SCIENCE IS AS SCIENCE DOES:

ALIGNING TEACHING PHILOSOPHY, OBJECTIVES, AND ASSESSMENT

by

MATTHEW D. KREHBIEL

B.A., Bethel College, 1999

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Curriculum and Instruction
College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2006

Approved by:

Major Professor
Lawrence Scharmann, Ph.D.

Copyright

MATTHEW D. KREHBIEL

2006

Abstract

In the current high-stakes assessment environment of NCLB and AYP, the type of assessment students are subject to is crucial. Good assessments make teaching to the test an effective, rather than deplorable, strategy that aligns learning objectives, classroom teaching, and student achievement. While science, this author's content area, isn't currently AYP mandated, it will be soon, and in the meantime, students are still required to take state and local assessments. This project seeks to effect change on a local level – specifically, with how the local CRT is written so that it can both reflect sound research-based assessment practice as well as an understanding of science as concept- and process-based rather than content- and fact-based.

Table of Contents

List of Figures

List of Tables

Preface

No Child Left Behind (NCLB) has been at the center of controversy in education since well before it was passed into law in 2001. The intention to educate *every* student inherent in the very name of this legislation is honorable and surely something to which passionate teachers aspire; furthermore, increasing the accountability of the public school system to reach this lofty goal seems desirable. However, lurking beneath the glossy surface of NCLB is an assessment monster. The only measure of whether or not students are being "left behind" is their performance on state assessments. The monstrosity is not that assessment is inherently a bad thing, even in a high-stakes environment. However, if the results of these assessments are going to be used to make important decisions about the futures of schools, they had better be incredibly well formulated. Poorly constructed assessment inhibits good teaching. It restricts teachers from developing innovative teaching methods that dig into the marrow of what we really want students to learn because it might be all for naught if the assessments don't cover such material. Good assessment, on the other hand, creates good instructional practice or at least allows it to be developed. If the test is really assessing student learning that matches the learning objectives, then "teaching to the test" is transformed from a deplorable, mind-numbing necessity into a coalescence of exemplary teaching methods and student achievement.

Judging from the reaction of many educators, the current state assessments leave much to be desired. A quick search of the Wilson Web educational full-text online database for NCLB will include results such as "NCLB's Murky Mess," "Is NCLB Falling Apart at the Seams?" "Are We Leaving our Children Behind?" "Clean up the Test Mess," "A Muddle of NCLB," "Preparing Teachers to Beat the Agonies of NCLB," and more along the same lines. Many teachers are downright hostile to what they see happening as teachers restructure their classes to fit flawed assessments. Bobby Ann Starnes, the executive director of Full Circle Curriculum and Materials, puts a clear voice to this perspective in a recent Phi Delta Kappan article when she writes, "…the narrow, restrictive programs being used to teach reading as a result of NCLB enforcement will have consequences beyond creating a generation of kids who hate [the content]. These programs require students to give quick right answers that fall within a very narrow range of possibilities. One of my scientist friends [….] points out that these programs will effectively 'select against' students' developing the traits most conducive to scientific thinking" (2006, p. 635). Anne Lewis, a national educational policy writer, adds, "The most critical fault of the policy making [….] is that the 'flexible' responses to the AYP threat of sanctions have involved game playing and dumbing down, not decisions that created better learning situations for students" (2006, p. 643). In other words, the drive for accountability has counter-intuitively led to lower standards and expectations for students.

It would be foolish for me, a high school science teacher, to pretend to be objective on this subject. Though the science assessments are not yet a part of Adequate Yearly Progress (AYP) expectations of NCLB, they are scheduled to be added to AYP by the end of this school year and science education has been indirectly affected by the shift of resources to the subjects (reading and math) that are already required to reach AYP (Starnes, 2006). I have personally watched colleagues in math and English work as hard as they can to get their students to improve their state assessment scores. Though the teachers at Junction City High School have made significant strides to raise students' scores, many of them have paid the heavy price of losing their passion for teaching. Time is at such a premium to cover the content necessary to prepare for the assessments that there is little opportunity for activities and projects that are expressions of a teacher's creativity and that provide opportunities for students to really embrace their education. As a science teacher I have simultaneously lamented the shift of resources away from science and rejoiced that I don’t have to limit my curriculum to a test that I think limits both me and my students and that I don't think is adequately assessing my students.

It is in this context—my dissatisfaction with the current state of assessments and my looking ahead to the addition of science to AYP—that I was evaluating the biology test that our district developed for the years in between the state assessment (Kansas state science assessments traditionally have been given every other year). I came to several realizations. The first realization was that this test was not focusing on what I thought was important for science students to learn. The second realization was how much the assessment affects how and what content is focused on in the classroom. I was reviewing the test to aid preparation of my semester review lesson plans when I came across the question, "To what order do humans, apes, and monkeys belong?" My immediate reaction to this question—which I actually said out loud—was, "Knowing how my students answer this question will tell me absolutely nothing other than whether or not they have memorized the fact that these three groups of animals belong in the order primates." I have absolutely no interest in running a classroom with the implicit expectation that assessment is based on how well you can memorize. I am much more interested in assessing if my students understand bigger-picture science, as in this case, the underlying concept of how and why scientists classify organisms. Determined that my students would be prepared for the test, I agonized over a way to review this content without just telling them that they needed to know this fact for the test, but to no avail. Depressed, I ended up re-wording the question and including it in the semester review Jeopardy game.

It was at this point that I decided that I needed to do something to improve the district test. Though affecting the state test seems beyond my influence, this local assessment was something right in front of me that I could do something about. Newly armed with better mental tools for evaluating assessment through my master's program and motivated with the thought of improving my teaching by better assessment, I set out to at least take the first steps in making this assessment what I think it needs to be. If a poor assessment can negatively affect what is presented in the classroom, a good assessment should do just the opposite.

This report details the first few steps in the journey to develop an assessment that challenges me to continually develop new ways to teach my students that science is as science does. Science *isn't* a catalog of facts that simply needs to be memorized; rather, it is an intricate web of process and inquiry that is defined and determined by how scientists do what they do. Science, in many ways, functions as a verb just as much as it does as a noun, and students of science will do well to learn this early on. Specifically, to draw once more on the monkey question, memorizing order isn't the process of taxonomy; more broadly, teaching scientific knowledge as a multiple-choice set of facts ignores the process of science. The process of science warrants more attention in our assessment of science students.

CHAPTER 1 - Giants' Shoulders

If I have seen further, it is by standing on the shoulders of giants.

--Sir Isaac Newton

It is humbling that Sir Isaac Newton, a man who made significant contributions to our current understandings of gravity, motion of objects, optics, and calculus to name a few, attributes his success to the accumulation of scientific knowledge by the scientists who preceded him. Rather than berating those who had proposed inferior explanations in the past, he credits them with enabling him to "see further." This is the very essence of the process of science: it relentlessly marches on to be more detailed, more explanatory, and more insightful—always building on what has come before even when it might seem to be tearing it down. The time has come for science teachers to stand on the shoulders of those who have been struggling to reform science education for decades and to see further.

During my undergraduate education, I spent a semester in Costa Rica studying tropical biology. There were so many experiences from this semester that reshaped my worldview and my perspective on biology, but one that stands out in this context is the trip that we took to the Insituto Nacional de Biodiversidad (INBio) in Santo Domingo de Heredia, Costa Rica. At INBio, their goal is to both gather knowledge about the biodiversity of Costa Rica and to promote its sustainable use. During our tour of their facility, our guide told us that nearly 80% of the insects that are brought in for identification have never before been named or described by scientists. This example demonstrates the overwhelming amount of information that we still do not know. This is not to say that we don't already know a lot, in fact, the amount that we do know is at least as overwhelming.

Yet, for some reason our biology classrooms still seem to end up as glorified exercises in memorization. Wolfgang Gräber, science researcher at the University of Kiel in Germany, member of the Leibniz-Institute for Science Education, and former high school science teacher, laments that science teachers generally fall to the left on the following continuums:

teacher centered----------student centered

teaching facts-------------teaching process

discipline oriented-------daily life oriented (Gräber, et al., 2001).

This is unacceptable—the sheer volume of what we know about the natural world is already impossible to internalize and there is still so much that we don't yet know. Furthermore, the history of science teaches us to be skeptical of what we know about science and to be continually searching for better explanations of what we know is around us. Yet this skepticism must be balanced with an awareness of what is already known. We want to take advantage of the giants that have gone before us, but we want to end up on their shoulders, not under the soles of their shoes. This balance can only be achieved if the sheer volume of facts that students are required to learn does not prevent students from actually understanding the concepts and processes of science. If students understand what makes science tick, they will not only be better able to apply it to their own lives, but they can also take science further than it has ever gone. We will prepare students who will not only climb on the shoulders of the giants before them, but who will be equipped to look further than we have ever seen before.

Where do we look?

Science educators also need to build on those that have come before us. Standing on a giant's shoulders doesn't help all that much if you're always looking down. We need to look out and up as we continue to improve science education while always keeping in mind what we are standing on and how it enabled us to get where we are today. The call to restructure science education to more explicitly focus on the important process of science is not a new one. The evolution of science education, like science itself, is always building on what we have done in the past and is never satisfied with where we are. Already in 1990, Richard Duschl, Professor of Science Education at Rutgers University, was writing of "…the need to explore the nature of science and scientific inquiry in science classrooms " (p. 5). These ideas have been around for a while; now it is time to take them to the next level. The term "lifelong learner" has been so often used in education for the last ten years that it seems like it is in every mission statement of every district in the United States, yet we are still just teaching students what they should know rather than how to find out what they want and/or need to know. I agree with Duschl that the scientific process is quite similar to the process of learning and that it offers a metacognitive opportunity to empower students to think for themselves (1990).

Duschl was certainly not the first or last one to suggest this shift in focus from detail-oriented memorization of content to a focus on concepts and process. F. James Rutherford and Andrew Ahlgren, in their book *Science for all Americans*, stress that it is essential that everyone be scientifically literate and, in their opinion, such literacy comes from focusing on the nature of science and the core concepts. According to Rutherford and Ahlgren, the brief, shallow instruction that has resulted from so much science content being crammed into a school year has led to many students not understanding the most important concepts in science or the very nature of scientific knowledge is acquired (1990). *Designs for Literacy*, an outgrowth of the American Association for the Advancement of Science's (AAAS) Project 2061, pushes these criticisms of the science curriculum further. AAAS claims that the curriculum content is obsolescent because it does not connect with the everyday life of most students. They claim that the curriculum has become "grossly overstuffed" with topics as entire new fields of science have sprung up like dandelions in the last fifty years. As these new topics are added to the curriculum, rarely is anything taken out. The result is obvious. If you have more topics to cover, but they have to be covered in the same amount of time, each topic gets a shallower treatment.

The shallow, rapid-fire survey of the subject matter limits a teacher's flexibility to work with students with different learning styles and with significantly different background knowledge (AAAS, 2001). A major struggle in teaching science is that students all come with preconceived notions about many of the topics that are covered in a biology class and we often don't have time to really dig out those misconceptions and restructure them if necessary. The result is that the new information, even if inordinately superior to the student's preconceived notions in both empirical support and explanatory power, doesn't stand a chance to be accepted.

The bottom line is, though seemingly ironic, we need to pare back the curriculum in order to take science to a higher level. When there are more content details to study than there is time to cover them, the science classroom becomes a boring, out-moded class for many students. When the content is stripped down to its core concepts, it allows for application of the scientific process to a novel situation instead of just recitation of past results. We need to spend less time talking about the Krebs cycle and more time investigating what is going on in the world around us and modeling how the scientific process can be a powerful tool for everyone to use.

CHAPTER 2 - Defining Science

F R O M Y O U
Poets don't invent poems
The poem is somewhere behind
It's been there for a long long time
The poet merely discovers it.

--Jan Skacel

To capture one definition of science, it would be very easy to replace "poet" with "scientist" and "poem" with "theories" and be done with it: Scientists don't invent theories/ The theory is somewhere behind/ It's been there for a long long time/ The scientists merely discovers it. However, for the sake of clarity, I will try to put into my own words the meaning of science that we science educators need to help students to discover. Any curriculum changes, objectives, or questions that I propose will be somehow affected by my own perspective on what science is and how it works, so it is imperative that it be explained prior to assigning objectives. It is the starting point for creating an assessment that is measuring what needs to be measured.

My frame of reference begins with the idea that all people have some natural scientific propensity—I have yet to see an infant or toddler who is not fascinated in some way with the world around them. People are naturally curious about their surroundings and search out or make up explanations for what they see. Our brains are constantly taking in, classifying, and analyzing information that we organize in order to keep it from overwhelming us. Stephen Jay Gould puts it this way: "We think that science is intrinsically hard, scary, or arcane, and that teachers can only beat the necessary knowledge, by threat and exhortation, into a small minority born with inborn propensity. No. Most of us are born with a love of science [.…] This love has to be beaten *out* of us if we are to fall by the wayside, perversely led to say that we hate or fear the subject" (2003, p. 117). Quite possibly, it is the reduction of science to a set of "facts" that has contributed to the love of science being "beaten out" of many students who claim to loathe the subject. What these students haven't been given a chance to realize, perhaps, is that science is a powerful tool with which to organize many types of information and to give us a better idea of what is happening in the natural world. It is a way of thinking that can help everyone in one way or another. Science is not just for scientists and it is important that all students regardless of gender, ethnicity, or socioeconomic status develop a solid understanding of science and when and how to use it. Having a scientifically literate population benefits everyone. Non-scientists benefit from having a logical method for problem solving that, among other things, can help prevent them from being scammed by those masquerading as scientists. Bruce Alberts, former president of the National Academy for the Sciences and one of the principal authors of *The Molecular Biology of the Cell*, has built a reputation for his commitment to science education. Regarding the importance of science education for all students, Alberts writes, "…one cannot help but be struck by the seemingly perfect match between creating a high school graduate capable of being a productive contributor to our economy and the type of inquiry-based education recommended for all students [. . . .] And precisely the same type of logical, problem-solving skills are badly needed by all citizens, if they are to make wise choices when confronted with the enormous number of personal, community, and national decisions that they will need to make in a complex democracy" (2003, p. 59).

