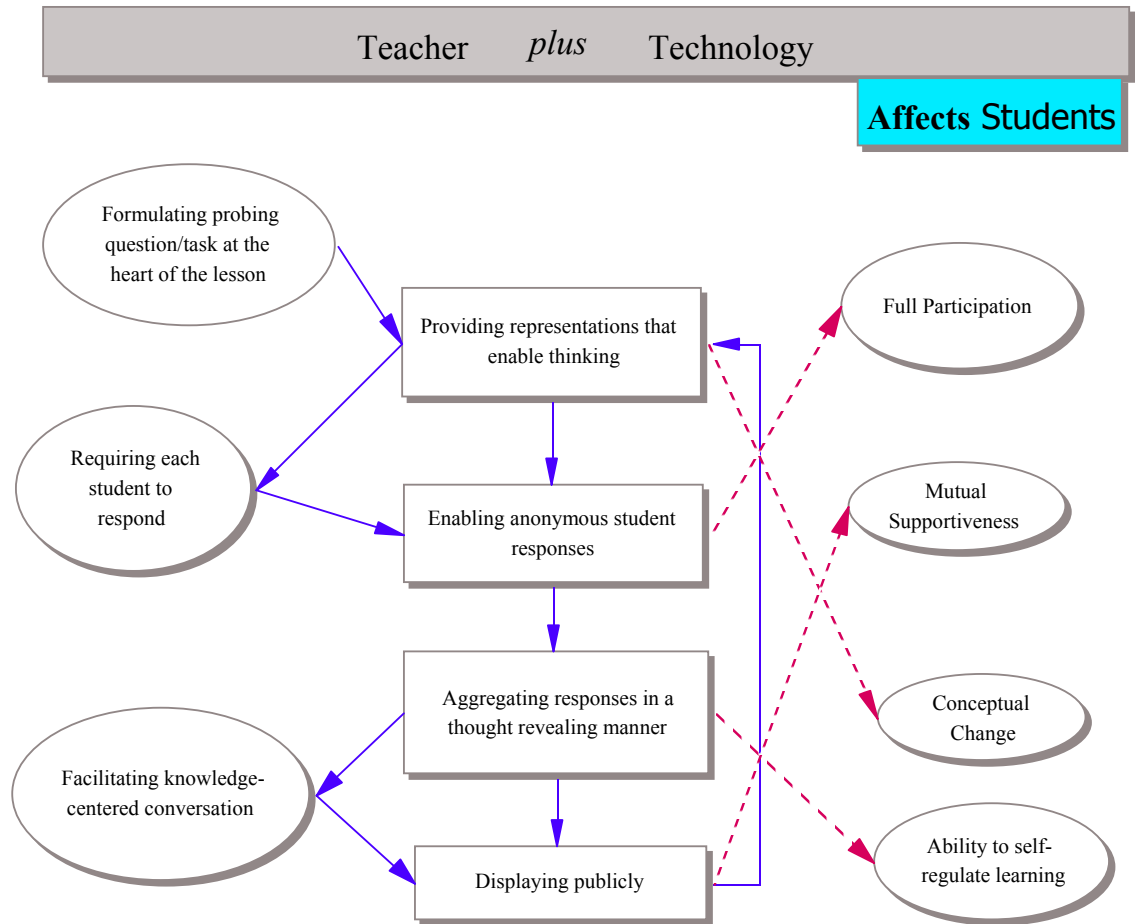


Figure 1: CATAALYST Pedagogy and Technology Integrate to Enhance Learning Adapted from Classroom response and communication systems: Research review and theory, by J. Roschelle, W. R. Penuel, and L. Abrahamson, 2004, p. 2. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

that enhances feedback delivery methods and capabilities. The focus of this study was to determine the impact on student achievement, as measured by learning gains, of these technology based feedback techniques.

Audience response systems are a group of emerging technologies that are being used to enhance instruction in the classroom and other training settings (Judson & Sawada, 2002; Simpson & Oliver, 2006). The primary components of each system are a

transmitter, a receiver, and computer software that collect and process data. Although in use from the early 1960s, response systems have recently proliferated, and have been designed by several manufacturers specifically for use in classrooms. The ARS system used in this study was the Classroom Performance System from eInstruction. Other manufacturers of audience response systems include Turning Technologies (TurningPoint), Qwizdom, GTCO Calcomp (InterWrite PRS), iClicker, Hyper-Interactive Teaching Technology (H-ITT), and many others (Barber & Njus, 2007).

The audience response system consisted of a wireless, radio frequency (rf) receiver connected to either a laptop or desktop computer which contained the database software specifically designed for the CPS receiver. The students responded to questions projected on to a screen via keys on handheld wireless transmitters (clickers). The receiver transferred the information to the computers' database where the responses were analyzed for accuracy, stored, and simultaneously displayed in a format chosen by the instructor. The software is versatile allowing for flexibility in number of questions, type of question (T F, multiple choice, free response), display (correct response only, graph), and storing options.

Theoretical basis of Audience Response Systems

The audience response system is based on sound constructivist principles. The principles of constructivism are based on a cognitive developmental framework asserting that knowledge is actively built rather than received, and that cognition is adaptive in character such that learning is constructed and organized via what is experienced (Jonassen, 1991). The constructivist principles assert, therefore, that learning will occur if the proper conditions for learning are available. These are learner-centered principles

rather than subject (topic) based, rooted in the research work of the Swiss psychologist Jean Piaget, the Russian psychologist Lev Vygotsky , and the American education reformer John Dewey (Lutz & Huitt, 2004). Piaget’s theory of cognitive development advocates that learners actively construct their understanding of the world, assuming that learners build upon prior experiences. New understandings are constructed by the learner as a result of new experiences that are incorporated into existing knowledge which is then adjusted to fit the new information (Lutz & Huitt, 2004). The process is thus adaptive in character. Vygotsky’s (1962) sociocultural cognitive theory also emphasizes this active construction of knowledge, but it stresses the role of language and the context of social relations (Lutz & Huitt, 2004). Dewey promoted the value of personal experience in learning (Lutz & Huitt, 2004). He rejected the notion of repetitive, rote memorization in the classroom and thought that the world was best understood through “directed living” where individuals constructed knowledge through interaction with their environment (Carvin, 2006).

These three constructivist perspectives and their learner-centered principles provide a theoretical framework for the use of audience response systems in the classroom. The framework aptly fits the ARS’ primary advantages of engaging, involving, and interacting with students (Simpson & Oliver, 2006), encouraging students to become actively involved in their own learning.

In science, constructivist learning stands in sharp contrast to the objectivist approach that views knowledge as something that can be imparted. Objectivists have a preference for the lecture format because they believe that instructors can pour knowledge into a student’s open head, close the head and have the students take a test on

the information (Jonassen, 1991; Leonard, 2002). Students who are at the level of understanding of the instructor are rewarded to the disadvantage of students who are at a lower level of understanding. Also, higher education administrators favor this approach for its low cost. In science, however, the objectivist learning approach is a dangerous one in light of how scientists acquire knowledge (Leonard, 2002). Much of science is through experiential learning, such as in a laboratory setting, where students have ample opportunities to process, interpret, and rationalize their experiences and the concepts they are exposed to. The ARS overcomes the objectivist approach by also engaging the students and provides new experiences through participation and discussions that lead to processing, interpreting, and rationalizing. There is ample experimental support in the literature for the constructivist learning approach showing higher student performance such that it warrants serious consideration in college science teaching, to the point that the approach has garnered the endorsement of the *National Science Education Standards* and the *Benchmarks for Science Literacy* (Leonard, 2002).

Definitions of Terms

Audience response system. A group of computer based technologies designed to collect and analyze feedback data from several people simultaneously.

CATAALYST. Acronym for Classroom Aggregation Technology for Activating and Assessing Learning and Your Students' Thinking.

Classroom performance system (CPS). Wireless audience response system manufactured by eInstruction, Inc. comprised of a receiver, response pads, and software for the computer.

Hawthorne effect. An observable phenomenon that demonstrates that when humans are aware that they are the subject of intervention, it affects the outcome of an experiment.

Normalized gains. The ratio of the actual average gain ($\% \text{posttest} - \% \text{pretest}$) to the maximum average gain ($100 - \% \text{pretest}$; Hake, 1998).

Delimitations of the Study

This study was delimited to biology classes at a 2-year college in South Florida. It was further delimited to courses with the same teaching objectives and instructors because the sample of this study was drawn only from biology classes with the same teaching objectives. Therefore, the findings of this study are generalizable to all students at the 2-year college. Other de-limitations on this study include the reliance on human assessments and on self-reported data.

Assumptions

This study was conducted with several assumptions. First, the study assumed that the audience response system would function successfully. Proper functioning of the technology is important for meaningful comparison between the audience response feedback groups and the non-audience feedback groups. It was also assumed that students in the treatment group had had minimal previous experience with audience response systems. It was assumed that students that have had previous experience with an audience response system may have an advantage over those that do not. The researcher also assumed that the students and instructors were receptive and willing to use a new technology in the classroom. The cooperation of students and instructors was essential for good data collection.

Summary

In summary, audience response systems are becoming a popular technology for providing classroom instruction as higher education experiences a shift toward a learner-centered, constructivist approach. This study was designed to measure the effects on student achievement of an audience response system in biology for science majors at a 2-year college in South Florida. Academic achievement was measured as normalized averaged learning gains (Hake, 1998).

The study is organized into the following chapters. Chapter 2 examines the relevant literature with respect to the efficacy and application of audience response systems in various topics, and in particular in the sciences. In addition, the history and relevant literature that supports the appropriate theoretical framework for accomplishing the objectives of the study is presented. The importance of viable instructional technology and perceptions to academic success are also examined. Chapter 3 outlines the research method and experimental design for the study. The data collection technique and experimental setup are addressed, as well as the validated instruments and data analysis methods. Chapter 4 presents the findings, analysis, and interpretation of the data, and Chapter 5 discusses the results with generalized conclusions and recommendations derived from the study.

CHAPTER II

LITERATURE REVIEW

This chapter describes the use of Audience Response Systems such as the Classroom Performance System (CPS) from eInstruction, and elucidates the increasing role of technology in transforming the 2-year college classroom to be more learner-centered, as well as knowledge-, and assessment-centered. The theoretical base that connects the pedagogy and technology with academic achievement is discussed along with the literature on the CPS technology and how it is utilized in instruction. The focus is on the use of CPS in the biological sciences and how it connects to the strategy of learning and performance in biology at the 2-year college level.

The interest in technologies that facilitate teaching and improve instruction is growing. Community colleges, along with other institutions of higher learning, have experienced a recent virtual explosion in the use of computers and associated technology. All indications are that, to one degree or another, these institutions and concomitant commercial enterprises are now willing to bet that technology will become a part of the educational enterprise. The technology tends to be expensive and its pedagogical value remains in question. This stems primarily from the difficulty in defining and measuring enhanced learning. The Audience Response System (ARS), also known as Classroom Response System (CRS), Electronic Response System (ERS), and “clicker” technology, is one form of the technology that promises cost efficiency in both high-enrollment and low-enrollment courses. It also has the potential to increase student learning. This technology, its integration into the 2-year college classroom, and its impact on science teaching, is described and reviewed in this section.

Audience Response Systems

With the rapid proliferation of new technology tools intended to enhance student learning, institutions of higher learning have been left scrambling for answers to such issues as incentives, cost assessment, pedagogical assessment, and intellectual property rights. The answers to these issues will determine the present and future shape of academe. In *How People Learn*, Branson et al. (National Research Council, 2000) describe audience response system technology and the pedagogy associated with it as one of the most promising innovations for addressing these issues, but in particular for changing the classroom from teacher-centered to more learner-, knowledge-, and assessment-centered.

The promise of the audience response system lies in its variety of uses and in particular the encouragement of active learning (Simpson & Oliver, 2006). Learning is active, as is illustrated in the Chinese proverb: “Tell me, I forget. Show me, I remember. Involve me, I understand.” Learning is also social. Piaget (1976) considered the role of social interaction in learning and argued, “social interaction is a necessary condition for the development of logic” (p. 80). Likewise, the “efforts to improve science education involve enhancing the social context of learning” (Preszler, Dawe, Shuster, & Shuster, 2007, p. 29). When learning is synthesized as active and social, it is interactive in other words. Audience response systems are interactive and as such promote interactive learning when incorporated into the classroom.

Audience response systems are wireless or wired technology in which students use keypads to answer questions presented in a lecture setting. The wireless systems were initially infrared (IR) and have been available for a while. There has been much

dissatisfaction with this form of the technology in large classrooms due to its line of sight limitations and high cost (Barber & Njus, 2007), but low-cost radio-frequency (RF) systems with a wider range of reception were more recently introduced. IR systems are more suitable for smaller classes although ARS vendors are more and more replacing it with lower-cost RF keypads.

One advantage of ARS is that students are able to get immediate feedback to their answers. The response is generally in the form of a histogram analyzing class results, or students may receive a signal on their individual keypads. The feedback thus obtained lends itself to reciprocal teaching through discussions of the correct/incorrect responses and to students learning from each other in a collaborative manner. The interactive system provides real-time information on the students' understanding, thus allowing the instructor to focus on misconceptions and concept areas that are confusing.

Although a good tool for taking attendance or quizzing students to test whether they are prepared for class, clickers are most effective at challenging students' understanding of the subject matter (Barber & Njus, 2007; Caldwell, 2007). This of course depends on good questions. "Do you understand this," followed by nodding heads or a "yes" reply does not guarantee understanding. As long as questions are carefully designed, clickers can reveal student misunderstanding (Caldwell, 2007). A good question helps determine whether the class has correctly grasped the concept. A good question also helps individual students recognize that they have a misconception. Conceptual questions, properly written, promote the most substantive discussions (Judson & Sawada, 2002). This is especially important in science where misconceptions get in the way of students' assimilation of ideas and making connections between concepts. To

substantiate this, Mintzes, Wandersee, & Novak (2001) describe the science classroom as being in need of a refocus on “meaningful learning and conceptual understanding” (p. 118) and that the goal of teaching science is “to help students assimilate well-integrated, strongly cohesive frameworks of interrelated concepts as a way of encouraging ‘real understanding’ of natural phenomena” (p. 118).

Other pedagogical applications of ARS include active learning (Barber & Njus, 2007; Judson & Sawada, 2002; Poulis, Massen, Robens, & Gilbert, 1998), assessment (Paschal, 2002), and peer instruction (Fagen et al., 2002; Knight & Wood, 2005). Caldwell (2007) cites examples of sustaining attention and breaking up the lecture, in view that the average human attention span is no more than 20 minutes.

History of audience response systems

Judson and Sawada (2002) date the use of electronic response systems in large lecture courses to the 1960s. The systems of the 1960s were hard-wired, and consisted of knobs or buttons mounted like telephone pads at the students’ seats. At the instructors’ station, gauges showed the results in percentage of students to each choice on a multiple choice question. More recent advances in the systems provide wireless and portability advantages along with graphing projection capabilities that allow the whole class to see the results. Although fundamentally different in architecture, the goal of both old and modern systems remains the same – instant feedback.

Research of the 1960s and 1970s using electronic response systems did not support an increase in student achievement as measured by exam scores (Judson & Sawada, 2002). Judson and Sawada (2002) cite a number of studies that emphasize incongruence in the lack of evidence of increased academic achievement and students’

endorsement of the system. In those studies, “positive attitude toward the class, feeling of the usefulness of the system, acceptance of the system, and feeling of increased understanding were all highly supported by the student survey data” (p.173).

Research from the 1990’s also show students support for the electronic response systems. However, “the issue of academic achievement remains open” (Judson & Sawada, 2002, p. 175). As Judson and Sawada (2002) put it:

In the 1990s investigations reporting student academic achievement were found only within the discipline of physics. The use of electronic response system was not a distinct characteristic among high achieving physics courses; however, electronic responses systems were viewed as one mechanism to elevate student interaction in large lecture halls. Among physics students, improved student achievement was detected when the pedagogy was distinguished as constructivist in nature, thus promoting interactive engagement among students. (p. 177)

There is an “ample converging evidence” (Caldwell, 2007, p. 13) that clickers improve student outcomes in exam scores or passing rates, comprehension, and learning. The literature also agrees, however, that “much of the research so far is not systematic enough to permit scientific conclusions about what causes the benefits” (Caldwell, 2007, p.13). Possible explanations range from the change in teaching methods (Judson & Sawada, 2002; Caldwell, 2007) to the “Hawthorne Effect” (Caldwell, 2007). Although interpreted in different ways, the Hawthorne effect is an observable phenomenon that demonstrates that when humans are aware that they are the subject of an intervention, it affects the outcome of an experiment: the test subjects are treated differently with the clickers and this special treatment brings about the improvement rather than the clickers.

Although difficult to rule out, a Hawthorne effect is less likely as exposure to the clicker system increases and the novelty is lost (Poulis et al., 1998). To further substantiate this, Simpson and Oliver (2006) make reference to a 1995 study by Halloran that “suggested the reverse [of the Hawthorne effect] may in fact be true: that students are initially resistant to the technology....but that they begin to learn more effectively once the novelty wears off“ (p.12).

Knigh and Wood (2005) conclude that generally the use of clickers either improves or does not harm exam scores; Simpson and Oliver (2006), however, find no consistent factors in the courses that use clickers that correlate with increased exam scores. Caldwell (2007) cites literature that asserts that for clicker research to advance, proper evaluation of various instructional methods require good standardized tests that assess student understanding of concepts. These can be found in physics, astronomy, and economics.

