POINT LOMA NAZARENE UNIVERSITY

A Revision of the Conceptual Inventory of Natural Selection

for

Middle School Students

A thesis submitted in partial satisfaction of the

requirements for the degree of

Master of Science

In General Biology

By

Patricia Lee Evans

Committee in charge:

Dr. Dianne Anderson, Chair
Dr. Ryan Botts
Dr. April Maskiewicz

2013

The thesis of Patricia Lee Evans is approved, and it is acceptable in quality and form for publication:

Chair

Point Loma Nazarene University
2013

Acknowledgement

I would like to thank Dr. Dianne Anderson for all her help, trust, and encouragement in revising the CINS for middle school students. Her insights and patience were invaluable. I would also like to thank April Maskiewicz for encouraging me to take a risk and pursue a thesis in the first place. Without her push, this project would never have gotten started. I would like to thank Dr. Ryan Botts for taking the time to assist with the statistical analysis of the data. I would like to thank Laurie Owen and Jodi Takei for allowing me to use their students for data collection. Finally, I would like to thank my husband who stood by my side without question through the long years of study.

Table of Contents

List of Figures and Tables

Abstract of the Thesis

A Revision of the Conceptual Inventory of Natural Selection

For Middle School Students

By

Patricia Lee Evans

Master of Science in General Biology

Point Loma Nazarene University, 2013

Dr. Dianne Anderson, Chair

The Conceptual Inventory of Natural Selection (CINS) developed by Dianne Anderson and Kathleen Fisher is a valuable tool to measure how well college students understand natural selection. However, despite its value at the undergraduate level, its high readability deemed it of little use at the middle school level. The purpose of this study was to make the CINS more equitable for middle school students. Changes in the form of word substitutions and shortened sentences were made to the CINS in an effort to lower its readability. The CINS was also divided into two shorter versions. Semi-structured interviews were conducted with 12 students to determine the outcome of these changes. The old and new versions of the CINS were administered to 325 students taking life science at a middle school in San Diego. Chi-squared analysis of the answer choices revealed that a significant number of students were guessing on seven of 20 questions on the original CINS in comparison to zero instances of significant guessing on the new versions. Bi-dimensional analysis supported this finding revealing that most students had specific reasons for their answer choices, whether in agreement with experts or not. It is hoped that the new CINS versions, now more accessible to middle school students, will be a useful tool in determining student alternative conceptions at this age, and help uncover the learning progression of students as they move from middle school to high school, and then to college.

Introduction

The theory of natural selection is considered to be a vital explanatory mechanism of the process of evolution. Ernst Mayr, a biologist and historical scholar, says of Darwin's work "…no one has influenced our modern world-view--both within and beyond science—to a greater extent than has this extraordinary Victorian" (Carroll, 2003, pg. 9). Biological theorist Michael Ghiselin states of Darwin's theory "…to learn the facts, one reads the latest journals. To understand biology, one reads Darwin" (Ghiselin, 1969, p. 232). So influential is the theory of natural selection, the National Research Council has recognized it as a unifying theory in biology education (NRC, 1996), and it has been described as a revolutionary idea that surpasses even the revolution of Copernicus (Gould, 2009). What makes it so unique is that it is "a view of what life is and how it came to be, and above all else, a view of who we humans are and how we came to be—"(Eldridge, 2005, pg. 8).

Given its importance, it is no surprise that the topic of natural selection has received much attention by science education researchers over the last several decades. Demastes, Good, and Peebles (1996) studied students' patterns of conceptual change in natural selection. Other researchers have examined how students' understanding of natural selection changes with different curricula and instructional strategies (Jensen & Finley, 1996). A myriad of alternative conceptions have been documented by a number of scholars (Bishop and Anderson, 1986; Clough & Driver, 1986; Cummins, Demastes & Hafner, 1994; Good, Trowbridge, Demastes, Wandersee, Hafner, & Cummins, 1992; Scharmann & Harris, 1992).

In 2002, Anderson, Fisher, and Norman developed the Conceptual Inventory of Natural Selection (CINS) in order to create a test that would provide "a simple yet effective means of identifying the frequency of some common misconceptions among large numbers of students" (Anderson et al. 2002, pg. 967). The CINS was originally designed for college students, and has been used by various teachers and researchers since its inception as a means of gathering pre and post-test data in research studies involving the understanding of natural selection. While the CINS has been successfully used with both college and high school students and has become a mainstay for educational research, surprisingly, it has not been tested with middle school students.

Given that middle school is the time when students are first introduced to natural selection, there is a need to include this younger group in studies involving the CINS. However, there are two major concerns with regards to administering the CINS to this age group. First, there is the issue of whether or not middle school students taking the CINS are able to understand the questions and answer choices of the test. The questions were designed for college readability (Anderson, et al., 2002). In addition, the CINS was designed in the style of Jungwirth in which long text is used in both the question header and item responses (Anderson, et al., 2002). This style may not be suitable for middle school students whose reading levels vary and are still developing. A pilot study conducted in the spring of 2010 by this author examined the above concern. The results suggest that the CINS in its original form is not suitable for middle school, whether the students are of a non-English Learner (non-EL) or EL designation. There is a definite need to revise the CINS for this age group. Second, it is not clear at what level middle school students comprehend the theory of natural selection as opposed to their older counterparts. Thus, the purpose of this study was to revise the CINS for middle school in a manner that made it equitable for middle school students, and which takes into consideration students' understanding of natural selection at the age of 11 or 12 years old.

Literature Review

Theoretical Perspective

Constructivism was born out of the work of the Swiss psychologist, Jean Piaget as described by Lawson (Lawson, 1994). At the core of constructivist thinking is the idea that students build on their existing knowledge. But students do not simply add new information to what they already know. There is an elaborate mental process that a student must undergo in order to acquire new knowledge. Piaget called this process "Psychological Equilibration" (Lawson, 1994, 136). When students are presented with new information, they first attempt to assimilate, or fit, what they encounter in the environment into their individual schema. Schemas are thought to be higher order cognitive structures that explain how old knowledge interacts with new knowledge in thoughts, memory, perception and language (McVee, 2005). If they cannot assimilate the new information, disequilibrium occurs, followed by accommodation in which they modify (or adapt) their schema to fit the evidence presented (Lawson, 1994).

Vygotsky, another psychologist, shared many of the same views as Piaget, however, Vygotsky placed a greater emphasis on social context, and, thus, is considered a social constructivist. Culture, tools, social interactions and language are viewed as important in the construction of knowledge. In fact, knowledge can be thought of as being co-constructed through these avenues of learning. An example is Vygotsky's description of language as a tool. Not only is language used to construct knowledge, but, in his book *Thought and Language*, Vygotsky describes the two as interdependent. Thought cannot be constructed without language and language cannot be constructed without thought (Howe, 1996). Language, as a tool, also plays a role in Vygotsky's Zone of Proximal Development. It is through the use of scaffolding language either by teachers or peers that students are able to move from one level of understanding to the next level (Howe, 1996). Thus, the use of language is critical when considering, either, the learning, expression of learning, or even the testing of learning of students in science.

Alternate Conceptions

In the process of constructing knowledge, students can often develop alternative conceptions or they may arrive at the classroom with alternative conceptions (Wenning, 2008). Alternative conceptions are ideas held by students that run counter to generally accepted scientific explanations. An alternative conception is considered authentic if it is held by a significant proportion of the students and is highly resistant to change (Anderson, et. al, 2002; Wenning, 2008). Alternative conceptions have been well researched in physics whereas alternative conceptions research in biology is fairly new, i.e. since the mid-80s. Wandersee, Mintzes, & Novak (1994) have "generated eight 'emerging' research-based claims relating to alternative conceptions in science" (Wenning, 2008, pg. 11). The claims are as follows:

Claim 1: *Learners come to formal science instruction with a diverse set of alternative conceptions concerning natural objects and events.*

Claim 2: *The alternative conceptions that learners bring to formal science instruction cut across age, ability, gender, and boundaries.*

Claim 3: Alternative conceptions are tenacious and resistant to extinction by conventional teaching strategies.

Claim 4: *Alternative conceptions often parallel explanations of natural phenomena offered by previous generations of scientists and philosophers.*

Claim 5: *Alternative conceptions have their origins in a diverse set of personal experiences including direct observation and perception, peer culture, and language, as well as in teachers' explanations and instructional materials.*

Claim 6: *Teachers often subscribe to the same alternative conceptions as their students.*

Claim 7: *Learners' prior knowledge interacts with knowledge presented in formal instruction, resulting in a diverse variety of unintended learning outcomes.*

Claim 8: *Instructional approaches that facilitate conceptual change can be effective classroom tools.*

The identification of alternative conceptions is critical to good teaching. If students build upon their existing knowledge, as constructivists believe, then teachers should first identify a student's ideas, and then provide learning experiences that further build upon that student's ideas or in the case of alternative conceptions, challenge them. Through the identification of alternative conceptions, teachers can also design intervention strategies so that obstacles are not created and advanced learning is not impeded later on (DeBoer, 2011). Research into the alternative conceptions of biology is still emerging as compared to the physical sciences (Tanner and Allen, 2005). Figure 1 shows the number of publications on alternative conceptions in each of the main fields of science (Tanner and Allen, 2005).