Before delving too deep into my definition of science, I want to be entirely clear that I would never venture to claim that I am describing the one true definition of science that will solve all the problems of education, nor am I claiming a definition that is unique. Rather, I am attempting to present a working definition whose validity comes from its usefulness in the everyday lives of my students. The last thing that I want to do at this point is to portray that there is only one definition of what science is that all scientists and philosophers of science agree on that we just need to explain to students. However, I also don't think that my perspective is really all that controversial—most of the disagreements about this topic are not relevant to the K-12 curriculum as long as we don't portray that the definition of science is any more static than the body of scientific knowledge. There are some core concepts that few would argue about that give science its unique perspective of the world. Scientific knowledge is tentative, empirically based, theory laden, requires creativity and imagination in its development, and is culturally and socially embedded. These characteristics of science would likely not be debated and are unlikely to change (Lederman, 2003).

Though there is widespread agreement on many of the core concepts of science, there is not one specific methodology that all scientists follow in their research. I still remember being taught the "5 Steps to the Scientific Method" in high school as though they were a rigid set of steps that all scientists followed all of the time. This false rigidity often limits those who choose science as a career to people who are good at following directions. Though this is definitely an important skill to have in scientific research, creativity, imagination and a willingness and ability to think outside the box are essential for any truly groundbreaking discoveries. Like the definition of science in the previous paragraph, there are those who might quibble about the details in the upcoming description of the scientific process, but most scientists and science teachers would agree with the core ideas.

The process of scientific thinking involves the evaluation and development of theories through structured observation of the world around us. This process involves asking questions, defining problems, experimentation, interpretation, analysis of data, making observations (directly and using precision instrumentation), and more. The starting and ending points for the process depend primarily on the specific situation that is being investigated and not on a prescribed list of steps. When using this process, scientists strive to collect accurate data. They use observations and measurements with their own senses and also instruments that allow them to reveal a world that is beyond our senses. Before starting experiments, scientists usually have a good idea of what type of results to expect. A useful hypothesis will not only suggest what data would support it, but also should make clear what data would refute it. When doing experiments, scientists attempt to change only one variable between trials in order to be better able to attribute any perceived results to this change. Scientists both influence the culture that they are in and are influenced by it through what they pay attention to and the problems that they choose to investigate. Developing theories and hypotheses is a creative, interpretive process that struggles to make sense of the natural world around us (Rutherford and Ahlgren, 1990).

The word "theory," used in the context of science, is so much more than just someone's guess about what is to happen. Theories are the core driving forces for discovery of scientific knowledge and they are the best explanations that we can currently make based on the available evidence. Even though theories are the nucleus of science, whether or not a theory is "true" is not even necessarily important. What is important for a theory is that it is empirically accurate and that it is useful. In other words, a theory needs to represent and explain the data that we already have and predict future discoveries or relationships (Losee, 2001).

It's not that these are radical definitions of science or the scientific process; however, it is important to focus on what science is and *how* science is as a starting point for the rest of the discussion of assessment. The purpose of this paper is not to spend an extensive amount of time arguing about whether or not these are the exact definitions, but rather to present solid working definitions and then use these definitions to revise assessment in such a way that it focuses more on core concepts and the scientific way of knowing and less on arbitrarily selected details that reveal little of student learning and do little to prepare students to learn after they leave the classroom. Now that these working definitions have been established, I can discuss the process of establishing the objectives for the JCHS biology curriculum and finally evaluate our current semester assessment.

CHAPTER 3 - Humans, Apes, and Monkeys

17. To what order do humans, apes, and monkeys belong?

A. primates	*C. australopithecines*
B. hominids	*D. Homo sapiens*

As I mentioned in the preface, this question was really the genesis of this entire report. As a question, it seems harmless enough and maybe not even a particularly unreasonable thing for a high school biology student to know, but this question has come, for me, to represent everything that is wrong with the 2006 Junction City (KS) High School Spring Semester Biology Criterion Reference Test (CRT) (see Appendix A for the entire test as it was originally written). The JCHS biology CRT is a locally written and administered test that is intended to measure student knowledge and understanding of biology. There is a fall and spring version of the test that connects with the content that is taught during those respective semesters. As I was preparing the semester review activities for this test, I read this question and I immediately had a visceral, negative reaction. Though it's not unreasonable to assume that a student would have been exposed to this fact during the year, it was not something that we had spent much time on during the year and I had no expectations that the students were to memorize the names of any orders. The question illuminates nothing of any underlying concept of taxonomy; it is only testing whether or not a biology student has memorized the specific name of the order for these organisms. Although this is a group of organisms that are pretty important to the world as we know it, if a student ever needs to know this information, he or she can look it up in their textbook or online. Furthermore, is this the only order that we expect students to know, or does it represent a one-question sample of a broader domain that we want students to understand? Because of the extremely specific nature of what is being asked, I could not think of any way to effectively prepare my students for the question without simply asking the same question on the review. I found myself in a situation I've never before been—contemplating teaching to a test.

This was extremely frustrating because it in no way fits with what I think is important to teach in the science classroom. I want my students to become science literate, not just walking encyclopedias. The Programme for International Student Assessment (PISA), an organization dedicated to the development of assessments that can be used for the development of international education policy, defines science literacy this way: "Science literacy is the capacity to use science knowledge to identify questions and to draw evidence-based conclusions in order to understand and help make decisions about the natural world and the changes to it through human activity" (Harlen, 2003). I was absolutely infuriated with how far this question seemed to be from what I believe needs to be taught in the science classroom. I ended up rewording the question for review and trying to hide it among similar questions, but the frustration that I felt in dealing with this question and several others on the test was enough to motivate me to get it fixed.

Setting the Objectives

The first step in preparing an effective assessment is to have well written objectives that clearly lay out what it is expected for students to learn. The Biology I curriculum at JCHS is tied closely to the state assessments due to the fact that the Kansas State Science Assessment is taken in the spring of the 10th grade year—the same year that most students at JCHS take Biology I. Though the Kansas Science Standards have been attacked for their stance on biological evolution, much of the standards, especially the general objectives, are closely tied to the National Science Education Standards (National Research Council 1996), which makes them a good starting point for determining the objectives for this course. The Kansas versions of the objectives that are covered during the spring semester at Junction City High School are listed in Figure 3-1(Kansas State Board of Education, 2006).

Figure 3-1: These objectives are selected from the 2006 Kansas State Standards for the spring semester of Biology I at Junction City High School.

1. **SCIENCE AS INQUIRY – The student will develop the abilities necessary to do scientific inquiry and develop an understanding of scientific inquiry.**
2. **HISTORY AND NATURE OF SCIENCE – The student will develop an understanding that science is a human endeavor and examine the characteristics and history of scientific knowledge.**
   a. **Benchmark 1: The student will develop an understanding that science is a human endeavor that uses models to describe and explain the physical universe.**

b. **Benchmark 2: The student will develop an understanding of the nature of scientific knowledge.**

c. **Benchmark 3: The student will understand science from historical perspectives.**

3. **LIFE SCIENCE – The student will develop an understanding of the cell, molecular basis of heredity, biological evolution, interdependence of organisms, matter, energy, and organization in living systems, and the behavior of organisms.**

    a. **Benchmark 3: The student will understand the major concepts of the theory of biological evolution.**

    b. **Benchmark 4: The student will understand the interdependence of organisms and their interaction with the physical environment.**

    c. **Benchmark 5: The student will develop an understanding of matter, energy, and organization in living systems.**

Figure 3-2: The three primary standards from the USD 475 K-12 Sex Education curriculum that are taught during Bio I.

1. Students will understand that contraceptive methods must be used consistently and correctly to be effective.
2. Students will realize that sharing accurate information about STDs and HIV/AIDS can help fight the spread of these diseases.
3. Students will review the process of contraception and will learn the function of hormones in the reproductive process.

In addition to the Kansas standards listed above, Unified School District 475 also requires that human reproduction be taught during Biology I as a part of the K-12 Sex Education curriculum (Figure 3-2). These two lists of objectives, though general, give guidance to what we want students to learn by the time they have finished the spring semester of this course. The "Science as Inquiry" and "History and Nature of Science" objectives listed are assessed in both semesters, but the selected Life Science objectives and Human Reproduction objectives are assessed only in the spring semester.

Connecting Objectives and Assessment

The objectives listed in Figure 3-1 and Figure 3-2 give good initial guidance as to what to focus on, but each one covers a rather broad area of content. For example, the umbrella benchmark, " The student will understand the major concepts of the theory of biological evolution," contains enough content that it could never be covered in a lifetime. Narrowing and specifying these objectives is a difficult balance of attempting to make sure all of the core concepts are included while still trying to prevent the curriculum from becoming overloaded with facts for students to simply memorize. I was helped in this decision by the indicators that are listed with each standard (especially ones that are highlighted as likely to be a part of the state assessment), but I also took into account my own teaching experiences and my philosophy of science and science teaching in the final wording of the objectives. This process of putting the standards in my own words gave me more ownership in the curriculum and a much more grounded sense of what we really want to make sure is incorporated into the curriculum and what is better thought of as optional. By the time that I was finished with this process, I also had a much better idea of what types of questions would fit in with each objective.

After these objectives were approved by the rest of those teaching biology at Junction City High School, the next steps were to evaluate whether the existing questions lined up with these objectives, to eliminate the questions that either didn't align with these objectives or were poorly written, and then to fill in the gaps that were either created by the addition of objectives or the deletion of poor questions. The process of coming to a consensus on the wording of the objectives was not without its moments of contention. It was sometimes difficult to discuss not including specific topics because teachers wanted to make sure that the areas of content that they were most interested in were included, but I think that we have developed a reasonable working set of objectives that will both allow freedom for individual teachers to put their passion into their classroom and ensure that students are adequately prepared for the next level.

Once the new objectives were established, it became possible to evaluate the questions on the Spring Semester Biology I CRT for both their alignment to the objectives and their ability to assess what they are intended to assess. Questions were evaluated with the following criteria: overall student performance on the item in question (percent correct) , point-biserial correlation, distractor choice frequency, and by simply reading the questions to identify any construct irrelevance. Since the spring test was taken online, item analysis data (Appendix B) was readily available through the Scantron® website (www.achievementseries.com). The item analysis was a significant help in identifying questions that were not performing as expected as well as questions that students were either having difficulty with or that everyone was getting correct. The point-biserial correlation coefficient compares a student's performance on the specific item relative to their overall performance on the test. Although item analysis doesn't directly address question validity, it does identify items that are performing peculiarly. Any item with a point-biserial of less than .20 was carefully evaluated for construct irrelevance including: whether or not they were aligned to the objectives that we were actually wanting to teach and whether or not the wording of the question or distractors made the question misleading or unnecessarily complex. A percent correct that was either above 90% or below 60% also warranted an extra look in item evaluation. If the percent correct was too high, the item was suspect because it may be too easy or may not actually indicate that much about student learning. If the percent correct was too low, it may indicate that there was either something that was confusing about the problem, or perhaps it was testing something that we weren't really covering in class. The values for closer examination were somewhat arbitrarily chosen to represent extremes that may indicate something was wrong with the question, but every question was given at least a read through because all the numbers could look good and the question might still not really be asking what we want it to ask (Hanna & Detmer, 2004); that is, student performance doesn't automatically reflect student knowledge.

The following sections will address the problems in the existing test grouped together by the four major areas of content: Evolution, Inquiry and the Nature of Science, Ecology, and Human Reproduction. Each area will be analyzed for its alignment to the newly refined objectives as well as for the validity of the questions. Question validity will be evaluated on a question-to-question basis primarily by re-reading through the each question, determining if it is connected to the objectives, and checking to make sure the question format won't confound our measure of student acheivement. Suggestions will be made for elimination and modification of questions and a table of specifications will be presented to guide the generation of replacement and additional questions for the test. Some suggestions will also be made for new questions, but in order to foster more teacher investment in the test, the biology teachers at Junction City High School will continue to collaborate to develop questions that fill out the table of specifications.

Evolution

The evolution objectives strive to focus on the concept of evolution while still allowing teachers the freedom to be innovative in how they present the content to their students. Some of the objectives, such as number one in Figure 3-3, are modified from the indicators that are listed in the Kansas state standards under the evolution benchmark (Kansas Science Education Standards, 2006, Standard 3, Benchmark 3), but others are my own additions. The second objective listed in Figure 3-3 is an example of a self-generated objective. I included this objective because students frequently enter the study of evolution with a host of misconceptions about evolution theory. In order to really get at what the theory *is* we actually need to identify what it isn't. Students are unlikely to disregard their previous framework for understanding evolution unless the inconsistencies and misinformation are explicitly addressed before or during the presentation of the new information (Fraser & Wahlberg, 1995).

Figure 3-3: The revised evolution standards for the spring semester Biology I CRT

1. Students will understand the major patterns in the types of evidence—the biodiversity and biogeography of life, the systematic similarities in anatomy and molecular chemistry (DNA, RNA, AA, etc), and the fossil record—that support the theory of evolution.
2. Students will discriminate between accurate statements about evolution theory and common misconceptions about evolution.
3. Students will evaluate descriptions of species change over time to determine which proposed mechanism for evolution that they represent.
4. Students will understand that organisms vary widely within and between populations and be able to identify sources of genetic variation.
5. Students will understand that taxonomy organizes organisms by their perceived evolutionary relationships.

The analysis of the questions that fit with this objective (numbers 1-20 in Appendix A) is shown in the table of specifications on the following page (Table 3-1). Each question was initially assigned to the objective that it was closest to without an initial determination of the value of the question. Then the set of questions that went with each objective was later analyzed to determine what needed to be dropped, modified, or added. It is immediately apparent in examining Table 3-1 that there is significant work to do in both realigning the questions to the objectives and increasing the number of higher-level questions that are included in this section. Half of the questions were aligned with the first objective, five to the fifth, and no more than two questions were aligned with any of the remaining objectives. Furthermore, Table 3-1 does not show the imbalance of questions connected to the fossil record as a type of evidence—seven of the ten questions that were aligned with this objective were related specifically to the fossil record with only one question connected to anatomical similarities and one connected to similarities in biochemical sequences. Though most of the questions were at least at the understanding level, this too was skewed by the questions aligned to the first objective.