The lure of technology has been the claim that it was going to bring change to the academy because it offers students the opportunity for interactivity (Young, 2004). The verdict is still out on this. Howard (2002) conducted a mixed methods study comparing traditional delivery courses with interactive telecourses. The theoretical framework for this study was primarily Astin’s concept of engagement which suggests that students learn more when they are actively engaged with the material, their instructor, and their classmates (Howard, 2002). The study was also based on research that suggests that critical thinking is fostered by student’s active participation in learning. The results showed increased participation in students at the receiving site of the telecourse, but almost complete silence when interacting with students and instructors at the origin site.

The results suggest that in courses of this type it is difficult to take advantage of the benefits of social relationships with teachers and other students. The study was limited in its small, non-random sample, thus precluding any generalizations other than that further research was needed. Although technology is not the focus of the study, Rendón's (1994) qualitative research on student peer engagement provides a framework that can be used for technology in the classroom as a validating tool. Using such constructivist, student-centered learning strategies in physics, Mazur (1997) and Crouch and Mazur (2001) have pioneered studies in peer instruction and the active learning associated with the technique, and have reported increases in student understanding. In biology, using clickers to promote active learning has had mixed results. Freeman et al. (2007), and Knight and Wood (2005) have used clickers as a prescribed active learning tool resulting in increased performance; similar studies in microbiology (Suchman et al., 2006) and physiology (Paschal, 2002) did not find any clear significant effect of clickers on students' performance on exams. Preszler, Dawe, Shuster, and Shuster (2007) studied the frequency of clicker usage over a broad range (freshman- to senior-level) of biology courses and reported that an increased use of the response system increased student learning. There is strong evidence of the positive opinion of students toward the use of clickers in biology; however, the empirical effects of clickers on student learning are still unclear.

Creative use of technology can maximize its ability to be inviting, engaged, and involved. Several studies (Everett, 2002; Horowitz, 1988; Poulis, et al., 1998) tout these very characteristics of the Classroom Performance System (CPS) – a computer software designed for content assessment. Poulis et al.(1998) and Everett (2002) are two studies

done in a higher education setting; Horowitz's (1988) is in a business setting. There is evidence that studies on the technology is also done at the high school level and even at professional schools such as a Law school (Caron & Gely, 2004).

Interactive audience response systems have evolved from wired systems that required a dedicated classroom to portable wireless systems. Classtalk, a wired system developed by Better Education, Inc., was a very common system used in research studies of the 1990s. Classtalk has since been discontinued and replaced by other wireless systems like the Q4 model from Qwizdom, Inc., Turning Point from Turning Technologies, Interwrite Personal Response System (PRS) from GTCO Calcomp, iClicker, and Hyper-Interactive Teaching Technology (H-ITT). Another popular system, and the one used for this study is the Classroom Performance System (CPS) from eInstruction, Inc.

Classroom performance system. The Classroom Performance System (CPS) is a product of eInstruction Inc. The technology is typically a handheld device that allows large groups of students to individually engage instructional content in real time in the classroom and obtain immediate feedback. The CPS consists of an infrared (IR), or radio wave (RF), transmitter ("clicker" pad), a receiver, and software that collects the individual student responses to instructor-posed questions and displays the results. Each student gets a transmitter, which is made up of eight buttons in the IR version. The RF version is significantly different with improvements in keypad design. It is this version that was used for this study and described herein. The keypad and its receiver are shown in Figure 2.

When a student transmits, both the response and the unique number of the pad are registered. The receiver transfers the information to the computer software where it is stored in a location specific to the transmitter number. The results can be projected to a screen showing all the pad numbers that have responded. Once all responses are received, the next question can be displayed and answered. The results can be displayed numerically or graphically, or exported in various formats. The new keypad introduced in 2006 has an LCD screen providing students with confirmation of their answers. The

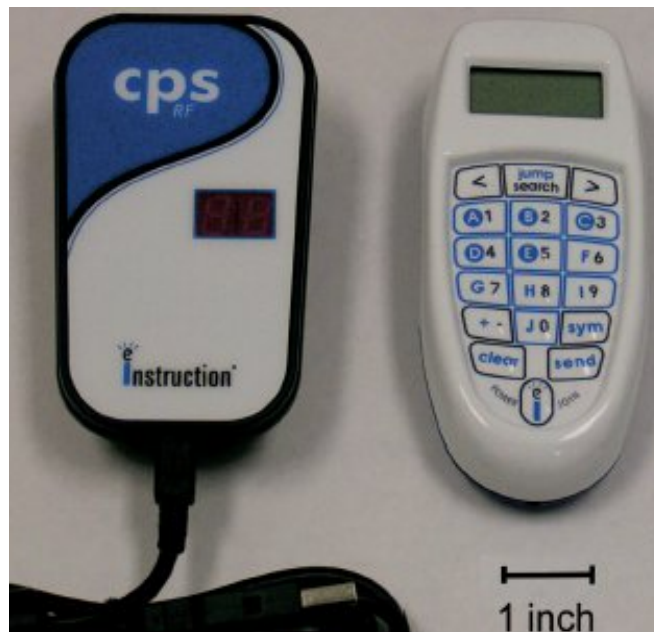


Figure 2. CPS receiver (left) and keypad

immediate feedback, the interactivity, and the fact that students can answer without the risk of embarrassment, makes this technology attractive.

As reported by Barber and Njus (2007) students do incur a cost for the “clickers” and there are additional charges for keypad registration. CPS is unique in charging for

registering the unit. The lifetime cap significantly reduces cost, however. CPS is fully integrated with PowerPoint. A CPS menu bar, which is new and an add-in to PowerPoint, allows for easy creation of question slides. PowerPoint slides can be imported into CPS, but they cannot be edited. It is possible to run CPS in parallel with PowerPoint. CPS offers a variety of ready-to-use curriculum packages. These include lessons, quizzes, and tests that can be downloaded right into CPS using a few simple graphic menus. A search function allows content to be easily located and downloaded (eInstruction, 2007).

Technology and audience response systems (ARS) in the 2-year college classroom. The use of technology in the classroom has been an indicator in the systemic shift in paradigms from the “instruction paradigm” to the “learning paradigm” (Barr & Tagg, 1995; Hanson, 2006; VanWagoner, Bowman, & Spraggs, 2005). As Allison and Scott (1998) state —“it is becoming clear that there is increased focus on the role of learning rather than teaching as the operational mind-set in higher education” (p. 71). The catalyst for this shift is in the Wingspread Group (1993) report of a “weeding out” of students from higher education by the current system and a lack of important skills in those who do graduate. This led to the belief by some that higher education as it presently functions does not provide for the goal of student learning; that there is a need to educate more people effectively and efficiently and with higher standards.

Instructional and computer technology advocates support the reforms by promoting awareness for the role technology plays in the curriculum and pedagogy (Allison & Scott, 1998; Brown, 2000; Evans, 1999; Lorenzetti, 2002; Merisotis & Phipps, 1999; Privateer, 1999; Stetson & Bagwell, 1999). The technology industry recognizes the opportunity this shift presents for new and creative instructional

technology. As it has been with textbook publishers, institutions of higher education are actively pursued by a variety of commercial and non-commercial enterprises that are willing to bet that technology will be a permanent part of education. Many have forged partnerships with 2-year colleges bringing technology into the classroom, and with it a new learning modality. Teaching with technology in the classroom is one way to address the learning styles of the students by matching the style with the appropriate instruction and experience (Merisotis & Phipps, 1999; Pascarella & Terenzini, 1998; VanWagoner, et al., 2005). Although the verdict is still inconclusive from “scientifically based research” specifications (Roschelle, Penuel, & Abrahamson, 2004, p. 3), ARS technology promises to accommodate students through interaction, active learning, increased communication and reasoning skills, and questioning and feedback (DeBourgh, 2008; Roschelle, Penuel, & Abrahamson, 2004).

The 1990’s ushered in the tail-end of the MTV generation and a more computer savvy population. Cell phones, mp3 players and Compact Disk are widespread, and notebook computers are now common. Institutions, in partnership with the technology industry, rushed to accommodate this growing genre and millions of dollars have been spent on “smart classrooms” that are packed with computerized projection systems, Internet ports at every seat, and video cameras with motion detectors that track the lecturer’s movements. The reason for this was to assist teaching. With a steep learning curve, large time commitments, and a resistance by older faculty, the initiative has had mixed results (Young, 2004). Through workshops on the use of technology and providing rewards for participation, institutions have made much progress but more needs to be done (Young, 2004).

Just having technology is no longer enough. The technology use has to go beyond the passive PowerPoint presentations to more creative uses for computer technology in the classroom. The 2-year college faculty is exhorted to “move quickly to using technology to increase learning in profound ways” (VanWagoner, *et al.*, 2005, p. 44). In fact, higher education is at a cross-road in which it essentially has to decide whether to continue the restrictive administrative approach to learning or embrace the meaningful change promised by instructional technology as a cognitive tool (Privateer, 1999). In support of the latter, ARS, like eInstruction’s Classroom Performance System, has been used to promote advanced reasoning skills in Nursing (DeBourgh, 2008), apparently improve learning, although this may be because of the novelty, as in the well-known Hawthorne effect, (Herreid, 2006; Simpson & Oliver, 2006), and encourage active learning (Simpson & Oliver, 2006).

Introductory science courses address several important goals in higher education. One critical goal is to offer science majors their first college experience in an area in which to focus their future careers and laying a strong foundation for additional coursework. Biology for science majors at the 2-year college level is a “gateway” course. It is the first of a two-semester sequence designed for the biology major or pre-professional student in medicine, dentistry, pharmacy, physical therapy, optometry or veterinary science. This course provides the student with an understanding of the basic principles of biology including an introduction to the scientific method, cell chemistry, cell biology, energy transformations, mitosis and meiosis, genetics and molecular biology. In this course the students are introduced to the principles of biology that will form the foundation to build on as they progress into the university system and their

future careers. This is one reason why it is unacceptable to either have a high failure rate or a weak foundation in a “gateway” course like biology for science majors. Other reasons include that it contributes to low graduation rates and extended time-to-graduation for the institution, and, given the student population comprising the 2-year/community college, underrepresented students and students of disadvantaged backgrounds are more likely to be affected. The 2-year/community college is poised to spend a considerable amount of its limited funds (Outcalt, 2003) on a teaching technique whose effects on learning are still unclear. The introductory lecture course is an integral part of the undergraduate college system and the problems that biology has vis-à-vis the failure of the lecture format and the loss of under prepared students make it imperative to determine if the ARS is a teaching technique that will increase achievement in the poorly prepared students characteristic of 2-year/community colleges. As such, it should be an effective way to increase recruitment and retention of minorities in the natural sciences.

In biology, the use of ARS (“clicker” technology) has had mixed results. Active learning increased performance through lower failure rates and higher total exam points in an introductory biology course, but this did not give concomitant increased scores in exams (Freeman et al., 2007). Freeman et al. were motivated by the high failure rate in the “gateway” introductory biology course. Using a linear regression model they obtained from performing a risk analysis on 3338 students who had started the introductory biology classes between autumn quarters of 2001 and 2004, and an ANOVA, they determined that the predicted scores for students in spring 2003, spring 2005, and fall 2005 were not significantly different. This gave student populations that were comparable for these quarters. All sections were taught by the same instructor.

The spring 2003 class was the equivalent of a pilot study in which active learning exercises were introduced. This included case history done in groups, think/pair/share, exam-question writing, papers, and discussions of exams questions from previous quarters. These exercises were not graded. There was no significant improvement in the failure rate. In the spring 2005 course, students registered blindly in two equal-sized sections (~345) taught back to back in the same classroom using identical notes. Students were randomly mixed in lab sections and required field trips. Two questions posed at the beginning of class reviewed the previous session and that day's material. After a modified Socratic style delivery, two more questions were posed at different times. These questions were designed to be difficult and if more than 60% of the responses were incorrect, the peer-instruction method was implemented, i.e. students discuss the question amongst themselves and then re-answered. In the ARS section, correct responses were allotted one grade point, tabulated as part of the grade, and kept secret. The second section used cards with either A, B, C, or D answers. These were held up when responding, were made public, and were not graded. Students from both sections also did online practice exams and study group practice exams to randomly fit one of the four designs: clickers + online practice, clickers + study group practice, cards + online practice, and cards + study group practice. All students took a common final exam and a common second midterm. The results indicated a statistically significant (Fischer's exact test, one-tailed, $p = 0.049$) drop in failure rate when compared to the spring 2003 class, and significant increases in exam scores (t test, $p < 0.001$), and significant increase in the second midterm average score (t test, $p < 0.001$). An ANCOVA statistic with the predicted grade as the covariate did not show any significant difference in the four

designs' effect in terms of boosting achievement. Other statistics show high-risk students in the clicker sections doing better than high risk students in the card section, although this was not a significant difference (ANOVA, $p = 0.080$). On the other hand low risk students in the cards section did better on the common exam than students in the clicker section (ANOVA, $p = 0.034$). Average attendance was significantly higher in the clickers section than in the cards section (paired t test, $n = 34$, $p < 0.0001$), and in the clickers section, attendance had a significant effect on predicting final grade ($R^2 = 0.24$, $n = 173$, $p < 0.0001$).

The two fall 2005 sections also had blind registration, were taught back to back in the same room, and used identical notes. The students in both sections took weekly graded exams online, and were required to purchase and register a clicker. In this design students from both sections were administered daily, in-class, multiple choice questions. In one section, the questions were graded for right/wrong responses as in spring 2005. In the other section, points were given for participation for each response regardless of the response's accuracy. All students took a common final exam, but each section took a different but equivalent midterm exam. The average exam scores were identical to spring 2005 and significantly different from spring 2003 (ANOVA, $p < 0.001$); thus the results replicated the spring 2005 results. The failure rate for this design, however, was no different than the failure rate for spring 2003 (Fisher's exact test, one-tailed, $p < 0.09$). ANCOVA statistic, with predicted scores as covariate, showed no difference in exam scores between sections, suggesting that both designs worked equally in improving exam performance. An ANCOVA gave significant differences in performance of students in

the clicker section that was graded than in the section where it was not ($p < 0.001$).

Attendance was again shown to be a significant predictor of final grade.

Freeman et al.'s (2007) study is one of many showing improved student performance through active learning designs in undergraduate science courses (Crouch & Mazur, 2001; Knight & Wood, 2005; Mazur, 1997). It is unique in several ways: in the emphasis on grading or public display of cards as active-learning exercises, and in distinguishing the performance between high- and low- risk students. The study highlights the benefits of highly structured active-learning environments for students in introductory biology classes as well as the benefits for students who are at high risk of failing the course. At the 2-year/community colleges, a high proportion of high risk students are attracted to or enroll in the introductory biology classes. The study provides insight for improving introductory biology courses to increase retention of high risk students in science majors. On another note, some might question the efficacy of having the same instructor teaching all the courses, and the temporal separation of the courses.

Knight and Wood (2005) reported increased learning through peer instruction and the use of clickers to promote cooperative learning in a developmental biology course. They conducted an experiment over two consecutive semesters by the same instructors. The Fall 2003 class was taught in the traditional way, while the Spring 2004 class used more student participation and cooperative problem solving. Pretest, posttest, and homework problems were used to estimate and compare gains between the two semesters. Higher gains and better conceptual understanding were found in the more interactive class. The interactive course format was repeated with similar results in Spring 2005 semester to assess reproducibility of the results. For completing the pretest,

students were rewarded with points, although it had no effect on grades. The pretest and posttest were identified as valid assessment tools of the conceptual understanding for concepts tested. The questions were similar but not identical for Fall '03 and Spring '04 classes with the latter designed to test more conceptual and less factual knowledge. The exams were substantially different and therefore could not be compared meaningfully. The results were compared based on normalized learning gains which allow valid comparison and averaging of learning gains for students of different pretest scores. The experiment was not ideally controlled but it still provided valuable insight into experimental design and to students' perceptions and their response to challenging topics. The effectiveness of the control can be questioned in two aspects: the temporal separation from the treatment groups, and the fact that both authors were present only in the interactive (treatment) group.

The results showed a slight negative correlation between the pretest performance and learning gains (Pearson correlation test coefficient: -.4 in Fall of '03 and -.1 in Summer of '04). The authors state, without providing the data, that various combinations of pretest and gains were represented. These include low pretest with high gain, low pretest with low gain, high pretest with high gain, and high pretest with low gain.

The average performance and standard deviation on the pretest were not significantly different in the two semesters: Traditional, 34%; interactive, 31%; indicating that the students of both groups were equally prepared. The average on the posttest was significantly higher in the interactive course, by 9% ($p = .001$, two-tailed t -test).

The results also showed that normalized learning gains were significantly higher for the interactive class. Normalized learning gain is defined as the actual gain divided by

the possible gain, expressed as a percentage $[100 \times (\text{posttest} - \text{pretest}) / (100 - \text{pretest})]$. The importance of normalization is that it provides valid comparison and averaging of learning gains for students with different pretest scores. The two courses showed a significant difference in average learning gains (16% difference; $p = .001$). More students in the interactive courses exhibited greater than 60% learning gains than in the traditional class. Both “A” and “B” students made higher gains in the interactive class, while “C” students did not show any gain.

The Spring '05 class had no curve and no 5% incentive on group grade, but had an additional group activity – concept mapping data from current articles with questions to answer as group in class. The results were similar to the Spring '04 class. This implies that the effects of interactive engagement seen between Fall'03 and Spring'04 are reproducible. The experiment was not ideally controlled but it still provided valuable insight into experimental design and to students' perceptions and their response to challenging topics. The effectiveness of the control can be questioned in two aspects: the temporal separation from the treatment groups, and the fact that both authors were present only in the interactive (treatment) group.

