

# **No Time for Venus Flytraps: Effects of End-of-Course Testing on Biology Curriculum in Two States**

**Julie F. Westerlund**  
Southwest Texas State University  
San Marcos, Texas

**Leslie K. Upson**  
University of Georgia  
Athens, Georgia

**James P. Barufaldi**  
University of Texas  
Austin, Texas

## **Introduction**

I could spend more time talking about plants and animals and odd little Venus Flytraps that the kids always want to talk about. We don't ever have time to talk about Venus Flytraps because we're trying to feed all of this other information to them.

Ms. Henderson<sup>1</sup>, High School Teacher, North Carolina

The kids will say to me, 'Can I do the fetal pig dissection?' And I say 'Well, we just don't have time anymore' I am supposed to cover this amount of stuff in 90 days and there is too much information.

Ms. Langworthy, High School Teacher, Texas

Is there time for Venus Flytraps, time for “why” and “how”, and time for curiosity in today's science classrooms? The National Science Education Standards (NSES) states that students will better understand and retain scientific concepts if they are provided with inquiry and investigative experiences (National Research Council (NRC), 1996). The NSES emphasizes that reform in science education must begin with scientific inquiry as stated in the following; “Students need to be able to “devise and carry out investigations that test their ideas” (NRC, 2000, p.xiii). The president of the National Academy of Sciences, Bruce Alberts, further reiterated that by stating that, “To improve education, educational systems need to exploit the natural curiosity of children, so that they maintain their motivation for learning not only during school but throughout life.” (NRC, 2000, p.xiii). However, in this era of science education reform, state science standards and multiple-choice standardized testing drive the curriculum often at the expense of scientific inquiry. (The National Center for Fair & Open Testing (NCFT), 1997; Madaus, 1991).

## **Effects of Science Education Reform in the United States**

---

<sup>01</sup> Teachers' names have been changed for confidentiality purposes

Today's science education reform efforts in the United States, that are primarily characterized by standards and multiple-choice standardized tests, have not produced significant improvement in secondary school science education (National Assessment of Education Progress (NAEP), 2000; Third International Mathematics and Science Study, (TIMSS), 1996). Improvement in science education, according to the NSES, means that more students must understand and retain scientific concepts (NRC, 1996). In order to detect an improvement in scientific knowledge in the United States, international and national science assessments have been conducted. These assessments have consistently ranked the United States as mediocre in comparison to those of other developed countries (Medrich & Griffith, 1992). For example, the TIMSS ranked the performance of United States eighth-grade students 17th out of the 41 countries studied in science (TIMSS, 1996). The TIMSS was repeated four years later as The Third International Mathematics and Science Study Repeat (TIMSS-R). The TIMSS-R found no change in eighth-grade science achievement in the United States and ranked the science achievement of United States eighth-grade students significantly below eleven other countries (NRC, 1999; National Science Teachers' Association (NSTA), 2001). The 2000 NAEP tests indicated that 82% of high school seniors were not performing at the proficient level in science. Furthermore, there were no significant changes in science scores in grades 4, 8 and 12 since students were tested by the NAEP, four years earlier, in 1996 (NSTA, 2001/2002). The NAEP and the TIMSS tests are composed of questions that are multiple-choice, constructed response and performance or hands-on based questions. Thus, both international and national standardized tests in science indicate little or no improvement in science education in the United States secondary school students.

### **Effects of Science Education Reform in other Countries**

Countries, such as Japan, that have shown high science achievement on international tests, conduct science instruction with an emphasis on scientific investigation and coverage of fewer topics. These practices parallel the recommendations of the NSES (NRC, 1996). The TIMSS 1999 results showed a positive correlation between high science achievement and the emphasis teachers placed on laboratory investigations (NSTA, 2001). For example, the emphasis on scientific investigations in the new 1990's National Curriculum in Britain and Wales may have contributed to the high achievement of British students on the TIMSS Performance Assessment. The TIMSS Performance Assessment was administered to 9th year students in 21 countries. This assessment focused on investigative, problem-solving and analytical skills based on laboratory activities (Harris, 1998). Students in England from the United Kingdom placed second out of nineteen countries (Harmon et al. 1997).

### **Comparison of Science Education in United States and in other Countries**

The educational systems and schools of Japan, Germany, and the United States have been extensively analyzed (NRC, 1999). The TIMSS study surveyed teachers and students, examined curricular materials, videotaped classes and conducted in-depth case study analyses. One key difference between countries studied was that the United States science curriculum covered a substantially larger number of topics, as reported by United States science teachers (NRC, 1999). Analysis of videotape revealed that U.S. teachers switch topics more frequently and refer less often to concepts that had been learned earlier. "This rapid movement from one topic to another suggests that U.S. instruction may be more superficial than in other countries, with students failing to acquire a deeper understanding of any particular topic." (NRC, 1999, p.11).

Furthermore, according to the TIMSS analysis, U.S. textbooks are substantially larger than textbooks in other countries. U.S. textbooks contain more review exercises, repeat more topics covered in earlier grades, and switch topics more frequently.

### **Improvement of Science Education Reform in the United States**

It appears through international comparisons of science programs, U.S. programs may be improved by 1) increasing the investigative or inquiry-based emphasis and 2) decreasing the number of topics covered so that instruction can promote a deeper understanding of science concepts. These improvements in science education are well known in the science education community as illustrated in the following quotation.

You can't expect to see big changes in student achievement when we haven't made big changes in the way students learn science. We know how to improve student achievement in science. Every student must be taught by a competent teacher with a strong background in the science they are teaching, and access to professional development to hone their skills in inquiry-based teaching...  
(Harold Pratt, NSTA President, NSTA 2001/2002)

However, instead of an increased emphasis on inquiry teaching and reduction in the number of science topics covered, many states respond to the NAEP results by adding or changing its science assessment tests. For example, Texas will require in 2003 that all students in grades 5, 10, and 11 pass standardized multiple-choice tests (Texas Assessment of Knowledge and Skills) in science in order to graduate from high school (NSTA 2001/2002). This response and approach to improving science education may simply repeat the pattern that has occurred in science education reform in the past decade (The National Center for Fair & Open Testing (NCFT), 1997; Madaus, 1991).

### **Scientific Literacy in the United States**

The improvement of science education and the promotion of scientific literacy are vital to the well being of the United States. Presently, scientific illiteracy is pervasive in this country as illustrated by the following:

One-half of the American public does not know the earth goes around the sun once a year and one-half believes that the earliest humans lived at the same time as the dinosaurs. (National Science Foundation, (NSF), 1996, p.8)

As illustrated in the quotation above, many United States citizens lack even the most rudimentary knowledge about scientific concepts that were established in the nineteenth century (NSF, 1996). U.S. citizens need to be scientifically literate not only to become a more informed voting citizenry and/or for a better appreciation of the natural world, but also for safety. Carl Sagan (1995) very aptly described why scientific literacy, "a candle in the dark", is essential for a safe modern society.

For much of our [human] history, we were so fearful of the outside world, with its unpredictable dangers, that we gladly embraced anything that promised to soften or explain away the terror. Science is an attempt, largely successful, to

understand the world, to get a grip on things, to get hold of ourselves, to steer a safe course. Microbiology and meteorology now explain what only a few centuries ago was considered sufficient cause to burn women [accused as witches] to death (Sagan, 1995, p.26).

If we as a country in the United States are to attain scientific literacy, efforts to improve science education in this country should be based upon science teaching techniques such as scientific inquiry, which has been used to improve scientific understanding in students (NRC, 2000).

### **Purpose of Study**

The purpose of this study was to compare teacher perceptions about recent science education reform in biology in North Carolina and Texas. The study focused on state-mandated biology end-of-course tests, which are the principal reforms, that North Carolina and Texas and many other states have implemented in an effort to improve biology education. Two research questions were addressed by this study: 1) In what ways, if any, do teachers in North Carolina, and Texas believe the biology end-of-course examinations have influenced the curriculum and teaching practices in their biology courses? 2) How do biology teachers' perceptions in North Carolina and Texas compare regarding the influence of biology end-of-course examinations on the biology curricula and their teaching practices?

### **Background**

#### **Prevalence of State-wide Standardized Tests in Science**

Today, the purpose of most standardized testing in science is to report individual achievement and to facilitate public accountability. Every year more states require students to pass science tests in order to graduate from high school. In 1999, 39 states had developed their own statewide science standards, mostly based on recommendations in the national standards documents, and 48 had statewide science tests to measure student achievement (Jerald & Boser, 1999). For example, in New York, one of the first states to begin standardized testing in science in the early 1900's, administers the Regents Examination. The Regents Examinations are comprehensive tests in 13 different subjects for grades 9 – 12 (Madaus, 1994).

In 26 states, K-12 standardized statewide science achievement tests are used to evaluate the effectiveness of public school science programs (Jerald & Boser, 1999). Each year more states are adding science achievement tests to their testing program and most states have developed their own tests in accordance with their state standards (Edwards, 1999). For example, New York, Texas and California, the most populous states, have developed their own unique tests (see Table 1, Madaus, 1994; Texas Education Agency, TEA, 1997; NSTA 2001/2002; California Department of Education, 2000; North Carolina Department of Public Instruction, NCDPI, 1999).