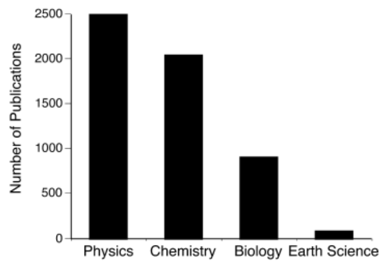


Figure 1. Analysis of number of publications (n=6,314) from the Students' and Teachers' Conceptions and Science Education (STCSE) database, compiled by Duit, 2004.

This is slowly changing as more research is being conducted on the alternative conceptions of biology students.

Natural Selection

While biology is behind chemistry and physics in terms of understanding alternative conceptions, one area of biology has advanced significantly with regards to this research field: the topic of natural selection. This is perhaps because natural selection is considered one of the most important scientific theories to have ever been developed and is central to the study of biology.

The theory of natural selection is an elegant theory that rests on five facts and three inferences, as described by Mayr (1982).

Fact 1: All populations have the potential to grow at an exponential rate.

Fact 2: Most populations reach a certain size, then remain fairly stable over time.

Fact 3: Natural resources are limited.

Inference 1: Not all offspring survive to reproductive age in part because of competition for natural resources.

Fact 4: Individuals in a population are not identical, but vary in many characteristics.

Fact 5: Many of the characteristics are inherited.

Inference 2: Survival is not random. Those individuals with characteristics that provide them with some advantage over others in that particular environmental situation will survive to reproduce, whereas others will die.

Inference 3: Populations change over time as the frequency of advantageous alleles increases. These could accumulate over time to result in speciation.

Although the theory of natural selection appears simple, students in non-major college biology classes often have difficulty with the concepts. Alternative conceptions of natural selection have been well documented with this age group and are listed in Table 1 from Anderson, et. al (2002).

Table 1
Scientific concepts and alternative conceptions addressed in CINS 2002 Version 3. Letters and numbers refer to question number and answer choices.

Topic	Scientific Concept	Alternative Conception
Biotic potential	All species have such great potential fertility that their population size would increase exponentially if all individuals that are born would again reproduce successfully (1C, 11B)	a) Not all organisms can achieve exponential population growth (11C) b) Organisms only replace themselves (1A, 11A) c) Populations level off (1B, 11D, 1D)
Population stability	Most populations are normally stable in size except for seasonal fluctuations (3B, 12A)	a) All populations grow in size over time (3A, 12B) b) Populations decrease (3D, 12C) c) Populations always fluctuate widely/ randomly (3C, 12D)
Natural resources	Natural resources are limited; nutrients, water, oxygen, etc. necessary for living organisms are limited in supply at any given time (2A, 14D)	Organisms can always obtain what they need to survive (2B, 2C, 2D, 14A, 14B, 14C)
Limited survival	Production of more individuals than the environment can support leads to a struggle for existence among individuals of a population, with only a fraction surviving each generation (5D, 15D)	a) There is often physical fighting among one species (or among different species) and the strongest ones win (5B, 15B) b) Organisms work together (cooperate) and don't compete (5A, 5C, 15A)
Variation within a population	Individuals of a population vary extensively in their characteristics (9D, 16C)	a) All members of a population are nearly identical (9A, 16A) b) Variations only affect outward appearance, don't influence survival (9B, 9C, 16B) c) Organisms in a population share no characteristics with others (16D)
Variation inheritable	Much variation is heritable (7C, 17D)	a) When a trait (organ) is no longer beneficial for survival, the offspring will not inherit the trait (7B, 17B) b) Traits acquired during an organism's lifetime will be inherited by offspring (7A, 17A) c) Traits that are positively influenced by the environment will be inherited by offspring (7D)
Differential survival	Survival in the struggle for existence is not random, but depends in part on the hereditary constitution of the surviving individuals. Those individuals whose surviving characteristics fit them best to their environment are likely to leave more offspring than less fit individuals (10C, 18B)	a) Fitness is equated with strength, speed, intelligence or longevity (10A, 10B, 18A, 18C, 18D) b) Organisms with many mates are biologically fit (10D)

(Continued)

Table 1

(Continued)

Topic	Scientific Concept	Alternative Conception
Change in a population	The unequal ability of individuals to survive and reproduce will lead to gradual change in a population, with the proportion of individuals with favorable characteristics accumulating over the generations (4B, 13B)	a) Changes in a population occur through a gradual change in all members of a population (4A, 13A, 17C) b) Learned behaviors are inherited (4C, 13C) c) Mutations occur to meet the needs of the population (4D, 13D)
Origin of species	An isolated population may change so much over time that it becomes a new species (8A, 20B)	a) Organisms can intentionally become new species over time (an organism tries, wants, or needs to become a new species) (8C, 8D, 20A, 20D) b) Speciation is a hypothetical idea (8B, 20C)
Origin of variation	Random mutations and sexual reproduction produce variations; while many are harmful or of no consequence, a few are beneficial in some environments (6B, 19C)	a) Mutations are adaptive responses to specific environmental agents (6C, 15C, 19D) b) Mutations are intentional: an organism tries, needs, or wants to change genetically (6A, 6D, 19A, 19B)

Learning Progressions

While it is clear which areas are troublesome for college students, it is less clear what alternative conceptions are held at the lower grades including high school and middle school (Liu & Lesniak, 2005; Nehm, 2006; NRC, 2001). No one is quite sure how student ideas develop over time in this domain, and other science domains, although we do have standards that dictate which ideas should be taught when. Some researchers are calling for the development of learning progressions that are "are empirically-grounded and testable hypotheses about how students' understanding of, and ability to use, core scientific concepts and explanations and related scientific practices grow and become more sophisticated over time, with appropriate instruction" (Corcoron, Mosher, & Rogat, 2009, pg. 8). These hypotheses describe how students are likely to master core concepts and they are based on research conducted on how students' learning actually takes place

(Corcoron, et. al, 2009). A number of important elements have been identified as being important to learning progressions. (Corcoron et. al, 2009, pg. 15):

> 1) Target performances or learning goals which are the end points of a learning progression and are defined by societal expectations, analysis of the discipline, and/or requirements for entry into the next level of education;
>
> 2) Progress variables which are the dimensions of understanding, application, and practice that are being developed and tracked over time; these may be core concepts in the discipline or practices central to scientific work;
>
> 3) Levels of achievement that are intermediate steps in the developmental pathway(s) traced by a learning progression. These levels may reflect levels of integration or common stages that characterize the development of student thinking. There may be intermediate steps that are non- canonical but are stepping stones to canonical ideas;
>
> 4) Learning performances which are the kinds of tasks students at a particular level of achievement would be capable of performing. They provide specifications for the development of assessments by which students would demonstrate their knowledge and understanding; and,
>
> 5) Assessments, which are the specific measures used to track student development along the hypothesized progression. Learning progressions include an approach to assessment, as assessments are integral to their development, validation, and use.

Although no such learning progression yet exists for natural selection, the work has begun. Catley, Lehrer, and Reiser (2005) were commissioned by the NRC to develop a learning progression on evolution. Because there is little research on how young children understand evolution, the development of the learning progression is still ongoing (Corcoron et. al, 2009). The NSF-funded AAAS Project 2061, while still in its infancy, has begun to create concept strand maps in order to show how students' understanding in science develops from Kindergarten through grade -12 (Atlas of Science Literacy, AAAS website, 2011). It is important to note these are hypothesized concept strands only. A sizable effort has been made by

Project 2061 to provide multiple choice questions regarding various biology topics to teachers to assess understanding of these topics. These AAAS Science Assessment questions are written much like CINS questions, with one correct answer and multiple distractors. The list of questions is impressive, over 600 to date, tested on over 2000 students. However, while these questions will undoubtedly provide valuable information to teachers, more comprehensive, domain specific assessments will still be needed to track hypothesized progressions. The CINS is a good starting place for assessing understanding in the domain of natural selection. It has been well tested at the college level. What is needed is a refined CINS for the lower grades. Among the many suggested starting points for research into domain specific learning progressions, Corcoron advocates the use of already existing assessments (Corcoron, 2009, pg. 38):

> They can be based on cross-sectional sampling of student performance by age or grade, using existing assessments, or assessments created to try to reveal aspects of understanding or performance thought not to be tapped by existing assessments. This sampling also can include more open-ended interviewing or observation of students at work (and collection of the work they produce).

By creating a CINS accessible to more grade levels, a better picture can be created of how students progress in their learning of natural selection.