Table 3-1:Table of Specifications for the existing evolution questions

	Question Level			
Objective	**Knowledge**	**Understanding**	**Application**	**Totals**
1. Types of evidence	3	6	1	**50%**
2. Misconceptions	0	0	2	**10%**
3. Mechanism for evolution	1	0	1	**10%**
4. Genetic variation	0	1	0	**5%**
5. Taxonomy	3	2	0	**25%**
Totals	**35%**	**45%**	**20%**	**100%**

This table does nothing, however, to analyze the validity of the questions assigned to each objective. A careful reading of the questions indicates that there are some additional concerns with this set of questions. For example, question #1 (Figure 3-4) shows that nearly 95% of the students answered this question correctly. Although this might seem like a good thing as it implies student understanding of the length of time the earth has been in existence, a careful look at the question indicates that the distractors are so extreme that they are not at all effective.

Figure 3-4: Question one from the original test

1. Based on the evidence, scientists think the Earth formed about
   - A. 11.8 billion years ago
   - **B. 4.6 billion years ago**
   - C. 680 million years ago
   - D. 12 million years ago

The question itself is basically testing the memorization of a fact that reveals little about what the student understands about the content. This question was originally chosen as an indicator of a student's awareness of time frame of earth's history, but the extremes of the distractors make it just as likely that the student will recognize the number 4.6 without even really connecting it to the billion years ago, and it is the billion, not the 4.6 per se, that we want students to have in perspective. The following question (Figure 3-5) is a suggested replacement that aims to make the distractors more effective and connect the number with the billions of years ago (which at least gets us one step closer to an understanding of the scale of the earth's history). In order to answer this question correctly, students must at least identify that it was billions of years ago. This is still a knowledge level question, but now it is more specific in what it is testing.

Figure 3-5: Question 1 revised

1. Based on the evidence from the dating of rocks, scientists think the Earth formed about
   - A. 4.6 thousand years ago
   - **B. 4.6 billion years ago**
   - C. 4.6 million years ago
   - D. 4.6 trillion years ago

The third question (Figure 3-6) is basically the same as the first one, but it refers to the earliest life rather than the formation of the earth. Students did not do quite as well on this question (76% correct), but the content overlap with the first question is significant and it seems that it is more important that students connect the relative time of this event rather than pick out the date from a list.

Figure 3-6: Question 3 original

3. Fossils of the earliest known life date back to about
   - A. 8 billion years ago
   - B. 5.2 billion years ago
   - **C. 3.5 billion years ago**
   - D. 10 million years ago

The following question will replace number three:

Figure 3-7: Question 3 revised

3. How soon do fossils of the earliest known life appear in the fossil record following the formation of the earth?
   - A. within a week of the formation of the earth
   - B. about 1,000 years after the formation of the earth
   - **C. slightly less than 1 billion years after the formation of the earth**
   - D. more than 2 billion years after the formation of the earth

As in question #1, students will likely recognize the number "3.5", which doesn't mean they understand the relative time of the origin of the earth and the origin of life; furthermore, if students answer "4.6" correctly in #1, it is almost a given that by process of elimination a test-wise student can determine that 3.5 is the only answer that makes sense. Incorporating relative time into the answer choices on the revision gets closer to measuring a student's content knowledge rather than their ability to take a test.

Question #2 (Figure 3-8) was closely examined because of its low point-biserial correlation coefficient (.04). Eighty-four percent of the students actually got this question correct, but the low correlation coefficient indicates that many of those who did miss this question were the students who did well on the rest of the test. It is not immediately apparent from the question what may have triggered this effect—perhaps it is something in the question and perhaps it has more to do with the instruction. However, I would assert that this question does not really assess significant student learning and should be replaced. Yes, students should know the answer to this question (and the % correct numbers show that most of them do), but writing good test questions is more than just providing a list of questions that students should know. It is about trying to assess what concepts the students understand and can apply. This question tells us only whether or not the students know what the earliest fossils of life were and nothing else and does not merit inclusion on the test.

Figure 3-8: Question 2 original

2. According to the evidence, the earliest life on Earth was
   A. algae　　C. protists
   B. fish　　**D. bacteria**

Question four (Figure 3-9) also shows a solid percent correct (79.55%), but the point-biserial is a marginal .20. Closer examination of this question reveals that there are multiple correct answers.

Figure 3-9: Question 4 original

4. What had to occur before there could be animals on land?
   A. The land had to cool off.　　C. Sea organisms had to evolve lungs.
   B. There had to be enough oxygen in the air.　　D. The land had to dry out.

In fact, the way that the question and distractors are worded, a reasonable case could be made for any of the distractors being the correct answer. There is nothing that clearly sets the correct answer apart from the others. Since there is nothing essential about the content of this question, it will be dropped.

Question number five* is a good example of a question that can look reasonable when glancing at the item statistics—76.70% correct; point-biserial =.30—but its value as a question is not obvious. The question asks students to identify the fossil record as the basis for the timeline of earth's history from among other types of evidence for evolution, but what does that really tell us other than that the students recognize that fossils are from the past—something that I would hope an elementary student would know. This question also needs to be dropped in favor of one that better assesses whether or not students understand the types of evidence that support evolution.

Number six, which asks students for a true statement about the earth's history, also looked good by the numbers (85.60% correct; point-biserial = .34), but the correct answer (C. Generally, life evolved from simple to complex.) is dangerously close to a common misconception about evolution—that it *always* moves from simple to complex and can't account for situations where species "lose" functions that their ancestors had (i.e. cave animals "losing" sight). Although the word generally takes into account the exceptions to this pattern, it's just too close to the misconception for me. To rectify this situation, all that is needed is to replace choice C with a different true statement about evolution. A possible replacement for choice C is: Most of the species found in the fossil record are less than 1 billion years old.

Most students have little difficulty connecting Darwin with evolution as they are asked to do in question number 7; however, this question does not really address any of our objectives and the overly high percent correct (93.75%) just doesn't justify having this fact as a part of the test. Again, taking this question off of the test does not mean that this isn't information that we expect students to know; it just means that it's not as important as knowing what the theory of evolution is.

---

*Questions henceforth that are not included in the text as "figures" and are instead referenced only generally can be read in their entirety in Appendix A.

Questions eight, nine, and ten (Figure 3-10) have terrible percent correct scores—57.10%, 49.15%, and 34.38% respectively—but I think they are all actually questions worth salvaging.

Figure 3-10: Questions 8, 9, and 10 originals

8. Which of the following is the BEST example of biological evolution? **A. Over many generations, horses have changed to be larger and have less toes.** B. Over hundreds of years, a redwood changes from a sapling to a giant tree. C. Individual rabbits grow more fur when placed in a cold environment. D. During its lifetime, a frog changes from a tadpole to a frog with lungs.
9. Which of the following is NOT consistent with concepts in the Theory of Natural Selection? A. There is natural variation within every population. **B. Traits acquired during an organism's lifetime are passed to offspring.** C. Well adapted individuals are more likely to reproduce. D. Conditions in the environment determine what traits are needed to survive.
10. What must be true for 2 populations to be considered separate species? A. They must live in separate places. **C. They must be reproductively isolated.** B. They must have different adaptations. D. They must NOT have a common ancestor.

The most frequently chosen distractor on number eight was D—During its lifetime, a frog changes from a tadpole to a frog with lungs. Students should have identified the "During its lifetime" part of the question and thrown out this question because it is talking about development rather than evolution. I think this answer choice could be improved by dropping the unnecessary words "with lungs," but otherwise this seems to be a good question for assessing misconceptions about evolution. Question nine asks students to identify which statement is not consistent with natural selection and it also seems to have well developed distractors. One change that might improve the selections would be to change answer choice A (the most frequently chosen distractor) from "There is natural variation within every population," to "There is genetic variation within and between populations of organisms." Removing the loaded word "every" in this way makes the statement more clear. Question ten was one of the few questions where the correct answer choice was not the most frequent answer selected—the correct answer "C" was chosen 121 times compared with answer 153 for choice D. The explanation for this lies in the confusing word choice in answer choice D. As can be seen in the question below, option "D" is unnecessarily worded as a negative. Answer choice D will be replaced with "They must be different size and color."

The next group of questions performed about as expected and did not seem to be testing something different than was intended. Number 11 should be dropped because it does not align with the new objectives, but questions 12 through 16 all are acceptable questions to use in the next version of the test.

Now we are down to number 17, the infamous human, ape, and monkey question that got this whole process started. It, along with number 20 that just asks students to identify the distinguishing characteristic of vertebrates, and question 18 that just asks what is included in a scientific name, need to be dropped from the test. These need to be replaced with a set of questions that more closely align with the objectives. The only taxonomy question to survive revision is number 19; however, the stem should be re-worded from "Biological classification of organisms is based on" to " What is the primary consideration in deciding how to classify organisms?"

Now that the most egregious errors of the evolution section of the test have been dealt with, it is time to establish what types of questions should be included in the new version of the test. This section of the test previously had 20 questions, but the new version will include more questions. We are going to shift some of the weight away from the human reproduction section of the test simply because it is the one that the least class time is spent on. The three other sections of the test will increase from 20 to 24 questions each in the new test version while the human reproduction section will be scaled back to 16 questions. This will change the total number of questions on the test from 80 to 88.

Table 3-2: Evolution table of specifications for setting up the new version of the test

	Question Level			
Objective	**Knowledge**	**Understanding**	**Application**	**Totals**
1. Types of evidence	2(3)*	3(2)	1(1)	**25%**
2. Evolution Theory/Misconceptions	0(0)	2(2)	1(1)	**13%**
3. Mechanism for evolution	1(0)	2(1)	2(0)	**21%**
4. Genetic variation	2(0)	2(1)	1(0)	**21%**
5. Taxonomy	1(0)	2(1)	2(0)	**21%**
Totals	**25%**	**46%**	**29%**	

*numbers in the parentheses indicate the number of questions that are currently assigned to each objective after the faulty questions have been removed

As mentioned earlier, these questions will be written through a combined effort of biology teachers at Junction City High School.

Inquiry and the Nature of Science

The inquiry and nature of science objectives (Figure 3-11) are key to assessing whether or not students understand the underlying process of how science works. They were written to challenge myself and my colleagues to more fully integrate the nature of science throughout the curriculum. Again, this list of objectives is a mix of those that are closely tied to the indicators in the Kansas Standards and those that try to bring together ideas from several indicators in a way that allows for flexibility in instruction, but challenges instructors to develop and/or implement innovative strategies for teaching science.

Figure 3-11: The revised inquiry and nature of science standards for the spring semester Biology I CRT

1. Students will design and evaluate investigations, including developing questions, gathering and analyzing data, and designing and conducting research.
2. Students will understand how scientific discoveries are communicated to other scientists and to the general public.
3. Students will be able to compare and contrast the terms observation, inference, theory, hypothesis, and law.
4. Students will understand how scientific knowledge has been and continues to be a process of building on previous research.
5. Students will distinguish between questions that can and cannot be answered with the scientific process.

This section had the lowest overall percent correct (Table 3-3) which, in my estimation, has three main reasons. The first main reason for the poor performance is that this area of assessment is the most difficult to write questions for on a multiple-choice test. On a multiple choice formatted test, these objectives are best assessed by having students apply the scientific process to given scenarios, but it is not an easy task to write a scenario that gives students all the information that they need in a way that is clear at multiple reading levels and that doesn't inadvertently test background knowledge. The second reason (which is probably related to the first reason) is that the questions in this section are the least aligned with the objectives. The third reason is we need to do a better job of incorporating the nature of science and similar types of questions throughout the year so that students are better prepared for the test.

Table 3-3: Average percent correct for the questions in each content area

Content area	Average percent correct
Evolution (1-20)	74.97%
Ecology (21-40)	70.37%
Human Reproduction (41-60)	76.72%
Inquiry and Nature of Science (61-80)	62.88%

The inquiry and nature of science section of the test was the last one to get put together when we last revised the test and it is immediately apparent in analyzing the questions that we were running out of steam on this section of the test. When the questions are aligned to the new objectives (Figure 3-4), 17 of the 20 questions fit with the first objective and most of those were specifically aimed toward interpreting graphs. While I don't deny that this is an important skill to teach and assess in science, it definitely should not account for that large of a portion of the test. The remaining three questions are aligned with the third objective. It is obvious that this area will require significant effort in generating new questions to be an effective assessment.

There are quite a few questions in this section that simply need to be dropped from the test. Both questions 61 and 62 need to be replaced because, in their asking where the scientific process starts, they imply a rigid scientific process that simply does not exist except in outdated and oversimplified textbooks. In reality the scientific process is much more flexible and an investigation can effectively "start" at any of the typically identified steps in the scientific method. Questions 73-75 and the accompanying diagram also need to be replaced. The phylogenetic tree diagram was created with imaginary organisms with the intent of avoiding assessment of background knowledge, but what we're left with is a rather abstract set of problems that tests nothing except for the ability to read a phylogenetic tree—a beneficial but certainly not an essential skill to be assessed. The scenario that goes with Questions 77-80 is inadequate in explaining how and why the data was collected. It is actually the type of problem that we need more of in this assessment, but it is not well written.

There are also a number of questions in this section that can be relatively easily fixed. Numbers 63 through 67, 69, and 72 all had a percent correct below 55%. Number 63 (Figure 3-12) shows a surprisingly low percent correct (54.83%) with answer choice B as the most frequently chosen distractor. I think the primary reason for this is how the word "proven" is used in this choice. It was written to imply that theories are not subject to change, but students often have difficulty with this because they associate "proven" with "supported." It is possible to re-word this choice so that this issue in semantics is avoided—it should be written, "It's a prediction that has been supported by a single experiment."

Figure 3-12: Question 63 original

63. Which of the following BEST describes the nature of a scientific theory?

A. It’s a possible answer to a question that is not yet tested.
B. It’s a prediction that has been proven by experimentation.
C. It includes many hypotheses and is based on much evidence.
D. It is just a guess about a problem, a starting point in scientific thinking.

There are several issues with number 64:

Figure 3-13: Question 64 original

64. Scientists usually design experiments

A. based on wild guesses
B. expecting to get certain experimental results
C. for the purpose of developing new tools
D. to make money

The first revision needed with this problem is that the correct answer needs to be changed in the answer key. The correct answer for this problem is listed as D! Although this could be a partial motive, it's not really what we're looking for here—the correct answer should be B. The ambiguity of choices C and D suggest that they should be replaced; however, I was only able to come up with one more effective distractor. Haladyna suggests that it is better in this situation to only offer three answer choices rather than add a useless distractor and in this case, that's the best option (1999).