There were no significant effects of clickers on students' exam performance in microbiology (Suchman et al., 2006) and physiology (Paschal, 2002) courses, although some encouraging trends emerged. In a combined study of six biology courses, Preszler et al. (2007) concluded that the use of response systems increased student learning and had a positive, but not significant, influence on students' performance on exam questions. None of these studies were conducted at the 2-year/community college level, although

the majority was freshmen level courses. This emphasizes the need for data using the ARS in biology courses at the 2-year/community college level.

Focus of ARS technology. The focus of ARS technology in the classroom should be on the production of intelligence, otherwise known as learning, which is a primary goal of higher education and an essential component of the 2-year/community college mission of meeting the needs of the community. Technology, computers in particular, generate a tremendous amount of information in a very short time. A fallout of this is that information quickly becomes outmoded (Privateer, 1999; Merisotis & Phipps, 1999). The focus on intelligence production is more practical, therefore, than the storage and recall of random and quickly outmoded information. Areas in which faculty have embraced the technology include writing classes with computer labs, computers in science lab to collect data, learning labs that help the under prepared student, Internet access in the classroom, electronic mail (e-mail), web-based instructions, and use of interactive CD-ROM's, distance learning through on-line classes, and more recently in the use of audience response systems like the Classroom Performance System.

The present status of ARS, and technology in general, and its uses in the classroom appears to be a contributing factor to the Dickensian paradox of “It was the best of times; it was the worst of times...” that institutions presently face. Educators and political leaders alike have enthusiastically embraced the use of technology, both for its perceived cost efficiency and its marketing potential. The result has been a rapid gain in momentum in instructional technology, proliferation of virtual universities, and lingering questions in the minds of researchers. Paramount among the questions is whether the ARS technology is delivering the learning it promises. Merisotis and Phipps (1999) and

Pascarella and Terenzini (1998) insist that technology has helped institutions keep this focus on the essential goals of teaching and learning, but it also unearths the very salient question: What is the best way to teach students? The “digital pedagogies” presented by the variety of technology emphasizes one underlying precept: not all information technologies and applications have the same cognitive impact (Evans, 1999; Nygren & Fisher, 1999; Privateer, 1999; Stetson & Bagwell, 1999). Yet, the institutional culture of higher education is that technology is a panacea or that computers are agents of universal change, and thus able to bring about institutional reform in the areas of course content and delivery. Add to this the popular idea that video games and computers alone actually help students learn, and a disconnection with the truth becomes apparent. Much of the research indicate that technology is not nearly as important as other factors, such as learning tasks, characteristics of the learner, time-on-task, student motivation, and the instructor. Technology cannot replace the human factor and invariably the instructor ends up being a technology jack-of-all-trades as content expert, design expert, motivator, mentor, and interpreter (Allison & Scott, 1998; Merisotis & Phipps, 1999; Pascarella & Terenzini, 1998).

In essence, the present status of technology is in the first of three phases described by Princeton historian Robert Darnton in the essay *The New Age of the Book* on electronic publishing (March 18, 1999 New York Review of Books; cited in Merisotis & Phipps, 1999). The three stages are: an initial phase utopian enthusiasm, a period of disillusionment, and a new tendency toward pragmatism. This first stage, an “infatuation with automated learning” (Privateer, 1999, p. 4), is of particular importance in the burgeoning distance learning agenda, as well as other agendas of the academic

community: it provides the motivation and excitement that drives the process. However, the traditional ideas on the art of teaching, i.e. pedagogy, are not compatible with innovational uses of technology. In most cases it is a situation of dressing the same doll in flashier, more expensive draperies. The doll remains the same, it only looks different. This is what Privateer (1999) seems to be saying: “higher education must understand and act upon the fact that the change extending the life of a controlling paradigm is not the same change that produces a new one” (p. 5). In order to harness the computer’s power to revolutionize learning, policy makers must first understand this. The research and literature review are in agreement that the academic community, including community colleges (Nygren & Fisher, 1999; Owen & Demb, 2004), has a lot to learn about how, and in what ways, technology can enhance learning (Allison & Scott, 1998; Merisotis & Phipps, 1999; Privateer, 1999). The 2-year/community colleges, however, seem to be greatest at risk. Their mission of responding to the educational needs of their community requires that they must adjust quickly to the way technology is embraced by the society. Moreover, as Nygren & Fisher (1999) explain, budget constraints coupled with rising enrollments make community colleges and state institutions fertile testing grounds. Faculty at most higher education institutions may view time invested in trying out educational innovations as detracting from tenure eligibility. This is generally not true for 2-year/community colleges. This places these institutions in the “lead” as innovators of technology, experimenting with programs without the benefit of long-term analysis (Owen & Demb, 2004), and with differential success (Evans, 1999; Gatz & Hirt, 2000). As “guinea pigs” or “miner’s canary”, the community college experience with technology and learning is thus a source of guidance for other institutions. There is a paucity of ARS

experiments in biology at the 2-year/community college level however, highlighting a gap that needs to be filled if the 2-year/community college is to act as this source of guidance.

The Challenges of Integrating ARS Technology into the 2-Year College Classroom

Cost assessment. The cost of the systems and the support needed is an issue in 2-year/community colleges. Computer technology has a very short generation span, and constant upgrades get expensive; personnel, training, lack of use, cost of software, and mismatch between faculty interest and training are among some of the more salient cost considerations (Daniels, 1993; Young, 2004; VanWagoner et al., 2005). 2-year/Community colleges are heavy investors in technology, having been convinced by hardware and software developers to purchase technology of questionable efficiency and function (Outcalt, 2003). Nevertheless, the availability of technology has been used as an indicator of success, but this has to be coupled with the quality of the technology, the networking infrastructure, and software availability (VanWagoner et al., 2005), and more importantly the need to shift from focusing on technological solutions to analyzing knowledge management (Outcalt, 2003). What are the appropriate units and time frames for measuring the costs of teaching? Studying the cost-implications of ARS, or any new forms of pedagogy, requires rigorous experiments, which in turn requires time and deliberation. The constant shifts inherent in the technologies make cost assessment difficult. Issues such as groups of students, classes, departments, the community college as whole are units of analysis that must be determined. Nygren and Fisher (1999) propose such obstacles are best confronted by proceeding with one model or another, work through the underlying difficulties and thus learn by doing.

Pedagogical assessment. It has been presented that evaluating the pedagogic value of technology has its own set of complex issues. Foremost among these is the question of how to best define and measure enhanced learning (Allison & Scott, 1998; Merisotis & Phipps, 1999; Nygren & Fisher, 1999; Privateer, 1999). Other issues include how to incorporate traditional assessment tools with the technology. Does the novelty of instructional technology, like the ARS, undermine the needs assessment or baseline surveys typical of many traditional tools? Word processors and emerging computer software, for example, do not lend themselves to assessment over a large timeframe since they were not needed before they were invented.

How adept are the students and faculty at using the tools? What is their technical competency? How portable across courses and departments, and how viable, is the technology? What counts as effective use of instructional technology? The issues with these questions include the novel nature of the technology which makes it difficult to establish control groups, benchmarks, and judge improvement (Nygren & Fisher, 1999). The research and literature review are in agreement that the academic community, including community colleges (Owen & Demb, 2004; Nygren & Fisher, 1999), has a lot to learn about how, and in what ways, technology can enhance learning (Allison & Scott, 1998; Merisotis & Phipps, 1999; Privateer, 1999). Due to budgeting pressures, these institutions may not wait for the experiments to be completed and at the least need to know what some of the answers may look like. This study helps to provide insight into some of these questions.

Intellectual property rights. Ownership of technological innovations constitutes another set of complex issues. Administrators hold the position that the institution's

ownership is embedded in the substantial investment in technological infrastructure while the faculty claims unique intellectual contributions. On the other hand, commercial enterprises are willing to bet that technology will be a permanent part of education, and as such actively pursue the rights to new products that may be commercially marketable. Ideally, the goal is to establish equitable access and fair distribution to all parties involved (Nygren & Fisher, 1999), but this depends on many different factors.

Incentives. Steep learning curves, demands on time for training and developing new materials, and perceived negative impact on tenure and job prospects are some reasons for faculty resistance to the use of technology. Incentives on the other hand include more interaction with students, exposure via web-based courses, better teaching methods, improved student access to information, and better faculty workload. Faculty, however, is interested in pedagogical gains while institutions are interested in cost-effectiveness (Allison & Scott, 1998; Nygren & Fisher, 1999; Privateer, 1999). By studying the effect of ARS, a relatively cost effective technology, on learning, this study contributes to information that assists in aligning the interests of both faculty and administration.

Theoretical Foundations

The use of technology as a learning tool is still lacking a theoretical basis that is integrative, coherent, and sophisticated (Owen & Demb, 2004; Pascarella & Terenzini, 1998). Theory provides the ability to build on the work of others and to identify and address the more significant questions. Without a guiding framework, researchers will be faced with the difficulty of replication and thus the ability to generalize the findings. According to Judson and Sawada (2002), investigators are emphasizing effective

pedagogical constructs supported by audience response systems and are influenced by the constructivist perspectives.

Constructivism. As Judson and Sawada (2002) suggest, much of the philosophical and theoretical rationale of the pedagogical value of audience response systems is rooted in constructivism. The concepts of constructivism relate to the mechanisms of learning and the nature of knowledge – it is basically a theory about how people learn. The characteristic assumption of the theory is that knowledge is constructed by the learner (Lutz & Huitt, 2004) through reflection such as questioning, exploring, and assessing what is experienced. Learners construct meaning as they learn, thus the construction of meaning *is* learning. The theory fits the shift from the “instruction paradigm” to the “learning paradigm” discussed earlier in which the teacher/educator shifts the emphasis from the subject material to the learner in an interactive, dynamic way that focuses on the notion that knowledge is not independent of the meaning the learner constructs. The approach borrows from many other practices in pursuit of its goal of helping students learn how to learn. As Judson and Sawada state it, “Affected by constructivist cognitive science and other epistemically compatible perspectives, investigators have highlighted the importance of collaborative discourse that allows students to negotiate meaning in science and mathematics classes” (p.173). Several researchers provide the theoretical underpinnings of the constructivist approach to the teaching/learning process, most notably John Dewey, Jean Piaget, and Lev Vygotsky (Lutz & Huitt, 2004). A brief overview of their ideas and contributions is presented below.

John Dewey. The philosophical root of constructivism is generally attributed to the education reformer John Dewey. Dewey promoted the value of personal experience in learning. He rejected the notion of repetitive, rote memorization in the classroom and thought that the world was best understood through “directed living” where individuals constructed knowledge through interaction with their environment (Carvin, 2006). As the foremost educator of his time, Dewey (1944) proposed radical reforms in pedagogical methods and curricula for preparing young people to live in a democratic society. Reflective assessments of personal experiences provide the learning necessary for successful living. He believed that thinking and doing were continuous, reciprocal activities rather than separate activities (Lutz & Huitt, 2004). Dewey had a profound impact on the progressive education movement setting the stage for later researchers (Lutz & Huitt, 2004).

Jean Piaget. To many in education, the work of Jean Piaget in constructivism is the standard. In this framework the individual goes from the youngster stages to those of the mature adult in a series of stages in logical reasoning. In Piagetian theory, learners assimilate new experiences to what they already know or they could accommodate their ideas to integrate new information. Assimilation and accommodation is a critical component of modern constructivism. The process is a continuous activity of self-construction that results in the building of knowledge (Lutz & Huitt, 2004) and can be used to construct curricula. One such design involves creating conceptual conflict and then helping the student resolve that conflict, as is seen in the well-known curriculum design “the learning cycle”: explore-explain-apply (Lawson, Abraham, & Renner, 1989). One version of The Learning Cycle involves exploring an idea by doing hands-on

activities, explaining the concept by connecting the hands-on experiences, and applying the concept in new, hands-on situations. The clicker technology lends itself to this approach in that the explore phase can be presented through demos, pictures and paragraphs, and questions. Concepts are then explained by the instructor or through student discussions. The apply phase is seeing how the concept works in new activities or case studies.

Lev Vygotsky. Another view of constructivism has been built on the work of L. S. Vygotsky in which society and culture impact learning. Vygotsky believed that social occurrences filter all higher mental functions that are then integrated into a person's thinking through the use of language (Lutz & Huitt, 2004). This idea that learning is primarily socio-linguistic is well accepted by language and reading educators but disputed among mathematics and science educators (Geer & Rudge, 2002; Lutz & Huitt, 2004). Nevertheless, it is generally agreed that Vygotsky provides the primary rationale for curricula that revolve around active student participation in negotiating and resolving meaning, i.e. learning. As Geer and Rudge (2002) put it: "All would agree the building of knowledge structures on the part of a student requires she or he be actively engaged in the process of learning" (p. 2). As a result, in Vygotsky's model, classroom discussions become a major focus. The CPS clicker technology, as used by Mazur (1997), Crouch and Mazur (2001), and others, and as it is used in this study, fits this model well in that classroom discussions and "convince-your-neighbor" activities are integral to its use.

This brief overview does not include the work of other researchers that have minor detractions from or that are major synthesis of the work of Dewey, Piaget, and Vygotsky. It does provide insight to the wide variety of epistemological and ontological

positions that currently exist in constructivism. Collectively they provide the foundation for the constructivist approach to learning (Lutz & Huitt, 2004), and perhaps a beginning definition of a constructivist classroom in which people are working together to learn. Such a classroom would have the elements of constructivism already mentioned. Lutz and Huitt (2004) provide a detailed presentation of the guiding principles for putting a constructivistic viewpoint into practice. The classroom would be a place where inquiry is conducted. The primary mode for engaging in resolving meaning would be discourse. The educator would understand the mental models of the students in order to help them learn and integrate new understanding. Assessment, measurement, and evaluation would be a natural part of the learning process through continuous feedback with active involvement of students in making judgments of learning. The latter would be combined with the teachers' judgment or that of experts when making decisions about grades.

The ARS technology incorporates these principles since, as has already been mentioned and is further emphasized below, it is inquiry based, providing continuous assessment with immediate feedback. It can be designed for discussion, participation and engagement. As constructivistic methods gain popularity, " what is needed is more work on both the validity of specific components or principles as well as methods of documentation that can accurately describe the benefits of this approach to student learning" (Lutz & Huitt, 2004, pp. 14 - 15). The ARS technology promises to provide a means to validate and document student learning.

ARS Outcomes Assessment

Primarily used in math and science courses, audience response systems have been providing immediate feedback to students and information of comprehension to

instructors for the past four decades (Judson & Sawada, 2002, Lowery, 2005). The research has been inclusive with results prior to the 1980s with ARS technology having no significant correlation with student academic achievement; post-1980 the results have been more positive. Judson and Sawada (2002) identify gains in conceptual understanding in physics courses using the ARS in a constructivist-oriented classroom. For the most part, however, the research studies are limited in scope (Lowery, 2005). Some of the criticism by Lowery (2005) include that the studies were for too short a time involving too few students, involved anecdotal findings, only one brand of ARS tested, and usually restricted to applications in math and science classrooms. One notable exception to these criticisms is Hake's (1998) study for physics. The study was a thorough statistical analysis of differences in pre/post test gains obtained for 14 traditional and 48 interactive-engagement physics courses involving 6542 students in high schools, colleges, and universities. The studies have collectively provided information on the impact of the ARS technology on student learning for areas of attendance and preparation, comprehension, participation, peer learning, retention, and student satisfaction (Caldwell, 2007). Several of the studies that have contributed this information in the area of science are analyzed in the following paragraphs.

Woods and Chui (2003) implemented the use of CPS in an attempt to address the problems of active participation and comprehension assessment in their large university classrooms (Woods in biology; Chiu in physics). They reported significant improvement in participation and interactivity with the technology acting as a positive stimulus to make the classes more interesting and lively. Attendance rates appeared to increase as assessed by a class survey consisting of four questions, one of which queries the impact

of the technology on the student's likelihood to attend class. The authors conjecture that this stems from the perception that the students' input matters and from the fact that the students' participation contributes to their final total points and can be easily tracked with the system. Woods and Chiu estimate CPS technology was instrumental in breaking up the monotony of a straight lecture and served to highlight the difficult vs. easy topics, the controversial topics, and student comprehension of the material. Woods and Chiu's study is a good example of what Lowery (2005) is critical about. It is mostly anecdotal with no reference to the class size, how data was collected or analyzed, and a brief sample for methods. The study references no other articles but yet is often cited when attendance is used as a motivating tool for learning (Lowery, 2005; Trees & Jackson, 2007). The authors themselves agree that more systematic studies are needed.

Burnstein and Lederman (2001) provide a more systematic approach to the use of wireless technology in the large lecture classes. The high attrition rate in their first-year science and engineering students at Illinois Institute of Technology was attributed to the ineffective and impersonal nature of the large class format rather than to the subject matter. They wanted to improve the lecture experience by involving the students more in the lecture. Over a five year period, starting in 1995, they implemented an interactive student response system consisting of two-way wireless keypads. Each student was assigned a keypad at the beginning of the class to use for entering responses during lecture. Multiple choice and "yes"/"no" questions were presented at relevant points in the lecture, as well as questions testing the students preparation for class. Peer instruction was encouraged for some questions. The students were timed for their responses and the results were projected for class viewing in the form of a histogram. The responses were

saved for analysis or for grading purposes. Students were awarded 10 points for a correct answer, 3 points for any answer, and 0 points if no answer was forwarded. Since the keypad questions could be used for understanding homework problems, less weight was given to the homework.