How Natural Selection is taught in our schools

The CINS assesses a student's knowledge of natural selection. A look at the standards across grade levels in California shows the progression of ideas taught to students in this domain (Appendix A). It is important to note these standards were formulated by selecting topics based on what was perceived to be a logical step by step advancement of topics through the different grade levels. They were based on current disciplinary knowledge and on the personal experience of teachers. There was no research to determine the validity of these progressions (Corcoron, et. al, 2009).

Middle school is where students are formally introduced to the theory of natural selection, but a close look at the standards as written in the California Science Framework (2004) reveals that students are exposed to key concepts of the theory in the lower grades (see Appendix A).

In grade seven, students are finally introduced to Darwin and the lines of evidence that led him to develop his theory of natural selection. During this same school year students are also introduced to genetics. In grade nine, students expand on their knowledge of genetics and the theory of Natural Selection.

Several questions arise when reading the frameworks. How well do students understand the standards? Do they reflect the learning progression of the students? What sort of alternative conceptions are found at the different grade levels? As will be discussed in the next section, concept inventories can help answer these questions.

Importance of Diagnostic tests

Conceptual inventories are research-based tests that measure how well students either understand given concepts or hold alternative conceptions. A good concept inventory can inform teachers which core concepts are misunderstood and also which mental constructs students cling to (D'Avanzo, 2008; Klymkowsky & Garvin-Doxas, 2008). Each question specifically targets one concept. The answers are usually comprised of one correct choice and a number of other choices meant to act as distractors (wrong answers) which either represent alternative conceptions or faulty reasoning (D'Avanzo, 2008). Concept inventories differ from other tests in that they circumvent test taking strategies that students may have. In addition, they are written in the language of the student. (Klymkowsky & Garvin-Doxas, 2008).

In 2002, Anderson, *et. al*, developed the CINS or Conceptual Inventory of Natural Selection. The CINS was the first validated conceptual inventory for biology (D'Avanzo, 2008). Other inventories and diagnostic tests include the Biology concept inventory (Klymkowsky & Garvin-Doxas, 2008), the Diagnostic Question Cluster: Genetics (Parker, et.al, 2008), the Cell division diagnostic tests (Williams, et. al, 2008), and the osmosis and diffusion diagnostic test (Williams, et. al, 2008), and the list keeps growing. The CINS test items focus on the five facts and three inferences considered crucial to the logic of natural selection (Mayr, 1982). Anderson et. al, also included test items on the origin of variation and the origin of species (Anderson, et. al, 2002). The test is made up of 20 questions, two for each of the above concepts. Answer choices are either the correct answer or are based on the alternative conceptions listed in Table 1.

What makes the CINS unique is not just its content, but the manner in which it was designed. To avoid the issue of the hypothetical scenarios used on other assessments, the questions of the CINS are based on authentic evolutionary studies (Anderson, et. al, 2002):

- Galapagos finches (Grant, Grant, & Petren, 2000, 2001)
- Venezualan guppies *Poecilia reticulata* (Endler, 1980)
- Canary Island lizards (Thorpe & Brown, 1989)

The CINS questions have been field tested numerous times and validated with semi-structured interviews. The CINS is a criterion-reference test designed to assess mastery of subtopics (Anderson, et. al, 2002). Its ease of use and ability to quickly diagnose areas of need is considered by many to be one of its strong points. Some, however, have pointed out weaknesses of the CINS. Nehm and Shonfeld (2008) have raised the issue as to whether or not the test is really suitable for college biology non-majors. They contend that many of the questions of the CINS have marginal discriminability and moderate difficulty which suggests that the test is very difficult for undergraduate non-majors (Nehm & Schonfeld, 2008). A pilot study conducted by this author in the spring of 2010 with middle school students suggests that there are issues with the readability of the CINS for the middle school level which might account for the discriminability issues seen in Nehm and Schonfeld's study.

The importance of readability of conceptual inventories

Research has shown a direct correlation between language proficiency and testing proficiency. Abedi, Hofstetter, and Lord (2004), have studied EL and non-EL learners and their ability to perform on tests in math. They point out that "research has documented amply the impact of students' language background on test performance. Students who lack proficiency in the language of the test consistently perform at lower levels" (Abedi, et. al, 2004, pg. 5). Presumably this is because there is a disconnect between the language of the student and the language of the test. Research has shown that reducing the linguistic complexity of a test is beneficial for EL students (Siegal, 2007). Abedi et. al, (2004) has also noted that changes in the language of the test can result in higher test scores.

Modifying tests for EL students is not simply a matter of changing a few words. The goal when making accommodations to a test involves decreasing the language factor, all the while keeping the content the same (Siegal, 2007). This is not easy, because language and content are intimately connected (Siegal, 2007). Learning science vocabulary is a complex process. As with language learning, learning vocabulary involves developing relationships among ideas and terms, in addition to their meanings (Lee & Fradd, 1998). Thus, in order to make assessments equitable, Siegal (2007) has suggested applying the McCes framework for equitable assessment. The McCes framework is as follows:

1. Match the learning goals of the original items and match the language of instruction; Be comprehensible for ELs, both linguistically and culturally;
2. Challenge students to think about difficult ideas, without watering down content;
3. Elicit student understanding; and
4. Scaffold the use of language and support student learning

Abedi has also made recommendations for equitable assessment. Abedi et al. (2004), examined the success of different types of accommodations for EL learners. Those accommodations were as follows: tests in student's native language, linguistic modification, extra time, published dictionaries, glossary and customized dictionaries, oral administration and other accommodations. The findings showed that linguistic modification provided the greatest benefit to EL students without affecting non-EL student performance. In 2005, Abedi, Courtney, Mirocha, and Leon, examined several means by which a test could be linguistically modified. The findings were as follows: use simpler words or high frequency words; use familiar words, use no colloquialisms; retain subject-verb-object structure in sentences and avoid clauses; remove expository material; and use present tense or active voice.

Given that the CINS was designed for college non-EL students, clearly the CINS is difficult for middle school students, both, EL and non-EL alike. A revision of the CINS is necessary for the middle grade levels. For instance, a CINS question that was troubling to students in a pilot study done by this author in 2010 is shown below.

3. Once a population of finches has lived on a particular island with an unvarying environment for many years,

    a. the population continues to grow rapidly
    b. the population remains relatively stable, with some fluctuations.
    c. the population dramatically increases and decreases each year.
    d. the population will decrease steadily.

In the pilot study, students were asked to underline words and phrases they did not understand. Those words and phrases were then coded as follows:

Key: *A=Words are long, and not high frequency words*
B=Words are short and not familiar
C=Sentences are complex syntactically. Sentences do not retain subject-verb-object structure in sentences. Sentences contain clauses.
D=Sentences contain expository material
E=Sentences do not use present tense or active voice

Within the question itself, students had particular trouble with the words "unvarying environment". For answer b, students had difficulty with the word "fluctuations" and the phrase "relatively stable, with some fluctuations". For answer d, students had difficulty with the word "steadily". The types of words that students had the most difficulty with were coded as A, i.e. long words that were unfamiliar.

Although the majority of middle school students are not EL, all students fall somewhere on the continuum of learning to read. The issues at play with EL students are the same for any student still learning to read text. At the middle school level, many students are still struggling with morphemes. Many of the words in the CINS question above contain what are described as morphemes, by linguists.

Morphemes are the smallest unit of a word that have meaning. Morphemes can be bound or free. A bound morpheme cannot stand alone. An example is the prefix “un” as in “unvarying”. Free morphemes can stand by themselves and the word “break” is an example (Dehaene, 2009, 22). The CINS is full of morphemes. The clearest examples are the following words: “vary”, “variable”, “unvarying”, and “variation”.

Dehaene utilized Frith’s description of three stages of reading that all children go through before becoming expert readers (Dehaene, 2009, p. 200-204). Table 2 shows these three stages of reading. The middle school students in the pilot study described above fell somewhere between the phonemic stage and the orthographic stage, i.e. they still had trouble with word length, especially morphemes of a base word, and had trouble with unfamiliar words.

Table 2
The three stages of reading as described by Dahaene.

Reading stage	**Characteristics of reading stages**
Pictorial Stage	Words are recognized as objects or faces
Phonemic Awareness	Children discover speech is made of phonemes, which can be combined to create new words
Orthographic stage	Word length begins to play less of a role. The number of times a word is encountered becomes important. Students read rare words more slowly and the number of neighbors has a dramatic impact.

Linguistic changes need to be made to the CINS if it is to be administered to middle school students. However, if linguistic changes are made, the scores of both non-EL and EL students would most likely improve. While having both groups improve on an assessment might be seen as a positive in the eyes of teachers, Siegal points out that some researchers view this as a validity problem with respect to EL scores. If modifications are to be made to an assessment, they should improve EL scores without affecting non-EL scores. Siegal, who has worked on revising life science items, states that such claims are based on an "unproven premise" (Siegal, 2007, pg. 876). Siegal makes the argument that high-performing ELs have often been shown to do poorly on tests because of a lack of content knowledge and not language ability. In fact, other researchers such as Solano-Flores and Trumbull (2003) feel that comparisons between ELs and non-ELs should not be the focus, but rather "language factors should be identified as sources of measurement error" (Siegal, p. 876). Indeed, if the CINS is to be used with middle school students in its present form, it will not give an accurate representation of what students actually know for either group. Thus, the research question for this study is as follows:

> Can CINS items be modified to make the test more equitable for middle school students without compromising the biology content?