Table 1  
Examples of Standardized Statewide Science Tests

State	Standardized Test
New York	1. 4 <sup>th</sup> Grade Elementary Science Program Tests (ESPETS) 2. Regents Examinations (graduation requirement)
Texas	1. 8 <sup>th</sup> grade Science Texas Assessment of Academic Skills (TAAS) 2. Biology End-of-Course Examination (EOC) 3. 5 <sup>th</sup> , 10 <sup>th</sup> , 11 <sup>th</sup> , Texas Assessments of Knowledge and Skills (TEKS) in Science (graduation requirement)
California	Golden State Examinations in Biology, Chemistry, Coordinated Science and Physics
North Carolina	End-of-Course Examination in Biology

### **Characteristics of Standardized Tests in Science**

Statewide K-12 state-developed science assessment tests share many similarities. For example, the majority of the questions are multiple-choice. Specifically, most state-developed, standardized science achievement tests have these characteristics: 1) multiple-choice, 2) short (fewer than 60 items), 3) few questions about science concepts, 4) administration time of one hour, 5) paper and pencil, 6) individualized and 7) administered twice a year (Gong, 1990). Only fourteen states use performance assessment in their statewide science examinations in a few grade levels. For example, California has a laboratory component in the Golden State Examinations in science. New York also has a laboratory component in its ESPETS. By contrast, Texas does not include performance assessment in any of the science-standardized examinations (TEA, 1994, 1996a). Also, most tests tend to be criterion-referenced in that students are expected to answer correctly a certain percentage of items that are based on state standards (NCFT, 1997, Gong, 1990). Lastly, portfolio assessment is not used in any statewide science assessment system (Jerald & Boser, 1999).

### **Consequences of Standardized Testing**

#### Positive

Although most science education research emphasizes the negative effects of standardized testing, some possible benefits may be gained from standardized tests:

- 1) The pressure placed on teachers as a result of standardized tests may improve the course content and strengthen the motivation of those teachers who are not doing a good job (Porter, 1989).
- 2) Multiple-choice tests, while not the only measure of achievement, are certainly the most cost-effective (Popham, 1987).
- 3) Nationally, students in grades 4, 8, and 12, are achieving higher scores in mathematics and slightly higher scores in reading (National Educators Goals Panel, 1999), which may be a result of teachers and administrators being focused on having students pass standardized tests. For example, African-American students in Texas

have shown significant improvement in writing skills and had the highest scores by African-Americans in the United States on the 1998 NAEP writing test (NAEP, 1998). Persistent efforts by Texas teachers to teach writing skills to students, because of an exit-level writing requirement imposed in 1985 by Texas' statewide accountability, may have contributed to the improvement in writing scores.

#### Negative

There are also well-known negative consequences of standardized testing on education which include:

- 1) The practice of "teaching to the test" causes a "narrowing of the curriculum". When there is public pressure to improve test results, schools and teachers are likely to emphasize, the material covered by the test to the exclusion of other topics of interest (Shepard, 1991; Herman & Golan, 1992). ). "Teaching to the tests" causes students to gain the "most elementary knowledge and skills and less of the deep understanding of even a few topics" (Stake, 1991 p. 246). This is demonstrated in the inability of test scores to generalize or transfer to other indicators of achievement. When a new testing program is brought into a state, scores tend to plummet in the initial years of testing since students have not been primed to answer questions on that exact test (Bracey, 2000). For example, Wisconsin fourth grade statewide science scores plummeted by 15 points on the Wisconsin Knowledge and Concepts Exam, in the 2000/2001 school year. One of the possible reasons for the drop was that new tests were administered in the 2000/2001 administration after two years of using the same test (Erickson, 2001). Again, this is an indication that students were being prepared for test questions rather than gaining in-depth knowledge of the concepts. Consequently, test scores that reflect higher-order thinking have been steadily declining (Darling-Hammond, 1991, 1994). In fact, the rote learning involved in "teaching-to-the-test" in the 1970s has been cited as one reason that U.S. students have ranked so low in international achievement tests (McKnight et al.,1987).
- 2) The emphasis of multiple-choice questions (NCFT, 1997) in standardized tests encourages the practice of rote memorizing, "cramming" of concepts and test-taking strategies in classrooms (Madaus, 1991). Cost and efficiency play instrumental roles in determining test format and content. Often, these tests measure what can be tested easily and cheaply by multiple-choice questions (Goodland, 1984; Madaus, 1988; Porter, 1989; McLaughlin, 1991).
- 3) The multiple-choice questions that appear on most widely used science achievement tests address only a low level of conceptual thinking (NSF, 1992, Morgenstern and Renner, 1984; Gong, 1990). Science assessments that use only multiple-choice questions may not adequately assess the in-depth knowledge of concepts that is stressed by the NSES. "Low level thinking questions" refer to recall and routine application questions. The NSES recommendations stipulate that student assessment includes questions, which assess high-level thinking. Such questions test for "active knowledge," meaning knowledge that is structured conceptually as opposed to "inert knowledge," or isolated bits of knowledge. According to the NSES, examinations should include open-ended and performance questions and portfolio assessments in order to assess knowledge that is structured conceptually or "active knowledge" (NRC, 1996, Gong, 1990). Furthermore, open-ended formats may provide a more valid assessment of student learning (Resnick and Resnick, 1992).

## **Science End-of-Course Tests**

End-of-course tests in science have been proposed as alternatives to broad-based science assessments that test science content learned over several grade levels. Broad-based science assessments have been criticized for having too few questions that cover an immense amount of content (Olson, 1999). End-of-course tests simply cover the content of one course. Therefore, it should be possible to develop subject-specific end-of-course tests with questions that probe a deeper level of understanding of the subject. End-of-course tests in the sciences have been in existence in North Carolina and Texas since 1987 and 1994 respectively. Currently, they are given in chemistry, biology, physical science, and physics (TEA, 2002; NCDPI, 1999). In Maryland, the state school board voted in December 1997 to begin administering “challenging end-of-course tests” in all core subject areas, including the sciences (Portner, 1999). In Virginia, end-of-course tests in the sciences were given for the first time in 1998 (Portner, 1999). Thus, there appears to be a growing trend towards end-of-course testing in the sciences.

In a literature search (1986 to present) of the Education Resources Information Center (ERIC), we found only two published articles which examined the effects of end-of-course testing in science. These were the Smith, Hounshell, Copolo and Wilkerson (1992) study of chemistry end-of-course testing in North Carolina and the Falk and Larson (1996) study of performance assessment components of the New York State Regents Examinations.

Smith et al. (1992) examined the impact of testing from the chemistry teachers’ perspectives through the use of a written survey and telephone interviews. The survey results revealed that more than 64% of the teachers were doing things to “specifically prepare students for the test.” Nearly 70% believed they had not changed their methods. However, 37% admitted to “teaching to the test” and 51% supported end-of-course testing. Telephone interviews with 10 teachers revealed that they were using more demonstrations in place of laboratories and that they were changing their instructional methods to adjust to the test. From these results, it is not clear whether and how end-of-course testing affected the chemistry curriculum. An approach that involves repeated interviewing of teachers may yield a more detailed database and thus, may help to elucidate those effects. Furthermore, the sampling population of the Smith et al. study was very small. Of 100 teachers who were mailed surveys, only 48 qualified teachers responded. The study did not describe any attempts to gather data from the non-respondents in order to check for bias in the respondent population. Lastly, the reliability and validity of the survey instrument was not mentioned.

Falk and Larson (1996) reported on the effects of a performance assessment component of the New York State Regents Examinations, the Regents Exam Options Project. As expected, many positive learning outcomes were reported when students were able to demonstrate their knowledge through projects, portfolios, and performances rather than the standard multiple-choice formats. These outcomes included: 1) changing the role of the teacher from “transmitter of information to facilitator of learning”, 2) more cooperative learning, 3) more inquiry-learning, 4) more problem-solving and use of higher-order thinking skills, and 4) more self-directed learning (Falk and Larson, 1996).

### **Significance of Study**

Our study addresses the need for more detailed information about the specific effects of end-of-course examinations on curriculum and instruction in biology classrooms, which can then be provided to policy makers, teachers, parents and students. This information is needed due to the increasing use of end-of-course examinations throughout the United States. Presently, only two studies, Smith et al. (1992) and Falk

and Larson (1996), specifically addressed the influence of end-of-course tests in the classroom. As indicated earlier, the Smith et al. study was primarily conducted using a survey that did not allow for teacher elaboration of these effects. And, the Falk and Larson (1996) study examined only the effects of the performance assessment components of the New York Regents Examinations. A literature search (1986 to present) using ERIC revealed that most studies about end-of-course testing concern student achievement scores rather than the effects of the end-of-course tests. Thus, the significance of our study is in three areas. First, it provides information through teacher perceptions about the influence of end-of-course examinations on curriculum and instruction. Secondly, it is unique in that it examines the influence of these examinations through case-study analyses conducted by researchers who were also practicing high school biology teachers. Case-study analyses of teachers may yield more detailed information about how EOC examinations have affected science curriculum. Thirdly, it is unique in that it compares the perceptions of biology teachers concerning end-of-course testing in North Carolina and Texas, states with the longest history of end-of-course testing.