Citing evidence that good preparation prior to lecture improved student performance, Burnstein and Lederman (2001) asked questions that referred to the assigned reading, recorded the results and used these results in the final grade. They also believed that students would more likely read the material prior to class if they knew they would be tested on it with keypad questions. As a result they would come to class better prepared.

The system in effect monitored the students' attendance and attention. Burnstein and Lederman (2001) found that "...when keypad scores count for greater than 15% of the term grade, there is a dramatic improvement in attendance that reaches 80 – 90% level and, in addition, the students make genuine attempts to prepare for the reading quizzes and remain alert throughout the lecture period" (p. 10). In addition, the flexibility of the system in facilitating the weaving of questions throughout the lecture and generating impromptu questions, allowed Burnstein and Lederman to ascertain, in real time, whether students recognized and understood concepts. They were thus able to repeat or modify topics to increase understanding.

Burnstein and Lederman's (2001) study was deficient in the absence of a control. The instructors were well aware of this and faulted the inertia of faculty in using and evaluating the technology. If more faculty start using the system "...we would be able to accumulate and evaluate data from classes taught with and without keypads by equally

competent faculty under controlled conditions. This would allow a quantitative assessment of the improvement achievable using a keypad classroom.” (p. 11)

Peer learning is one method of instruction that benefits from clickers (Caldwell, 2007). Both this form of active learning and interactive student response systems in the classroom have attracted a lot of interest for they have been demonstrated to result in higher learning gains and exam scores over the traditional content-based lecture (Caldwell, 2007; Hake, 1998; Mazur, 1997). Mazur (1997) is the most referenced publication in articles discussing active learning and interactive student response system. In 1991, Mazur began developing his now well-known “Peer Instruction” (PI) strategy in physics classes at Harvard. For 10 years, Mazur showed continuous improvement in pretest/posttest gains by students in the Force Concept Inventory (FCI) (Hestenes, Wells, & Swackhamer, 1992) using his PI pedagogy (Mazur, 1997; Roschelle, Penuel, & Abrahamson, 2004). One format that peer learning uses is the ConcepTest format in which tests of conceptual understanding are alternated with mini-lectures. In this format the students spend a large portion of class time working or discussing problems in small groups. The students are thus involved in the teaching process and focus their attention on underlying concepts.

Peer Instruction lectures uses the ConcepTest format. The ConcepTest has the following general format (Mazur, 1997, p. 10):

- | | |
|---|--------------|
| 1. Question posed | 1 minute |
| 2. Students given time to think | 1 minute |
| 3. Students record individual answers (optional) | 1- 2 minutes |
| 4. Students convince their neighbors (peer instruction) | 1- 2 minutes |

5. Students record revised answers (optional)
6. Feedback to teacher: Tally of answers
7. Explanation of correct answer 2+ minutes

The instructor decides on the basis of the responses whether to move on to the next topic, as when a high percentage of students choose the correct answer, or to provide more explanation, discussion, or a review if the percentage is too low. This approach helps clarify any confusion and provide a more positive effect on learning outcomes (Caldwell, 2007; Mazur, 1997). The convince-your-neighbor step of the ConcepTest increases both the percentage of correct responses and student confidence (Mazur, 1997). Mazur explains that it is easier to change the mind of someone who is wrong than when the person is correct for the right reasons. It is also conjectured that students seem to be able to explain concepts to each other more effectively than their teachers can; similarly, Caldwell (2007) cites studies showing that the more able or knowledgeable students do help those who are less knowledgeable achieve a higher level of understanding.

Mazur (1997) uses the Force Concept Inventory (Hestenes, Wells, & Swackhamer, 1992) and the Mechanics Baseline Test (Hestenes & Wells, 1992) to assess student learning. Both tests assess students' conceptual understanding of Newtonian mechanics. The FCI is a multiple choice test designed to test students' understanding of the Newtonian concept of force as opposed to choices indicating preconceptions or misconceptions. Mazur used this test as a pre-test and a post-test. The Mechanics Baseline (MBT) is a problem solving test administered after instruction. Hestenes and Wells (1992) indicate that the FCI tests students without formal training in mechanics and elicits their misconceptions about the subject; the MBT emphasizes concepts

requiring formal knowledge of mechanics. Using both tests is emphasizing both conceptual understanding and problem-solving skills. Similar approaches in biology courses have shown significant improvement in learning gains over traditional lecture-only approaches (Knight & Wood, 2005).

Mazur (1997) provided evidence that peer learning appears to work. The pre- and post- FCI scores for learning gains increased from 8% in 1990 with the traditional lecture method to 22% in 1995 using Peer Instruction. Similarly, the MBT scores increased from 67% in 1990 with the conventional method to 76% in 1995 using Peer Instruction.

Accountability, preparation, and feedback are essential components of formative assessment (Caldwell, 2007). Peer Instruction, coupled with a clicker system, provided Mazur (1997) with the tools to easily and more readily engage students in frequent formative assessment. Mazur used the clicker system Classtalk (Better Education, Inc) as the interactive response system for his students to answer the ConcepTest questions. The responses were received by the instructor's computer, collated and projected as a class result for the class to view and discuss. One added advantage of the process is the anonymity it confers, a factor that is purported to increase participation and produce an engaged, active learning environment (Trees & Jackson, 2007).

Mazur's (1997) work has set the groundwork for much of the research in physics and other disciplines involving active learning and interactive student response systems. One drawback Mazur recognized with incorporating Peer Instruction strategy is the difficulty of covering the material typically covered in a traditional lecture. The students were therefore made more accountable with assigned reading material to be completed before class, as well as material that may not have been covered in class. He frequently

gave homework problem and exam questions on such topics and gave reading quizzes that counted toward the final grade. Accountability has subsequently been transferred to Web-based assignments (Crouch & Mazur, 2001). This reduces the time required for reading quizzes and provides more time for interactive teaching. Clickers provide an efficient way to hold students accountable for pre-class preparation. It can be used to give the quizzes and to automatically do the grading, thus freeing the instructor of that task. Knight and Wood (2005) highlight the clickers use for brief quizzes on assigned readings or homework to encourage preparation, thus making time available for more productive ways than “coverage”.

Peer learning approaches in physics heavily emphasizes conceptual understanding over numerical problem solving. The latter is essentially removed in Peer Instruction. To correct this Mazur (1997) gives homework assignments and offers problem-solving sessions in order to develop problem-solving skills. Mazur also uses a mixed examination approach in which conceptual essay questions as well as standard textbook problems are used to promote increased emphasis on concepts. Mazur provides a formula sheet on exams to discourage memorization and allow students to focus on the meaning of the equations.

In the area of physiology, Paschal (2002) examined the effectiveness of the traditional approach of lecturing students followed by homework and out of class reading assignments with a wireless communication system. In this approach, homework was eliminated and the “clicker” technology was used to provide instant feedback on classroom activities and for quizzing students on the reading assignment. Paschal reported 100% participation and a beneficial rapid formative assessment of the students.

In addition, the study indicated that the learning of physiology concepts was at least, if not more, effective with the instant feedback mechanism provided by the “clicker” over the traditional teaching techniques, including homework.

Paschal’s (2002) study was conducted over two separate fall semesters with the same instructor. Both groups had identical meeting days and times, and classroom. The instructor was experienced in the content material but was a novice on the use of the wireless communication system. The control group consisted of 69 students taught in the traditional way, whereas the test group consisted of 63 students who received the modified instruction. In both years the study was conducted for the first eight 75-min class period, covering the same content topics, guest lecturer, and reading material. In the test group, however, textbook reading selections were more focused.

The traditional instruction method of Paschal’s (2002) study included lectures with online notes provided via a course Website, three home-work assignments, and reading assignments. The homework assignments were graded by both a graduate teaching assistant (TA) and an undergraduate grader. The TA drafted the answer key which was then revised by the instructor. The graded homework assignment was returned within a week after it was submitted and the answer key was posted online. Homework assignments accounted for 15% of the semester grade.

The modified instruction method included a personal response system (PRS), in-class activities with multiple choice questions, four in-class quizzes on reading assignments similar to the assignments of the control group, and no homework. The PRS gathered, analyzed, and reported the student responses to in-class questions. The feedback was instantaneous. As with the control group, lecture notes were available

online. Quiz answer keys for the test group were also posted online. The quizzes accounted for 14% of the semester grade.

Two exams (A and B) were generated by the instructor with similar content and similar styles. To assess the students' performance, questions were randomly selected (coin toss) from exam A and exam B and allocated between each of the two fall semester exams. This gave two different exams with the same number reasoning, calculations, and memory questions. The exam was given in the ninth period and graded by the instructor using the same criteria for each semester. The exam grade was 20% of the semester grade in each group.

In the test group, the instructor returned to the traditional method of instruction after the exam. This covered six class periods covering different topics of the class via lectures, homework, reading assignments, with no quizzes. At the end, the students were surveyed to determine their completion of the reading assignments and for their preference of instructional style.

After obtaining approval from the local Institutional Review Board, the researcher gathered information regarding student grade point averages and number of credits earned for all but two (one in each group) students for whom the data was not available. This information was deemed necessary to determine whether prior achievement affected performance in the course. The two groups were then compared.

An unpaired (2-sample) *t*-test assuming unequal variances was conducted on the results of the two exams to determine if the mean scores were significantly different in each group. The test group scored higher ($70.1 \pm 10.7\%$) than the control group ($66.6 \pm$

12.6%) on the exam but were not significantly different from each other. The test group, however, had fewer students earning scores less than 50% than did the control group.

The unpaired *t*-test assuming unequal variances determined that there were no significant differences between the two groups in terms of preparedness and prior achievement. The survey results showed that 47.5% of students completed all readings in time for class when quizzes were given compared to 33.3% completion for the module after the modified instruction period when no quizzes were used, and 0% completion for the last module of the class when again no accountability tool was used. The survey on instructional style indicated that a majority of students preferred the combination of reading assignments with quizzes, in-class exercises, and no homework. This was the method used with the test group. The second popular choice was no quizzes on reading assignments with in-class exercises and homework.

Paschal (2002) identifies one confounding effect that may account for the lack of significant difference between the higher scores of test group and the scores of the control group. The test group was tested in the fall semester of 2001, thus the Sept. 11, 2001 terrorist attacks that occurred on the day of one of the class periods and quizzes in the study could have lowered student performance and reduce the significance. In addition, 9/11 memorial activities and walkouts organized by the Student Government Association disrupted another class period in the study. Other potential confounding effects include the instructor's inexperience with the PRS system and a possible bias of the exam design toward the traditional method.

Paschal (2002) discussed the benefits of the PRS system including the instant feedback that helps the instructor and student identify misconceptions that can then be

clarified; the anonymity of the respondents promoted 100% participation; and the active engagement of the students. In the test group the students were made more accountable for the reading and therefore were more prepared in class. After relaxing the accountability, the students were less likely to read the assigned material.

The test group had reduced workload due to no homework assignment and therefore had more time to do the reading assignments. The TA also had reduced workloads. This allowed reassignments to create quizzes, answer keys, and PRS files under the instructors' supervision. This is similar to creating homework assignments and answer keys as in the traditional method. Initially the instructors' workload increased due to the need to create new in-class material, but these are reusable in subsequent classes. The PRS system setup was also an issue in that it uses 5-10 minutes of class time for setup and breakdown. This issue is now addressed with newer models of wireless communication equipment, portable units and laptop computers, as well as permanent installations of the PRS equipment.

Paschal (2002) concludes that the results indicate the use of PRS coupled with in-class quizzes is at least as effective, if not more effective, than traditional learning activities such as lectures and homework in the learning of systems physiology concepts. When a major confounding factor is accounted for, the results using the modified instruction seems more effective.

The study leaves some questions unanswered. There is no indication whether the author is the instructor of the class or whether the instructor is someone else. This information would help clarify the confounding effect of exam or instructional style bias. Another question is that of whether the number of students in each group is a large

enough sample size, and if so there is no analysis presented for effect size or statistical power. The experimental design also puts the results into question since the control and the test groups vary in more than one variable. The control group includes lectures, reading assignments that are not quizzed or graded, and homework that are graded. The test group includes lectures with the PRS component, reading assignments that are quizzed on, and no homework. Paschal (2002) does address that the effect of eliminating the homework cannot be isolated and cites relevant literature that is consistent with the results when concluding that the results show "... essentially equivalent performance when homework was replaced with in-class learning activities and quizzes with instant PRS feedback..." (p. 307). Nevertheless, one can be critical of the experimental design in terms of the multiple variables (interventions) tested in this study.

Preszler, Dawe, Shuster, and Shuster (2007) provide a more thorough study of the effects of student response systems on student learning and attitudes over a broad range of biology courses. Student learning was determined by performance on exams and student attitudes gleaned from a survey of students in six biology courses that included lower-division courses and upper-division courses. Students' performance was compared on exam questions across the spectrum of lectures with low, medium, or high number of in-class questions with the student response system. The researchers tested the prediction that the student exam scores will increase with the increased use of clicker questions in corresponding lectures.

The study was conducted in the fall of 2005 in six courses that used wireless student response units. The response units (clickers) used was the Classroom Performance System (CPS) available from eInstruction (<http://www.einstruction.com/>).

The clicker questions were varied within the courses identified as low, medium, and high frequency use, and the performance on exam questions corresponding to clicker use frequency was compared. Student perception of clicker use in biology courses was surveyed and analyzed for variations among courses, among students who studied individually versus those who studied cooperatively, and students' opinions based on their course grade.

The six biology courses were a diverse set of courses that included four lower-division courses and two upper-division courses. The courses were diverse in their demographic characteristics as well, and included courses with non-majors, issue-based approach, science majors and pre-professional and specific microbiology and biochemistry majors. The instructors of the courses also had varied experience with clicker use, ranging from first time users to prior experience in one course and in three courses prior.

Using the constructivist theory of learning as their theoretical framework, Suchman, Uchiyama, Smith, and Bender (2006) evaluated the impact of a classroom response system (CRS) in a microbiology course and concluded that there was no difference in learning based on the CRS although the scores for the CRS user group were significantly better in questions from both CRS and non-CRS lectures. The experiment used two large sections of a microbiology course, referred to as section A and section B, to evaluate the effectiveness of the technology in increasing student learning, confidence, attendance, and in clarifying student's misconceptions. The CRS was used as a quizzing tool. The questions for both groups were the same, coming from instructor's common

lecture content. The questions were categorized into two groups: some without CRS use, and some from lecture and CRS questions.

The general microbiology course used for this experiment was comprised of a majority of students (80%) being upperclassmen, but of the 30 different majors represented only 10% were microbiology majors. The course required both general biology and second semester chemistry as prerequisites and was taught each semester with a senior faculty responsible for each section. The enrollment was 143 and 84 in section A and section B, respectively.

Both sections used the CRS technology to quiz the class at the beginning of the period on material from the previous class. Section A used the CRS technology to pose questions and collect responses throughout the lecture. This occurred nineteen times throughout the semester and was accompanied by group discussions and the chance to re-answer the question after discussion. The answers had no credit assigned to them, while in Section B the students were quizzed fifteen times throughout the semester without the use of the CRS technology, and the correct responses were assigned an extra credit point. The quiz days were not announced in this group. The extra credit point in Section A was obtained by buying the personal data transmitter (PDT). Both sections were taught on Mondays, Wednesday, and Fridays with Section A taught from 9:00 to 9:50 and Section B from 12:10 to 1:00 p.m.

Suchman et al. (2006) used three different data sources for analysis. This included an instructor's journal in which daily records were kept of the number of CRS questions used, the number of times students engaged in discussions, and the number of times questions were modified by the instructor. The second source was the data collected by

the computer program, and the third was comprised of students' performance on four formal examinations. All the exams contained 10 to 14 content questions that were identical in each section; five to seven of these questions were based on the lecture delivery for each section, and the remaining five to seven questions were based on the different delivery methods of each section, Section A using the CRS technology and group discussions, and Section B a lecture only method. The instructor for Section A wrote the questions for the first two examinations, and the instructor for Section B wrote the remaining questions for the last two examinations; in each case the instructors were in agreement that the questions addressed the course content.

The statistic used to analyze student performance on the examinations was an ANOVA in which the fixed effects were section, exam, and question type; questions with type and exam were random effects. The question type was described as either covered in both classes by lecture or by CRS only as in Section A. The level of significance was set at $p \leq 0.05$. Analysis of perceptions of learning, confidence, and study habits were analyzed using a survey in which students rated statements on their CRS experience on a scale from 1 to 5, with 5 being *strongly agree* and 1 being *not applicable*. The differences in the survey result were analyzed with Mantel-Haenszel chi-square analysis at the $p \leq 0.05$ level of significance. Statistically significant differences between the two sections were obtained in questions dealing with greater interaction and participation, and those dealing with student confidence in their learning, knowledge, and ability to connect with the key concepts of the course. The authors indicate that this agrees with the research and literature involving learning styles, involvement, and engagement as ways to make meaning of new information.

Suchman et al.'s (2006) experiment showed that in Section B, where the CRS was used only for initial quizzes on the previous lecture material, the instructor modified the lecture only once due to responses from the CRS questions, whereas in Section A the instructor modified the lecture plan 84 times. Another result of the experiment evidenced the student interaction increasing when the CRS technology was used throughout the lecture. The authors concluded that the "frequency of CRS use during a lecture appears to impact formative course improvements" (p. 9). In Section A, 95.3% of the students bought the PDT, whereas 87.2% did in Section B. Students in Section A participated by using their PDT 84 times more than students in Section B, and attendance as monitored with the CRS technology was about the same for both sections each day (72.8% versus 74.8%). No attendance information was gathered for students who did not purchase a PDT.