Methodology

Research Design

A sequential mixed method design (Creswell & Plano Clark, 2011) was used to revise the CINS for use with both EL and non-EL middle school students as shown in Figure 2. The first phase of the study was a qualitative exploration. The CINS was revised from its original form into two tests of 10 questions each. Headers, questions and answers were altered to read at a middle school level using the Dale-Chall readability formula. Questions were reorganized to mirror similar concepts on each of the two test versions and questions were corrected for concept overlap. Next, semi-structured interviews were conducted with middle school students using the revised version of the CINS. Feedback obtained in the interviews was used to revise the CINS, further. Finally, in the quantitative phase of this study, the original test was reduced to two tests of 10 questions each and administered along with the newly revised versions of the test to 325 students as a post test. The post test scores were compared to determine if there was improvement in the overall revision of the test and to see if there was a statistically significant difference in student performance between the new version of the test and the original version.

Study Site and Participants

The study was carried out during the spring of 2012 at a large suburban middle school, grades sixth through eighth, in Southern California. The student participants were all in the seventh grade and varied in age between 11 and 13 years of age. The school has an ethnically diverse population: out of approximately 1100 students, 7% were African American, 22% Asian, 24% Filipino, 23% Hispanic, 1% Pacific Islander, 17% Caucasian and 2% two or more races. When the study was conducted, 10% of the student population was considered English Learners and 21% were reclassified as Fluent English Proficient (Accountability Progress Reporting, 2011). Eleven classes of seventh grade Life Science taught by three different teachers participated in the study. Teacher one, the author of this study, taught five classes. Teacher two taught five classes, while teacher three taught one class. All students took the test as part of regular class activities, but a smaller group of students were recruited to volunteer to be interviewed. The study was carried out using a protocol and procedures approved by the Point Loma Nazarene University Institutional Review Board (Appendix L).

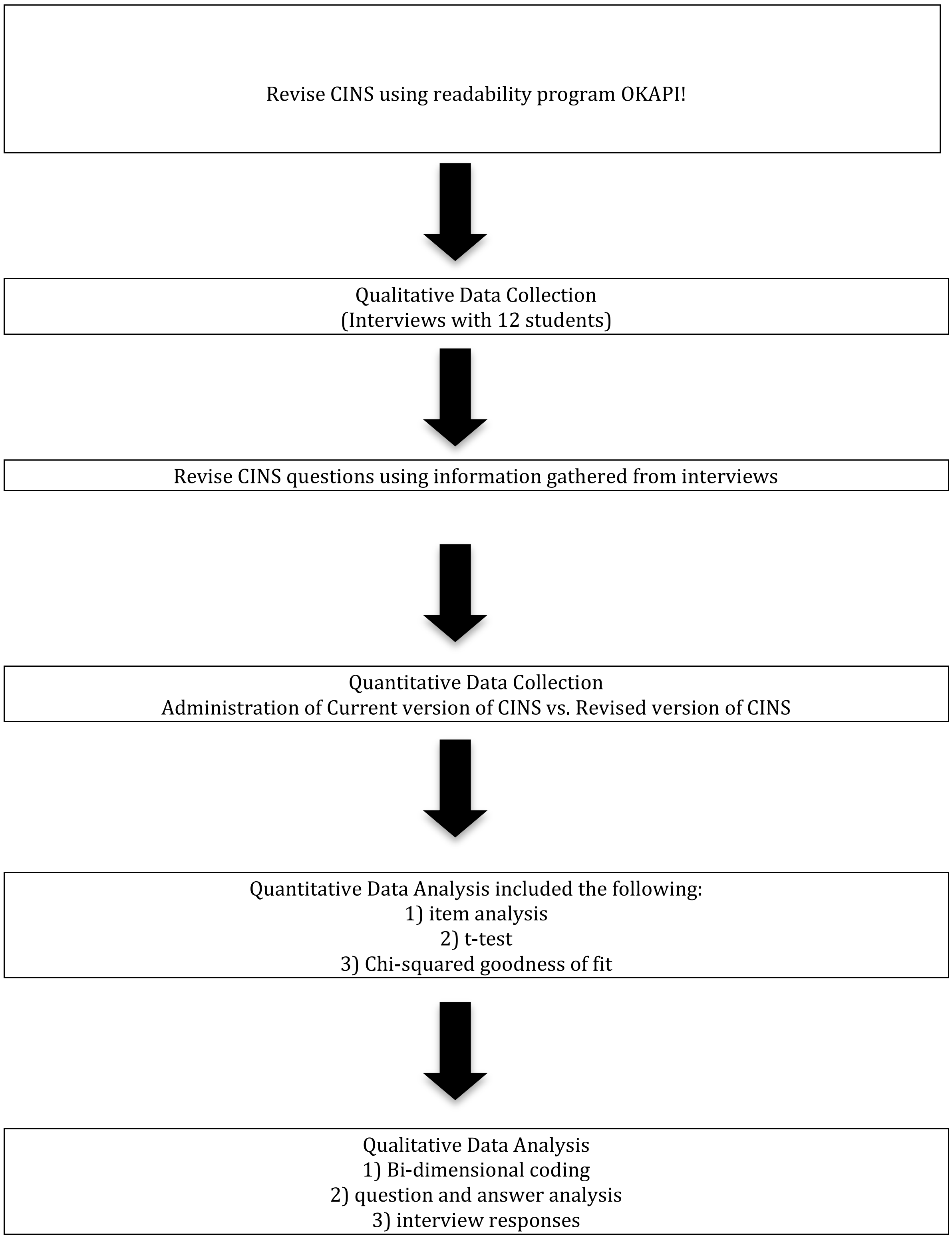


Figure 2. Research design included both qualitative and quantitative phase

Revision of CINS for middle school

A number of readability formulas exist that allow a researcher, or teacher to determine the readability of a text. For this study, the CINS was revised using the Dale-Chall readability formula developed in 1948. It looks at two factors, sentence length, and the number of words that are not on a list of 3000 familiar words (Brown, 2010). It is the most commonly used readability formula. In this study, the headers, questions, and answers were run through the Dale-Chall readability formula available on the website, OKAPI!, (http://www.lefthandlogic.com/htmdocs/tools/okapi/okapi.php). Okapi! is an internet application designed to help users create curriculum that is age appropriate in terms of reading level. Users have the choice of using the Spache readability program which is commonly used up to grade three, and the Dale-Chall readability program, appropriate for grades four and higher. Headers, questions and answers on the original CINS were individually plugged into the readability program to first determine the reading grade level of each component of the test (see Appendix B). Words that were underlined by the program as being difficult were swapped for more age appropriate words when possible. At times, sentence length had to be shortened in order to lower the reading level. Some words such as *evolutionary*, *population*, and *random*, although targeted as difficult, were not substituted because it was determined to be a key word for the subject of natural selection. Words such as “characteristic” were swapped for the word “trait” due to the high number of syllabi in the former. Words and sentence length were swapped until the majority of the test header, questions and answers were determined to be seventh to eighth grade level by the OKAPI program. In order to make it easier to read, the headers were changed from paragraphs to bullets and some information was removed.

In addition to reworking the words and sentence structure of the test, an effort was made to streamline the new CINS in terms of question flow. The original CINS questions tested 10 concepts on each half of the test, but within each half, the concepts were randomly arranged. Based on a previous pilot study using the original CINS with middle school students, there was strong consensus from teachers, students and researchers that the original CINS was too long. Thus, a new goal then arose to split the CINS into two equal tests, a Version A and a Version B, The questions and answers followed a logical progression and mirrored each other both in the concepts addressed and the answer choices. Thus, question one on each test was now *a biotic potential* question, question two was a *stable population* question, and so on, and the answer choices of each question on Version A also aligned with the answer choices for each question on Version B. The new middle school versions of the CINS were designed to follow the sequence of facts and inferences put forth by Mayr (1982). Anderson et al.'s (2002) "origin of variation" question was placed at the beginning of the variation questions. By splitting the test into two equal halves, sequencing bias between the two halves was eliminated (see Appendices C and D for question rearrangement). Because the original CINS was partitioned into three separate header and question answer sections, the context of some of the questions on the middle school versions was altered to fit the header that preceded a particular question. Thus, for example, what used to be a guppy question may have been changed to a lizard question. For ease of grading and analyzing data, a concerted effort was made to ensure that the answers to each question matched the corresponding version questions and answers, thus the answer keys are now identical (see Appendix G, H, and K for revised middle school CINS and answer key).