### **History of Science Education Testing in Texas and in North Carolina**

#### **Science Education Testing in Texas**

Due to legislation enacted in the early 1980's by the Texas Legislature, Texas student performance is assessed today through standardized testing. In April of 1992, the Texas Board of Education included testing in the form of criterion-referenced end-of-course examinations. The purpose of these tests was to give the state additional information about the effectiveness of instructional programs and were based on the statewide curriculum, the Texas Essential Elements (TEA, 1993). Also, according to Senate Bill One, end-of-course tests in English II, Algebra I, Biology I and United States History were intended to replace the state's graduation test, the Texas Assessment of Academic Skills (TAAS) (TEA, 1995a). However, these tests did not replace the TAAS and are expected to be phased out in 2003. The TAAS testing program, renamed Texas Assessment of Knowledge and Skills (TEKS) will be expanded in 2003 to include testing in science in the 5<sup>th</sup>, 10<sup>th</sup> and 11<sup>th</sup> grades (TEA, 2002). Presently, school scores on the end-of-course examinations are on the annual TEA report card for a school as "Report Only" academic excellence indicators and are not used to determine the performance status of a school.

The Biology End-of-Course examination (EOC) is a traditional, multiple-choice, machine-scorable test of 42 questions. Students may take as long as needed to finish the test. Beginning in 1997, the people of Texas were able to view previously administered Biology EOCs. The design and content of the Texas Biology EOC (TEA, 2002) has been evaluated using the NSES Assessment Standard B and the NSES Life Science Content Standards for Grades 9 -12 (NRC, 1996). Results indicated that the Texas Biology EOC did not comply with standards from NSES in areas concerning coverage of biology concepts, depth of knowledge required to answer questions, and format of questions. For example, on a representative test, the Spring 1997 Biology EOC, only 31% of the questions or 13 questions required knowledge of biology to answer the questions. Furthermore, the purpose of 12 of the 13 questions was to ascertain whether students had knowledge of a fact, rather than understanding of a concept. Only one question of those



13 questions or 1%, required a higher level of conceptual thinking than knowledge level. Also, the Biology EOC is all multiple-choice and does not meet the NSES recommended format of open-ended questions (Westerlund & West, 2001, NRC, 1996).

### **Science Education Testing in North Carolina**

The “ABC’s of Public Education in North Carolina” specify standardized tests, which are used to judge individual schools with respect to the achievement of their students. Tests are administered at the high school level in the following subjects: Algebra I; Biology I; Economics, Legal and Political Systems; English I; English II and U.S. History. Tests are based on the state curriculum, the Standard Course of Study. In science these include, physical science, chemistry and physics. Teachers and support staff may receive monetary rewards based upon the scores received by students at their schools. Teachers at schools that meet expected growth receive bonuses of \$750 and teachers at schools with exemplary growth receive bonuses of \$1500. Schools that are low performing are assigned an assistance team which help the schools and principals develop strategies to improve test scores (NCPDI, 1999). The 1997-1998 school year was the first in North Carolina wherein high schools operated under the “ABC’s of Public Education Accountability Model” (NCPDI, 1999).

The end-of-course test in biology is one component in a complicated formula used to determine a composite score for each high school. There are 95 multiple-choice questions on the biology end-of-course test. Students are given 95 minutes to complete the test. Students may use an additional fifteen minutes if extra time is needed (NCDPI, 1998). The state has released Biology Pretest (Form A), which contains 44 multiple-choice questions and one open-ended question. The actual North Carolina end-of-course examinations in biology, which have not been released to the public, are very similar to the biology pretest (personal communication, 6/21/99, Heather Koons, NCDPI, Accountability Division). All questions on both North Carolina biology pre-tests and end-of-course tests are evaluated and labeled according to the thinking skill required to answer the question. In the North Carolina Pretest of Biology, Form A, 31% percent of the questions are labeled as representing the lowest level of thinking skill or “knowledge level” And, 80% of the questions required some knowledge of biology to answer the questions (NCDPI, 1999).

### **A Comparison of North Carolina and Texas Biology EOC Tests**

The North Carolina and Texas Biology EOC examinations have similarities and differences (see Tables 2 and 3) (TEA, 1995a, Westerlund and West, 2002, NCPDI, 1999).

Table 2

Similarities between North Carolina and Texas End-of-Course Examinations

State	Content	Type of Questions	Accountability	Requirement for graduation
N.Carolina	State-wide Curriculum, "Standard Course of Study"	Multiple-Choice	"High stakes", part of school's annual report card	Not required
Texas	State-wide Curriculum, "Essential Elements"	Multiple-Choice	"High stakes", part of school's annual report card	Not required

Table 3

Differences between North Carolina and Texas End-of-Course Examinations

State	Number of items	Coverage of biology content	Higher – order (above knowledge level) thinking questions in biology content	Time allotted for completion	Monetary Awards for EOC scores	Year of Implementation
N. Carolina	95	80%	33%	110 minutes	Bonuses for faculty	1987
Texas	42	31%	1 %	Unlimited	None	1994

**Theoretical Framework of Study**

**Heuristic Inquiry: A Phenomenological Approach to Research**

The design of this study is that of qualitative phenomenological research. This type of research focuses on the question "What is the structure and essence of experience of this phenomenon for these people?" (Patton, 1990 p.69). The phenomenon is the biology EOC examination and the people are the biology teachers.

The design is a type of phenomenological qualitative research called heuristic inquiry. Heuristic research is characterized by the discoveries, personal insights and reflections of the researcher. In order for a study to be considered a heuristic study, two components must be present. "1) The researcher must have personal experience and

intense interest in the phenomenon under study. 2) Others (those being studied), or co-researchers must share an intensity of experience with the phenomenon" (Patton, 1990 p.71). Unlike traditional phenomenological studies; 1) Heuristic research does not assume the detachment of the researcher from the phenomenon being studied. The researcher is very connected to the research (Douglass & Moustakas, 1985).

At the time of this study, we (Westerlund and Upson) were biology teachers like those who were interviewed and we had direct experience with the biology EOC examinations in Texas and in North Carolina. We examined teacher perspectives on the EOC examinations both as researchers and as biology teachers. Our ability to understand and interpret other teachers' perspectives was strengthened by our shared experience as biology teachers. Douglass & Moustakas (1985, p.51) describe this enhancement of understanding by stating "When we know a thing from our experience, its meanings can be recognized in others, without the typical accountings and explanations."

In qualitative studies such as this, the quality of the data is also strengthened when the researcher is a natural part of the setting of the study. In the Texas component of the study, the "interviews" were as natural as a spontaneous conversation between two biology teachers. The researchers' experiences as biology teachers helped to establish a peer-level rapport with the participant biology teachers. This relationship allowed the collection of the richest possible responses from the teacher participants in the study. Thus, the credibility of the data was enhanced through the use of this heuristic approach (Patton, 1990).

## Methods

### Overview

The research was conducted in a qualitative manner with a heuristic approach. A *purposeful, maximum variation* sample (Patton, 1990, p.172) of five Texas biology teachers from two schools and nine North Carolina biology teachers from three schools whose students took the biology end-of-course examinations was used in the study (see Table 4). Four of the Texas teachers and three of the North Carolina teachers taught prior to the implementation of the biology end-of-course examination. The sample was *purposeful* in that cases were selected that would yield the most information about the research questions (Patton, 1990). Both teachers who had taught before the implementation of the examination (1987 in North Carolina, 1994 in Texas) and relatively new teachers whose only experience included teaching with the end-of-course examination were included in the sample. The sampling was considered of *maximum variation* because the sites selected varied considerably in terms of academic performance on standardized tests and student ethnicity.

Five Texas biology teachers were sampled. There were three from Roosevelt<sup>2</sup> High School and two from Marshall High School. The three from Roosevelt included one of the author-researchers. The five Texas teachers provided data in various ways over a two-year interval.

Nine North Carolina biology teachers were sampled. Four biology teachers were from Anderson High School, two biology teachers from Bluffton High School and three biology teachers from Connor High School. The nine North Carolina teachers provided data primarily through interviews.

---

<sup>2</sup> Names of schools have been changed for confidentiality.

All of the teachers in this study have been profiled in detail in order to better understand their approaches to teaching biology. Their profiles are mentioned briefly (see Appendix A) due to space limitations.

### **Site Selection**

#### **Texas**

The settings of the Texas component of the study were Roosevelt and Marshall High Schools. Most of the study was conducted at Roosevelt High School. Marshall High School was selected as part of the maximum variation sampling because it differed considerably from Roosevelt High School in terms of its passing rates on the TAAS test and its student profile (see Table 4). In addition, Marshall was a larger school with approximately 1000 more students than Roosevelt. In the 1994-95 school year, there were 2394 students enrolled at Marshall H.S. and 1338 enrolled at Roosevelt H.S. according to the 1994-95 TEA School Report Card (TEA, 1996b).