The data on responses in both sections on questions answered at the beginning of the class show very little difference (78.2% versus 82.0% average); however, students in Section A improved their score by 22% after they were allowed to discuss and to re-answer the question. This result, the authors contend, is in agreement with the constructivist theory of learning that purports knowledge is actively constructed and that learning occurs in a social setting.

Students in the lecture section augmented with the CRS technology performed better in both types of examination questions ($p < 0.0001$), but the results were not significantly different in learning based on CRS question participation. The section and exam were statistically different ($p < 0.05$), but between questions types (CRS questions versus non-CRS question) there was no significant difference even when analyzed by

section, exam, or both. These results did not support the authors predictions that Section A students would perform better on CRS questions and both sections would perform the same in non-CRS questions. As a result the authors conclude that the data did not support an increase in learning based on the use of CRS questions. The better performance in all questions by the students in Section A did suggest that the technology, when used as more than a quizzing tool, “may translate to greater learning of those topics not covered by CRS questions” (p. 10).

Suchman et al.’s study does generate some questions. First, the two sections, presumably the control (non-CRS) and the experimental group, have different instructors. This introduces questions regarding confounding effects of teaching styles, personalities, etc. Other factors to be taken into account include the time of day, and/or class demographics. Moreover, the article does not specify if any of the authors were the actual instructors. The authors contend that with both sections having comparable results with questions at the beginning of class, both using the CRS technology, the performance of both sections should have been expected to be the same, rather than the CRS section outperforming the non-CRS section in all questions. In addition, the authors conclude that although no gains in student-learning were obtained, the data supports an expanded use of CRS beyond the quizzing activity to improve students’ learning and confidence.

CHAPTER III

METHODOLOGY

This study investigated the interaction in the college classroom in a quantitative manner. The quantitative design employed a pretest-posttest, non-equivalent groups experimental design. The participants were students in six biology sections into which they self-selected. The sections were of adequate size and representative diversity. The learning gains were assessed using conceptual questions from concept inventories, administered on a pre-test and post-test basis. An ANCOVA statistic was used for analyzing the data.

Experimental Design

Trochim (2002) described the non-equivalent groups design (NEGD) as the most frequently employed design in social science research. Garcia (2006) cited the appropriate literature supporting the use of NEGD's, namely the Jackard and Becker study of 1997, as similar to the pretest-posttest randomized experiment without the randomized feature. The choice of groups designed as treatment and control groups are assumed to be similar. This is achieved by selecting two comparable classes (Trochim, 2002) that exhibit as many similarities as possible for valid comparisons between treatment and control group. The best comparison, nevertheless, is still obtained from completely random assignment (Jackard & Becker, 1997). Complete randomization was difficult in this study because the subjects tended to self select through course-time and teacher shopping. The self-selection bias was controlled by offering several sections of biology for science majors (BSC1010) lectures, some using the Classroom Performance System (CPS) and others without. The instructors were experienced biology instructors at

the research site. The instructors received training in the use of the CPS technology. The Instructional Technology (IT) department of the institution provided training in the use of the CPS technology which provides all the skills necessary for the instructors to be able to effectively use the technology in class. The training included software download and installation, instrument activation, question construction, recording of responses, data storage and retrieval, and accessing reports. The instructors each used the CPS technology for a semester in two different classes each as part of the training.

Data Collection

The biology concepts inventory exam was administered to CPS and non-CPS groups on a pre- and post-course basis. As an incentive to do their best on the pretest, the participants were advised that the score on the pretest would be applied to any low grade obtained in the unit tests of the course. In addition, encouraging the best efforts on the pre-test was enhanced by having the instructors address the participants and explain that because of the difficulty students have in the course, the faculty, staff, and institution were interested in identifying instructional tools for improvement (Garcia, 2006).

Subjects

The participants of the study were students in six sections of biology that self-enrolled into the sections and that were taught by two instructors who agreed to cooperate in the research. One instructor taught four sections, two of which were randomly selected by a coin toss as treatment groups, those sections using the CPS technology, and two as the control group, those sections not using the CPS technology. A second instructor taught two separate sections. This instructor also selected the treatment group in a random draw using a coin toss, and by default the second class became the control group.

The population from which the subjects were obtained was students enrolled in sections of a biology lecture course for science majors at a large 2-year college in South Florida. The college offers hundreds of career and technical programs leading to associate's degrees, bachelor's degrees, and advanced certification. It serves transient students from a college or university, offers continuing education courses, and fast track career and technical programs for jobs in a variety of fields. As a stepping stone, the college boasts over 100,000 graduates who have gone on to professional students in fields such as law, health, education, as well as have community and business leaders. The college has a high enrollment (over 67,000 for the 2010 – 2011 academic year), is highly ranked, affordable, and conveniently accessible via three campuses and several learning centers. Its students are from 150 different countries, reflecting the demographics of South Florida. Students are also diverse in age, from first time in college high school graduates, to adult students returning for a career change or to update credentials and develop new or refresh old skills. The population is known to be racially and ethnically diverse, as well as of a wide age range

Procedure

The following procedure was discussed with the instructors participating in the study and agreed upon for consistency. In addition the instructors agreed to have the researcher visit their classes at random and without prior notice to assess fidelity of the agreed procedure. This was done twice for each of the six classes throughout the semester.

The CPS formatted sections had approximately 200 “clicker questions” authored by the instructors that covered the major topics in the chapter summaries at the end of

each chapter of the text. The students were asked to respond to 8 -10 questions each class. The first question at the start time of the class served the additional purpose of taking attendance. Midway through the class another 6 - 8 questions were asked, and the remaining 1 - 3 questions are asked at the end of class. Following each question the class performance on the question was displayed and discussion ensued if less than 70% of the students got the correct answer. The students earned 2 points for each correct response, 1 point for an incorrect response and 0 points for unanswered questions, whether the latter were due to absence, tardiness, or inattention. The first two classes were used to practice with the clickers (10 to 12 questions each) before starting to record the grades. The 20 – 24 points were used to compensate for one or two “excused” class absences that may occur during the semester. The daily questions were printable and made available to be used as a study guide for the unit exams.

The non-CPS classes took attendance manually and were given daily quizzes of 10 points each administered at random during class on material that had already been covered. The quizzes were from the same set as those used for the CPS classes. Two quizzes were dropped or used to replace one or two “excused” absences. The quizzes were made available for use as a study guide for the unit exams. This was designed to keep the two classes as similar as possible and reduce double-exposure or time-on-task concerns as the source of improvement in student achievement.

Instrumentation

The independent variable in the study was the use of the CPS technology to deliver instruction in the biology classes. The dependent variable was the achievement in learning gains as measured with a concept inventory consisting of 49 questions. The

concept inventory questions were administered on a pretest and posttest basis. The pretest was the covariate in the data.

Concept inventories. Usually individual instructors make up their own final exam questions. This results in some variation and runs the risk of the instructor teaching to the questions. Different kinds of learning assessments are available in the form of questions that are multiple choice, matching, fill-in-the-blank, true-false, and essays, as well as interviews, surveys, and standardized concept inventories. The correlation of averaged final exam scores has been used to justify validity by some conceptual instruments (Smith et al., 2008) but this is problematic in that it is easy to teach to a test once the particulars of the test have been established (Klymkowsky, Underwood, & Garvin-Doxas, 2010). The relative value of each of the formative and summative learning assessment mentioned is still contested (Anderson, Fisher, & Norman, 2002; Brewer, 2004; Odom & Barrow, 1995; Udovic, Morris, Dickman, Postlethwait, & Wetherwax, 2002). Several biology educators have developed, or are in the process of developing, concept inventories that help assess learning gains, namely the Biology Concept Inventory (BCI; Garvin-Doxas, Klymkowsky, & Elrod, 2007; Garvin-Doxas & Klymkowsky, 2008; Klymkowsky, Underwood, & Garvin-Doxas, 2010), the Genetics Concept Inventory (Elrod, 2007), the Genetics Concept Assessment (Smith, Wood, and Knight, 2008), Introduction to Molecular and Cell Biology Concept Assessment (IMCA; Shi, 2010), and the Diffusion and Osmosis Diagnostic Test (Odom & Barrow, 1995). As suggested by Libarkin (2008), the study used questions for analysis that have been tested for validity and/or reliability and that focus on diagnosis and assessment. Klymkowsky et al. (2010) established face validity of the BCI questions through student interviews and

content validity through discussion of the questions with subject matter experts, but does not provide any reliability information, whereas Odom & Barrow (1995) determined that their diagnostic test had a Pearson r reliability of .74 based on the Spearman-Brown formula. The instrument, used with permission, was the 49 questions that make up the molecular and cellular approach of the BCI (Klymkowsky, et al., 2010) and diagnostic test of Odom & Barrow (1995). Together these cover the material of the biology course for science majors. As pretest and posttest measures for learning gains (Hake, 1998), the questions can be used to assess student learning in a variety of ways using different modes of instruction (Smith et al., 2008). The impact of the independent variable can be determined from the distribution of learning gains of each student (Knight & Wood, 2005).

Quantitative Data Analysis

The study used a pre-test/post-test format in which the post test was the dependent variable measuring how achievement is influenced by the instructional method. The pre-test was used as the covariate for the data analysis. Therefore the statistic that was used is the analysis of covariance (ANCOVA). The ANCOVA is more commonly used to control extraneous variables and as a means of increasing power (Gay & Airasian, 2000). Examples of variables controlled by using ANCOVA include pretest performance (Gay & Airasian, 2000). Possible sources of variation within the participants that may influence the post-test scores include subjects having recently had a biology class, technology phobia, or adeptness. By using covariance it is possible to reduce the variation in the post-test scores caused by these variables. This is desirable in order to increase the confidence that any achievement gains shown in the posttest are due to the

instructional treatment. In this study every attempt was made to make the samples as similar as possible, but randomization does not guarantee that groups were equated on all variables (Gay & Airasian, 2000). In this study the groups were assigned to treatment groups by each instructor performing a random draw, but the participants may not randomly select the group. The ANCOVA may still be used in this quasi-experimental design, but with data interpretation done more cautiously (Gay & Airasian, 2000).

Any variance within-groups can also be reduced by the ANCOVA, thus increasing the power of the test (Gay & Airasian, 2000). Statistical power relates to the ability to reject a false hypothesis, or the ability to avoid a Type II error (β). If the study has little statistical power, thus increasing the chance of committing a Type II error, it is possible that any achievement gains from the instructional treatment will be missed. In this case a valuable tool for increasing learning may be discarded when it should not be. Increasing sample size also increases power, but the Biology classes are capped at 40, and the possibility of attrition exists, thus the ANCOVA was useful here. By accounting for variation in different groups the ANCOVA decreases sample error and increases the power of the test.

The data were analyzed at the significance level (α) of .05 in a one-tailed test. The one-tailed test was chosen because the results from the treatment group are expected to be better than those of the control group (Paschal, 2002). The α level of .05 represents the probability of rejecting a true null hypothesis, or finding a difference in the treatments when there is in reality none (Type I error). It represents the criterion for the test and tells the researcher where to start rejecting the null hypothesis. Making a Type I error is considered to be more serious than a Type II, and is therefore more important to avoid. In

this study such a mistake could validate CPS technology as an instructional tool and may be dangerously misleading if they are relied upon to influence institutions to unnecessarily spend the scarce funds they have on equipment, support, training, and personnel.

It stands to reason that increasing power is in the best interest of the stakeholders. Decreasing α for a more stringent test, say to .01, makes it more difficult to reject the null hypothesis and therefore more likely to make a Type II error. Since this is not a case of life and death as in, for example, medical or pharmaceutical tests, $\alpha = .05$ is adequate. Using a one-tailed test also increases power. Increasing effect size can also increase power. In this case the medium effect size of .50 was used. With a sample class size of 40, an α of .05, and a one-tailed test, this effect size gives a power of between .90 and .95 (Hinkle, et al., 2003, table C.11). Decreasing variance can be done by increasing sample size or using better measures. The sample class size was set at 40 in this study, which leaves the ANCOVA as the testing tool. The SPSS computer program was used to do the ANCOVA.

All necessary paperwork and permissions as dictated by IRB policies of Florida International University were obtained and all ethical concerns, such as anonymity and confidentiality were addressed.

CHAPTER IV

RESULTS

This chapter presents the results of the study. The study examined the efficacy of one audience response system (ARS), the Classroom Performance System (CPS or “Clicker”) technology, in teaching Biology for Majors to community college students and to assess the impact on these students’ achievement in exam scores measured as learning gains. This chapter provides a description of the sample, and the tests associated with the corresponding hypothesis.

Description of the Sample

The original sample consisted of 226 participants in six sections of a Spring 2012 college biology course at a large community college in South Florida. The final sample consisted of 172 study participants, comprised of 41.9% Hispanic, 24.4% Black, non-Hispanic, 24.4% White, non-Hispanic, 7.6% Asian or Pacific Islander, and 1.7% as other.

Data on ethnicity and other demographic characteristics, such as GPA, sex, academic performance in science-based courses, and prior experience with the Classroom Performance System (CPS) were collected using a questionnaire. Sixty-one percent (61%) of the participants reported they were sophomores and 20.9% were juniors. Ninety-four point two percent (94.2%) of all study participants were in the 18 to 30 age range.

An attrition of 54 participants constituted a 23.9% decrease in the number of participants who completed the course and the study. Although 89% of the participants reported that the course was a required course for their program of study, this attrition is not atypical for such a course. Table 1 provides the distribution of study participants by

sex for the Control and Treatment groups at the time of administration of the pretest and posttest. In both groups there was an attrition rate of 22% for males, whereas for the females the control had a 23% attrition, and the treatment group showed a 28% attrition.

Table 1

Distribution of Study Participants by Sex

| Sex | Control | | Treatment | |
|--------|---------|----------|-----------|----------|
| | Pretest | Posttest | Pretest | Posttest |
| Female | 73 | 56 | 67 | 50 |
| Male | 36 | 28 | 49 | 38 |
| Total | 110 | 84 | 116 | 88 |

Research Question

In this study, the research question under investigation was:

Will students who participate in a Classroom Performance System (“Clicker” technology) format in a biology course for majors have higher achievement scores measured as learning gains than students in the same course that is not formatted for delivery with Classroom Performance System?

To answer the research question, a sequential, multi-step approach to analyses was conducted. This included measurement of learning gains and performance of statistical analyses involving an independent- samples *t*-test and analysis of covariance (ANCOVA).

Learning Gains. To provide an estimate of the average effectiveness of a course in promoting conceptual understanding, the average normalized gain for the control and treatment groups was calculated using the following, as proposed by Hake (1998):

$$\frac{\text{Posttest score} - \text{Pretest score}}{1 - \text{Pretest score}}$$

Examination of the average normalized gains for the treatment and control groups indicate performance of the treatment group of 0.32 units above the performance of the control group (Table 2).

Table 2

Normalized Gains

| Control | Treatment | Difference |
|---------|-----------|------------|
| 2.12 | 2.44 | 0.32 |

Independent-samples *t*-test. To determine whether difference in observed normal gains between the experimental and control groups is statistically significant, an analysis involving an independent-samples *t*-test was conducted to evaluate differences in pretest scores for the control and experimental groups. In a similar study on audience response systems, Conoley (2005) utilized this statistical test to examine differences between groups of high school students on their achievement in agriscience courses. Huck (2012) outlined the theoretical implications and general procedures for conducting an independent-samples *t*-test.

The *t*-test was not significant at the .05 level of significance, $t(170) = -1.813$, $p = .07$, suggesting that there was no significant difference on the posttest scores for the study participants assigned to the control and experimental groups.

Analysis of Covariance (ANCOVA). Since the study design did not include random assignment of participants to groups, it was very likely that the treatment and control groups in the study were non-equivalent on the important variables of interest.

An analysis of covariance (ANCOVA) is the appropriate statistical test to conduct to account for other initial differences between groups which cannot be accounted for by an independent-samples *t*-test; the ANCOVA design is a more robust test for instances in which random assignment to groups was not possible and increases the statistical power of the study (Huck, 2012).

Three one-way ANCOVA were conducted. The purpose of the ANCOVA design was to control statistically for initial differences in the study participants which may have been present and which can confound posttest differences between the two groups.

The independent variable is the instruction delivery mode consisting of two levels: CPS or non-CPS. In the first ANCOVA, the dependent variable is the posttest scores and the covariate is the pretest scores. The second ANCOVA included the following additional covariates: sex, GPA, status, and experience with clickers (CPS).

Before conducting the ANCOVA, the homogeneity-of-slopes assumption was tested; the tested null hypothesis is that the population slopes are homogeneous. Evaluation of the homogeneity-of-slopes assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly at the .05 level as a function of the independent variable, $F(1, 170) = 1.532, p = .217$. The interaction between the covariate and the factor indicate that population slopes do not differ, $F(1, 171) = .009, p = .926, \text{partial } \eta^2 < .001$. Based on this test, the interaction between the covariate and the factor was removed from the analysis.

The results of the first ANCOVA was non-significant, $F(1, 169) = .240, \text{MSE} = 3.708, p = .63$. The strength of relationship between the instructional delivery mode and the dependent variable is weak, as assessed by a partial η^2 , with instructional delivery

mode accounting for .1% of the variance of the dependent variable, holding constant study participants' pretest scores. Table 3 presents the source table for the ANCOVA.