Some concern had been raised with regard to some questions containing more than one concept. The revised versions of the CINS were carefully scrutinized to avoid concept overlap as was the case with the natural resources and limited resource questions, previously. For example, the original question stem on the natural resources question, question two, read as follows:

2. Finches on the Galapagos Islands require food to eat and water to drink. How does this fact impact the population?

It was unclear from the stem whether or not the question was about natural resources or populations. The question was changed to read as follows:

3. Finches on the Galapagos Islands require food to eat and water to drink. Which statement is true about the birds and the available resources?

The focus of the new question is now clearly on natural resources.

Interviews

In order to determine how well the middle school versions of the CINS measured students' understanding of natural selection, the CINS-MS was reviewed by twelve students post-instruction in semi-structured interviews. Students were recruited for interviews in all classes mentioned above by asking students if they would like to participate in a short 20-30 minute interview after school in a campus classroom. I explained to the students that they would be answering questions about some test items and that the interviews would be audiotaped. It was also explained to the students that their grade would not be affected by the interview and they would need to get consent from their parent. Students who participated in the interviews were given a ten dollar gift card at the end of the session (see Appendix L for assent and consent forms).

Students were asked to perform several tasks as part of the interviews. The first task was a think aloud where students were asked to explain their answers to the following questions (Dwyer, 2011):

Think aloud questions

1. Please tell me which answer you would choose and why.
2. Please tell me why you did not choose each of the other answers.
3. Did you understand what the question is asking?
4. Did you have an issue with any of the words within the question? If so, what words would you take out or change?

In addition to the above questions, students were asked to perform a card sort as described by Anderson et. al (2002):

Card sort

Instructions:
''Please arrange these cards on the table so that the words that are closely related to each other are close together, and those that are unrelated to each other are far apart. If there are any terms that are unfamiliar to you or have nothing to do with natural selection, put them aside.''

adaptation	need
competition	offspring
fitness	population of rabbits
gene	random
individual rabbit	survival
mutation	variation
want	

As part of a third task, students were asked to perform an interview about instances as described by Anderson, et. al (2002):

Interview about instances

Instructions:

"I am going to show you several pictures now. Please tell me whether each one is an example of one or more of the terms on this card, then explain your answer."

Terms on card:
competition
variation within a species
variation between species
survival

Photos shown during post instruction interviews:
1. Cheetah camouflaged in grasses
2. Four distinctly different monkey species
3. Deer pulling the last leaves off branches during a snowstorm
4. Hundreds of birds nesting on a rocky beach
5. Parrot prying open a pine cone with its beak

Use of the interview data to inform revisions of the CINS

Following the interviews, a few changes were made to the revised CINS. Changes that were made were done because students gave consistently similar responses which warranted further attention. Alterations were made to questions nine and ten, and students were re-interviewed on those two questions. A few minor changes were made to other questions to improve readability such as question two (see Appendix C-H for the full set of revisions):

2. A population of finches lives on an island for many years where there are predators. What will most likely happen to the population if everything remains the same?

The word "remains" was changed to "on the island stays" because some students questioned what was meant by the words "remains the same".

2. A population of finches lives on an island for many years where there are predators. What will most likely happen to the population if everything on the island stays the same?

Upon completion of the interviews, 11 classes of seventh grade life science students took either the revised middle school version of the test, Versions A or B, or the original CINS, which had been broken up into two Versions C or D, of 10 questions each (see appendices G, H, I, and J). The test was administered in a random fashion; as students arrived to class, they randomly chose a card from a basket that listed the version of the test they were to take. We already knew from the previous pilot study that the original version would be difficult for the students, thus, the number of students taking the original CINS was reduced from previous studies, while still maintaining a large enough sample for valid statistical tests. The rest of the students took the revised CINS. Those that took the original CINS were taught by one teacher in the study. Those that took the revised CINS were taught by three teachers. 126 students took Version A, 118 took Version B, 34 took Version C and 35 took Version D.

Analysis of the quantitative data

The mean scores of the students who took the revised CINS and the original CINS were compared using a two-tailed independent samples t-test. Item analysis and Chi-squared analysis were completed on both the revised CINS and original CINS test data in order to determine the effectiveness of the revised distractors. Item analysis was done to identify which correct answers or distractors were being chosen by students on a particular item. Chi-squared analysis was performed on the answer choices in order to determine whether students were guessing on items. Chi-squared analysis looks for deviations between what would be the observed frequency and the expected frequency. In this case, given four answer choices, if students were guessing, the expected frequency would be 25%. Significant deviations from this expected frequency would indicate that students either were strongly choosing the correct answer, *if* choosing the correct answer, or strongly choosing a particular distractor. The less students' guessed, the more effective the items. Pearson correlation coefficients were run on questions 1-10 for each test to determine whether there was any perceived concept overlap by students. For example, a strong Pearson correlation between two questions would indicate that students perceived some connection between the two concepts.

Analysis of interview data

Analysis of interview data was conducted using a bi-dimensional coding matrix. Interview data was examined for both concept knowledge and depth of knowledge. The ten concepts the CINS was designed to test (Anderson, et. al, 2002) were given number designations 1-10 as shown in Table 3. Depth of understanding of a particular concept was coded on a letter scale of A-G, where A concurs with expert thinking and G indicates student did not address the question (Table 4). The total number of responses that fell between expert and novice for all concepts was tallied and recorded as totals at the bottom of a bi-dimensional coding matrix similar to the one shown in Table 5. The bi-dimensional coding matrix was adapted from Hogan and Fishkeller (1996, pg. 951).

Table 3
Numbered concepts of CINS

Concept number	Concept on CINS
1	Biotic Potential
2	Stable population
3	Natural resources
4	Limited survival
5	Variation
6	Origin of variation
7	Variation inherited
8	Differential survival
9	Change in population
10	origin of species

Table 4
Modified coding scheme

Modified Coding Scheme		
Code	Code name	Descriptor
A	Compatible Elaborate	Statements concur with the expert proposition and have sufficient detail to show thinking behind them and/or recur throughout the transcript in the same form.
B	Compatible Sketchy	Statements concur with expert proposition, but essential details are missing. Often represent correct guess among choices provided, but no ability to explain why choice was made.
C	Compatibility/Incompatibility	Makes sketchy statements that concur with proposition, but are not elaborated, and makes sketchy statements that disagree. Contradictory statements are often found in two parts of the transcript in response to different questions or tasks on the same topic.
D	Incompatible sketchy	Statements disagree with proposition, but very few details or logic given, and do not recur throughout transcript. Often seem to be responses given just to say something, a guess.
E	Incompatible elaborate	Statements disagree with proposition and students provide details or coherent, personal logic backing them up. Same or similar

		statement/explanations recur throughout transcript.
F	Nonexistent	Used when students respond "I don't know" or do not mention the topic when asked a question calling for its use.
G	No evidence	Used when a topic was not directly addressed by a question and students did not mention it within the context of response to any question.

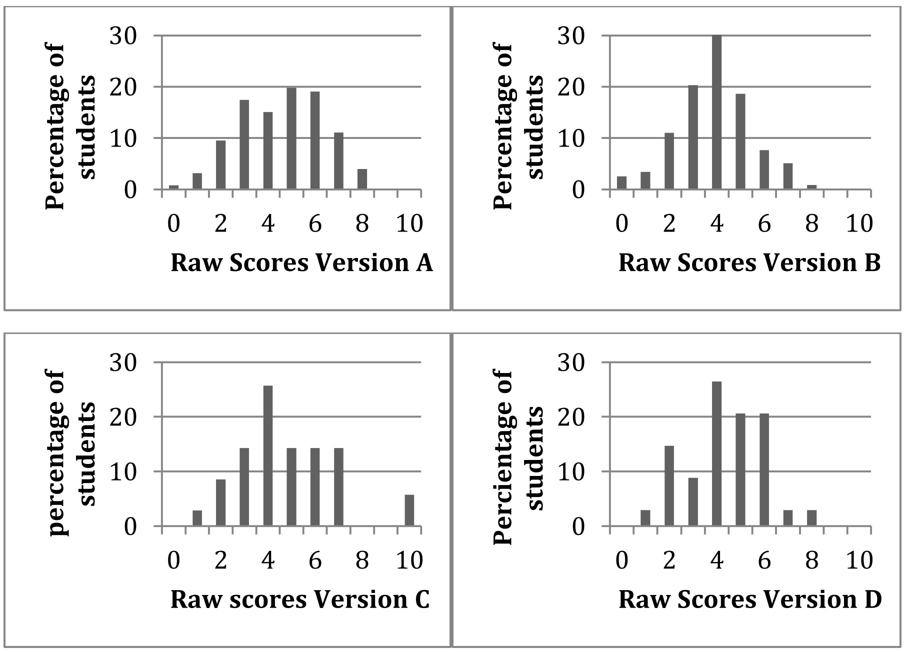
Table 5
Modified Bi-dimensional coding scheme matrix.