The question that will replace number 64 will look like this:

Figure 3-14: Question 64 revised

64. Scientists usually design experiments
    A. based on wild guesses.
    B. expecting to get certain experimental results.
    C. to support their religious beliefs.

Questions 65 and 69 indicate that our students do not understand the difference between independent and dependent variables. On question 65, the frequency of response indicates that students knew that it was either independent or dependent, but then they were essentially guessing after that—there were 157 students that selected the correct answer (A), while 112 students selected B. The question in this case is pretty clearly worded and I think that there needs to be some change in either the instruction or review of this topic. I discussed this with several of the biology teachers at Junction City High School and they were frustrated that the pneumonic, " The dependent variable depends on the independent variable," that they used in class didn't seem to work. My experience is that students understand what these two variables are, but they simply forget which is which. We either need to develop a new pneumonic, or perhaps an entirely different teaching strategy that incorporates more inquiry activities that force students to identify and/or define these two types of variables. Question 69 asks students to identify the independent variable on a graph, but the answer choices are not worded in the same way as the information on the graph. The correct answer was listed as temperature, but the graph just had a legend that identified the separate lines on the graph as hot and cold. The most frequently chosen distractor in this case was "time" (chosen by almost 45% of students) which was the label on the x-axis. The identification of the x-axis variable as the independent variable is likely due to a graph-intensive curriculum that has been adopted by the math department in an attempt to meet AYP. Students' strong association of the x-axis with the independent variable suggests that it would be a good idea to start the year with graph interpretations as a way to improve student understanding of the difference between how the terms are used in science and mathematics. The confusion on this question can be alleviated without compromising the question's validity by simply changing the stem so it is asking for the dependent variable and then changing the answer key accordingly since the dependent variable was already one of the choices (C).

An interesting dilemma presented itself on question 66. Neither myself nor the teachers that I asked were able to convincingly argue whether the correct answer was A or C. If the teachers can't agree what the correct answer is, how can we expect students to figure it out? The simple answer to this situation is simply to remove one of the contentious distractors. After correcting a typo and changing one of the answer choices, the revised question becomes:

Figure 3-15: Question 66 revised

66. Jan and Dean are discussing why a plant in the classroom died. They wondered if moisture, insects or lighting might be the problem. What step in scientific problem solving are Jan and Dean doing?

A. forming a hypothesis C. analyzing their data
B. performing an experiment D. developing a theory

Questions 67-70 are connected to a scenario that describes a student experiment that needs some revision so it is clear what the data mean. The following reorganization of the scenario should help reduce student confusion:

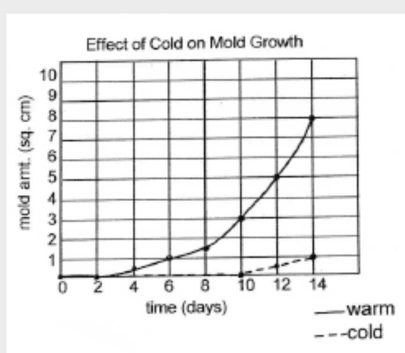
Figure 3-16: Revised scenario for questions 67-70

In biology class, Jodie learned how cold temperatures slow yeast growth. Since bread mold is related to yeast, Jodie wondered if colder temperatures also slowed down bread mold growth. She moistened two pieces of bread and put each in a plastic bag. She put one bag in the refrigerator and the other bag in a warm corner of the kitchen. Every 2 days Jodie checked the bread and measured the amount of area covered by the mold. She recorded her observations in chart below.

Measurements of Mold Growth

time (days)	warm bread amount of mold (sq. cm)	cold bread amount of mold (sq. cm)
0	0	0
2	0	0
4	0.5	0
6	1.0	0
8	1.5	0
10	3.0	0
12	5.0	.5
14	8.0	1.0

Jodie then used this data to create a line graph, which helped her to better analyze her data.



Although this scenario may not represent the thought process of a typical teenager, it does have the basic components of an experiment and has data that should be clear. The first question that goes with this scenario, number 67, asks students to identify the problem Jodie is investigating. Almost as many students chose the distractor A (117) as chose the correct answer (180) for a percent correct of 51.14%. Upon re-reading choice A, "What conditions are needed for the bread to mold?" I think it's a tricky distractor, though it's right on the borderline between effective and tricky. Jodie is trying to find out one of the conditions that is necessary for bread mold growth, but not all the conditions. The stem does ask for the "BEST" problem, but I suggest that A should be changed to, "What are all the conditions needed for bread to mold?" to alleviate confusion.

Questions 68 and 70 are fine as they are and 69 was addressed earlier, so now we turn our focus to number 71. The low point-biserial correlation (.16) brought this question to my attention. The distractor selection was relatively even (A-61; B-21;C-39) which is a decent indicator that it's the stem that is confusing students. To fix this issue, the stem can simply be re-worded from "Which of the following would be the BEST way to confirm Jodie's results?" to "Which of the following would be the BEST way to provide more support for Jodie's result that bread mold grows slower at colder temperatures?"

Question 72 (Figure 3-17) has one of the lowest point-biserial correlations (.05) and one of the lowest percent correct scores (32.67%) of the entire test. Interestingly enough, the answer key in this case was incorrect in identifying the correct as "C" when it should have been "B." This is likely to affect the point-biserial, but it won't change the percent correct as the same number of students selected B as selected C. The poor performance on this question may indicate that we are not teaching what an inference is, but it also is likely an indicator of a poorly written question.

Figure 3-17: Question72 original

72. Which of the following is the BEST example of an inference?
    A. A fish is a vertebrate.
    B. Mammals and reptiles have a common ancestor.
    C. Fossils are often found in sedimentary rock.
    D. Burning releases CO_2 to the atmosphere.

Looking at the question, there is too much assumed in terms of background knowledge and it can be easily fixed by changing the question so that it is tied to the same scenario as questions 68 through 71. The new question can be seen in Figure 3-18 on the following page.

Figure 3-18: Question 72 revised

72. Which of the following is an inference that can be supported by Jodie's data?
    A. Cold temperatures do not affect mold growth.
    B. Mold grows faster in warmer temperatures.
    C. Warm bread mold grows twice as fast as cold bread mold.
    D. The mold grew to a size of 1.5 sq.cm by day 8.

This new question eliminates the background knowledge issue and still assesses whether or not students understand what an inference is. To round out the analysis of the inquiry of science questions, as mentioned earlier, 73 through 75 and 77 through 80 need to be replaced with new questions; 76 can be left as is.

This section has received a pretty thorough gutting, but it was also in the most need of revision. Some of the twenty questions that we started with have been revised, but the total number of questions in this section has also been whittled down to eleven. The table of specifications below (Table 3-4) provides a guideline for the generation of new questions. This guideline reflects the fact that questions in this area of the test are necessarily higher-level questions. As mentioned earlier, this is one of the most essential areas of science teaching, but it is also one of the more difficult areas to assess—especially in the multiple-choice format that the district prefers. It will be my suggestion to the rest of the biology teachers to develop a performance assessment complete with rubrics for evaluation to use for assessment of students' science process understanding in addition to this section of the test. This performance assessment would probably be some sort of semi-structured inquiry activity where students would be required to apply the scientific process to solve a problem.

Table 3-4:Inquiry and Nature of Science table of specifications

	Question Level			
Objective	**Knowledge**	**Understanding**	**Application**	**Totals**
1. Investigation Evaluation(8)	0(0)*	3(3)	5(5)	**31%**
2. Communication (4)	2(0)	2(0)	0(0)	**15%**
3. Theory, law, hypothesis, etc (6)	1(1)	2(1)	3(1)	**23%**
4. History of Science (4)	2(0)	2(0)	0(0)	**15%**
5. Scientific Questioning (4)	0(0)	2(0)	2(0)	**15%**
Totals	**19%**	**42%**	**38%**	

*numbers in the parentheses indicate the number of questions that are currently assigned to each objective after the faulty questions have been removed

Ecology

Figure 3-19: The revised ecology standards for the spring semester Biology I CRT

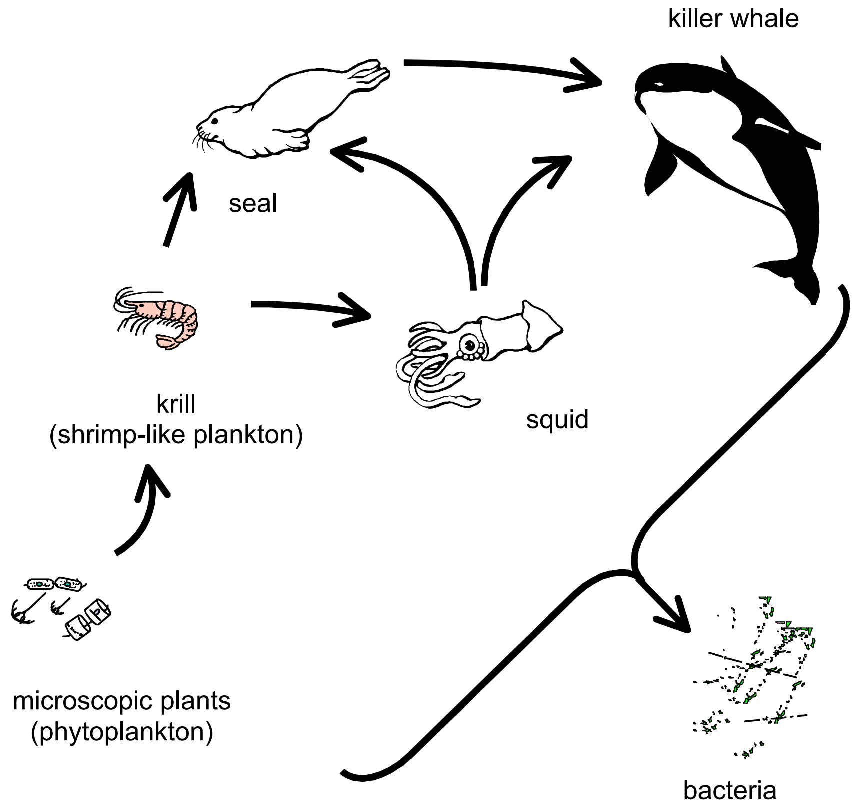
1. Students will understand atoms and molecules on the Earth cycle among living and nonliving components of the biosphere.
2. Students will understand energy enters ecosystems (primarily from the sun) and is transformed and expended through trophic levels, but not recycled.
3. Students will understand the interactions of living and non-living factors that regulate population levels. (carrying capacity, competition, symbiosis, etc)
4. Students will understand how organisms release biochemical energy from food molecules through cellular respiration.
5. Students will identify major causes and effects of current environmental issues.

The ecology section of the test is much better aligned to the standards and the questions seem to have fewer blunders in general than either of the two previous sections; however, I think it also has too few higher-level questions. Questions 21,22, 24-27, 35, 36, 38, and 40 all seem to be performing in a way that makes sense and the wording on the stems and distractors is at least adequate.

The first question that deserves discussion is number 23, which asks students to identify the abiotic components in a diagram. In this case, distractors A and B contain a mixture of biotic and abiotic factors while C contains only biotic factors and the correct answer (D) contains only abiotic factors. The most frequently chosen distractor (selected 73 times—more often than the other two distractors combined) was C. This indicates that students are mostly able to distinguish between abiotic and biotic factors, but that they forget which term goes with which group. If this is a concept important enough to assess, then we need to do a better job of helping students to distinguish between these terms.

Sometimes the diagram that is intended to aid student performance actually makes things worse. This was the case with the food web (see Appendix A) that was connected with questions 29 through 32. It's a struggle to fit decomposers into food webs because they tend to make things messy with lots of extra arrows that confuse the relationships between the other organisms in the web. On the other hand, the essential role of decomposers in ecosystems is so often overlooked that it's not good to leave them completely out of the picture. The revised diagram below (Figure 3-20) solves this problem by having the decomposing bacteria off to the side with a bracket that clearly indicates that they are taking energy from all of the organisms in the web.

Figure 3-20: Modified diagram for questions 29-32



Another confusing part of the original diagram is that the producers do not tie into the web in any way other than to the decomposers—which doesn't make sense if they are the base of the food web. This issue is also resolved in the new diagram.

Now that the diagram issues have been cleared up, we can address the issues with the questions that are connected to the diagram. Both 30 and 31 should improve with a better diagram, but in both cases, the answer choices are still too ambiguous. Number 30 had the worst percent correct on the entire test (21.88%). More than half of the students chose D as the correct answer. This question is almost entirely about just reading the graph and though the graph has been improved, I think this question should be replaced with a higher-level question, or at least a question that is testing more of the content and less of the graph reading. Though the percent correct on 31 was high this question was deemed so confusing by the biology teachers prior to the test that we gave the students the correct answer (with administrative permission) since it was too late to change the answer key in the online test. To make this question work, the answer choices need to add in whether the changes in population are direct or indirect. The following example fixes this issue with the correct answer remaining as "D":

Figure 3-21: Question 31 revised

31. Which of the following describes an impact of a sudden drop in the seal population?
  A. The krill population would decline because they have fewer predators.
  B. The killer whale population would decline unless they shifted to another food source.
  C. The phytoplankton would decline because the herbivore population would increase.
  D. All of these are likely to occur.

I attribute the poor performance (67.61% correct and .17 bi-serial coefficient) on number 32 to the confusing diagram that will hopefully be replaced with the new diagram; however, it would probably be a good idea to make sure that we are being very careful in explaining how decomposers fit into the transfer of energy between trophic levels because it's somewhat of a variation from the pyramid model that we use for producers and consumers.

The pyramid diagram that is connected with questions 34-36 seems to be adequate for our purposes, but several of the questions connected to it need revision. Number 34 ends up basically being a test of the vocabulary word "biomass." We could get at the underlying concept by rewording the stem to say "total mass of living organisms" rather than "biomass." Then we would have a better idea if they understood that there have to be many more producers than consumers in a stable ecosystem, which is more important to assess than the definition of biomass. This concept is important to resource use issues, but a very meaningful discussion can be had without using the term biomass. This is not to say that we shouldn't be teaching students important vocabulary throughout the semester, but with limitations on the number of questions that we can have on a cumulative test, testing the underlying concept is more important. Number 35 is primarily a graph reading question that doesn't address any of the objectives in this section of the test. It needs to be replaced.