Table 4  
Characteristics of High Schools and Teachers

High Schools	State	Teachers (yrs. exp.)	Urban or Rural	Number of Students	Student Ethnicity and % on Free and Reduced Lunch	Passing rates on exit test
Roosevelt	Texas	Meyers, 11 yrs. Langworthy, 8 yrs. Tucker, 3 yrs.	Urban	1338	46.8% AA <sup>3</sup> 32.6% His <sup>4</sup> 18.5% Whi <sup>5</sup>  51.1% FrLun <sup>6</sup>	37.9% passed TAAS test
Marshall	Texas	Spencer, 10 yrs. Taylor, 4 yrs	Urban	2394	5.8% AA <sup>4</sup> 21.7% His <sup>5</sup> 69.6% Whi <sup>6</sup>  9.3% FRLun <sup>7</sup>	64.9% passed TAAS test
Anderson	N. Carolina	Henderson, Murray, Drummond, Underwood, (1-4 yrs)	Rural and Urban	1275	50.1% AA <sup>4</sup> 0.0% His <sup>5</sup> 48.2% Whi <sup>6</sup>  32.03% FRLun <sup>7</sup>	no exit test
Bluffton	N. Carolina	Ford 20 yrs, Bartholomew, 2 yrs	Urban	1216	67.1% AA <sup>4</sup> 0.0% His <sup>5</sup> 31.1% Whi <sup>6</sup>  36.79% FRLun <sup>7</sup>	no exit test
Connor	N. Carolina	Wilson, 24 yrs, Douglas, 6 yrs Adams, 24 yrs	Rural	871	27.1% AA <sup>4</sup> 4.5% His <sup>5</sup> 64.6% Whi <sup>6</sup>  28.38% FRLun <sup>7</sup>	no exit test

The TEA Report Card for the 1993/94 School Year stated that Roosevelt High School had only 37.9% of its students passing the TAAS whereas Marshall High School had 64.9% of its students passing the TAAS. Also, the TEA Report Card for the 1993/94 School Year indicated that the ethnic membership for the two schools was different. The

<sup>3</sup> African-American

<sup>4</sup> Hispanic

<sup>5</sup> White

<sup>6</sup> % Free and Reduced Lunch

Roosevelt High School student population consisted of 46.8% African-American, 32.6% Hispanic and 18.5% white. The Marshall High School student population consisted of 5.8% African-American, 21.7% Hispanic and 69.6% white (TEA, 1995b). Roosevelt High School classrooms met the criterion of high-minority classrooms (above 60%) as defined by the National Science Foundation (NSF, 1992). Also, 51.1% of Roosevelt students were economically disadvantaged whereas only 9.3% of students at Marshall were economically disadvantaged (TEA, 1995b). The two high schools were very different, each with problems and challenges.

Roosevelt and Marshall High schools were also selected because the sites were ideal for conducting qualitative research. One of the researchers, Westerlund, was a full-time biology teacher at Roosevelt High School for 7 1/2 years. As a result of this connection: 1) There was easy access, since one of the researchers was there teaching, observing and conversing with teachers. 2) There were two other biology teachers at Roosevelt High School who were willing to participate in the study. Mr. Tucker was new to the profession and Ms. Meyers had 10 years of experience. 3) The researcher already had a long-term, trusting relationships with the biology teachers at Roosevelt High School, and this is essential for a good qualitative study. 4) Power conflicts were not present between the teacher-researcher, and the subjects of the research because of the relationship. This further increased the credibility of the study, and thereby generated high quality data.

Marshall High School was selected as a second site because of its suitability for qualitative research. There were two biology teachers and both were willing to participate in the study. Ms. Spencer had 10 and Ms. Taylor had 4 years of teaching experience. The researcher had shared a professional relationship with the two Marshall teachers prior to the study. In the past, the researcher had helped them with a district biology teachers' workshop and had exchanged ideas on teaching human genetics to students.

### **North Carolina**

The settings of the North Carolina component of the study were Anderson, Bluffton and Connor High Schools. Of the three schools, Bluffton High School was the only high-minority school (above 60%). Though located in the same county, the three schools had different characteristics (see Table 4). The schools differed in population and setting. Anderson High School was located in an area that draws students from both rural and urban settings and had a population of 1148 students. Ms. Henderson, Mrs. Murray, Mrs. Drummond, and Ms. Underwood all taught at Anderson H.S. and had between 1 to 4 years teaching experience. The Anderson High School teachers were new teachers that taught after the implementation of the biology-end-of-course examination. Bluffton High School was located in an urban area and had a population of 1116 students. Mrs. Ford who had twenty years of experience, and Mr. Bartholomew, who had two years of experience taught at Bluffton H.S. Connor High School was located in a rural area and had a population of 827 students. Mr. Wilson with twenty years of teaching experience, Mr. Douglas with six years of teaching experience and Mr. Adams with twenty-four years of teaching experience all taught courses at Connor H.S.

## Data Collection

### Texas

Four data sources were used for the collection of data. To validate results, multiple sources of data were used.

1) Interviews and Discussions - The primary data sources were personal in-depth interviews, discussions during lunchtime or after school and spontaneous conversations. All interviews and discussions were audiotaped and later transcribed. Of Patton's three interviewing styles, the *informal conversational interview* was used (Patton, 1990, p.281). The format of the interviews and discussions was very open-ended so that questions could be directed in whatever direction was appropriate and informative. These interactions were more in the form of dialogue than structured conversations. In heuristic research, these dialogues flow naturally. "One is encouraged to permit ideas, thoughts, feelings and images to unfold and be expressed naturally" (Douglass & Moustakas, 1985, p.46). Written notes were not taken during the interview or discussion, thereby a feeling of a natural conversation was present. Instead, all interactions were audiotaped and later transcribed. After the interview or discussion, field notes on the interview were audiotaped and later transcribed.

2) Oral Journals - Another major data source was oral journals. The Roosevelt High School (TX) teachers were provided with mini-cassette tape recorders. As they went through their daily routines, in the classroom or driving to-and-from school, they recorded any thoughts they had about the biology end-of-course examinations. The data gathered as transcripts from these oral journals were valuable because they were totally free of researcher input or bias.

3) Open-Concerns Inventory - A third data source was an open-concerns inventory. This inventory was used in order to gather data about the written expression of teachers' perceptions about the examination.

4) Personal Observations - A fourth data source was the tape recording of any events relating to the EOC over the normal course of a school day. Observations of teachers and spontaneous discussions with teachers became part of that record.

### North Carolina

Interviews were used as the primary source of data for this research. The interviews were conducted using a semi-structured format (Bogdan and Biklen, 1998). The researcher followed a list of general questions, which allowed the main themes of the study to be addressed; however, additional questions were asked on the responses of the participants. The list of questions was analyzed and modified by three professors in the science education department at North Carolina University<sup>7</sup>, and their comments and concerns were considered in the development of the final list of questions. Interviews were conducted on an individual basis, were taped and later transcribed. Follow-up interviews were done for clarification and verification of the participants' comments and thoughts (Bogdan and Biklen, 1998).

---

<sup>7</sup> Pseudonym for blind review purposes.

### **Limitations of the Study**

The study examined the perceptions of five Texas biology teachers and nine North Carolina biology teachers. A limitation of the study is its ability to generalize to other populations of biology teachers. The study was intended to be exploratory research to discover previously unknown teacher perceptions about the effects of biology end-of-course examinations. A small-scale study was selected so that the study could be an in-depth examination of teacher perceptions.

### **Data Analysis**

The researchers from both states used inductive analysis (Patton, 1990) to evaluate the qualitative data sets from Texas and North Carolina. The researchers used codes as tools to see similarities and differences in the varied perceptions of the teachers. In this study, the codes were examined to evaluate the effects of biology end-of-course examinations on the curriculum and instruction in teachers' classrooms.

In order to develop the codes, the transcripts were read several times by the researchers from each state. Statements from the transcripts that represented similar ideas, or that contained similar key phrases or terms, were grouped together under one code. The coding process meant, "looking for 'recurring regularities' in the data" (Patton, 1990, p.403). Coded statements were highlighted with a specific code color to specify coded text. After the coding phase was completed, the researchers collaborated in examining the data, comparing the codes, and redefining the codes.

The credibility of the study was enhanced through the use of analyst triangulation and triangulation of data sources (Patton, 1990). Analyst triangulation required that the researchers independently read and code the data. Similar patterns or codes in the data that were seen by both researchers were defined as being present in the perceptions of the teachers. Analyst triangulation was also employed by having the participant teachers read the raw transcripts of the interviews, discussions and oral journals. Each of the teachers confirmed that the transcriptions of their remarks were accurate and did not dispute any comment contained within the data. Triangulation of data sources occurred when data from the interviews, discussions, oral journals, and concerns inventory questions were compared and contrasted to determine if the teachers expressed the same views in different data sources.

### **Results**

Overall, six themes in the data were apparent to the researchers, that concerned the ways in which the end-of-course test influenced curriculum and instruction. These themes are titled as short quotes from the teachers that represent the gist of the themes. Each theme is also illustrated below using excerpts from the teachers. Using this approach, the reader may better understand the ways in which end-of-course testing has affected the biology curriculum and instruction in North Carolina and Texas.



## 1. “Keeping on task”

It causes me to be more on task because I know that I am going to be evaluated. If you're (the students) not doing well, they are going to look at that test. So it makes me stay on task more and it keeps me on pace to keep going.  
Mr. Bartholomew

That the biology end-of-course tests helped teachers “keep on task” was mentioned by two of the North Carolina teachers. As Ms. Henderson stated, “I think for some teachers, it has probably helped them keep on task. The Texas teachers did not mention the test as being helpful in keeping the teachers on task.