	Concept on CINS	A	B	C	D	E	F	G
1	Biotic Potential							
2	Stable population							
3	Natural resources							
4	Limited survival							
5	Variation							
6	Origin of variation							
7	Variation inherited							
8	Differential survival							
9	Change in population							
10	origin of species							
Total	---------------------------							

Results

Quantitative results: overall test scores

Descriptive statistics. Post test scores were analyzed and the percentage of students who achieved a given score out of a possible 10 points was plotted against raw scores for the new middle school versions of the CINS (Versions A and B) and the original CINS (Versions C and D), and are presented below in Figures 3-6. The results for the new versions of the CINS show normal distributions both for Version A and for Version B suggesting the newer versions of the CINS might be more reliable indicators of student achievement. The mean for Version A was 4. 58 out of 10, with a median of 4.70 and an SD of 1.8, where n=126. The mean for Version B was 3.90, with a median of 3.92 and an SD of 1.58, where n=118 (see Table 6). In contrast to Versions A and B, histograms of the percentage of students versus raw scores of Versions C and D do not show typical normal distributions, but rather, exhibit some skewing to the right suggesting the older versions may be less reliable instruments. The mean for Version C was 4.80, with a median of 4.44, and an SD of 2.04, where n=35. The mean for Version D was 4.35, with a median of 4.39, and an SD of 1.61, where n=34 (see Table 6).



Figures 3, 4, 5 and 6 (upper left, upper right, lower left, lower right). Histograms of the distribution of raw student scores on Versions A, B, C, and D, respectively. A raw score of 10 reflects that a student answered all 10 questions correctly.

Table 6
Mean, median, SD and n for each version of the CINS, where
A and B are the new middle school Versions and C and D are the original Versions.

	Version A	Version B	Version C	Version D
Mean	4.58	3.90	4.80	4.35
Median	4.70	3.92	4.44	4.39
SD	1.80	1.56	2.04	1.61
n	126	118	35	34

Inferential statistics

Two tailed, independent samples t-tests were performed on four pairs of CINS results in an effort to determine if there were significant differences between any of the test versions. Thus, comparisons was made between scores on both middle school revisions A and B, the scores of Version A and the original Version C, the scores of Version B and the original Version D, and finally, the scores of both original Versions, C and D. Using a $p<0.05$, the only comparison that yielded a significant difference was the comparison made between the means of Version A and B where $p=0.00253$(see Table 7), indicating the two tests are not equivalent. For all other comparisons there was no significant difference between students' performance on the versions. The smaller sample size of Versions C and D, could account for this latter outcome.

Table 7
T-test results for Versions A and B, Versions A and C, Versions B and D, and Versions C and D where $p<0.05$.

Comparisons	p-value
Versions A and B	0.00253*
Versions A and C	0.53545
Versions B and D	0.14325
Versions C and D	0.31366

*Significant at $p<0.05$

Quantitative results

Analysis of individual questions. In order to determine if the middle school revision of the CINS had answered the question, "Can CINS items be modified to make the test more equitable for middle school students without compromising the biology content", we first determined if there were differences in scores on the individual versions of the exams. Next, using Chi-squared analysis, we analyzed the randomness of responses on individual items. Lastly, we looked at inter-item connectedness using Pearson's correlation.

Standard item analysis was run on the CINS-MS Versions A and B, and on the original CINS, Versions, C and D to determine differences in scores. The results are presented in Tables 8-11. Item analysis of CINS-MS tests A and B reveal that questions 1, 2 and 4 were the least problematic for most of the students. The majority of students were able to choose the correct answers on both versions with the percentage correct ranging from 40%-76%. Differences were seen with question 3 between the two versions. Students performed better on Version A than B (Table 8). For question five, the majority of students were able to choose the correct answer regardless of version, although there was a substantial difference seen in the total number of correct responses between the two versions, i.e. fewer students chose the correct answer on Version A than for Version B (Tables 8 and 9). We see a similar trend occurring with question 7. Students performed better on questions 9 and 10 on Version B. Perhaps the most interesting finding was that the majority of students did not choose the correct answers on questions 6 and 8 and this finding was consistent on both Versions A and B:

Table 8
Standard item analysis for the CINS-MS Version A

#	concept	Correct Group Responses			Point Biserial	Correct answer	Response frequencies			
		Total	Upper 27%	Lower 27%			A	B	C	D
1	Biotic Potential	76.98%	88.24%	70.59%	0.27	C	4	10	*98	15
2	Stable population	72.22%	91.18%	47.06%	0.42	B	8	*91	17	10
3	Natural resources	44.44%	64.71%	8.82%	0.45	A	*56	34	19	17
4	Limited survival	55.56%	91.18%	14.71%	0.58	D	17	10	29	*70
5	Variation	48.41%	76.47%	17.65%	0.41	D	7	43	15	*61
6	Origin of variation	7.94%	14.71%	5.88%	0.05	B	79	*10	23	14
7	Variation inherited	43.65%	67.65%	14.71%	0.42	C	7	23	*55	41
8	Differential survival	21.43%	38.24%	2.94%	0.36	B	57	*27	19	23
9	Change in population	38.89%	55.88%	20.59%	0.38	B	17	*49	14	46
10	Origin of species	48.41%	82.35%	29.41%	0.45	A	*61	7	17	41

*indicates correct answer

Table 9
Standard item analysis for the CINS-MS Version B

#	concept	Correct Group Responses			Point Biserial	Correct Answer	Response frequencies			
		Total	Upper 27%	Lower 27%			A	B	C	D
1	Biotic Potential	61.86%	81.25%	28.13%	0.44	C	9	7	*73	28
2	Stable population	49.15%	75.00%	21.88%	0.31	B	14	*58	25	20
3	Natural resources	37.29%	50.00%	18.75%	0.33	A	*44	54	10	9
4	Limited survival	42.37%	75.00%	21.88%	0.45	D	7	41	19	*50
5	Variation	74.58%	87.50%	53.13%	0.29	D	2	13	14	*88
6	Origin of variation	15.25%	28.13%	3.13%	0.36	B	35	*18	61	3
7	Variation inherited	63.56%	93.75%	37.50%	0.49	C	11	21	*75	10
8	Differential survival	11.86%	31.25%	0.00%	0.33	B	58	*14	6	39
9	Change in population	18.64%	31.25%	9.38%	0.32	B	30	*22	17	48
10	Origin of species	15.25%	21.88%	12.50%	0.22	A	*18	10	47	42

*indicates correct answer

Standard item analysis was also performed on Versions C and D, but because the original sequence of test questions was rearranged and some answers were significantly changed during revision, the questions on Versions C and D did not correspond to Version A and B questions and had to be rearranged for comparison. The raw data is shown in Tables 10 and 11.

Table 10
Standard item analysis for the original CINS, Version C.

#	***Concept***	***Correct group responses***			***Point Biserial***	***Correct Answer***	***Response Frequencies***			
		Total	Upper 27%	Lower 27%			A	B	C	D
1	Biotic Potential	62.86%	66.67%	33.33%	0.27	C	3	5	*22	5
2	Stable population	71.43%	100.00%	55.56%	0.31	A	*25	4	2	4
3	Natural resources	62.86%	100.00%	33.33%	0.42	B	6	*22	6	1
4	Change in population	17.14%	33.33%	11.11%	0.42	B	13	*6	3	13
5	Limited survival	60.00%	100.00%	33.33%	0.52	D	2	3	9	*21
6	Origin of variation	8.57%	33.33%	0.00%	0.63	B	11	*3	12	9
7	Variation inherited	51.43%	77.78%	22.22%	0.47	C	1	6	*19	10
8	origin of species	51.43%	88.89%	22.22%	0.44	A	*18	2	9	6
9	Variation	57.14%	88.89%	44.44%	0.59	D	1	5	8	*20
10	Differential survival	37.14%	55.56%	11.11%	0.48	C	10	8	*13	3

*indicates correct answer

Table 11
Standard item analysis for the original CINS Version D.

#	Concept	Correct group responses			Point Biserial	Correct answer	Response frequencies			
		Total	Upper 27%	Lower 27%			A	B	C	D
1	Biotic Potential	70.59%	88.89%	66.67%	0.26	B	5	*24	1	4
2	Stable population	38.25%	55.56%	22.22%	0.43	A	*13	1	14	6
3	Change in population	29.41%	33.33%	22.22%	0.10	B	5	*10	7	12
4	Natural resources	26.47%	44.44%	11.11%	0.16	D	5	19	1	*9
5	Limited survival	52.94%	66.67%	33.33%	0.32	D	0	5	11	*18
6	Variation	73.53%	100.00%	22.22%	0.67	C	1	4	*25	4
7	Variation inherited	20.59%	33.33%	11.11%	0.21	D	9	10	8	*7
8	Differential survival	26.47%	55.56%	11.11%	0.45	B	11	*9	2	12
9	Origin of variation	47.06%	66.67%	0.00%	0.49	C	8	5	*16	6
10	Origin of species	50.00%	88.89%	22.22%	0.37	B	0	*17	1	15

*indicates correct answer

The original CINS was designed as a 20 question test and although there were two questions per concept, there was no specific order to those concepts. In order to draw clearer comparisons between the original and the new versions, the items for both tests, C and D, and their percent correct were rearranged to match Versions A and B based on the question content (Table 12). For example, on the original CINS, question three signified with a *3, was a *stable population* question, whereas on the new CINS, question 2 became the *stable population* question.