Table 3

Summary of Analysis of Covariance with Pretest as the Covariate

| Source | Sum of Squares | df | Mean Square | F | Sig. |
|----------------|----------------|-----|-------------|--------|-------|
| Between groups | 3.71 | 1 | 3.71 | .24 | .63 |
| Covariate | 1652.27 | 1 | 1652.27 | 107.14 | <.001 |
| Within groups | 2606.16 | 169 | 15.42 | | |
| Total | 4262.14 | 171 | | | |

Figure 3 and Table 4 illustrate the unadjusted and adjusted means for the experimental and control groups in the study.

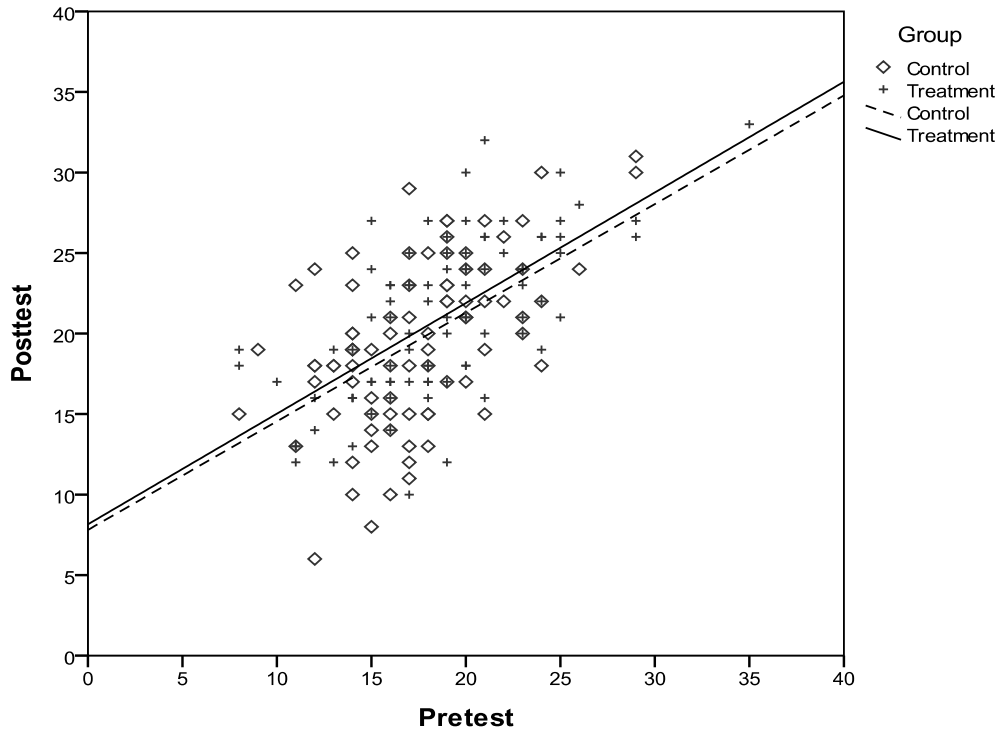


Figure 3. Differences in slopes and between estimated group means on posttest scores for values on the covariate.

Table 4

Mean Pretest Scores and Unadjusted and Adjusted Mean Posttest Scores

| Group | Pretest | | | Posttest | | |
|-----------|---------|-------|------|-----------------------|------|---------------------|
| | N | M | SD | Unadjusted Means M | SD | Adjusted Means M |
| Control | 84 | 16.17 | 4.03 | 18.51 | 5.07 | 18.95 |
| Treatment | 88 | 17.36 | 4.60 | 19.67 | 5.00 | 19.25 |

A second ANCOVA was conducted in which pretest scores, sex, GPA, status, and experience with clickers (CPS) served as covariates. The homogeneity-of-slopes assumption indicated that the relationship between the covariates and the dependent variable did not differ significantly at .05 level as a function of the independent variable, $F(1, 170) = 3.651, p = .058$. The interaction between the covariate and the factor indicate that population slopes do not differ, $F(2, 134) = .252, MSE = 4.352, p = .777, \text{partial } \eta^2 = .004$.

The result of the ANCOVA was not significant, $F(1, 134) = 1.8, MSE = 31.06, p = .18$. Inclusion of the covariates – sex, GPA, status, and experience with clickers – account for an additional .9% of the variance of the dependent variable, holding constant study participants' pretest scores. The source table for the ANCOVA is presented as Table 5.

Table 5

Summary of Analysis of Covariance with Other Variables as Covariates

| Source | Sum of Squares | df | Mean Square | F | Sig. |
|----------------|----------------|-----|-------------|-----|-------|
| Between groups | 31.06 | 1 | 31.06 | 1.8 | .18 |
| Covariates | 2124.09 | 36 | 59 | 3.4 | <.001 |
| Within groups | 2312.08 | 134 | 17.25 | | |
| Total | 4467.23 | 171 | | | |

Note. Pretest, sex, GPA, and Experience with Clickers are the other covariates.

A third ANCOVA was conducted in which ethnicity served as a covariate.

The homogeneity-of-slopes assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly at .05 level as a function of the independent variable, $F(1, 168) = .022, p = .881$. The interaction between the covariate and the factor indicate that population slopes do not differ, $F(1, 169) < .001$, $MSE = .002, p = .993$, partial $\eta^2 < .001$. As a result of this test, the interaction between the covariate and the factor was removed from the analysis.

The result of the ANCOVA was not significant, $F(1, 167) = 2.77, MSE = 68.76, p = .098$. Inclusion of ethnicity as a covariate in the analysis accounted for 1.6% of the variance in the dependent variable. The source table for the ANCOVA is presented as Table 6.

Table 6

Summary of Analysis of Covariance with Ethnicity as Covariate

| Source | Sum of Squares | df | Mean Square | F | Sig. |
|----------------|----------------|-----|-------------|------|------|
| Between groups | 68.76 | 1 | 68.76 | 2.77 | .098 |
| Covariate | 21.87 | 1 | 21.87 | .88 | .005 |
| Within groups | 4143.50 | 167 | 24.81 | | |
| Total | 4213.14 | 169 | | | |

Ancillary Analyses

Two independent-samples *t*-test were conducted to evaluate differences in learning gains between groups in the study taught by different course instructors.

For participants in the treatment and control groups taught by the first professor, the test was non-significant, $t(58) = -0.094$, $p = .925$, and the results were counter to the research hypothesis. Participants in the control group ($M = 19.57$, $SD = 5.77$) on the average scored slightly better than participants in the treatment group ($M = 19.44$, $SD = 5.27$) on the posttest. The 95% confidence interval for the difference in means ranged from -2.99 to 2.72.

For participants in the treatment and control groups taught by the second professor, the test was significant, $t(110) = 2.469$, $p = .015$. Participants in the treatment group ($M = 21.75$, $SD = 4.59$) on the average scored 2.33 units better than participants in the control group ($M = 19.53$, $SD = 4.97$). The 95% confidence interval for the difference in means ranged from .44 to 4.02. For this group of participants, an analysis of covariance (ANCOVA) was then implemented to statistically control for differences in scores on the pretest.

Evaluation of the homogeneity-of-slopes assumption indicated that the relationship between the covariate and the dependent variable did not differ significantly at the .05 level as a function of the independent variable, $F(1, 110) = 2.238, p = .137$. The interaction between the covariate and the factor indicate that population slopes do not differ, $F(1, 108) = .115, p = .735$, partial $\eta^2 < .001$. Based on this test, the interaction between the covariate and the factor was removed from the analysis.

The results of the ANCOVA were nonsignificant, $F(1, 109) = 2.73, MSE = 40.84, p = .101$. The strength of relationship between the instructional delivery mode and the dependent variable is weak, as assessed by a partial η^2 , with instructional delivery mode accounting for 2.4% of the variance of the dependent variable, holding constant study participants' pretest scores. The source table for the ANCOVA is presented as Table 7. Although non-significant results were obtained for all the analyses, classroom observations throughout the study showed little difference in attendance between the six classes. Although the instructors had markedly different styles, there wasn't much difference observed in the course content covered neither by instructor nor by study group. The CPS groups seemed more engaged than the non-CPS groups and exhibited a more competitive atmosphere. It was noticed that much more time was used in analysis of topics and question discussions in classes with the CPS than with the non-CPS. No problems with the clickers were observed.

Table 7

Summary of Analysis of Covariance with Instructional Delivery Mode as the Covariate

| Source | Sum of Squares | df | Mean Square | F | Sig. |
|----------------|----------------|-----|-------------|-------|-------|
| Between groups | 40.84 | 1 | 40.84 | 2.73 | .101 |
| Covariate | 887.68 | 1 | 887.68 | 59.40 | <.001 |
| Within groups | 1628.81 | 109 | 14.94 | | |
| Total | 2557.33 | 111 | | | |

Summary

The study had a 23.9% attrition rate. Despite the fact that the course is a required course for the majority of study participants, it is not uncommon for this course to have such a percentage of students unable to complete the course. A comparison of the distributions of participants in the treatment and control groups indicate similar, general composition by gender of the groups. Institutional data gathered at the study site indicate that about 40% of students enrolled in biology for science majors classes, and that are taught by full time instructors, do not persist to the next level, either through poor performance or through attrition.

Preliminary analyses of performance by group on the posttest, represented by examination of average learning gains, indicate that participants in the treatment group performed slightly better than those in the control group. Further examination, however, revealed that the average performance of participants in the treatment group was not statistically different than participants in the control group, as provided by an independent-samples *t*-test. Further analyses using analysis of covariance (ANCOVA) models, in which the pretest and other variables of interest served as covariates, does not support the hypothesis that participants exposed to Classroom Performance System (CPS

or “Clicker”) technology have higher achievement scores than students in the same course that is not formatted for delivery with Classroom Performance System technology.

Ancillary analyses involving treatment and control groups allocated by course instructor provide no evidence of a relationship between participants’ performance and teaching approaches. While for one professor, an independent-samples *t*-test revealed a significant difference between the treatment and control groups, this statistically significant result becomes nonsignificant when pretest scores were controlled in an Analysis of Covariance (ANCOVA) model.

CHAPTER V

DISCUSSION

This chapter summarizes the major findings of the study, draws appropriate conclusions based on results, discusses implications of study results, and provide direction for future research.

Discussion of the Findings

In light of the current trends in education to facilitate learning in the classroom through the use of innovative technology, in line with learners' ease of access to and familiarity with technological advances, it has become pivotal to evaluate the efficacy of classroom technological tools that have been promoted to advance student learning.

This study examined whether participants exposed to Classroom Performance System (CPS or "Clicker") technology have higher achievement scores than students in a different class of the same course but which is not formatted for delivery with Classroom Performance System technology.

The study implemented a non-equivalent control group design. Six sections of a biology course were assigned to either the treatment or control groups, with two course instructors each of whom taught sections of the treatment and control groups. Such assignment to each of the groups was fitting as study participants had already registered for the different sections of the course. An examination of the characteristics, such as sex, GPA, and ethnicity of the study participants who failed to complete the course and considered as having dropped out of the study to those who completed the course indicated no differences that caused the groups to become dissimilar in characteristics of

interest. Differences between participants in the treatment and control groups were statistically controlled in the analyses.

An initial analysis using average normalized learning gains, as suggested by Hake (1998), indicated that the achievement scores for study participants in the treatment group were 0.32 units larger than for those for study participants in the control group. While it may be concluded that such gain in achievement can be attributed to the implementation of the CPS (or “Clicker”) technology in the treatment group, above and beyond non-implementation of such technology as is the case in the control group, results from an independent-samples *t*-test indicate that such a difference is non-significant. This leads to the conclusion that the data did not support an increase in learning based on the use of the CPS technology. This is in agreement with Suchman et al. (2006) and Paschal (2002) who both found positive but no significant effect of clickers on performance on exams in microbiology and physiology courses respectively. Knight and Wood (2005), however, used a normalized gains analysis in their study on peer instruction (PI) and reported an increased learning with the use of clickers. It should be noted that the group that showed the increase also had an additional activity involving concept mapping. Knight and Wood’s (2005) experiment was critiqued for not having ideal controls and this leads to the question whether other confounding effects like time-on-task could be influencing the results.

In the present study, further results from an Analysis of Covariance (ANCOVA), in which performance on a pretest was controlled, supported non-significant results for the efficacy of CPS in promoting achievement scores. CPS does not account for a significant increase in achievement scores beyond a traditional approach to learning.

Ancillary analyses to determine if teaching style accounted for differences in achievement in the treatment and control groups provided non-significant results. Independent samples t-tests for the groups delineated by course instructor showed significant results for one group but not for the other. However, after controlling for differences in performance on the pretest, results were nonsignificant. Teaching style did not influence the efficacy of CPR technology in promoting student learning.

Unlike studies conducted by Abrahamson (1998), Conoley (2005), Dufresne et al. (1996), and Horowitz (1988), the null hypothesis in this study failed to be rejected. After a semester of participating in sections of 2-year college biology for science majors course incorporating the Classroom Performance System (CPS) technology, students did not have difference in examination grades than students in sections that do not incorporate the CPS technology.

Strengths of the Study

CPS technology has been promoted as a tool to enhance student learning because of the level of interaction that can occur between the teacher and students in providing immediate feedback about the degree of comprehension about the discussed concepts and ideas. Based on students' responses, teachers can make immediate classroom decisions about continuation of delivery of content or review or make changes to the present delivery to increase students' probability of acquiring the necessary knowledge; teaching can be immediately adapted to accommodate learning and students can build on prior knowledge by exploring the information in the CPS questions, explaining the concepts via discussions or by the instructor, and applying the information in new activities. This

conforms to a learner-centered approach (Barr & Tagg, 1995), grounded in a constructivist framework.

The research design included several strengths:

Fidelity of treatment. The treatment was implemented as designed. Prior to implementation of the treatment, the researcher trained the course instructors in proper implementation of CPS technology in the treatment groups and non-implementation in the control groups. Class observations during the study ensured that the groups received the designated treatments. In addition, participants' prior knowledge on use of CPS was statistically controlled for in the analysis.

Lack of differential attrition. Examination of the characteristics of the study participants who drop out of the study to those who remained indicated no differences that caused the groups to become dissimilar in characteristics of interest.

Lack of interaction of selection and maturation. Since study participants in the treatment and control groups were of the same age group and the course was a requirement for the majority of them, the lack of significant results in the study cannot be attributed to different rates of maturation.

Power analysis. An a priori power analysis, in which there was a .95 probability of correctly rejecting the null hypothesis when it is false, was conducted. Hence, an adequate sample size for the study to ensure that the null hypothesis could be appropriately rejected when false was determined. This study had sufficient power to reject a false null hypothesis.

Limitations of the Study

Studies that implement quasi-experimental research designs, as the present study on CPS technology and student achievement, suffer from a number of threats to internal validity – that CPS technology is the only cause in higher student achievement, having ruled out alternative explanations for such causality.

The primary issue with quasi-experimental studies such as this one is the lack of random assignment to the treatment and control groups, which result in varying degrees of non-equivalence in the groups on the variables of interest (Shadish, Cook, and Campbell, 2002). Despite the ability to statistically control for initial differences between the groups, it is virtually impossible to know which variables are important variables, collect data on them, and control for those differences in the analysis. As a result, such studies usually provide inconclusive evidence about the true relationship between the independent and dependent variables.

The present study may have several potential validity issues:

Selection bias. Initial group differences, not statistically controlled by the ANCOVA models, may account for the lack of significant differences in the dependent variable - students' achievement scores. Since study participants self-selected into the six sections of the course and each of the sections were conveniently designated as either the treatment or the control groups, it is likely that differences on other variables, not accounted for in the analysis, can account for lack of differences in achievement scores.

Treatment strength. Despite the researcher's maintenance of fidelity of the treatment, the duration, frequency, and intensity of the treatment may have been provided at inadequate levels to positively influence differences between groups. Established

treatment protocols to ensure strength and intensity of the treatment, stipulating the precise conditions under which the treatment should be delivered, may need to be revisited and adapted. One such protocol would address the role of active learning exercises in the study. Freeman et al. (2007) utilized such a protocol in which the active learning included case studies, think/pair/share, reviewing old exam questions, and discussions. Different experimental designs, namely clicker + online practice, clickers + study group practice, cards + online practice, and cards + study group practice were compared. The results were positive in that a reduced failure rate and statistically significant increases in exam scores were observed. An ANCOVA statistic with the predicted grade as a covariate did not show any significant difference in terms of boosting achievement. Other studies in which increased learning was reported included protocols that required additional work such as concept mapping (Knight & Wood, 2005), assigned readings (Burnstein & Lederman, 2001; Mazur, 1997; Paschal, 2002), and/or homework (Mazur, 1997; Paschal, 2002). However, when implementing protocols like these, care should be taken to account for the effect of time-on-task, something none of these studies addressed. Mazur (1997) and Paschal (2002) also utilized longitudinal studies comparing results over several years or multiple semesters. Except for Mazur (1997), who used the Force Concept Inventory (FCI) in physics, these studies used instructor generated exams to assess learning gains and in longitudinal studies these run the risk of teaching to the test.