## 2. “A standardization of curriculum”

To several North Carolina teachers, the standardization of the curriculum as a result of the end-of-course test was considered an advantage. Mr. Douglas believed that the objectives covered on the test gave teachers, “A guideline of things that we need to teach, a series of topics that need to be covered.” Mrs. Drummond concurred by stating, “I think it has probably had some benefits as far as making sure people are teaching basically the same things.” Mr. Bartholomew also believed that the test brought positive outcomes for students, as stated in the following; “Now students probably get a better overall gist of what biology is about because there are more areas that they are looking at. It is better for the student.” According to Mr. Bartholomew, students do benefit from the test because, “The students get a maximum knowledge of what the state wants them to know. Before there was a test, probably, teachers could just give them what they wanted them to know.” Mr. Adams expressed that the EOC “forced a standardization of curriculum which was probably good in ways” but he also liked diversity of curriculum. He felt that end-of-course testing prevented teachers from teaching the areas of biology in which they were the most knowledgeable, as indicated in the following excerpt.

Each one of us are specialists, each one of us has particular fields of study that we like, and that we really like to impart more information to the kids. I love plants, I love teaching plants. I love teaching ecology. But currently I just skim over both of them because there are no plant questions on the end-of-course test, maybe a leaf. You take what you like and you just PSHHT right by it simply because you have to cover the goals and objectives of the state. And I don't like that aspect of it. We are all put into the same bucket.  
Mr. Adams

Mr. Adams mentioned that without an end-of-course test, he would be able to enjoy teaching biology more because he could spend more time teaching topics that he enjoys. Furthermore, he would be able to impart that love of biology to his students so that they could see how he has carried that love into his own life.

I would just cover things that I enjoy more and I feel like it would make me better in the classroom because the more you enjoy a particular aspect about a

job you do, (the better you are). It is an aspect that I think these kids need to carry out with them out of the class. I feel like a student in my classroom should carry with them something into his life, not to college, not to the next level, but into his life.

Mr. Adams

One Texas teacher, Mr. Tucker, also saw a benefit to having a unified curriculum. He stated, "I think there are some beneficial things to the exam. For instance, it does provide a unified curriculum for the entire state." Furthermore, he mentioned the need to eliminate the district curriculum and replace it with the EOC curriculum and expressed frustration with having to teach both.

When I first came to the district, I was given a book and a 3 ring binder about 4 inches thick and they said, 'here's your curriculum for the entire biology course'. And then, all of a sudden I went to a conference and they said 'by the way, here are the (EOC) objectives' and they are not aligned with the district objectives or with the national curriculum."

Mr. Tucker

Texas teachers at the high-minority high school, Roosevelt High, appeared torn between teaching the curriculum for the end-of-course test and teaching the district curriculum. Ms. Langworthy felt that the administration wanted teachers only to follow the EOC objectives.

The associate principal said, 'Don't take this as a dictate from me but basically we should pretty much throw out the Biology I curriculum and totally focus everyday on those objectives on that (EOC) specification booklet; everything needs to come out of those objectives. So that is what happens at a school like Roosevelt.

Ms. Langworthy

However, Texas teachers at Marshall High School did not appear to change their district curriculum. As Ms. Taylor stated, "They can say, 'take the EOC' but it doesn't have any bearing on my class because unless they give me a curriculum to go with that EOC, I am going to use my curriculum." Mr. Tucker mentioned that he deliberately changed his curriculum. "That is where our curriculum directly goes against this one (the EOC). I have thrown out a lot of things... putting my curriculum basically towards this test." North Carolina teachers did not distinguish between two separate biology curricula, as did the Texas teachers.

### **3. "Moving at the state's pace rather than the student's pace"**

Teachers in both North Carolina and Texas felt they needed to cover many topics to prepare their students for the end-of-course examinations. As a result, the pace of the course was faster than they would have preferred. Mr. Tucker described the pace in his biology course as: "There is a pressure to try to cover, I hate to say cram in, at least present, the information to kids in a hurried type manner as opposed to a natural sequence

the way the course should go.” The “natural sequence” in which students are taught topics that interest them and at an appropriate pace was also expressed by Ms. Murray in the following excerpt.

I feel like I have to rush through a lot of material just so I can have them exposed to it instead of going into the stuff that is really interesting like genetics, and maybe doing more on certain topics in which I think they would be interested. So, I guess the speed and trying to make sure that you cover everything and you know it is more quantity instead of quality at this point.

Ms. Murray

Both Ms. Meyers and Mr. Wilson expressed the added frustration of having to move quickly through the curriculum in classes that are composed of students at different ability levels.

You feel the need to make every moment a teaching moment, which isn't all-bad but grades are dipping because some of the kids can't keep up. I've got kids here with an IQ of 70 in with kids with an IQ of 120. It is not all bad, but when you are outpacing the kids.

Ms. Meyers

If we didn't have the EOC then I feel like we could move along at the classes' pace. I would be more free to experiment and move along at the students' pace rather than the state's pace. I feel like I have to get a certain amount of material covered in a certain amount of time because of the EOC, whether or not I feel like they've mastered it. Because of mainstreaming, we've got people that have below average I.Qs and we've got people that have above average I.Qs and that makes it kind of tough.

Mr. Wilson

Mr. Douglass stated that as a result of the faster pace through the curriculum, he was not able to teach the concepts in detail, and learning of concepts was superficial.

I feel I am not doing as good a job as I could because we're not able to go into depth and talk specifically about certain things, that we just have to sort of brush over it and just get a sketch of it instead of a detailed analysis. A lot of it is superficial coverage because you have so much to do and so little time to do it in.

Mr. Douglass

The need to cover many topics due to the end-of-course examination which resulted in “outpacing the kids” (Ms. Meyers) or “presenting information to kids in a hurried type manner” (Mr. Tucker) existed in both the Texas and North Carolina biology courses.

#### 4. “Teachers have become more adept at teaching a test than teaching knowledge”

Biology end-of-course practice tests and other techniques were used in both North Carolina schools and in Roosevelt High School in Texas to specifically prepare students for the test. Ms. Langworthy prepared practice tests for all of the Roosevelt biology students and in the following excerpt, she described her ambivalent feelings about the practice tests.

Well, my students took the biology EOC today. With 3 practice exams, I had them ready for this test. I figured out what was going to be on the test. Is this valuable to their education? No! It is simply preparing them for a standardized test. It's not getting them to know about abstract biological thought. It's not getting them to think about their world or make them scientifically literate. It's trying desperately to get them to pass this standardized test so Roosevelt HS doesn't look bad and my teaching doesn't look bad.

Ms. Langworthy

The preparation and use of practice tests also occurred in North Carolina as indicated in the following excerpt.

I'll have on Wednesday, Thursday, Friday, and Monday, four and a half days of just literally just like cram review. Um, what I'm doing right now is I made up what I call end-of-course practice test and it covers everything from the first six weeks and it has diagrams and questions that are asked like they are on the end-of-course test. You know, look at this picture and tell me this about it.

Ms. Henderson

The value of preparing students for the test or “teaching to the test” was summed up by Mr. Wilson in the following excerpt.

I think teachers have become more adept at teaching a test than teaching knowledge. You know, if I knew what the questions were on the EOC, I would just teach them. And I don't think that is right. We've gotten away from a lot of hands-on stuff that I thought benefited the students and now we are not doing it because we're more concerned with whether or not they are going to pass a simple test at the end of the year. I think that is wrong. It's wrong.

Mr. Wilson

Other teachers in North Carolina chose not to use practice tests and instead taught their students relaxation techniques as illustrated in the following excerpt.

Before I give a test I teach them things like, sit up, relax, go someplace else mentally that's pleasant just for a few minutes so you can feel the tension going out of you. If you want to put a piece of candy or something in your mouth to

relax that is fine. If I see them, like pondering, I get them to move on, move on. And that's the only thing I do with my kids. That other stuff, I don't bother with it.  
Ms. Ford

Teachers at Roosevelt High School in Texas also taught graphing and table interpretation skills to prepare their students for the end-of-course tests. Mr. Tucker felt that these skills were important for them and he may not have included them in his lessons without the end-of-course test. He stated, "A positive for me on the EOC is the fact that I am now teaching graphing quite a bit. That is good for them to know that."

##### **5. "We used to do leaf projects"**

A major perception in both North Carolina and Texas teachers was their frustration in not being able to conduct hands-on activities such as projects and laboratory and field investigations with their students due to lack of time and coverage of information for the end-of-course test. Their frustration is seen in the following excerpt.

I would like to do more hands-on stuff, but because of lack of time and so much stuff that I have to cover, I don't have as much time as I would like to do those kinds of things. I would do more experimentation, more outside activities and more projects if we didn't have the test.  
Mr. Bartholomew

Mr. Wilson described a hands-on project that he does not do any more due to the EOC.

We used to do leaf projects and we used to do, believe it or not, an ecology lab where we would go down and mark off a square, you know a yard square and you write down what is found in each little square.  
Mr. Wilson

Both Ms. Henderson and Ms. Langworthy mentioned feeling "pressured" by the test which prevented their use of laboratory activities in the classroom as indicated in the following excerpts.

You feel pressured to cover absolutely as much as possible and so sometimes it gets pretty boring because you end up doing worksheets and notes trying to cover the information when you could be doing other things. But those other interesting lab activities take more time.  
Ms. Henderson

Ms. Langworthy mentioned in her comment a specific laboratory exercise that she feels has been eliminated due to the pressure of the test.

With the pressure of the test, it makes it very difficult to do the laboratory work. One kid who wanted to dissect fetal pigs said, 'We never do labs, I thought this was a lab course.' It made me feel really bad.  
Ms. Langworthy

However, Texas teachers at Marshall High School indicated that their use of laboratory activities had not changed as a result of the test. It appears from the following excerpt that the Marshall HS biology curriculum had not been affected by end-of-course testing.