Table 12
Percentage of students who chose the correct answer on each version of the CINS. Each test version was administered to a different group of students. The sample sizes for each version were n=126, 118, 35, and 34 for Versions A,B,C, and D, respectively.

Q#	Concept	Version A N=126	Version B N=119	Version C N=35	Version D N=34
1	Biotic Potential	76.98%	61.86%	62.86%[*1]	70.59%[*1]
2	Stable population	72.22%	49.15%	62.86%[*3]	38.24%[*2]
3	Natural resources	44.44%	37.29%	71.43%[*2]	26.47%[*4]
4	Limited survival	55.56%	42.37%	60.00%[*5]	52.94%[*5]
5	Variation	48.41%	74.58%	57.14%[*9]	73.53%[*6]
6	Origin of variation	7.94%	15.25%	8.57% [*6]	47.06%[*9]
7	Variation inherited	43.65%	63.56%	51.43%[*7]	20.59%[*7]
8	Differential survival	21.43%	11.86%	37.14%[*10]	26.47%[*8]
9	Change in population	38.89%	18.64%	17.14%[*4]	29.41%[*3]
10	origin of species	48.41%	15.25%	51.43%[*8]	50.00%[*10]

*indicates question number on original CINS.
Highlighted questions indicate majority of students chose alternative conceptions rather than the correct answer on these items.

Breaking the data down further, a side by side comparison between similar tests, i.e. Version A and C, and Versions B and D, shows the items where students were choosing alternative conceptions over the correct answers. Items where the relative majority of students had chosen the correct answer were determined by the plurality of the response frequencies in tables 8-11 and designated with a (+). Items where the majority of students had chosen an alternative conception were designated with a (-). The results are shown in Table 13.

Table 13
Side by side comparison of Versions A and C shows where students chose distractors (-) over correct answer choices (+) as determined by the plurality of the response frequencies in tables 12 and 13.

Concept	A	C-original CINS
Biotic Potential	+	+
Stable population	+	+
Natural resources	+	+
Limited survival	+	+
Variation	+	+
Origin of variation	-	-
Variation inherited	+	+
Differential survival	-	-
Change in population	+	-
origin of species	+	+

*highlighted concepts indicate a significant difference in scores on question pairs when analyzed using a test for two proportions.

Table 13 shows that most students chose the correct answer on the first five concepts, the concept of *variation inherited*, and the concept of *origin of species* for both Versions A and C. Similarly, most students missed the concepts of *origin of variation* and *differential survival* on both versions. The one concept for which there was a notable change was the *change in populations* question. Most students chose the correct answer on Version A versus Version C. A test for two sample proportions performed on this question pair confirms there was a positive significant difference in Version A over Version C on the *change in populations* question, where p=0.05, and the p-value for a two tailed test was 0.016.

Side by side comparisons of Versions B and D yielded more convoluted results. Students missed the same number of concepts on each version but the concepts that were missed shifted (Table 14). Most students missed the concept of *stable population* on Version D while the majority of students performed well on this concept on Version B. Alternatively, students performed poorly on the concept of *variation inherited* on the original CINS, but not on the new version. *Origin of species* on the original version of the test produced better results than the same concept on Version B. A test for two proportions performed on the above question pairs indicates that there was positive significant difference in scores on the concept of *variation inherited*, where p=0.05 for a two tailed test and the calculated p-value was 3.50×10^{-5}, indicating significant improvement on this question. Alternatively, there was a negative significant difference in scores on the concept of *origin of species*, with a p-value of 1.35×10^{-5}, indicating further revision is necessary. Interestingly, no significant difference was found for the question pairs of *stable population* or *origin of variation*.

Table 14
Side by side comparison of Versions B and D shows where students chose distractors (-) over correct answer choices (+) as determined by the plurality of the response frequencies in tables 14 and 15.

Concept	B	D-original CINS
Biotic Potential	+	+
Stable population	+	-
Natural resources	-	-
Limited survival	+	+
Variation	+	+
Origin of variation	-	+
Variation inherited	+	-
Differential survival	-	-
Change in population	-	-
Origin of species	-	+

*highlighted concepts indicate a significant difference in scores on question pairs when analyzed using a test for two proportions.

Chi-squared goodness of fit analysis. In light of the fact that there was not a significant difference in the means of the new versions versus the original versions of the CINS, but there was a significant difference between the means of Versions A and B, we wanted to know if there was a difference in how students were choosing their answers on the different versions. Chi-squared goodness of fit analysis was performed on the answers chosen by students on each version to determine if students were guessing. A null hypothesis, $H_{o,}$ would predict the percentage of students who chose each of the four possible answers to approximate 25% if students were guessing on the questions. We report findings here that are significant at the 0.001 level. With p=0.001, Chi-squared values below 11.35 indicated guessing on the part of the students. Values above 11.35 indicated students were favoring either the correct answer, or particular alternative conceptions as their answer choices as shown in Table 15.

Table 15
Chi-squared good fit value when p = 0.001. Cutoff Chi-squared value is 11.35.

Question	Concept	Chi-squared Version A	Chi-squared Version B	Chi Square Version C	Chi-squared Version D
1	Biotic Potential	181.87	96.55	27.05*1	35.17*1
2	Stable populations	151.26	40.07	28.65*3	12.11*2
3	Natural resources	31.84	55.06	40.54*2	21.05*4
4	Limited survival	66.19	38.57	26.14*5	18.70*5
5	Variation	55.90	152.47	23.00*9	37.29*6
6	Origin of variation	98.31	61.34	5.57*6	8.82*9
7	Variation inherited	41.17	97.94	17.51*7	2.70*7
8	Differential survival	28.53	57.30	6.08*10	6.47*8
9	Change in population	31.90	18.96	8.77*4	2.70*3
10	Origin of species	55.14	32.14	15.85*8	26.70*10

*indicates question number on original CINS
Shaded cells indicate questions on which students guessed according to Chi-squared analysis.

Table 15 shows Chi-squared goodness of fit analysis of Versions A and B indicating that students did not guess on any of the questions since Chi-squared values for all questions on these two versions were greater than 11.35. Conversely, questions *6, *10, and *4 on Version C fell below the cutoff value of 11.35 with values of 5.57, 6.08, and 8.77, respectively. Similarly, question *9, *7, *8 and *3 on Version D also fell below the cutoff with values of 8.82, 2.70, 6.47, and 2.70, respectively. Thus, results indicate that students were guessing on seven (35%) of the original 20 CINS questions.

Pearson's correlation. Since the new CINS and original CINS were split up into two separate versions and administered to different students, a Pearson's correlation was performed on each set of 10 questions to look for possible connections between items. A strong Pearson's correlation indicates that students who answer one question correctly are also answering another question correctly.

Version A shows some degree of correlation between the following questions. Questions 3 and 4 or the concepts *natural resources* and *limited survival* show a strong correlation, i.e. >0.5 (Cohen, 1988), as do the concepts *differential survival* and *origin of species* (Table 16). Questions 1 and 9 or the concepts *biotic potential* and *change in populations*, show a significant correlation, >0.3 (Cohen, 1988), much like the concepts *stable populations* and *natural resources* (Table 16). Interestingly, the only real correlation of interest on Version B was that between the concepts *biotic potential* and *variation inherited*, with a value of 0.333 (Table 17). Version C revealed four significant correlations between the following pairs of concepts, *natural resources* and *variation*, *change in populations* and *origin of variation*, *change* *in populations* and *variation inherited*, and *limited survival* and *origin of species* (Table 18). Version D also revealed three significant correlations for the following pairs of concepts, *limited survival* and *variation*, *variation* and *origin of variation*, and *origin of variation* and *origin of species* (Table 19). There was one weak negative correlation of significance on Version A between the concepts *variation* and *differential survival* and on Version D between the concepts of *biotic potential* and *limited survival* with values of -0.347 and -0.389, respectively. Negative correlations indicate that if the student answered one question correctly, they answered the other question incorrectly.

The number of correlations >0.3 for Version A was similar to those of Versions C and D, although the concept pairs differed. Versions A and C had four correlations, while Version C had 3. Version B only had one correlation. A number of reasons might explain these results. The questions might have some overlap in ideas, have some related ideas, or share the same amount of difficulty.