Question 28 had a dismally low percent correct (46.02%). This question is a prime example of a question that should only have two distractors. The extra distractor (choice D) does not fit with the other distractors and its forced addition to maintain a format of four answer choices adds in an unnecessary level of confusion. The unique nature of this distractor ended up with being selected as the correct answer almost twice as many times (89) as the two other distractors combined (100). This becomes a better question that more accurately tests what we want it to test if we remove choice D and leave this question with just three answer choices.

Coming in a close second to last in terms of percent correct (28.98%), Question 37 has more than one correct answer and tests a trivial bit of content so it needs to be dropped.

The final question in need of revision in this section of the test is question 39. Even though neither the percent correct (70.45%), nor the point-biserial correlation (.25) are terribly low, it is testing a trivial tidbit that asks students to choose between sulfur dioxide and sulfuric acid as what is made when the gases from burning sulfur-rich coal combine with water vapor. While I would hope that students would answer this correctly, I am more interested assessing student understanding and application of the causes and effects of major environmental issues than a nit-picky fact like this.

The ecology section on the new version of the test will contain 24 questions and will assess a broader and more balanced cross-section of this area of content. The numbers in parentheses in Table 3-5 indicate that what is left of this section of the test is weighted toward trophic levels and the knowledge types of questions. Hopefully the table of specifications will help to rectify this situation.

Table 3-5: Ecology table of specifications

	Question Level			
Objective	**Knowledge**	**Understanding**	**Application**	**Totals**
1. Nutrient cycles(5)	1(1)*	2(2)	2(0)	**21%**
2. Trophic levels (6)	2(3)	2(4)	2(2)	**25%**
3. Populations (4)	1(0)	2(1)	1(0)	**17%**
4. Photosynthesis/respiration (5)	1(0)	2(0)	1(0)	**21%**
5. Environmental Issues (4)	0(0)	2(1)	2(0)	**17%**
Totals	**21%**	**42%**	**33%**	

*numbers in the parentheses indicate the number of questions that are currently assigned to each objective after the faulty questions have been removed

Human Reproduction

This section of the test is required by the district (not the state) to be a part of the Biology I curriculum. Students also get information about human reproduction in their health classes, but in biology we focus more on the anatomy and physiology of reproduction and how it is affected by STDs and controlled through methods of contraception. This is the smallest unit of the semester both in terms of the amount of content and in terms of class time that is allotted to teach it. The box below (Figure 3-22) summarizes the objectives as they are taught during this unit. Since this unit is so much smaller than the rest of the units that are covering during the semester, the number of questions that will be allocated to this area will be reduced from 20 to 18. This will come primarily from reducing the number of questions connected to the identification of reproductive anatomy. Prior to test revision, labeling reproductive structures on the test was accounting for 10 of the 80 questions.

Figure 3-22: The revised human reproduction standards for the spring Biology I CRT

1. Students will explain how current methods of contraception interfere with the biology of reproduction.
2. Students will distinguish between STDs that are only treatable and those that are curable based on the biology of the pathogen.
3. Students will identify male and female reproductive structures on diagrams.
4. Students will determine the functions of the major structures involved in the male and female reproductive systems.

Since we will be reducing the total number of questions as we work through this section of the test, it is helpful to look at the table of specifications (Table 3-6) up front. This will help guide the decision-making as the questions are evaluated.

Table 3-6: Human Reproduction table of specifications

	Question Level			
Objective	**Knowledge**	**Understanding**	**Application**	**Totals**
1. Contraception(4)	1(1)*	2(2)	1(0)	**25%**
2. STDs (4)	1(3)	2(4)	1(2)	**25%**
3. Reproductive Structures (4)	4(0)	0(1)	0(0)	**25%**
4. Reproductive Physiology (6)	0(0)	3(0)	3(0)	**25%**
Totals	**38%**	**38%**	**25%**	

*numbers in the parentheses indicate the number of questions that are currently assigned to each objective after the faulty /extra questions have been removed

Ten of the first twelve questions in this section of the test are matching the names of reproductive structures with labels on a diagram. To fit in with the table of specifications, this will be reduced to two male and two female structures. This means that questions 42, 44, 45, 49, 51, and 52 will be removed. There is nothing wrong with these questions, but this objective was simply taking up too much of the test. The same diagrams can still be used on the test and the questions should be retained for use future development of alternate forms of the test.

Next we need to examine the questions that are related to reproductive physiology. Question 46 is connected to this objective, but it is assessing trivial content so it will be dropped in favor of the following question:

Figure 3-23: Question 46 replacement

46. What type of cells are produced by the male and female reproductive organs?
   A. genetically different cells with half as much genetic information as a body cell
   B. genetically different cells with the same amount of genetic information as a body cell
   C. genetically identical cells with half as much genetic information as a body cell
   D. genetically identical cells with the same amount of genetic information as a body cell

Question 47 is also a male reproductive physiology question and it doesn't need any revision. Numbers 53-55 are female reproductive physiology questions that need little revision. Question 54 had a lower than expected percent correct (66.19%), but it is a solid question and the primary distractor in this case represents a common misconception that fertilization usually happens in the uterus, so this question will be left unchanged and we will need to try harder to address this misconception in class. The stem on question 55 can be improved if it is changed so that it is a question rather than a partial statement for completion. This stem will be changed from "Eruption of an ovum from a follicle is called" to "What is the name for the release of the egg from a follicle?" Question 56 is unnecessarily confusing as demonstrated by the 52.84% correct and a point-biserial correlation coefficient of only .19. In this case it will be easiest to drop the question and replace it with a new one.

Questions 57 and 58 are connected to the contraception objective. Number 57 can remain as it is, but the stem on number 58 is unnecessarily confusing because of the double-barreled nature of the stem. Some of the answer choices work with one item but not the other, which turns this question into an exercise in test-taking ability rather than really testing the content. This problem is likely the cause of the low percent correct (59.66%), but can be fixed by changing the stem from "Birth control pills and depo provera both" to "What do the following methods of contraception have in common: depo provera, birth control pills, the vaginal ring, the patch?" By including more types of contraceptives in the group, the overall pattern should be clearer and the same answer choices can still be used.

The two remaining questions are both aligned with the STD objective. Both questions show a relatively low percent correct (59-52.27%; 60-63.64%), but the reasons for the low scores are different. The word order for the stem on Question 59 (Figure 3-24) could be improved and several answer choices are not clear.

Figure 3-24: Question 59 original

59. Sexually transmitted diseases such as gonorrhea and chlamydia, which are caused by bacteria

A. are usually curable
B. sometimes clear up on their own, without treatment
C. are more common in females than males
D. all of these

The awkward phrasing in the original question stem as seen above could be improved by changing it to the following question, "Which of the following statements about bacterial STDs is true?" Answer choice B was the most commonly selected distractor, in part because it's actually true. When the distractor was originally written, we were trying to imply that the disease was cured without treatment, but as written, it could actually happen. It is possible that the symptoms may be greatly reduced or disappear for a short period of time, but the bacteria has not been killed. The following answer choices eliminate the issue of two correct answers by re-wording all of the answer choices to:

A. They can usually be cured with antibiotics, but antibiotic resistant strains are on the rise.
B. They are treatable with drugs to reduce symptoms, but they cannot be cured.
C. Infection with bacterial STDs can be prevented by getting vaccinated prior to sexual activity.
D. They are significantly more common in men than in women.

By looking back at the numbers in parentheses in Table 3-6 it is clear that this section of the test is the closest to being completed, which makes sense considering that it has the fewest questions and the highest proportion of lower level questions.

Filling in the Gaps

By the time that this report is published, there will have been at least one work session of the biology teachers from Junction City High School to work on filling in the gaps that are now apparent in the tables of specifications. The district has agreed to voucher pay all the biology teachers for at least two 6.25 hour work days to make this test a reality. The combined effort of the entire group of teachers will go a long way to making sure that the questions fit with what is being covered in all of our classes and will hopefully provide opportunity for good discussions of how we can all improve our teaching in the upcoming years.

CHAPTER 4 - Where to go From Here

This test is not yet completed, but it has a new direction and new life. I am optimistic that, with the combined work of the biology teachers at Junction City High School, this test will develop into a meaningful assessment of student understanding of both the process of science and the essential concepts of biology as defined by the objectives in this report.

However, even after the tables of specifications are filled with well conceived questions, I think there are further steps that need to be taken with this test to make it an even more useful tool. In order to really be able to measure student performance, test banks need to be developed for each question and the test needs to shift from a criterion-referenced test to a norm-referenced test. Criterion reference tests and mastery levels are really best for assessment of explicit and masterable domains, whereas norm-referenced tests are better suited for expansive and developmental domains (Hannah and Dettmer, 2004). Biology is certainly expansive and developmental, so it would make logical sense to move toward having a norm-referenced test. A norm-referenced test, if carefully and thoughtfully written, would make it easier to compare student understanding from year to year. The sheer volume of quality questions that are necessary to make a norm-referenced test work is daunting, but in the end, I think the effort will be well rewarded.

Why all this hassle about assessment? Writing great test questions is certainly not an easy or quick task; but we must not be satisfied with a mediocre assessment—it often leads to mediocre teaching. A poor assessment constricts a teacher's options for creating dynamic and effective lesson plans. It restricts the possibilities of digging beneath the content and really teaching students how to think instead of what to think. If tons of tidbits of content are a part of the assessment, then it is only fair to the students that all of the content is surveyed and reviewed prior to assessment. Sometimes it gets to the point—especially if the content on the assessment is trivial, obscure, or covers too broad of a spread—that classes end up being exercises in regurgitation.

Some might say, what is the harm in that? Memorization is a good skill to develop and should benefit students. Furthermore, they will have been exposed to a wide variety of content that will prepare them for the next level of their educational journey—learning how science works. The glaring hole in this logic, however, is in the title of this report—science is as science does. This statement can be read in two complementary ways. What makes science unique is the process of thought that is demonstrated by doing science. The only real way to understand the value is to *do* science. In this I do not mean following the directions to get a result that you already know will come. I mean approaching real life problems and digging in to the roots of the issue with the process of science. If students understand the value of this process and can apply it in their lives, then we have gone a long way in promoting their continued growth and learning regardless of what they decide to do in life. *And* we have better prepared our future scientists to look up and out from the shoulders of the giants that we are standing on. Conversely, if we allow our science classes to become so overwhelmed with content that we only have enough time to do cookie cutter labs and memorize the steps of the scientific method, in a very real way science for our students becomes a tired exercise in remembering factoids and nothing more.

As long as there is still a push for high stakes assessment, teachers need to take advantage of the opportunity to ask their districts, states, and the federal government for time and money to truly develop effective assessments so they can teach to a *good* test. The very future of science may depend on it; after all, science literacy is as science teachers do.

References

Alberts, B. (2003). On creating a "scientific tempor." In S.P. Marshall, J.A. Schepper, & M.J. Palmisano (Eds.), Science literacy for the twenty-first century. (pp. 57-61). Amherst, NY: Promethius Books.

American Academy for the Advancement of Science. (2000). Designs for Science Literacy. New York: Oxford University Press.

Duschl, R. A. (1990). Restructuring science education: The importance of theories and their development. New York: Teacher's College.

Fraser, B. J., & Wahlberg, H. J. (Eds.). (1995). Improving science education. Chicago: The University of Chicago Press.

Geller, M. J. (2003). Opening the doors of science. In S.P. Marshall, J.A. Schepper, & M.J. Palmisano (Eds.), Science literacy for the twenty-first century. (pp. 23-30). Amherst, NY: Promethius Books.

Gould, S. J. (2003). Drink deep, or taste not the Pierian Spring: Musings on the teaching and learning of science. In S.P. Marshall, J.A. Schepper, & M.J. Palmisano (Eds.), Science literacy for the twenty-first century. (pp. 23-30). Amherst, NY: Promethius Books.

Gräber, W., Nentwig, P., Becker, H.J., Sumfleth, E. Pitton, A., Wollweber, K., Jorde, D. (2001). Scientific literacy: From theory to practice. In Behrendt, H., Dahncke, H., Duit, R., Gräber, W., Komorek, M., Kross, A., & Reiska, P. Research in science education – Past, present, and future. (pp. 49-60). Netherlands: Kluwer Academic Publishers.

Hanna, G. S. & Dettmer, P.A. (2004). Assessment for effective teaching: Using context-adaptive planning. Boston, MA: Pearson Education.

Haladyna, T.M. (1999). Developing and validating multiple choice test items (2nd ed.). Mahweh, NJ: Lawrence Erlbaum Associates.

Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA Project. In Behrendt, H., Dahncke, H., Duit, R., Gräber, W., Komorek, M., Kross, A., & Reiska, P. Research in science education – Past, present, and future. (pp. 49-60). Netherlands: Kluwer Academic Publishers.

Lederman, N. G. (2003). Scientific inquiry and the nature of science as a meaningful context for learning in science. In S.P. Marshall, J.A. Schepper, & M.J. Palmisano (Eds.), Science literacy for the twenty-first century. (pp. 85-95). Amherst, NY: Promethius Books.

Lewis, A.C. (2006). Clean up the test mess. Phi Delta Kappan, 87 643-644.

Losee, J. (2001). A Historical Introduction to the philosophy of science (4^{th} ed.). New York: Oxford University Press.

Kansas State Board of Education (2006, February 14). Kansas science education standards. Retrieved June 6, 2006 from the Kansas State Department of Education website: http://www.ksde.org/outcomes/sciencestd.doc

National Research Council (1996). National Science Education Standards. Washington, DC: National Academies Press.

Rutherford, F. J. & Ahlgren, A. (1990). Science for all Americans. New York: Oxford University Press.

Starnes, B.A. (2006). On nerds, science education, and horror films. Phi Delta Kappan, 87, 634-635.

Trefil, J. (2003). Two modest proposals concerning scientific literacy. In S.P. Marshall, J.A. Schepper, & M.J. Palmisano (Eds.), Science literacy for the twenty-first century. (pp. 150-160). Amherst, NY: Promethius Books.