I don't really know how to prepare them beyond doing what we do, we do hands-on activities, we do lab activities... we try to go in-depth on topics so that they have an understanding of it so that they can think when presented with a problem.  
Ms. Spencer

## **6. Student interest in biology**

### **“Why can't we do that? Why can't we go outside as much?”**

Teachers from both North Carolina and Texas mentioned that their biology courses could not be responsive to student interests due to end-of-course examinations. As indicated in the following, teachers could not teach topics that may interest students, such as Venus Flytraps, due to time restraints and the need to devote time to end-of-course test topics.

Today a couple of girls said, 'We want to grow germs'. I mentioned last week about the inoculating loop [for the EOC] and I saw that some of them were distraught that they just saw a picture of it, and that they don't actually use it. But you know, its just squeeze, squeeze, squeeze and all the stuff we have to know.  
Ms. Langworthy

Ms. Henderson mentioned that the students' prior knowledge about science is not being tapped since class time is not spent on topics that might interest students.

I don't get to spend time on things like animals and plants that the kids might truly be interested in and have some prior knowledge about because I'm too busy trying to teach. Okay, well, the cell has these fourteen things in it and you have to know every single one.  
Ms. Henderson

Ms. Henderson described, in the following excerpt, the effect of not being able to discuss topics that interest students.

They're not seeing how science affects them. We are not giving them the chance to go down an avenue that later in their lives might interest them because we're so bogged down with, you know, we have to do this.  
Ms. Henderson

Mrs. Underwood mentioned, in the following excerpt, how interested her biology students were about the investigations in her non-EOC environmental science class and her inability to exploit that interest.

My biology students will see me getting ready for my environmental science class (not end-of-course tested) or they will see something on the board and will say ‘Why can’t we do that? or Why can’t we go outside as much? or Why don’t we do that? It (EOC) has kept me from being as creative as I would like to be. I would rather teach them about plants and animals and things that they are interested in.  
Mrs. Underwood

## **Discussion**

We have determined from our study of teacher perceptions in North Carolina and Texas that biology end-of-course testing has specific effects on biology curriculum and instruction. Some of the effects appear to be beneficial to the biology classroom. These include helping the teachers to keep on task and having a unified biology curriculum so that students “get a better overall gist of what biology is about” (Mr. Douglass). However, most of these effects are detrimental to students being able to understand biological concepts according to the NSES (NRC,1996). Furthermore, the learning environments that result from these effects are incongruent with those of classrooms in countries whose students have demonstrated excellence in their understanding of science (TIMSS, 1996, NRC, 1999).

### **Effects**

The pace of the day-to-day biology instruction is increased by the presence of an end-of-course test. Although two North Carolina teachers thought it was beneficial to “keep on task”, teachers from both states felt that that they had to “move at the state’s pace rather than the students’ pace” to cover all the topics that could appear on the test. Teachers’ comments about covering these topics in a “hurried manner” (Mr. Tucker) or “rush through a lot of materials” (Ms. Murray) specifically indicate the increased pace and increased number of topics in EOC courses. These teacher perceptions add further support to the TIMSS results that stated U.S. teachers cover too many topics. The TIMSS results specifically addressed the “rapid movement from one topic to another” (NRC, 1999, p.11) in U.S. classrooms, that result from some form of standardized testing in science.

There is less detailed analysis and more superficial coverage of topics in biology courses with end-of-course tests. A remark such as “We just have to sort of brush over it and just get a sketch of it instead of a detailed analysis” (Mr. Douglass) directly points out the inability of teachers and students to study biology in detail. Furthermore, teachers are impaired in using their judgments about student understanding to make changes in the presentation of concepts as illustrated in the following comment. “I have to get a certain amount of material covered in a certain amount of time because of the EOC, whether or not I feel they’ve mastered it” (Mr. Wilson). These comments about the effects of end-

of-course testing may provide reasons for the TIMSS finding that “U.S. instruction may be more superficial than in other countries, with students failing to acquire a deeper understanding of any particular topic (NRC, 1999, p.11).

“Teaching to the test” and/or test preparation activities are emphasized in courses with biology end-of-course testing. The most startling practice of this in both states, in terms of end-of-course testing, is the preparation and use of practice tests. For example, a comment from North Carolina teacher was “I made up what I call an end-of-course practice test” (Ms. Henderson). A remark from a Texas teacher was “With 3 practice exams, I had them ready for this test.” (Ms. Langworthy). Practice tests were not used at Marshall High School, which is a low-minority high school. Instead, Ms. Spencer mentioned, “We do hands-on activities, we do lab activities, and we try to go in-depth on topics so that they have an understanding of it so that they can think when presented with a problem.” Her curriculum was not affected by the biology end-of-course test, a phenomenon that is typical in classrooms of primarily white students (Madaus, 1994). Her students always performed well on the end-of-course test so there was no reason to change her curricula or instruction.

Teachers in courses that have biology end-of-course tests may set aside time in the biology curriculum to have students practice taking a test. This use of class-time for test taking skills rather than biology content and scientific thought appeared to disturb some of the teachers in our study, as indicated by the following statement about practice tests. “It’s not getting them to know about abstract biological thought. It’s not getting them to think about their world or make them scientifically literate. It’s trying desperately to get them to pass this standardized test.” (Ms. Langworthy) That teachers “teach to the test” is a well-known consequence of standardized testing (Shepard, 1991; Madaus, 1991; Herman & Golan, 1992). For the most part, if students need to take a test, teachers will emphasize concepts that have been on previous tests or the test objectives which causes students to develop only a superficial understanding of the subject (Stake, 1991). However, the overall long-term effects of this change in emphasis in the curriculum, as shown in this study, may cause U.S. students to fail to develop a deep understanding of biological topics. This is revealed later in mediocre performances on national and international science assessments (NRC, 1999, NSTA, 2001, Bourque, Champagne, Chrisman, 1997). Furthermore, that U.S. students have only superficial understanding of science is also revealed when students take brand-new standardized tests in science and their achievement scores dramatically drop in comparison with previous years (Bracey, 2000).

In courses wherein students undergo end-of-course tests, laboratory, field investigations and hands-on activities are de-emphasized in order to allow time for coverage of more topics to prepare students for tests. As Mr. Bartholomew said, “I would like to do more hands-on stuff, but because of lack of time and so much stuff to cover, I don’t have as much time as I would like to do those kinds of things.” Teachers reported that they have their students complete “worksheets and notes” or “written work” to cover the information in the time they are given. Comments such as Mr. Wilson’s that “We used to do leaf projects” indicates that a shift in curricular emphasis has occurred from hands-on activities to more vocabulary building activities. This shift is contrary to what is known about effective science teaching. The NSES specifies that students need to be provided with inquiry experiences in order to better understand and retain scientific



concepts (NRC, 1996). Furthermore, science teachers in other countries that use inquiry experiences and laboratory investigations to teach scientific concepts have students that perform well on international science assessments such as TIMSS (NSTA, 2001).

Students are less able to explore their interests in science in courses with end-of-course testing. As pointed out in the opening quotations, Venus Flytraps and fetal pigs, which interest students, are generally omitted due to time constraints. One North Carolina teacher described the level of interest in her class as follows. "Sometimes it gets pretty boring because you end up doing worksheets and notes trying to cover the information when you could be doing other things." (Ms. Henderson) "Doing worksheets and notes" to learn new vocabulary does not encourage curiosity about scientific issues and actually makes students less motivated to learn about science (NRC, 2000). The following excerpt describes science education that is driven by wide coverage of concepts.

Science classes of this type treat education as if it were preparation for a quiz show or a game of trivial pursuit. Most students are not interested in being quiz show participants. They fail to see how this type of knowledge will be useful to them in the future. They therefore lack motivation for this kind of "school learning." (NRC, 2000, p.xii)

Students in science courses with end-of-course test are less likely to conduct scientific inquiry activities because test preparation activities are often the major focus. The inquiry process as defined by the NSES is as follows:

Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. (NRC, 1996, p.23).

Teaching students science through scientific inquiry instead of through worksheets allows students to "deepen their understanding of scientific knowledge through observing and manipulating conditions in the natural world." (NRC, 2000, p.13). It is based upon what students already know and are interested in and what they learn using the inquiry process. Effective science teaching begins with scientific inquiry. (NRC, 1996). In this process, students are able to connect with science by asking their own questions and exploring answers to those questions. As one North Carolina teacher said about students in her end-of-course biology class, "They're not seeing how science affects them. We are not giving them the chance to go down an avenue that later in their lives might interest them." (Ms. Henderson) Germs, plants, such as Venus Flytraps, and animals may not be on the state curriculum or on the end-of-course objectives but students may develop a deeper understanding of scientific concepts and interest in science through their explorations of them.

### **Implications for Science Education**

This study is grounded in the voices of biology teachers. It is through their separate and collective voices that we hear their concerns about the effects of biology end-of-course testing on curriculum and instruction.