Implications of the Study

This study contributes to the extant literature on the efficacy of Audience Response Systems to promote student learning. Following Conoley's (2005) advice to

expand research on Audience Response Systems by replication in different classroom environments, this study sought to examine the efficacy of such systems in a college biology course environment. The lack of statistically significant results in the study, in contrast to results from other studies, suggests that implementation of the treatment needs to occur under specific circumstances conducive to obtaining significant results. This suggests that these specific circumstances, not generalizable to broader settings, may impede the acceptance of Audience Response Systems as a technological tool for enhancing student learning. The technology is generally used to enhance teaching and learning. The Audience Response System promised to provide a means to validate and document student learning. Real-time feedback, stored students' responses for analysis, and a focus on conceptual instruction are practical ways in which the technology enhances teaching. The empirical effects of this technology on student learning are still unclear and the study adds to the body of research that indicate that it does not provide any significant learning advantage over traditional course delivery or low tech flash cards (Freeman et al., 2007; Lasry, 2008). Lasry (2008) states it more succinctly as: "The pedagogy is not the technology by itself (p. 244)." In some measure these results may lead to questioning whether clicker technology adheres truly to the constructivist framework as proposed throughout the literature. Future studies may be needed that take a closer look at whether the classroom interactivity resulting from the clicker technology is really constructivist.

The implications extend to higher education administrators, faculty, and students. This study adds to the body of research that puts into question the efficacy of ARS technology and costs associated with adopting it. The capital expense necessary to

purchase the equipment – clickers and related hardware – may not be available, and requiring the students to make the financial commitment may not be possible or desirable. It is wise, then, for administrators to take into consideration whether or not ARS increases student outcomes, as well as consider the availability of other options for formative assessment. Cheaper options using Smartphone capabilities, such as SMS texting, are presently available, e.g. Polleverywhere.com. As new technologies develop and the options to administrators, faculty, and students become more diverse, it is important that the technologies be evaluated to determine if perceived benefits outweigh the costs.

Directions for Future Research

Despite the lack of significant results in this study, the evaluation of the theoretical concept associated with CPS technology in other settings, in different academic courses with participants from different age groups should be considered for further studies. Of particular interest would be the implementation of a similar study as this one using younger students as seen in Conoley's (2005) study consisting of a sample of 61 high school students. Does the need to work by the older 2-yr college students affects the results as opposed to the younger high school students who might not have a need to work and can better focus on their studies? Can a difference in performance between the two age groups be attributed to a difference in comfort level with technology between the two?

Future research studies are recommended that address the relationship of the pedagogy and the technology. As suggested by Lasry's (2008) fundamental statement on the pedagogy not being the technology, such research can study the differences in using

review questions, those that merely have students recall material recently learned, as opposed to thinking questions. The latter may involve questions that require critical analysis, synthesis, or application of the material.

Another consideration for future research may involve topic specific studies. This study utilized a concept inventory of questions that cover a range of topics with the pretest and posttest assessment being several months apart. Campbell (2008) reported no significant differences between the use of an Audience Response System and conventional groups on retention and transfer exam questions. Since the concept inventory questions per se are not addressed in ARS questions, retention and transfer may be a problem after several months. In topic specific studies, pretests can be administered at the beginning of a topic unit and the posttest at the end of the unit, which can be a time span of only a few weeks. Topic specific concept inventories that have been developed, or are in the process of being developed, include the Genetics Concept Inventory (Elrod, 2007), the Genetics Concept Assessment (Smith, Wood, & Knight, 2008), and the Introduction to Molecular and Cell Biology Concept Assessment (IMCA; Shi et al., 2010).

Summary

As classroom technological tools become available and promoted as tools to enhance student learning, it becomes important to evaluate their efficacy in the task they are marketed for. Classroom Performance Systems, based on Audience Response Systems technology, is such one tool.

The results of this study provide information that CPS technology may have the potential to influence student learning, as indicated by a small increase in learning gains

than in non-CPS environment, but such gains are not statistically large enough to suggest that such technology merits general classroom adaptation. Implementation of such systems may be limited to specific circumstances or conditions and age groups, as suggested by Conoley (2005); further research is needed to determine those specific conditions or circumstances under which CPS technology can positively influence student learning.

REFERENCES

- Abrahamson, A. L. (1998, June). *An overview of teaching and learning research with classroom communication systems*. Paper presented at the Samos International Conference on the Teaching of Mathematics, Village of Pythagorion, Samos, Greece.
- Anderson, D., Fisher, K., & Norman, G. (2002). Development of the conceptual inventory of natural selection. *Journal of Research in Science Teaching, 39*, 953-978.
- Allison, R. D., & Scott, D. D. (1998). Faculty compensation and obligation: The necessity of a new approach triggered by technology integration. Kamala Anandam (Ed.), *New Directions for Community Colleges, 101*, 69-78.
- American Association of Community Colleges. (2009). *Community college trends and statistics*. Retrieved from <http://www.aacc.nche.edu/AboutCC/Trends/Pages/default.aspx>
- Barber, M., & Njus, D. (2007). Clicker evolution: Seeking intelligent design. *CBE Life Science Education 6*(1) 1-8.
- Barr, R. B., & Tagg, J. (1995). From teaching to learning: A new paradigm for undergraduate education. *Change, 27*(6), 12-25.
- Brewer, C. A. (2004). Near real-time assessment of student learning and understanding in biology courses. *BioScience, 54*, 1034 -1039.
- Brown, S. (2000). The 21st century classroom. *Techniques (Association for Career and Technical Education, 75*(7), 22 -25.
- Burnstein, R. A., & Lederman, L. M. (2001). Using wireless keypads in lecture classes. *The Physics Teacher, 39*, 8-11.
- Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE Life Sciences Education 6*(1), 9-20.
- Campbell, J. E. (2008). Increasing learning in college lecture classes: Is it just a click away? *Dissertations Abstracts International, B 68*(10) AAT 3285821.
- Caron, P. L., & Gely, R. (2004). Taking back the law school classroom: Using technology to foster active student learning. *Journal of Legal Education, 54*, 551-554.

- Carvin, A. (2006). EdWeb: Exploring technology and school reform. Retrieved from <http://www.edwebproject.org/constructivism.dewey.html>.
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53, 445-459.
- Classtalk Classroom Communication System (CCS). (2007). <http://www.bedu.com/classtalk.html>
- Conoley, J. W. (2005). *Impacts of an audience response system on student achievement in high school agriscience courses*. (Doctoral dissertation, North Carolina State University, Raleigh, NC). Retrieved from <http://repository.lib.ncsu.edu/ir/bitstream/1840.16/5468/1/etd.pdf>
- Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970-977.
- Cuban, L. (1986) *Teachers and machines: The classroom use of technology since 1920*. New York, NY: Teachers College Press.
- Cushman, S. (1995) Make new professors, but keep the old. *New Literary History*, 26.3, 613 -621.
- Daniels, S. (1993). The impact of technology on teaching. *Work Study*, 42(3), 15-16.
- DeBourgh, G. A. (2008). Use of classroom 'clickers' to promote acquisition of advanced reasoning skills. *Nurse Education in Practice*. 8, 76-87
doi:10.1016/j.nepr.2007.02.002.
- Dewey, J. (1944) *Democracy and education*. New York: Macmillan. Retrieved May 2008, from <http://www.ilt.columbia.edu/publications/dewey.html> (Originally published in 1916).
- Dufresne, R. J., Gerace, W.J., Leonard, W.J., Mestre, J.P., and Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, 7, 3-47.
- Elrod, S. (2007). Genetics concept inventory. Retrieved March 15, 2009 from <http://bioliteracy.colorado.edu/Readings/papersSubmittedPDF/Elrod.pdf>
- Evans, R. (1999). Serving modern students in a modern society at the community college: Incorporating basic technological literacy. *T.H.E Journal*, 27(3), 102-106.
- eInstruction. (2007, product brochure). www.einstruction.com

- Everett, M. D. (2002). Classroom response system: An evaluation at an easy-access regional university. Retrieved October 16, 2006, from <http://einstruction.com>
- Fagen, A. P., Crouch, C. H., & Mazur, E. (2002) Peer instruction: Results from range of classrooms. *Physics Teacher*, 40, 206-209.
- Flowers, L., Pascarella, E. T., & Pierson, C. T. (2000). Information technology use and cognitive outcomes in the first year of college. *The Journal of Higher Education*. 71(6), pp. 637- 667.
- Freeman, S, O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., et al. (2007). Prescribed active learning increases performance in introductory biology. *Cell Biology Education*, 6, 132-139.
- Garcia, E. (2006). Supplemental instructions, study habits, and the community college student. *Dissertations Abstracts International*, 67(05) AAT 3217569.
- Garvin-Doxas, K., & Klymkowsky, M. W. (2008). Understanding randomness and its impact on student learning: Lessons learned from building the biology concept inventory (BCI). *CBE Life Science Education*, 7(2), 227-233.
- Garvin-Doxas, K., Klymkowsky, M. W., & Elrod, S. (2007). Building, using, and maximizing the impact of concept inventories in the biological sciences: Report on a National Science Foundation-sponsored conference on the construction of concept inventories in the biological sciences. *CBE Life Science Education*, 6, 277-282.
- Gatz, L. B., & Hirt, J. B. (2000). Academic and social integration in cyberspace: Students and email. *The Review of Higher Education*, 23(3), 299-318.
- Gay, L.R., & Airasian, P. (2000). *Educational research: Competencies for analysis and application* (6th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Geer, U. C. & Rudge, D. W. (2002) A review of research on constructivist-based strategies for large lecture science classes. *Electronic Journal of Science Education*, 7(2), 1-22. Retrieved February, 14, 2008, from <http://wolfweb.unr.edu/homepage/crowther/ejse/ejsev7n2.html>
- Hake, R. R. (1998). Interactive engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses." *American Journal of Physics*, 66(1), 64-74. Retrieved February 14, 2008, from <http://www.physics.indiana.edu/~sdi/ajpv3i.pdf>
- Hanson, C. M. (2006). From learning to education: A new paradigm for the community college. *Community College Review*, 34(2), 128-138.

- Herreid, C. F. (2006). "Clicker" cases: Introducing case study teaching into large classrooms. *Journal of College Science Education*, 36(2), 43-47.
- Hestenes, D., & Wells, M. (1992). A Mechanics Baseline Test. *The Physics Teacher*, 30, 159-166.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158.
- Hinkle, D. E., Wiersma, W., & Jurs, S.J. (2003). *Applied statistics for the behavioral sciences* (5th ed.). Boston, MA: Houghton Mifflin.
- Horowitz, H.M. (1988). *Interactivity in a classroom environment*. Paper presented February 24, 1988 at the Sixth Conference of Interactive Instruction Delivery for the Society of Applied Learning Technology (SALT).
- Howard, J. R. (2002). Do college students participate more in discussion in traditional delivery courses or in interactive telecourses? *The Journal of Higher Education* 73(6), 764 – 780.
- Huck, S. W. (2012). *Reading statistics and research*. Boston, MA: Pearson Education.
- Jackard, J., & Becker, M. A. (1997). *Statistics for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.
- Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology Research and Development*, 39(3), 5–14.
- Judson, E., & Sawada, D. (2002). Learning from past and present: Electronic response systems in college lecture halls. *Journal of Computers in Mathematics and Science Teaching*, 21(2), 167–181.
- Kearsley, G. (1996). Education as usual: Comments on Chris Dede's article. *The American Journal of Distance Education*, 10(2), 55–58.
- Klymkowski, M. W., Underwood, S. M., & Garvin-Doxas, R. K. (2010). *Biological Concepts Instrument (BCI): A diagnostic tool for revealing student thinking*. Retrieved December 18th, 2012 from <http://arxiv.org/abs/1012.4501>.
- Knight, J. K., & Wood, W. B. (2005). Teaching more by lecturing less. *Cell Biology Education*, 4, 298–310.
- Lasry, N. (2008). Clickers or flashcards: Is there really a difference? *The Physics Teacher*, 46, 242–244.

- Lawson, A. E., Abraham, M. R., & Renner, J. W. (1989). *A theory of instruction: Using the learning cycle to teach science concepts and thinking skills* [Monograph number one]. Kansas State University, Manhattan, KS: National Association for Research in Science Teaching.
- Leonard, W. H. (2002) How do college students best learn science: An assessment of popular teaching styles and their effectiveness. *Innovative Techniques For Large-Group Instruction*. Arlington, VA: National Science Teachers Association Press.
- Libarkin, J. (2008, October). *Concept inventories in higher education science*. Manuscript prepared for the National Research Council Promising Practices in Undergraduate STEM Education Workshop 2; Washington, DC.
- Lorenzetti, J. P. (2002). Training online for teaching online in Baltimore. *Distance Education Report*, 6(17), 3–4.
- Lowery, R. C. (2005). *Teaching and learning with interactive student response systems: A comparison of commercial products in the higher-education market*. Paper presented at the Annual Meeting of the Southwestern Social Science Association and its Affiliates, New Orleans, LA.
- Lutz, S. T., & Huitt, W. (2004). Connecting cognitive development and constructivism: Implications from theory for instruction and assessment. *Constructivism in the Human Sciences*, 9(1), 67–90.
- Mason, B. J., & Bruning, R. (2001). Providing feedback in computer-based instruction: What the research tells us. Retrieved November 13, 2010 from <http://dwb.unl.edu/Edit/MB/MasonBruning.html>
- Mazur, E. (1997). *Peer instruction: a user's manual*. Upper Saddle River, NJ: Prentice Hall
- Merisotis, J. P., & Phipps, R. A. (1999). What's the difference? Outcomes of distance vs. traditional classroom-based learning. *Change*, 31(3), 12–17.
- Merril, J. (1987). Levels of questioning and forms of feedback: Instructional factors in courseware design. *Journal of Computer-Based Instruction*. 14(1), 18–22.
- Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2001). Assessing understanding in biology. *Journal of Biology Education*. 35(3), p. 118–124.
- National Research Council. (2000). *How people learn: brain, mind, experience and school*. Washington, DC: National Academy Press.

- Nygren, T. I., & Fisher, S. (1999). Cost effective uses of technology in teaching: current issues and future prospects. *The Andrew Mellon Foundation*. Retrieved October 9, 2007, from <http://www.ceutt.org/CEUTTx98.htm>
- Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32(1), 45–61.
- Outcalt, C. L. (2003). The uses of technology in community colleges. *ERIC Clearinghouse for Community Colleges, EDO-JC-03-09*. Retrieved October 8, 2006 from <http://www.gseis.ucla.edu/ccs/digests/digest0309.htm>
- Owen, P. S., & Demb, A. (2004). Change dynamics and leadership in technology implementation. *The Journal of Higher Education*, 75(6), 636-666.
- Pascarella, E. T., & Terenzini, P.T. (1998). Studying college students in the 21st century: Meeting new challenges. *The Review of Higher Education*, 21(2), 151-165.
- Paschal, C. B. (2002). Formative assessment in physiology teaching using a wireless classroom communication system. *Advances in Physiology Education*, 26(4).
- Piaget, J. (1976). *The grasp of consciousness: Action and concept in the young child*. Cambridge, MA: Harvard University Press.
- Poulis, J., Massen, C., Robens, E., & Gilbert, M. (1998). Physics lecturing with audience paced feedback. *American Journal of Physics*. 66, 439–641.
- Preszler, R. W., Dawe, A., Shuster, C. B., & Shuster, M. (2007). Assessment of the effects of student response systems on student learning and attitudes over a broad range of biology courses. *CBE Life Sciences Education*, 6(1), 29–41.
- Privateer, P. M. (1999). Academic technology and the future of higher education: Strategic paths taken and not taken. *The Journal of Higher Education*, 70(1), 60-79.
- Rendón, L. I. (1994). Validating culturally diverse students: Toward a new model of learning and student development. *Innovative Higher Education*, 19(1), 33–51.
- Roschelle, J., Penuel, W. R., & Abrahamson, L. (2004). *Classroom response and communication systems: Research review and theory*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth Cengage Learning.
- Shi, J., Wood, W. B., Martin, J. M., Guild, N. A., Vicens, Q., & Knight, J. K. (2010). *A diagnostic assessment for introductory molecular and cell biology*. *CBE Life Sciences Education* 9(4), 453–461.
- Simpson, V., & Oliver, M. (2006). Using electronic voting systems in lectures. Retrieved June 30, 2007, from <http://www.ucl.ac.uk/learningtechnology/assessment/electronicvotingsystems.pdf>
- Smith, K. S., Wood, W. B., & Knight, J. K. (2008). The genetics concept assessment: A new concept inventory for gauging student understanding of genetics. *CBE Life Sciences Education* 7, 422–430.
- Stetson, R., & Bagwell, T. (1999). Technology and teacher preparation: An oxymoron? *Journal of Technology and Teacher Education*, 7(2), 145–152.
- Suchman, E, Uchiyama, K, Smith, R, & Bender, K. (2006). Evaluating the Impact of a Classroom Response System in a Microbiology Course. *Microbiology Education*, 7, 3–11.
- Trees, A. R., & Jackson, M. H. (2007). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, 32(1), 21–40.
- Trochim, W. K. (2002). *Research methods knowledge base*. Available on-line at <http://trochim.human.cornell.edu/kb/>
- Trumbach, A. (2006). *An empirical study of factors influencing inter-organizational information systems implementation: A case of the real estate industry*. (Unpublished doctoral dissertation), Nova Southeastern University, FL.
- Udovic, D., Morris, D., Dicman, A., Postlethwait, J., & Wetherwax, P. (2002). Workshop biology: Demonstrating the effectiveness of active learning in an introductory biology course. *BioScience*, 32, 272–281.
- VanWagoner, R. J., Bowman, L. S., & Spraggs, L.D. (2005). The significant community college. *Community College Review*, 33(1), 38–50.
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: Harvard University Press.