Appendix A - Spring Semester Biology I CRT as taken in 2006

1. Based on the evidence, scientists think the Earth formed about

   A. 11.8 billion years ago
   B. 4.6 billion years ago
   C. 680 million years ago
   D. 12 million years ago

2. According to the evidence, the earliest life on Earth was

   A. algae
   B. fish
   C. protists
   D. bacteria

3. Fossils of the earliest known life date back to about

   A. 8 billion years ago
   B. 5.2 billion years ago
   C. 3.5 billion years ago
   D. 10 million years ago

4. What had to occur before there could be animals on land?

   A. The land had to cool off.
   B. There had to be enough oxygen in the air
   C. Sea organisms had to evolve lungs.
   D. The land had to dry out.

5. What kind of evidence is the timeline of Earth's history based on?

   A. fossils
   B. homologous structures
   C. DNA and amino acids
   D. chemical analysis

6. Which of the following is TRUE based on the timeline of Earth's history?

   A. Humans have been on Earth for most of Earth's history.
   B. The first reptiles appeared before the first fish.
   C. Generally, life evolved from simple to complex.
   D. Changes in dominant life forms came just before mass extinctions.

7. Who wrote the Theory of Natural Selection?

A. Dalton
B. Malthus
C. Mendel
D. Darwin

8. Which of the following is the BEST example of biological evolution?

A. Over many generations, horses have changed to be larger and have less toes.
B. Over hundreds of years, a redwood changes from a sapling to a giant tree.
C. Individual rabbits grow more fur when placed in a cold environment.
D. During its lifetime, a frog changes from a tadpole to a frog with lungs.

9. Which of the following is NOT consistent with concepts in the Theory of Natural Selection?

A. There is natural variation within every population.
B. Traits acquired during an organism's lifetime are passed to offspring.
C. Well adapted individuals are more likely to reproduce.
D. Conditions in the environment determine what traits are needed to survive.

10. What must be true for 2 populations to be considered separate species?

A. They must live in separate places.
B. They must have different adaptations.
C. They must be reproductively isolated.
D. They must NOT have a common ancestor.

11. The process by which a population changes to have more beneficial traits is called

A. adaptation
B. mutation
C. isolation
D. diversification

12. Mutation and sexual reproduction are important to the process of evolution because they

A. help a population grow larger
B. increase variation within a species
C. eliminate the weaker individuals
D. result in the evolution of new species

13. The occurrence of similarities in amino acid sequences in different species is evidence that

A. these species evolved in the same habitat
B. these species eat the same food
C. these species have a common ancestor
D. they are actually all part of the same species

14. A population of bacteria has gone from being 90% treatable with penicillin to being only about 50% treatable. How does this observation support the Theory of Evolution?

A. It shows that this species of bacteria has changed over time.
B. It compares the molecules of different organisms.
C. It's an example of one species evolving from another.
D. It supports evidence of bacteria in the fossil record.

15. An evolutionary biologist looking for similar bones in organisms of different species is looking for

A. embryos
B. fossils
C. molecular evidence
D. homologous structures

16. When fossils are found in different layers of undisturbed rock, it is likely that

A. the organisms represented have become extinct
B. the fossils are at least 1.5 billion years old
C. the older fossils are in the upper layers
D. the older fossils are in the lower layers

17. To what order do humans, apes, and monkeys belong?

A. primates
B. hominids
C. australopithecines
D. Homo sapiens

18. Scientific names for organisms always consist of the

A. class and genus names
B. order and family names
C. family and species names
D. genus and species names

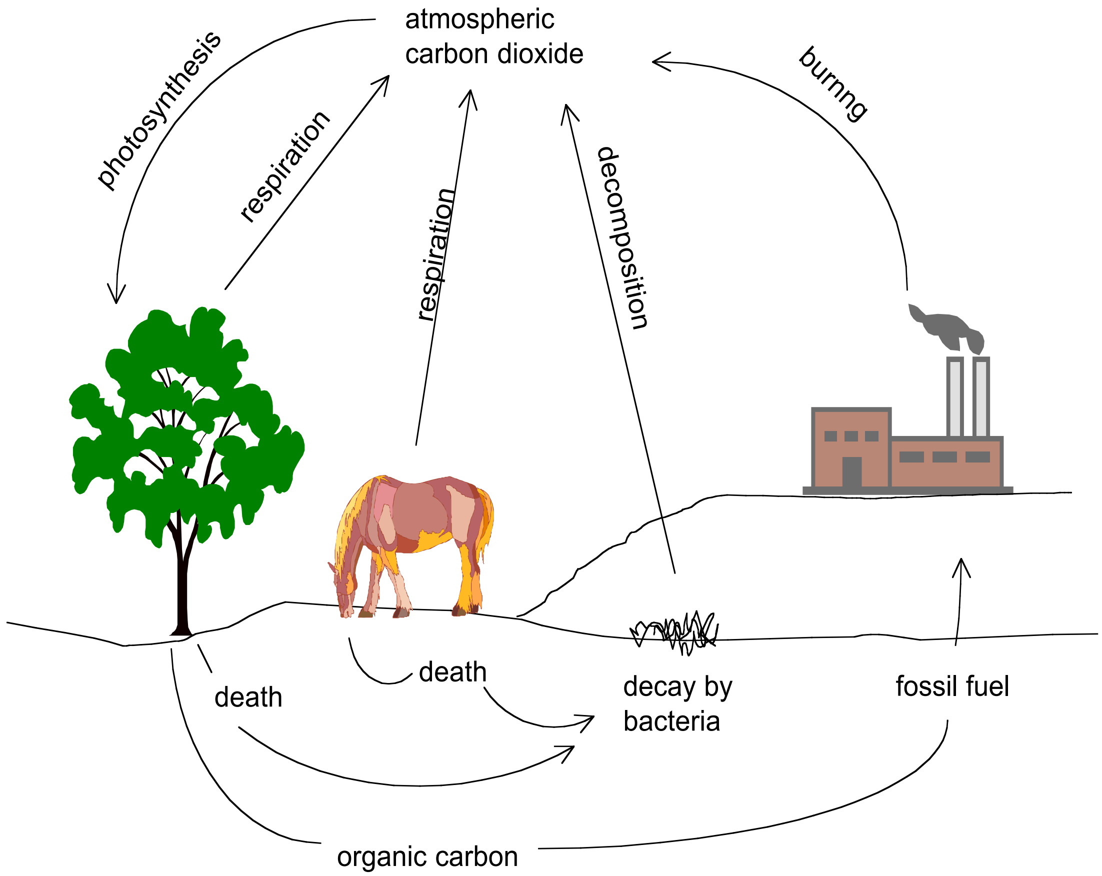
19. Biological classification of organisms is based on

A. food-getting strategies
B. evolutionary relationships
C. where the organisms live
D. similarities in behaviors

20. What is the distinguishing feature of all vertebrates?

A. complex brain
B. jointed appendages
C. backbone
D. multicellularity

For 21-24 refer to this diagram.



21. What organism in this diagram is a decomposer?

A. horse
B. tree
C. bacteria
D. fossil fuel

22. What organism in this diagram is a producer?

A. tree
B. factory
C. bacteria
D. horse

23. Abiotic components in this diagram include

A. air, grass, and bacteria
B. the factory, tree, and fossil fuel
C. horse, tree, and bacteria
D. air, soil, and smoke

24. Based on this diagram, what is the main process that removes carbon dioxide from the air?

A. decomposition
B. respiration
C. photosynthesis
D. burning

25. How does removing large areas of rainforest impact the carbon cycle?

A. Increased respiration causes an increase in atmospheric carbon dioxide.
B. Increased photosynthesis causes a decrease in atmospheric carbon dioxide.
C. Decreased photosynthesis causes an increase in atmospheric carbon dioxide.
D. Decreased photosynthesis is balanced out by increased decomposition.

26. Which of the following BEST describes the relationship between producers and consumers?

A. Consumers eat producers.
B. Producers parasitize consumers.
C. Producers and consumers compete within an ecosystem.
D. Producers obtain their energy from consumers.

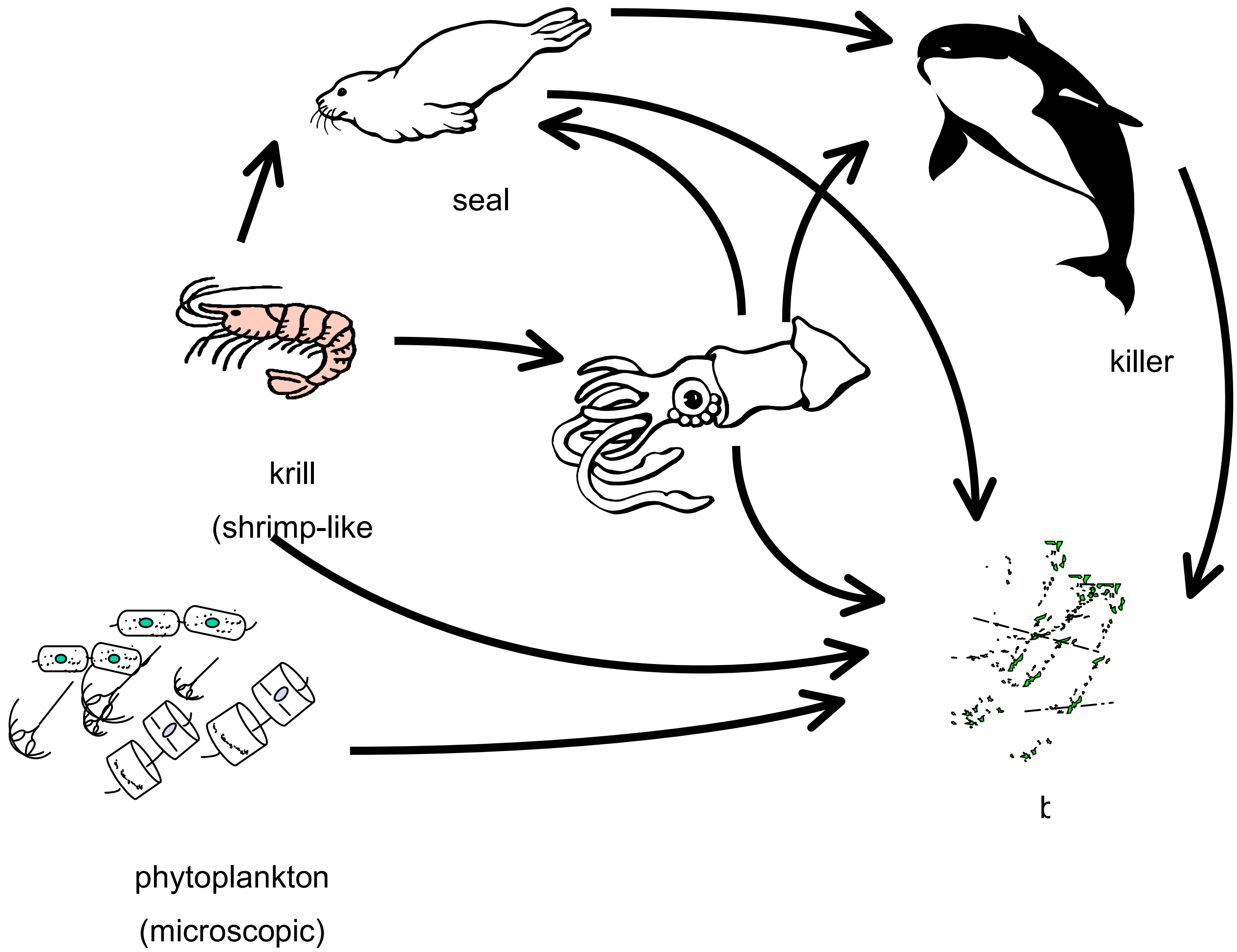
27. All of the following cycle within an ecosystem EXCEPT

A. water
B. nitrogen
C. energy
D. carbon

28. Which of the following is TRUE about matter and energy in the biosphere?

A. Both matter and energy are cycled within the biosphere.
B. Matter is cycled, but there must be continuous input of energy to the biosphere.
C. Energy is cycled, but there must be continuous input of matter to the biosphere.
D. As matter moves through the biosphere, total energy is increased

For items 29-32, refer to this diagram.



29. What organism is the producer in this web?

A. krill
B. phytoplankton
C. bacteria
D. squid

30. Which of the following is TRUE about organisms shown in this web?

A. The seal and killer whale are competitors.
B. Krill and bacteria are both eaten by seals.
C. Squid is a tertiary (top level) consumer in this web.
D. none of these

31. What would be the impact of a sudden drop in the seal population?

A. The krill population would probably increase.
B. The squid population would probably decrease.
C. The killer whale population would probably decrease.
D. All of these are likely to occur.