It is the belief of the researchers, after extensive analysis of our data, that these end-of-course tests, with their associated sanctions, are affecting the biology curriculum and instruction in ways that are contrary to effective science teaching practices as determined by the NSES (NRC, 1996) and the TIMSS (NRC, 1999). Reports from these national and international studies stress the importance of scientific inquiry, laboratory and field investigations and instruction that is geared towards students developing a deep understanding of science. (NRC, 1996, NRC, 1999). In particular, biology courses with end-of-course testing; 1) cover too many topics, 2) move at a rapid pace that is determined by the number of topics rather than student understanding, 3) replace biology curriculum instructional time with practice tests and other test preparation activities, 4) diminish emphasis on laboratory and field investigations and 5) are not based on scientific inquiry or student interests in biology.

It might be true that an end-of-course test may force some teachers who are not “keeping on task” to improve. But at what cost is that improvement? Those struggling teachers should be helped to improve their practices without adversely affecting those teachers that are already teaching science effectively. Due to the importance placed on end-of-course tests by the state and local school districts, teachers become forced to focus their creative energies on improving their students’ performance on one multiple-choice test. A more productive and beneficial use of those energies would be to help foster deep understanding and appreciation for science in their students. Even when teachers realize that some students have not mastered certain concepts, they must continue to race forward in an attempt to cover all of the material. Furthermore, teachers are forced to ignore the interests of their students in an effort to produce high end-of-course test scores. Students’ questions about the biology of things living near them such as the native North Carolina Venus Flytraps have to be put aside.

End-of-course testing in science, as determined by our study, is not compatible with effective science teaching. Science teachers, as stated in national science education standards such as the NSES, must be allowed to engage their students in science through the students’ interests in their local environment and surroundings. As indicated earlier, Falk and Larson (1996) reported many positive gains when students are actively engaged in their learning. Students also must be allowed to encourage the exploration of student questions and be allowed to set the pace of the course based upon student understanding (NRC, 1996). This approach may not produce high end-of-course test scores but it will make our students more scientifically literate and hopefully, better citizens.

## References

- Bogdan, R. C., and Biklen, S. K. (1998). Qualitative research in education: An introduction to theory and methods. Needham Heights, MA: Allyn and Bacon.
- Bourque, M.L, Champagne, A.B. Chrissman, S. (1997). 1996 Performance Standards; Achievement results for the nation and the states. Washington, D. C: National Governing Board
- Bracey, G. A. (2000, October). The 10<sup>th</sup> Bracey report on the condition of public education. Phi Delta Kappan. 133 – 144.
- California Department of Education (2000). *Golden State Examinations*. Retrieved 2000, from <http://www.cde.ca.gov/statetests/>.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. Harvard Educational Review (64) 1, 5-30.
- Darling-Hammond, L. (1991, November). The implications of testing policy for quality and equality. Phi Delta Kappan, 220-225.
- Douglass, B.G. & Moustakas, C. (1985). Heuristic inquiry: The internal to know. Journal of Humanistic Psychology (25)3, 39-55.
- Edwards, V. B. (1999). Quality Counts '99. Rewarding results, punishing failure. (Education Week/Pew Charitable Trust Report). 85. Bethesda, MD.
- Erickson, D. (2001, June 22). Test numbers perplex officials. Wisconsin State Journal, A1, A7.
- Falk, B. & Larson, J. (1996). An invitation to invention: Top-down support for bottom-up reform of assessment in New York State (Report No. EA 027 403). New York: National Center for Restructuring Education, Schools, and Teaching. (ERIC Document Reproduction Service No. ED 393 158)
- Gong, B. (1990, April). Current state science assessments: Is something better than nothing? Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.
- Goodland, J.I. (1984). A place called school: Prospects for the future. New York: McGraw-Hill.
- Harmon, M., Smith, T., Martin, M., Kelly, D., Beaton, A., Mullis, I., Gonzales, E., & Orpwood, G. (1997). Performance assessment in IEA's Third International Mathematics and Science Study (TIMSS). Chestnut Hill, MA: Boston College.

- Harris, S. (1998). Achievement in practical science activities at age 13 in England. Education in Science. 10-13.
- Herman, J. & Golan, S. (1992). Effects of standardized testing on teaching and learning CSE Technical Report 334. National Center for Research on Evaluation, Standards and Student Testing, University of California at Los Angeles.
- Jerald, C. & Boser, U. (1999, January). Taking stock. Quality Counts '99 (Education Week/Pew Charitable Trust Report) 81. Bethesda, MD.
- McKnight, C., Crosswhite, F., Dossey, J., Kifer, E., Swafford, S., Traver., K., Cooney, T., (1987). The underachieving curriculum: Assessing U.S. school mathematics from an international perspective. Champaign, IL. Stipes.
- McLaughlin, M. (1991, November). Test-Based accountability as a reform strategy. Phi Delta Kappan 248-251.
- Madaus, G. F. (1988). The influence of testing on curriculum in *Critical Issues in Curriculum* 87<sup>th</sup> Yearbook of the National Society for the Study of Education. Chicago, IL: University of Chicago Press.
- Madaus, G. F. (1991, November). The effects of important tests on students. Implications for a national examination system. Phi Delta Kappan , 226-231.
- Madaus, G.F. (1994). A Technological and historical consideration of equity issues associated with proposals to change the nation's testing policy. Harvard Educational Review 64 (1). 76-95.
- Medrich, E.A., & Griffith, J.E. (1992). International mathematics and science assessments: What have we learned? (Report No. NCES 92-001) Washington, D.C.: U.S. Department of Education.
- Morgenstern, C.F., & Renner, J.W. (1984). Measuring Thinking with Standardized Tests. Journal of Research in Science Teaching, 21 (6), 639-648.
- National Assessment of Educational Progress (NAEP) (1998). *NAEP 1998 National and State Writing Summary Data Tables for Grade 8 Students Data Weighted Percentages and Average Scale Scores*. Retrieved from, <http://nces.edu.gov/nationsreportcard/TABLE/SDTTOOL.HTM>
- National Assessment of Educational Progress (NAEP) (2000). *Science 2000 Major Results*. Retrieved 2000 from, <http://nces.ed.gov/nationsreportcard/science/>
- The National Center for Fair & Open Testing [NCFT]. (Summer, 1997). *Testing our children: A report card on state assessment systems*. Retrieved 2000, from <http://fairtest.org/states/intro.htm>.
- National Educators Goals Panel. (1999). *All Indicators for the Nation*. Retrieved 2000, from [http://www.negp.gov/cgi-bin/broker.exe?\\_program=goalsprg.pg150.sas&\\_service=goals&\\_debug=0](http://www.negp.gov/cgi-bin/broker.exe?_program=goalsprg.pg150.sas&_service=goals&_debug=0)

- National Research Council (NRC). (1996). National Science Education Standards. Washington, D.C.: National Academy Press.
- National Research Council (NRC) (1999). Global perspectives for local Action. Using TIMSS to improve U.S. mathematics and science education. Washington, D.C.: National Academy Press.
- National Research Council (NRC). (2000). Inquiry and National Science Education Standards: A guide for teaching and learning. Washington, D.C: National Academy Press.
- National Science Foundation. (1992). The Influence of Testing on Teaching Math and Science in Grades 4-12. (Report: Center for the Study of Testing, Evaluation and Educational Policy). Chestnut Hill, MA.
- National Science Foundation. (1996). Chapter 7: Science and technology: Public attitudes and public understanding, Science and Engineering Indicators 1996, 3-21. Washington, D.C: U.S. Government Printing Office.
- National Science Teachers' Association (NSTA) (2001, May, June). TIMSS-R Benchmarking results a 'Mixed Bag' for eighth-grade achievement. NSTA Reports! 12 (6) 1-3.
- National Science Teachers' Association (NSTA) (2001/2002, December, January). NAEP 2000 Science Scores Show no Significant Change Since 1996. NSTA Reports! 13 (3) 1-3.
- National Science Teachers' Association (NSTA). (2002, February, March). News. NSTA Reports! 13 (4) 25.
- North Carolina Dept. of Public Instruction (NCPDI) (1999). *The 2000-20001 North Carolina Testing Program Overview*. Retrieved 2000, from <http://www.ncpublicschools.org/accountability/testing/policies/TestProg0001.htm>
- Olson, L. (1999, January). Shining a spotlight on results. Quality Counts '99 (Education Week/Pew Charitable Trust Report).Bethesda, MD. 8.
- Patton, M.Q. (1990). Qualitative evaluation and research methods. (2nd Ed.) Newbury Park, CA: Sage Publications.
- Popham, J. (1987, May). The merits of measurement-driven instruction. Phi Delta Kappan. 679-683.
- Porter, A. (1989). External standards and good teaching: The pros and cons of telling teachers what to do. Educational Evaluation and Policy Analysis. Vol. 11, No. 4, pp. 343-356.
- Portner, J. (1999, January). Reality Strikes. Quality Counts '99 (Education Week/Pew Charitable Trust Report).Bethesda, MD. 181.