- Wingspread Group on Higher Education. (1993). *An American imperative: higher expectations for higher education*. Racine, WI: The Johnson Foundation, Inc. (ED 364 144).
- Woods, H. A. & Chui, C. (2003). Wireless response technology in college classrooms. Retrieved October 16, 2006, from http://www.mhhe.com/cps/docs/CPSWP_WoodsChiu.pdf
- Young, J. R. (2004). When good technology means bad teaching. *The Chronicle of Higher Education*, 51(12), A31–A32. Retrieved from <http://chronicle.com/weekly/v51/i12/12a03101.htm>

APPENDICES

APPENDIX A

Pre-/Post-Test

Please do not write on this assessment. On the answer sheet provided, include the ID provided to you, today's date and Pretest (or Posttest, as the case may be) as the subject. For each of the questions choose the best response and circle it on the answer sheet provided. Answer the questions to the best of your knowledge. The number of correct responses on the pretest will be added to your total points at the end of the semester as extra points. Non-participants will be offered a regular quiz of similar point value as extra points.

1. Many types of house plants droop when they have not been watered and quickly "straighten up" after watering. The reason that they change shape after watering is because
 - A. Water reacts with, and stiffens, their cell walls.
 - B. Water is used to generate energy that moves the plant.
 - C. Water changes the concentrations of salts within the plant.
 - D. Water enters and expands their cells.
2. In which way are plants and animals different in how they use energy?
 - A. Plants use energy to build molecules; animals cannot.
 - B. Animals use energy to break down molecules; plants cannot.
 - C. Animals use energy to move; plants cannot.
 - D. Plants use energy directly, animals must transform it.
3. How can a catastrophic global event influence evolutionary change?
 - A. Undesirable versions of genes are removed.
 - B. New genes are generated.
 - C. Only some species may survive the event.
 - D. There are short term effects that disappear over time.
4. There exists a population in which there are three distinct versions of the gene A (a1, a2, a3). Originally, each version was present in equal numbers of individuals. Which version of the gene an individual carries has no measurable effect on its reproductive success. As you follow the population over a number of generations, you find that the frequency of a1 and a3 drop to 0%. What is the most likely explanation?

- A. There was an increased rate of mutation in organisms that carry either a1 or a3.
 - B. Mutations have occurred that changed a1 and a3 into a2.
 - C. Individuals carrying a1 or a3 were removed by natural selection.
 - D. Random variations led to a failure to produce individuals carrying a1 or a3.
5. Natural selection produces evolutionary change by
- A. changing the frequency of various versions of genes.
 - B. reducing the number of new mutations.
 - C. producing genes needed for new environments.
 - D. reducing the effects of detrimental versions of genes.
6. What makes DNA a good place to store information?
- A. The hydrogen bonds that hold it together are very stable and difficult to break
 - B. The bases always bind to their correct partner.
 - C. The sequence of bases does not greatly influence the structure of the molecules.
 - D. The overall shape of the molecule reflects the information stored in it.
7. What is it about nucleic acids that make copying genetic information straightforward?
- A. Hydrogen bonds are easily broken.
 - B. The binding of bases to one another is specific.
 - C. The sequence of based encodes information.
 - D. The shape of the molecule is determined by the information it contains.
8. It is often the case that a structure (such as a functional eye) is lost during the course of evolution. This because
- A. it is no longer actively used.
 - B. mutations accumulate that disrupt its functions.
 - C. it interferes with other traits and functions.
 - D. the cost to maintain it is not justified by the benefits it brings.
9. When we want to know whether a specific molecule will pass through a biological membrane, we need to consider
- A. the specific types of lipids present in the membrane.
 - B. the degree to which the molecule is water soluble.
 - C. whether the molecule is actively repelled by the lipid layer.
 - D. whether the molecule is harmful to the cell.

10. How might a mutation be creative?
- A. It could not be; all naturally occurring mutations are destructive.
 - B. If the mutation inactivated a gene that was harmful.
 - C. If the mutation altered the gene product's activity.
 - D. If the mutation had no effect on the activity of the gene product.
11. An allele exists that is harmful when either homozygous or heterozygous. Over the course of a few generations the frequency of this allele increases. Which is a possible explanation? The allele
- A. is located close to a favorable allele of another gene.
 - B. has benefits that cannot be measured in terms of reproductive fitness.
 - C. is resistant to change by mutation.
 - D. encodes an essential protein.
12. In a diploid organism, what do we mean when we say that a trait is dominant?
- A. It is stronger than a recessive form of the trait.
 - B. It is due to more, or a more active gene product than is the recessive trait.
 - C. The trait associated with the allele is present whenever the allele is present.
 - D. The allele associated with the trait inactivates the products of recessive alleles.
13. How does a molecule bind to its correct partner and avoid "incorrect" interactions"?
- A. The two molecules send signals to each other.
 - B. The molecules have sensors that check for "incorrect" bindings.
 - C. Correct binding results in lower energy than incorrect binding.
 - D. Correctly bound molecules fit perfectly, like puzzle pieces.
14. Once two molecules bind to one another, how could they come back apart again?
- A. A chemical reaction must change the structure of one of the molecules.
 - B. Collisions with other molecules could knock them apart.
 - C. The complex will need to be degraded.
 - D. They would have to bind to yet another molecule.
15. Why is double-stranded DNA not a good catalyst?
- A. It is stable and does not bind to other molecules.
 - B. It isn't very flexible and can't fold into different shapes.
 - C. It easily binds to other molecules.
 - D. It is located in the nucleus.

16. Lipids can form structures like micelles and bilayers because of
- A. their inability to bond with water molecules.
 - B. their inability to interact with other molecules.
 - C. their ability to bind specifically to other lipid molecules.
 - D. the ability of parts of lipid molecules to interact strongly with water.
17. A mutation leads to a dominant trait; what can you conclude about the mutation's effect?
- A. It results in an overactive gene product.
 - B. It results in a normal gene product that accumulates to higher levels than normal.
 - C. It results in a gene product with a new function.
 - D. It depends upon the nature of the gene product and the mutation.
18. How similar is your genetic information to that of your parents?
- A. For each gene, one of your alleles is from one parent and the other is from the other parent.
 - B. You have a set of genes similar to those your parents inherited from their parents.
 - C. You contain the same genetic information as each of your parents, just half as much.
 - D. Depending on how much crossing over happens, you could have a lot of one parent's genetic information and little of the other parent's genetic information.
19. A mutation leads to a recessive trait; what can you conclude about the mutation's effect?
- A. It results in a non-functional gene product.
 - B. It results in a normal gene product that accumulates to lower levels than normal.
 - C. It results in a gene product with a new function.
 - D. It depends upon the nature of the gene product and the mutation.
20. Imagine an ADP molecule inside a bacterial cell. Which best describes how it would manage to "find" an ATP synthase so that it could become an ATP molecule?
- A. It would follow the hydrogen flow.
 - B. The ATP synthase would grab it.
 - C. Its electronegativity would attract it to the ATP synthase.
 - D. It would be actively pumped to the right area.
 - E. Random movement would bring it to the ATP synthase.

21. You follow the frequency of a particular version of a gene in a population of asexual organisms. Over time, you find that this version of the gene disappears from the population. Its disappearance is presumably due to
- A. genetic drift.
 - B. its effects on reproductive success.
 - C. its mutation.
 - D. the randomness of survival.
22. Consider a diploid organism that is homozygous for a particular gene. How might the deletion of this gene from one of the two chromosomes produce a phenotype?
- A. If the gene encodes a multifunctional protein.
 - B. If one copy of the gene did not produce enough gene product.
 - C. If the deleted allele were dominant.
 - D. If the gene encoded a transcription factor.
23. Gene A and gene B are located on the same chromosome. Consider the following cross: $AB/ab \times ab/ab$. Under what conditions would you expect to find 25% of the individuals with an Ab genotype?
- A. It cannot happen because the A and B genes are linked.
 - B. It will always occur, because of independent assortment.
 - C. It will occur only when the genes are far away from one another.
 - D. It will occur only when the genes are close enough for recombination to occur between them.
24. Sexual reproduction leads to genetic drift because
- A. there is randomness associated with finding a mate.
 - B. not all alleles are passed from parent to offspring.
 - C. it is associated with an increase in mutation rate.
 - D. it produces new combinations of alleles.
25. How is genetic drift like molecular diffusion?
- A. Both are the result of directed movements.
 - B. Both involve passing through a barrier.
 - C. Both involve random events without regard to ultimate outcome.
 - D. They are not alike. Genetic drift is random; diffusion typically has a direction.

The following questions consist of 12 pairs of questions that examine your knowledge of diffusion and osmosis. Each question has two parts: a multiple-choice response followed

by a multiple-choice reason (e.g. 26 and 27 are paired like this). On the answer sheet please bubble in one answer for each question.

26. Suppose there is a large beaker full of clear water and a drop of blue dye is added to the beaker of water. Eventually the water will turn a light blue color. The process responsible for blue dye becoming evenly distributed throughout the water is:
- A. osmosis
 - B. diffusion
 - C. a reaction between water and dye
27. The reason for my answer is because:
- A. The lack of a membrane means that osmosis and diffusion cannot occur.
 - B. There is a movement of particles between regions of different concentrations.
 - C. The dye separates into small particles and mixes with water.
 - D. The water moves from one region to another.
28. During the process of diffusion, particles will generally move from:
- A. high to low concentrations
 - B. low to high concentrations
29. The reason for my answer is because:
- A. There are too many particles crowded into one area; therefore, they move to an area with more room.
 - B. Particles in areas of greater concentration are more likely to bounce toward other areas.
 - C. The particles tend to move until the two areas are isotonic, and then the particles stop moving.
 - D. There is a greater chance of the particles repelling each other.
30. As the difference in concentration between two areas increases, the rate of diffusion:
- A. decreases
 - B. increases
31. The reason for my answer is because:
- A. There is a less room for the particles to move.
 - B. If the concentration is high enough, the particles will spread less and the rate will be slowed.
 - C. The molecules want to spread out.
 - D. There is a greater likelihood of random motion into other regions.
32. A glucose solution can be made more concentrated by
- A. adding more water
 - B. adding more glucose

33. The reason for my answer is because:
- A. The more water there is, the more glucose it will take to saturate the solution.
 - B. Concentration means the dissolving of something.
 - C. It increases the number of dissolved particles.
 - D. For a solution to be more concentrated one must add more liquid.
34. If a small amount of sugar is added to a container of water and allowed to set for a very long period of time without stirring, the sugar molecules will:
- A. be more concentrated on the bottom of the container
 - B. be evenly distributed throughout the container
35. The reason for my answer is because:
- A. There is a movement of particles from a high to low concentration.
 - B. The sugar is heavier than water and will sink.
 - C. Sugar dissolves poorly or not at all in water.
 - D. There will be more time for settling.
36. Suppose you add a drop of blue dye to a container of clear water and after several hours the entire container turns light blue. At this time, the molecules of dye:
- A. have stopped moving
 - B. continue to move around randomly
37. The reason for my answer is because:
- A. The entire container is the same color, if they were still moving, the container would be different shades of blue.
 - B. If the dye molecules stopped, they would settle to the bottom of the container.
 - C. Molecules are always moving.
 - D. This is a liquid: if it were solid the molecules would stop moving.
38. Suppose there are two large beakers with equal amounts of clear water at two different temperatures. Next, a drop of green dye is added to each beaker of water. Eventually the water turns light green (see Figure 1). Which beaker became light green first?
- A. Beaker 1
 - B. Beaker 2

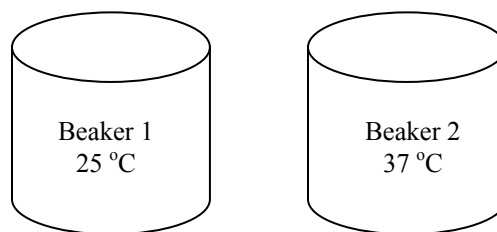


Figure 1

39. The reason for my answer is because:
- The lower temperature breaks down the dye.
 - The dye molecules move faster at higher temperatures.
 - The cold temperature speeds up the molecules.
 - It helps the molecules to expand.
40. In Figure 2, two columns of water are separated by a membrane through which only water can pass. Side 1 contains dye and water; side 2 contains pure water. After 2 hours, the water level in side 1 will be:
- higher
 - lower
 - the same height

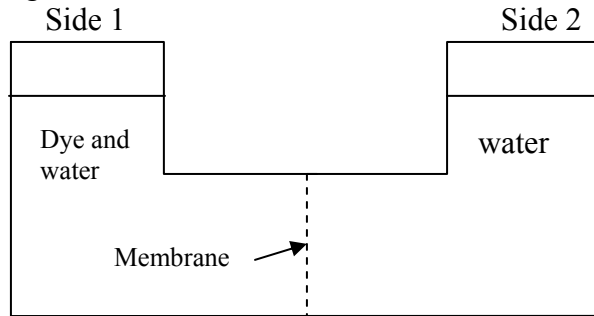


Figure 2

41. The reason for my answer is because:
- Water will move from the hypertonic to hypotonic solution.
 - The concentration of water molecules is less on side 1.
 - Water will become isotonic.
 - Water moves from low to high concentration.
42. In Figure 3, side 1 is _____ to side 2.
- hypotonic
 - hypertonic
 - isotonic

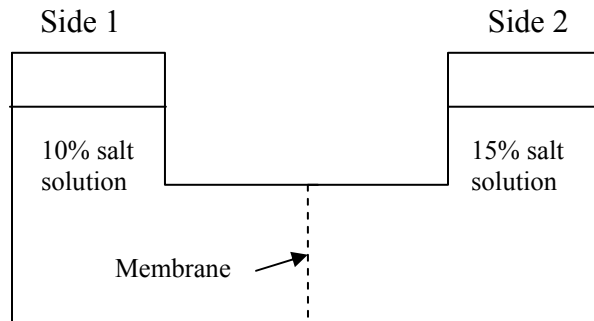


Figure 3

43. The reason for my answer is because:
- A. Water is hypertonic to most things.
 - B. Isotonic means “the same.”
 - C. Water moves from a high to a low concentration.
 - D. There are fewer dissolved particles on side 1.
44. Figure 4 is a picture of a plant cell that lives in freshwater. If this cell were placed in beaker of 25% saltwater solution, the central vacuole would:
- A. increase in size.
 - B. decrease in size.
 - C. remain the same size

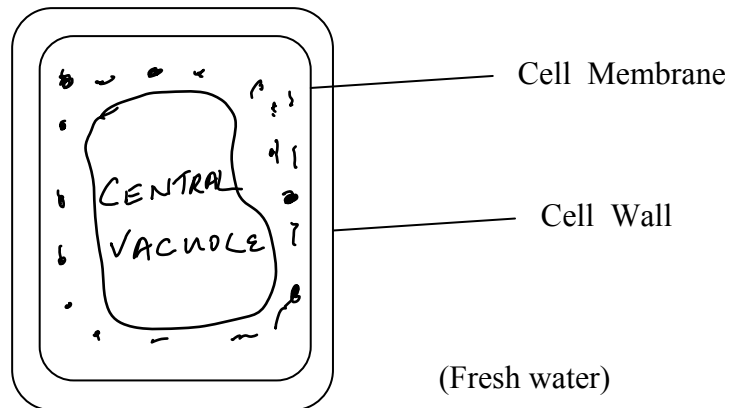


Figure 4

45. The reason for my answer is because:
- A. Salt absorbs the water from the central vacuole.
 - B. Water will move from the vacuole to the saltwater solution.
 - C. The salt will enter the vacuole.
 - D. Salt solution outside the cell cannot affect the vacuole inside the cell.
46. Suppose you killed the plant cell in Figure 4 with poison and placed the dead cell in a 25% saltwater solution.
- A. Osmosis and diffusion would not occur.
 - B. Osmosis and diffusion would continue.
 - C. Only diffusion would continue.
 - D. Only osmosis would continue.
47. The reason for my answer is because:
- A. The cell would stop functioning.
 - B. The cell does not have to be alive.

- C. Osmosis is not random, whereas diffusion is a random process.
 - D. Osmosis and diffusion require cell energy.
48. All cell membranes are:
- A. semipermeable
 - B. permeable
49. The reason for my answer is because:
- A. They allow some substance to pass.
 - B. They allow some substance to enter, but they prevent any substance from leaving.
 - C. The membrane requires nutrients to live.
 - D. They allow ALL nutrients to pass.

VITA

NILO ERIC MARIN

Born, Benque Viejo, Cayo, Belize

- | | |
|-------------------------------|--|
| 1971 – 1973 | A.S. Biology and Math St. John's Junior College Belize City, Belize |
| 1976 – 1978 | B.S. Biology University of Michigan Ann Arbor, Michigan |
| 1981 – 1983 | M.S. Physiology Southern Illinois University Carbondale, Illinois. |
| 1973 – 1975 | Math and Science Instructor Sacred Heart College (High School) San Ignacio, Cayo, Belize |
| 1975 – 1976 | Math and Biology Instructor St. John's College – High School Belize City, Belize |
| 1978 – 1981 1983 – 1985 | Biology Instructor St. John's Junior College Belize City, Belize |
| 1986 – 1994 | Biologist/Consultant The Lake Doctors, Inc. Fort Lauderdale, Florida |
| 1987 – 1994 1994 – Present | Adjunct Biology Instructor Biology Professor Broward College Davie, Florida |

PUBLICATIONS

Marin, N. E., J. H. Myers, and R. H. Browning. *Cardiac supersensitive response to isoproterenol following withdrawal from chronic propranolol infusion.* THE PHARMACOLOGIST 24:203, 1982