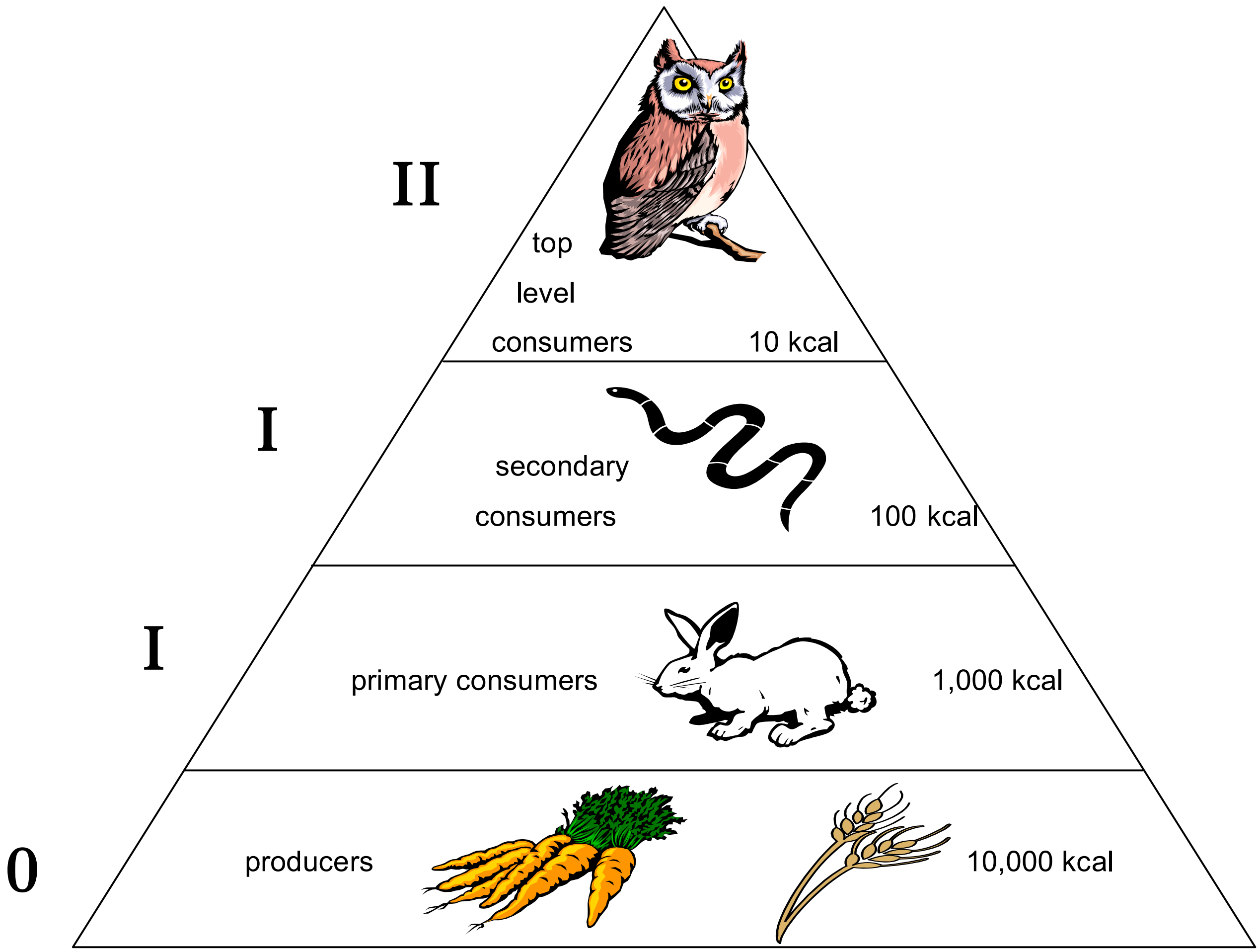
32. Which organism is at the highest trophic level in this web?

A. phytoplankton
B. seal
C. bacterial
D. killer whale

33. What is the original source of energy for all organisms in the biosphere?

A. plants
B. decomposers
C. the sun
D. all biotic matter

For items 34-36, refer to this diagram.



34. Refer to the diagram above. In a stable ecosystem, which trophic level would have the greatest biomass?

A. III
B. II
C. I
D. 0

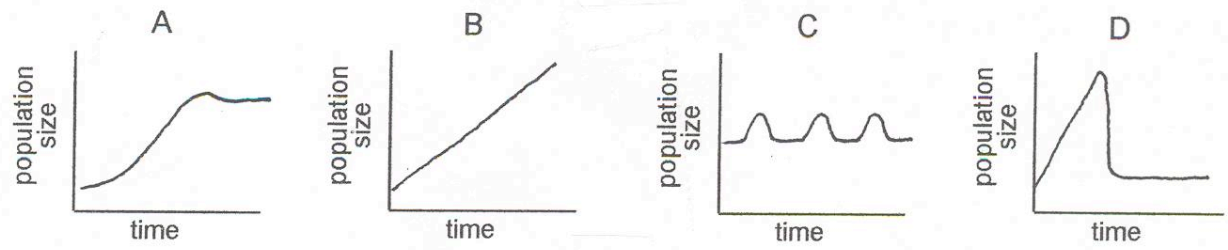
35. Refer to the diagram above. If there were 800 kcal of energy available at level I, how much energy would be available at level II?

A. 8 kcal
B. 80 kcal
C. 800 kcal
D. 8,000 kcal

36. Why is an energy pyramid this shape?

A. Producers are the most complex organisms.
B. As you go down a level, energy is lost.
C. Energy is lost in moving up the pyramid.
D. Only big animals can be at the top.

37. Which of the following illustrates the population growth pattern of species in a balanced ecosystem?



38. Which of the following work(s) to limit a population's size?

A. availability of food
B. amount of water
C. disease
D. all of these

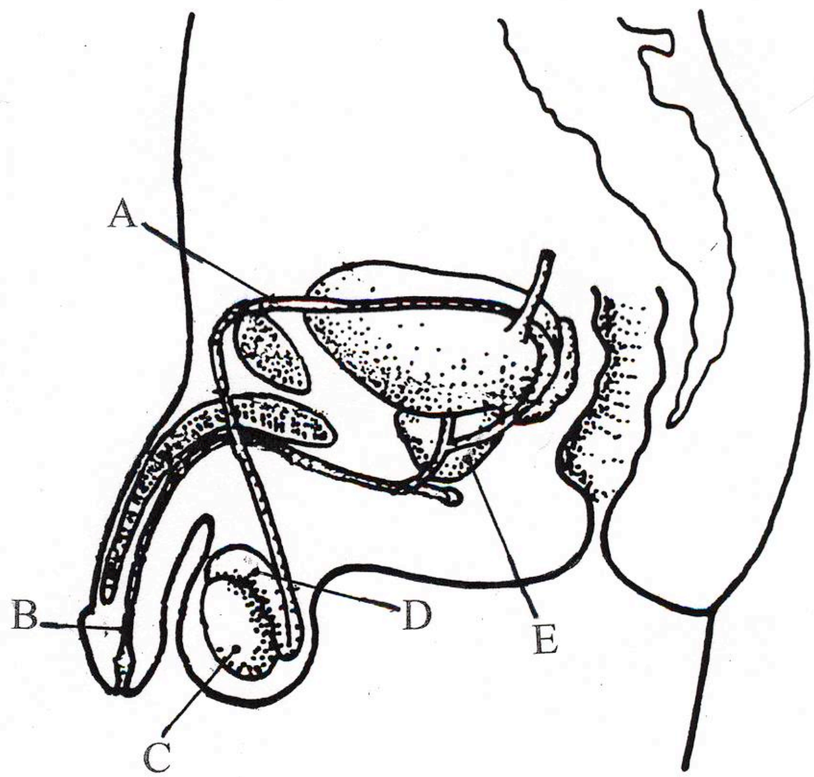
39. Burning sulfur-rich coal releases gases that combine with water vapor in the atmosphere to make

A. sulfur dioxide
B. sulfuric acid
C. ozone
D. all of these

40. Scientists have detected a thinning of the ozone layer. Why is this of major concern?

A. The ozone layer blocks UV radiation that is damaging to life on Earth.
B. Damage to the ozone layer is the main cause of global warming.
C. Ozone in this layer is used by plants in the process of photosynthesis.
D. The layer prevents the escape of other vital gases in Earth's atmosphere.

For 41-45, refer to this diagram.



41. Which letter in the diagram indicates the testes?

42. Which letter in the diagram indicates the urethra?

43. Which letter in the diagram indicates the vas deferens?

44. Which letter in the diagram indicates the prostate gland?

45. Which letter in the diagram indicates the epididymis?

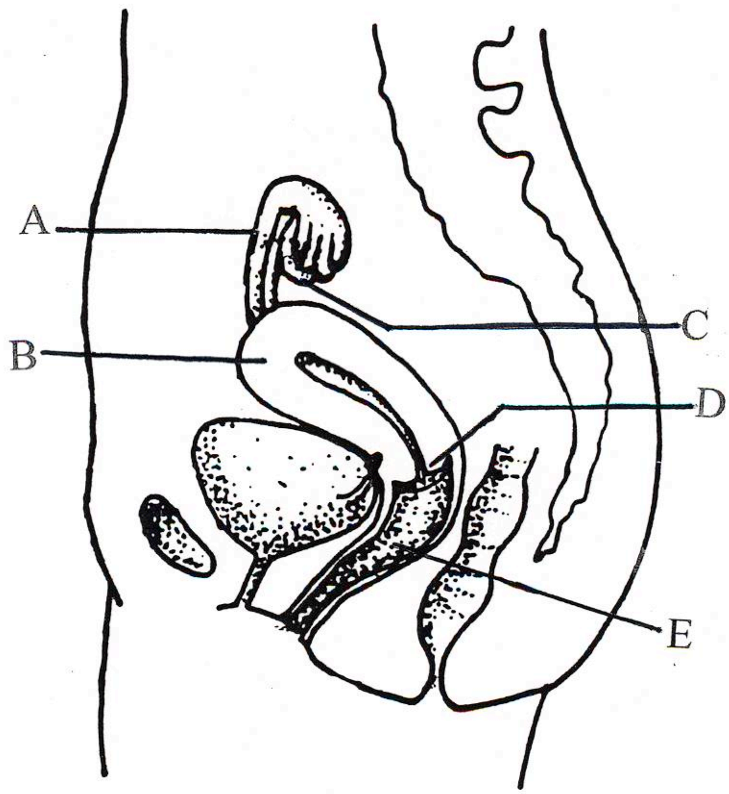
46. In what order does sperm pass through the structures?

A. scrotum, bladder, prostate gland, seminal vesicles, urethra
B. testes, epididymis, vas deferens, ejaculatory duct, urethra
C. testes, ejaculatory duct, prostate gland, vas deferens, urethra
D. ejaculatory duct, testes, epididymis, seminal vesicles, urethra

47. Where are the sperm produced?

A. epididymis
B. testes
C. seminal vesicles
D. vas deferens

For 48-52, refer to this diagram.



48. Which letter in the diagram indicates the uterus?

49. Which letter in the diagram indicates the oviduct (fallopian tube)?

50. Which letter in the diagram indicates the vagina?

51. Which letter in the diagram indicates the ovary?

52. Which letter in the diagram indicates the cervix?

53. What is the function of the ovary?

A. connect the oviducts to the uterus
B. increase blood flow to the uterus
C. protect the uterus
D. provide ova and hormones

54. In what structure does fertilization normally occur?

A. vagina
B. ovary
C. uterus
D. oviduct (fallopian tube)

55. Eruption of an ovum from a follicle is called

A. menstruation
B. emission
C. ovulation
D. fertilization

56. The uterus is maintained in a state of readiness to receive an embryo by

A. the menstrual cycle
B. ovulation
C. gametogenesis
D. the ovarian cycle

57. Which of the following is a TRUE statement about methods of contraception?

A. Oral contraceptives work by killing the developing embryo.
B. Abstinence is the only method of contraception that is 100% effective.
C. Condoms offer little protection against the spread of STDs.
D. A vasectomy stops the formation of sperm in the male.

58. Birth control pills and depo provera both

A. must be taken daily in order to prevent pregnancy
B. are barrier methods of birth control
C. use hormones to prevent ovulation
D. are also effective in the preventing STDs

59. Sexually transmitted diseases such as gonorrhea and chlamydia, which are caused by bacteria

A. are usually curable
B. sometimes clear up on their own, without treatment
C. are more common in females than males
D. all of these

60. Which of the following is/are TRUE about STDs?

A. All STDs can be treated.
B. Sometimes there are no symptoms that a person has an STD.
C. Some STDs can lead to infertility.
D. all of these

61. Scientific problem solving usually begins with a(n)

A. hypothesis
B. question
C. experiment
D. theory

62. The way to test a hypothesis is by

A. asking questions
B. drawing conclusions
C. developing a theory
D. doing an experiment

63. Which of the following BEST describes the nature of a scientific theory?

A. It's a possible answer to a question that is not yet tested.
B. It's a prediction that has been proven by experimentation.
C. It includes many hypotheses and is based on much evidence.
D. It is just a guess about a problem, a starting point in scientific thinking.

64. Scientists usually design experiments

A. based on wild guesses
B. expecting to get certain experimental results
C. for the purpose of developing new tools
D. to make money

65. The variable that is measured at the end of the experiment to determine results is the

A. dependent variable
B. independent variable
C. experimental control
D. evaluative factor

66. Jan and Dean are discussing why a plant in the classroom died. They wondered if moisture, insects or lighting might be the problem. What step in scientific problem solving are Jan and Dean are doing?

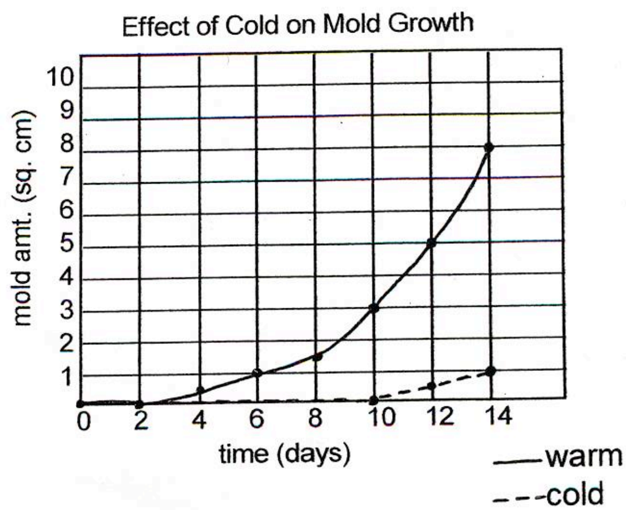
A. looking for a question to investigate
B. performing an experiment
C. forming a hypothesis
D. developing a theory

Refer to this scenario for items 67-70.

Moldy Bread- After learning about yeast in biology class, Jodie wanted to study bread mold, a related
organism. Knowing that fungus does well in a warm, moist environment, Jodie decided to see if cold temperature would slow the growth of bread mold. Jodie moistened 2 pieces of bread and put each in a plastic bag. He put one bread bag in the refrigerator and one in a warm corner of the kitchen. Every 2 days, Jodie checked the bread, and measured the amount of mold. His observations are shown below.

Measurements of Mold Growth

time (days)	warm bread amount of mold (sq. cm)	cold bread amount of mold (sq. cm)
0	0	0
2	0	0
4	0.5	0
6	1.0	0
8	1.5	0
10	3.0	0
12	5.0	.5
14	8.0	1.0



67. Which of the following BEST describes the problem Jodie is studying?

A. What conditions are needed for bread to mold?
B. Does mold grow better where it is warm?
C. Does cold temperature slow the growth of mold?
D. Can mold grow inside a sealed container?