- Resnick, L.B. & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B.R. Gifford & M.C. O'Connor (Eds.), Changing assessments: Alternative views of aptitude, achievement and instruction, (pp. 37-75). Boston: Kluwer
- Sagan, C.A. (1995). Science as a candle in the dark. The demon-haunted world. New York: Random House.
- Shepard, L.A. (1991, November). Will national tests improve student learning? Phi Delta Kappan. 232-238.
- Smith, P.S., Hounshell, P.B., Copolo, C., and Wilkerson, S. (1992). The Impact of End-of-Course Testing in Chemistry on Curriculum and Instruction. Science Education, 76 (5), 523-530.
- Stake, R.E. (1991, November). The teacher, standardized testing and prospects of revolution. Phi Delta Kappan. 243 – 247.
- Texas Education Agency (TEA). (1993). Statewide Accountability System. An overview of the accreditation procedures as revised by Senate Bill 7. (Report: The Office of Accountability). Austin, TX: Author.
- Texas Education Agency (TEA) (1994). Biology I objectives and measurement specifications (Report: Division of Student Assessment). Austin, TX: Author.
- Texas Education Agency (TEA). (1995a). Briefing book on Senate Bill One 74<sup>th</sup> Texas Legislature (Report Texas Education Agency) Austin, TX: Author.
- Texas Education Agency (TEA). (1995b). Texas Education Agency Academic Excellence Indicator System 1993-94 Campus Profile. (Report: Division of Performance Reporting, Office of Policy Planning and Information Management). Austin, TX: Author.
- Texas Education Agency. (1996a). Table 4: Blueprint of 1996-2000 Accountability Systems. (Report: Office of Accountability). Austin, TX: Author.
- Texas Education Agency (TEA). (1996b). Texas Education Agency Academic Excellence Indicator System 1994-95 Campus Profile. (Report: Division of Performance Reporting, Office of Policy Planning and Information Management). Austin, TX: Author.
- Texas Education Agency. (1997). Texas Student Assessment Program Technical Digest For the Academic Year 1996-1997. Austin, TX: Author
- Texas Education Agency (TEA). (2002). *Texas Education Agency - Student Assessment Division* Retrieved 2000 from, <http://www.tea.state.tx.us/student.assessment/>

Westerlund, J.F., & West, S.S. (December, 2001). The Use of the National Science Education Standards to Critique a Standardized High School Biology Examination. Electronic Journal of Science Education, 6, (2). Retrieved 9/02 from, <http://unr.edu/homepage/crowther/ejse/westerlundetal.html>.

Third International Mathematics and Science Study. [TIMSS] (1996). TIMSS Report, National Center for Education Statistics, Washington, D.C.: United States Department of Education.

## **Appendix A Teacher Profiles**

### Roosevelt High School Teachers (TX)

The Roosevelt teachers were Mr. Tucker, Ms. Meyers, and Ms. Langworthy. Ms. Meyers and Ms. Langworthy had taught biology for 11 and 8 years respectively whereas Mr. Tucker had taught biology for 3 years.

Ms. Meyers was a favorite biology teacher at Roosevelt and had been voted Teacher-of-the-Year three times by the students. She was noted for connecting the big ideas in science and trying to emphasize the relevance of science to students' daily lives. She preferred to teach abstract ideas like protein synthesis on a basic level and then link it to other concepts later, such as the HIV life cycle or digestion.

Ms. Langworthy, who held an advanced degree in genetics and taught genetics classes at the local community college, was noted for being innovative in using different laboratory and field activities to get the students interested in biology. She conducted lengthy projects to encourage student interest in science such as use of Wisconsin Fast Plants, the development of a large garden behind the school, and biotechnology experiments such as extraction and gel electrophoresis of cat teste DNA.

Mr. Tucker believed that biology should be fun so that his students would want to come to class and learn. He was well known for his excellence as a storyteller. Students were mesmerized as he related his experiences from hospital work to his travels through Costa Rican jungles, to the beaches of Oaxaca, Mexico. In his second year of teaching, he changed his emphasis from doing laboratories to using controversial discussions or debates so that students could relate scientific concepts to their own experiences.

### Marshall High School Teachers (TX)

The Marshall High School teachers were Ms. Spencer who had taught for 10 years and Ms. Taylor who had taught biology for 4 years.

Ms. Spencer believed that biology teaching meant making connections to the real world. She stated, "I rarely use the book. I am very much into articles and what is happening in the news...how the principles in biology apply to everyday life so kids can see how it affects them in their world." She was a leader in the district biology teaching community and was one of six teachers elected to rewrite the district curriculum, which was theme-based and not textbook-based. She led a six-day institute for other biology teachers in the district so they could become familiar with the curriculum.

Ms. Taylor's biology teaching was characterized by her commitment to science fair projects. Often, her students spent long hours after school with her working on their experiments. Her approach to teaching biology was based on questioning after an introduction of basic concepts. She stated, "Many times, if a student asks a question, I respond with another question leading them to the answer. This can be frustrating for students because they are used to being given information." She preferred to do actual laboratory and field investigations instead of paper activities. For example, she had a drilling company open a cave beneath the high school so that her students could explore it.



### Anderson High School Teachers (NC)

Ms. Henderson was a first year teacher who felt that it was important to enrich the science curriculum with topics that were relevant to students such as the “odd little Venus Flytraps”. She was excited by her new career and repeatedly reported that she “loves her job.” One of Ms. Henderson’s principle beliefs was that it was necessary to develop an atmosphere in her classroom where the needs of all learners were met. She felt that it was important to use a diversity of teaching methods and assessment strategies to meet this goal. Incorporating activities that allowed students to get outside and scheduling guest speakers were strategies that she used to reach her students and their interests. She recognized that all of her students would not become scientists and therefore felt that they should have a background that allowed them to be knowledgeable about environmental and other issues that would directly impact their decisions as citizens. She wished that she could do more laboratory activities but felt pressured to teach the vocabulary that would be on the test. She did not know whether she would continue teaching biology.

Ms. Murray was a four-year veteran science teacher who had spent her career at one school. She believed that students should develop a depth of science knowledge in topics that interested them, instead of a superficial breadth of knowledge about a myriad of topics. She believed that group work and outside resources were vital to success in a science classroom. Inquiry teaching methods were favorites of Ms. Murray.

Ms. Drummond was a first year teacher excited by the possibilities of her new career. Her primary objective was to make her biology classes relevant and interesting for the students. Plants and animals were topics that she felt would be of particular interest to her students. She wanted to have her students involved in many hands-on activities. She felt that classrooms where the teacher constantly lectured and relied on worksheets were boring.

Ms. Underwood was a participant in the Teach for America Program. Although she had two years of experience teaching environmental education, this was her first teaching experience in a public school setting. She wanted to expose her students to the political and societal influences that affected science and how scientific information was used. Making science relevant to the lives of her students and including topics that interest them were important objectives for Ms. Underwood. She liked to use projects to allow her students to pursue scientific topics that were of interest to them.

### Bluffton High School Teachers (NC)

Ms. Ford was a twenty-year veteran teacher who had spent her entire career in the same school teaching different levels of biology and anatomy and physiology. She felt that teaching was an important job and wanted to see all teachers striving to achieve teaching excellence with their students. In her teaching, she liked to include videos and media articles that focused on current events in science. Helping her students learn to think was an important objective for this teacher. In addition, she liked to assess the level of knowledge of her students and then help them to master areas where their content knowledge was weak. She liked to include many experiments that built on the background knowledge that her students possessed.

Mr. Bartholomew was a second year teacher with experience in biology, physical science, and health and physical education. In his class, he focused on the application of

scientific knowledge and helping students be able to “use what they know.” He used hands-on activities and projects to help reach this desired end.

#### Connor High School Teachers (NC)

Mr. Wilson was a twenty-year veteran teacher who had experience in K-12 education. He has taught biology, anatomy and physiology, health, physical education, and physical science during his career. At his current school, he was voted favorite teacher by the senior class on several occasions. Mr. Wilson was concerned about how the lives of his students outside of school affected what occurred during the school day. He used a diversity of methods including, projects, dissections, and labs to help his students learn the material. He also included current event topics in his science classes to increase student interest in the subject.

Mr. Douglas was a six-year veteran science teacher. He felt that he had established good rapport with his students and was able to gear instruction to topics of interest for his students. Mr. Douglas liked to implement many experimental approaches to teaching in his classroom. He strived to continually modify his teaching practice and was interested in using in-depth, and hands-on techniques in his science teaching.

Mr. Adams’ professional career spanned 30 years, which included 6 years in the private sector and 24 years teaching science. His vast experience included teaching biology, ecology, botany, and physical science at both the high school and junior high levels. Mr. Adams liked to relate his science teaching to the “real life” of the students. He believed in diversifying his curriculum to address both the interests of the students as well as the knowledge that he possessed of his own favorite topics in science.

#### *About the Authors...*

**Julie F. Westerlund** has eight years of experience as a high school science teacher and nine years as a college and university science teacher in Texas. Dr. Westerlund’s research interests include inquiry-based science teaching, standardized testing, earth science education, and science teacher professional development.

**Leslie K. Upson** has four years of experience as a high school science teacher in North Carolina. She is currently a doctoral student in science education at University of Georgia at Athens.

**James P. Barufaldi**, is a Ruben E. Hinojosa Regents Professor in Education and the director of the Science Education Center at The University of Texas at Austin. Dr. Barufaldi has authored or co-authored more than 60 articles, books, chapters, and book reviews and has presented more than 300 workshops, papers, and seminars throughout the U. S., and in countries such as Portugal, Russia, Japan, Israel, Bermuda, Iran, Barbados, Costa Rica, Mexico, Australia, Belize, Canada, Iceland, Finland, Panama, Taiwan, and Trinidad. Barufaldi’s special areas of interests are curriculum design, implementation, evaluation, and science teacher education. He is currently investigating the process of building successful collaboratives in the science education community and variables, which may contribute to high intensity, sustained collaboration.