68. Which of the following BEST describes Jodie's hypothesis?

A. If bread is kept cold, then mold won't grow very fast.
B. If mold is refrigerated, then all of the mold will die.
C. If you want to grow mold, then keep it in a warm place.
D. If mold is cold, then it can't reproduce.

69. What is the independent variable in Jodie's experiment?

A. time
B. amount of mold growth
C. amount of water
D. temperature

70. Which of the following is true about the data?

A. Mold in the cold environment grew faster than mold in the warm environment.
B. Both warm and cold mold stopped growing on day 14.
C. Mold in a warm environment shows linear growth.
D. Mold growth in the warm environment was fastest on days 12-14.

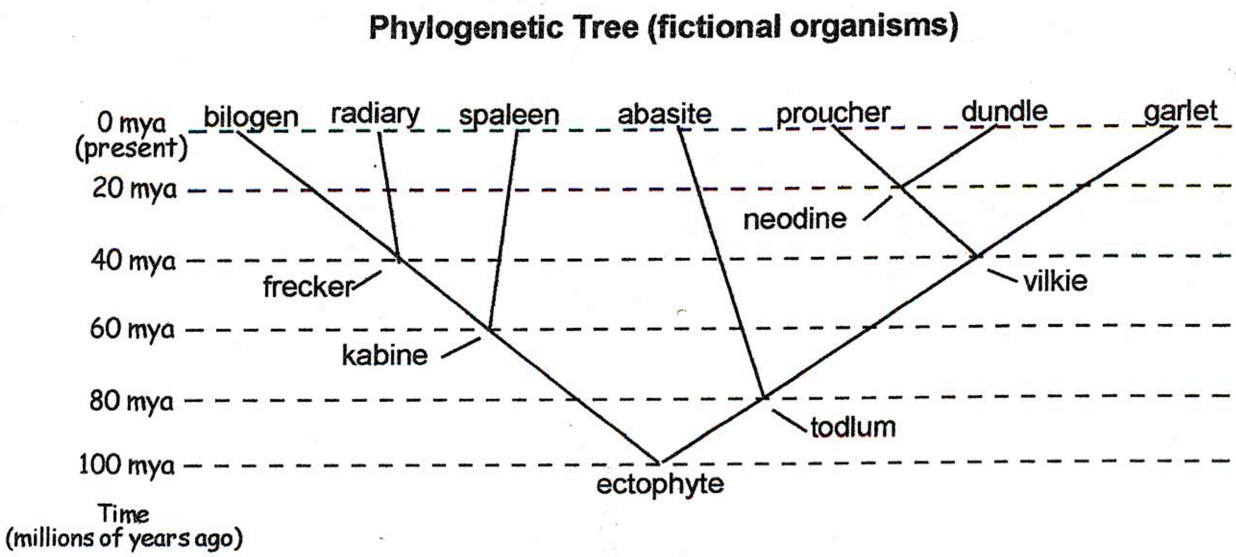
71. Which of the following would be the BEST way to confirm Jodie's results?

A. Repeat the experiment using 2 different kinds of bread.
B. Repeat the experiment using bread, cheese, and an orange.
C. Repeat the experiment the same way several more times and compile the data.
D. Repeat the experiment without putting the bread into baggies.

72. Which of the following is the BEST example of an inference?

A. A fish is a vertebrate.
B. Mammals and reptiles have a common ancestor.
C. Fossils are often found in sedimentary rock.
D. Burning releases CO_2 to the atmosphere.

For 73-75 refer to this diagram.



73. Of the following, which is most closely related to a radiary?

A. spaleen
B. kabine
C. bilogen
D. dundle

74. Which pair of organisms have the most recent common ancestor?

A. bilogen and spaleen
B. abasite and garlet
C. radiary and bilogen
D. proucher and dundle

75. The most recent ancestor of the proucher and the garlet lived about

A. 100 million years ago
B. 60 million years ago
C. 40 million years ago
D. 20 million years ago

76. Which of the following is TRUE about scientific theories?

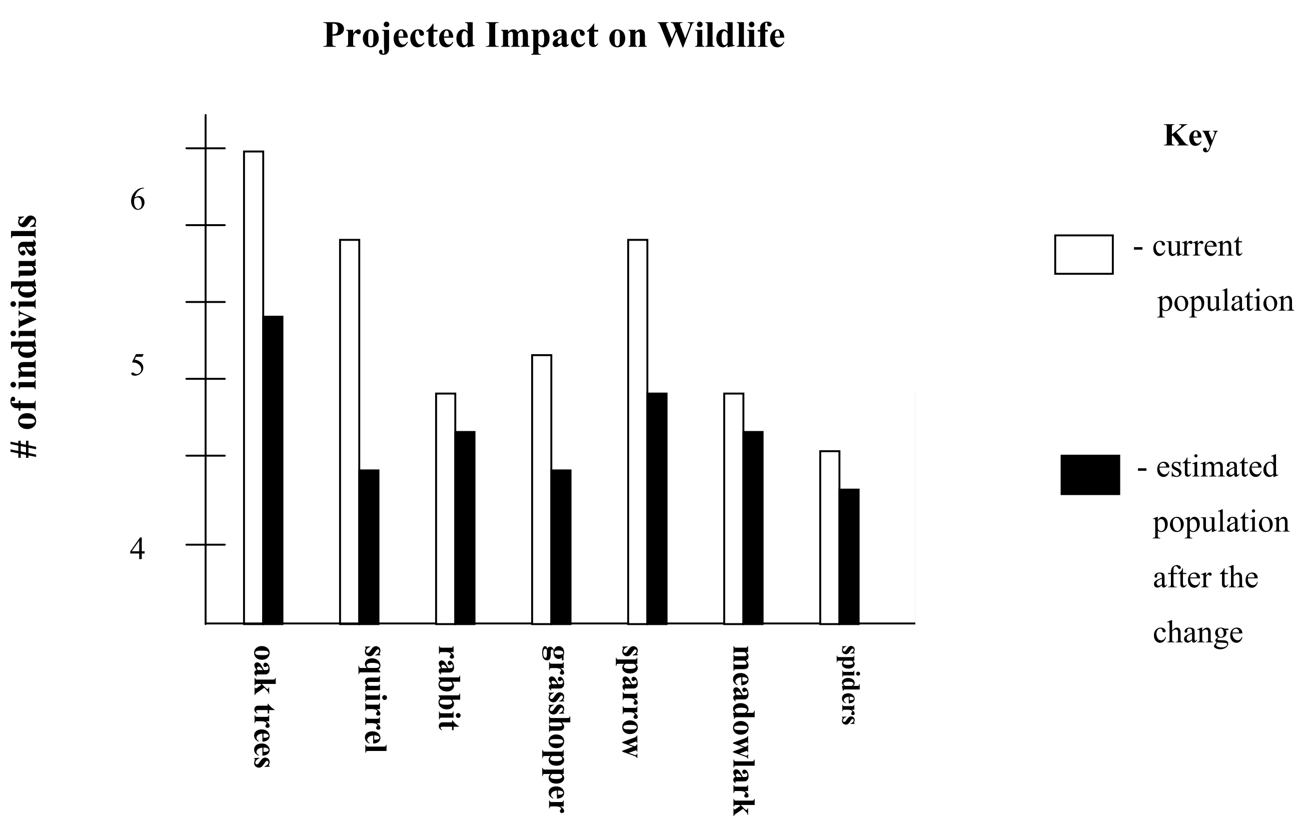
A. If no evidence is found to refute a theory, it becomes a fact.
B. There can be only one correct theory for any particular concept.
C. A scientific theory may be changed to explain new evidence.
D. none of these

For items 77-80 refer to the information below.

A major American retail chain is considering opening a new store on the eastern edge of town. The site is currently 400 acres of deciduous forestland.

If the city engineers agree to this change, over 150 species of plant life, animals, birds, and insects would be affected in the area. The following graph represents the projected impact of the change on several of the dominant species in the area.

Using data from the graph, answer the following questions.



77. What is the estimated change in the population of maple trees?

A. 2
B. 75
C. 220
D. 150

78. Which species would experience the least change in population as a result of building the retail chain?

A. sparrows
B. oak trees
C. rabbits
D. grasshoppers

79. What question is this study trying to answer?

A. Will building this store affect the survival of species in this area?
B. Would building this store be good for the economy?
C. What kinds of wildlife live in the area where the store will be built?
D. Is this the best location for the new store?

80. Which is the most likely hypothesis that could be formed based on the estimated data?

A. If building the shopping center harms squirrels, then all squirrels will become extinct.
B. If the new shopping center is built, then there will be lower numbers of organisms in the area.
C. If the new shopping center is built carefully, then the environment will be improved by it.
D. If the new shopping center brings more business to the area, then new jobs will be created.

Appendix B - Spring 2006 Item Analysis Details

Table B-4-1: Item Analysis Data for USD 475 2006 Spring Semester Biology CRT

Item No.	Correct Answer	Percent of Students with the Correct Answer	Percent of Students with the Incorrect Answer	Point-Biserial Correlation Coefficient	Answer Choice Response Frequency				
					A	B	C	D	E
1	B	94.60%	5.40%	0.36	10	333	6	3	
2	D	83.81%	16.19%	0.04	11	6	40	295	
3	C	75.85%	24.15%	0.33	17	34	267	34	
4	B	79.55%	20.45%	0.2	5	280	59	8	
5	A	76.70%	23.30%	0.3	270	27	34	21	
6	C	85.80%	14.20%	0.34	9	16	302	25	
7	D	92.61%	7.39%	0.19	5	3	18	326	
8	A	57.10%	42.90%	0.43	201	22	46	83	
9	B	49.15%	50.85%	0.47	92	173	52	35	
10	C	34.38%	65.63%	0.38	14	64	121	153	
11	A	77.27%	22.73%	0.36	272	59	7	14	
12	B	60.51%	39.49%	0.46	65	213	16	58	
13	C	86.65%	13.35%	0.28	15	10	305	22	
14	A	86.08%	13.92%	0.4	303	13	30	6	
15	D	70.17%	29.83%	0.43	5	70	30	247	
16	D	87.78%	12.22%	0.3	17	17	9	309	
17	A	59.94%	40.06%	0.34	211	58	16	67	
18	D	75.28%	24.72%	0.28	48	9	30	265	
19	B	72.44%	27.56%	0.33	8	255	46	43	
20	C	93.75%	6.25%	0.27	11	9	330	2	
21	C	81.82%	18.18%	0.41	48	8	288	8	

22	A	87.50%	12.50%	0.31	308	16	2	26	
23	D	62.50%	37.50%	0.5	30	29	73	220	
24	C	78.98%	21.02%	0.34	13	46	278	15	
25	C	77.27%	22.44%	0.33	29	28	272	22	
26	A	79.26%	20.74%	0.38	279	13	21	39	
27	C	68.18%	31.82%	0.37	17	83	240	12	
28	B	46.02%	53.69%	0.36	43	162	57	89	
29	B	60.80%	39.20%	0.36	50	214	74	14	
30	A	21.88%	78.13%	0.17	77	54	37	184	
31	D	83.81%	16.19%	0.12	36	6	15	295	
32	D	67.61%	32.39%	0.17	28	12	74	238	
33	C	88.92%	11.08%	0.29	20	7	313	12	
34	D	65.34%	34.66%	0.28	82	14	26	230	
35	B	79.83%	19.89%	0.29	31	281	5	34	
36	C	84.38%	15.63%	0.4	15	38	297	2	
37	A	28.98%	71.02%	0.21	102	125	118	7	
38	D	92.90%	6.82%	0.26	11	2	11	327	
39	B	70.45%	29.26%	0.25	46	248	9	48	
40	A	80.97%	19.03%	0.32	285	52	5	10	
41	C	92.33%	7.67%	0.26	4	6	325	16	1
42	B	82.67%	17.33%	0.49	19	291	6	10	26
43	A	82.39%	17.33%	0.4	290	9	2	26	24
44	E	80.97%	19.03%	0.49	15	10	2	40	285
45	D	75.28%	24.72%	0.49	31	20	13	265	23
46	B	85.51%	14.49%	0.41	13	301	30	8	
47	B	88.35%	11.65%	0.19	17	311	18	6	
48	B	74.43%	25.57%	0.55	14	262	16	46	14
49	A	82.10%	17.90%	0.43	289	39	13	8	3
50	E	96.31%	3.69%	0.21	1	3	3	6	339
51	C	83.52%	16.48%	0.43	33	11	294	13	1

52	D	78.41%	21.59%	0.5	15	37	17	276	7
53	D	77.56%	22.44%	0.47	52	13	14	273	
54	D	66.19%	33.81%	0.39	8	39	72	233	
55	C	71.02%	28.98%	0.32	70	16	250	16	
56	A	52.84%	47.16%	0.19	186	70	15	81	
57	B	88.92%	11.08%	0.33	10	313	13	16	
58	C	59.66%	40.06%	0.33	91	45	210	5	
59	A	52.27%	47.73%	0.22	184	10	32	126	
60	D	63.64%	36.36%	0.31	5	101	22	224	
61	B	65.06%	34.94%	0.37	93	229	7	23	
62	D	86.65%	13.35%	0.39	26	12	9	305	
63	C	54.83%	45.17%	0.3	31	96	193	32	
64	D	47.73%	52.27%	0.21	24	137	23	168	
65	A	44.60%	55.40%	0.4	157	112	42	41	
66	A	35.51%	64.49%	0.34	125	12	160	55	
67	C	51.14%	48.86%	0.37	117	47	180	8	
68	A	79.26%	20.45%	0.35	279	9	44	19	
69	D	30.11%	69.89%	0.11	146	84	16	106	
70	D	63.35%	36.36%	0.28	13	25	90	223	
71	C	65.63%	34.38%	0.16	61	21	231	39	
72	C	32.67%	67.33%	0.05	56	115	115	66	
73	C	76.99%	23.01%	0.25	34	43	271	4	
74	D	75.57%	24.43%	0.43	16	14	56	266	
75	C	74.72%	25.00%	0.33	38	6	263	44	
76	C	73.30%	26.70%	0.41	15	21	258	58	
77	C	70.17%	29.55%	0.42	64	24	247	16	
78	A	77.56%	22.44%	0.35	273	25	28	26	
79	A	69.32%	30.40%	0.3	244	37	30	40	
80	B	82.67%	17.05%	0.44	18	291	30	12	