Table 16
Pearson Correlation for Version A

	Biotic Potential	Stable population	Natural resources	Limited survival	Variation	Origin of variation	Variation inherited	Differential survival	Change in population	origin of species
Biotic Potential	1									
Stable population	-0.074	1								
Natural resources	-0.0272	0.384	1							
Limited survival	-0.0272	0.277	0.589	1						
Variation	-0.209	0.053	0.038	-0.129	1					
Origin of variation	0.000	0.000	0.000	0.000	0.000	1				
Variation inherited	-0.039	0.205	-0.091	0.258	-0.133	0.000	1			
Differential survival	0.102	-0.138	0.102	0.200	-0.387	0.000	0.197	1		
Change in population	0.308	0.277	0.235	0.176	7.16E-18	0.000	7.16E-18	0.200	1	
Origin of species	0.127	-0.032	0.055	0.117	-0.220	0.000	0.296	0.502	-1.96E-17	1

Highlighted values show significant correlation, >0.3, or >(-0.3) between concept pairs.

Table 17
Pearson Correlation for Version B

	Biotic Potential	Stable population	Natural resources	Limited survival	Variation	Origin of variation	Variation inherited	Differential survival	Change in population	origin of species
Biotic Potential	1									
Stable population	0.159	1								
Natural resources	-0.052	0.178	1							
Limited survival	-0.043	0.043	0.106	1						
Variation	0.098	0.159	-0.178	-0.043	1					
Origin of variation	0.018	-0.018	-0.205	0.027	0.180	1				
Variation inherited	0.333	0.071	0.109	0.220	0.063	-0.039	1			
Differential survival	0.229	-0.012	0.182	-0.085	0.012	-0.1440	0.200	1		
Change in population	0.264	-0.119	-0.293	0.171	-0.025	0.2888	0.205	0.071	1	
origin of species	0.223	-0.223	-0.101	0.102	-0.080	-0.044	-0.150	0.098	0.232	1

Highlighted values show a significant correlations, >0.3, between concept pairs.

Table 18
Pearson Correlation for Version C

	Biotic Potential	Natural resources	Stable population	Change in population	Limited survival	Origin of variation	Variation inherited	Origin of species	Variation	Differential survival
Biotic Potential	1									
Natural resources	0.156	1								
Stable population	0.128	0.156	1							
Change in population	-0.112	-0.209	-0.112	1						
Limited survival	-0.289	-0.015	0.202	0.230	1					
Origin of variation	0.244	0.200	0.244	0.400	0.260	1				
Variation inherited	0.106	-0.220	-0.135	0.436	0.169	0.085	1			
origin of species	-0.378	-0.091	0.106	0.281	0.408	0.293	0.291	1		
Variation	0.154	0.336	0.154	0.100	0.099	0.276	0.230	0.230	1	
Differential survival	0.245	0.109	0.120	0.112	0.289	0.182	0.135	-0.106	0.089	1

Highlighted values show a significant correlations, >0.3, between concept pairs.

Table 19
Pearson correlation for Version D

	Biotic Potential	Stable population	Change in population	Limited resources	Limited survival	Variation	Variation inherited	Differential survival	Origin of variation	Origin of species
Biotic Potential	1									
Stable population	0.109	1								
Change in population	-0.008	-0.109	1							
Natural resources	-0.051	-0.197	-0.094	1						
Limited survival	-0.349	0.256	0.091	-0.235	1					
Variation	0.051	0.197	-0.051	0.057	0.369	1				
Variation inherited	0.009	0.048	-0.169	0.024	-0.102	-0.024	1			
Differential survival	0.094	0.213	-0.094	0.093	0.164	0.208	0.024	1		
Origin of variation	0.091	0.106	-0.220	-0.031	-0.055	0.432	0.102	-0.164	1	
origin of species	0	-0.181	-1.43E-17	0.066	-0.117	0.066	-0.072	0.066	0.353	1

Highlighted values show a significant correlations, >0.3 or >(-0.3) between concept pairs.

Qualitative results:

An analysis of the data would not be complete without a serious examination of the qualitative results. Bi-dimensional analysis was used to determine the depth of knowledge of student conceptions. Analysis of students' feedback of the questions and wording was performed to determine if the questions and choice of words were in line with grade level OKAPI! predictions. Interview scores were plotted against test scores to determine the validity of the test as a whole, and a look at student responses was used to shed light on the reasons for students' choices.

Bi-dimensional Analysis. Twelve students were interviewed using think-aloud protocols with either Version A or Version B. Six were interviewed with Version A and six were interviewed with Version B. The interviews were analyzed for depth of knowledge using the bi-dimensional coding scheme modified from Hogan and Fishkeller (1996, pg. 951) described earlier. The totals for depth of knowledge categories A-G were tallied and are shown below (Tables 20 and 21). The majority of the responses of the twelve students interviewed were either in line with expert views, backed with sufficient detail to warrant a classification of compatible/elaborate, or disagreed with expert views, but were based on extensive, logical, although incorrect, explanations (Incompatible/elaborate). Totals for depth of knowledge categories A and E on Version A were 111 of 234 total responses and 62 of 234 total responses, respectively. Only a few responses, 8 of 234 were non-existent, meaning students were unable to justify their answers. The same trend was seen for Version B. Totals for depth of knowledge categories A and E were 82 of 214 total responses and 78 of 214 total responses, respectively.

Table 20
Bi-dimensional coding scores for revised CINS Version A.
A =Compatible/Elaborate, B=Compatible sketchy, C=Compatibility/Incompatibility, D=Incompatible sketchy, E= Incompatible elaborate, F=Non-existent, and G=No evidence. Data is from six interviewees.

Question Number	CONCEPT	A	B	C	D	E	F	G
1	Biotic Potential	16	2	1	0	5	0	0
2	Stable population	3	1	1	2	16	1	0
3	Natural resources	15	4	1	0	3	0	0
4	Limited survival	19	1	0	1	2	1	0
5	Variation	14	3	1	2	3	1	0
6	Origin of variation	10	2	1	4	5	2	0
7	Variation inherited	12	5	1	2	3	2	0
8	Differential survival	0	2	1	2	14	1	0
9	Change in population	8	0	0	9	5	0	0
10	origin of species	14	3	0	1	6	0	0
	Total	**111**	**23**	**7**	**23**	**62**	**8**	**0**

Table 21
Bi-dimensional coding scores for revised CINS Version B.
A=Compatible/Elaborate, B=Compatible sketchy, C=Compatibility/Incompatibility, D=Incompatible sketchy, E= Incompatible elaborate, F=Non-existent, and G=No evidence. Data is from six interviewees.

Question number	CONCEPT	A	B	C	D	E	F	G
1	Biotic Potential	13	2	0	0	4	2	0
2	Stable population	8	2	0	5	7	2	0
3	Natural resources	5	0	0	1	13	1	0
4	Limited survival	10	1	0	2	11	0	0
5	Variation	13	1	0	0	7	2	0
6	Origin of variation	7	0	0	2	11	1	0
7	Variation inherited	8	0	0	4	6	1	0
8	Differential survival	5	1	1	7	10	0	0
9	Change in population	6	3	1	2	4	3	3
10	origin of species	7	1	1	3	5	0	3
	Total	**82**	**11**	**3**	**26**	**78**	**8**	**6**

Student feedback on questions and wording. During the interviews, students were asked the following questions:

1) Did you understand what the question is asking?
2) Did you have an issue with any of the words within the question?

In response to the first question, most students, answered in the positive (Table 22). Out of 116 recorded responses, 104 were affirmative indicating they understood the question.

Table 22
Number of responses of students who did and did not understand the questions on Version A and B.

Student	Did understand questions	Did not understand questions
1	9	0
2	10	0
3	10	0
4	8	2
5	7	1
6	8	1
7	9	0
8	10	0
9	7	3
10	8	2
11	9	1
12	8	2
Total	**104**	**12**

Students one-six were interviewed with Version A
Students seven-twelve were interviewed with Version B

In response to question two, most of the students answered in the negative (Table 23). Out of 114 recorded responses, 107 responses indicated they had no trouble with the words in the questions.

Table 23
Number of responses of students who did and did not understand the words on Version A and B.

Student	No trouble with words	Trouble with words
1	9	0
2	10	0
3	10	0
4	10	0
5	7	1
6	9	0
7	9	0
8	10	0
9	6	3
10	9	1
11	9	1
12	9	1
Total	**107**	**7**

Students one-six were interviewed with Version A
Students seven-twelve were interviewed with Version B

Analysis of interview scores. In order to obtain an interview score that reflected the total number of responses by students for each item, the bi-dimensional data was converted to a point system. Responses that were scored as A were given a point value of +2, B a value of +1, and C,F or G responses were scored as 0. A D response was scored as (-1) and an E response was scored as (-2). Responses for each item were scored and then averaged. The averages were then totaled to obtain a total interview score that fell within a range of -20 to +20. The scores of students who were interviewed and took the final revision of the test are listed below (Table 24). Interviewed students took the same version of the test that they were interviewed with. To avoid contamination of the data due to priming, the scores of the interviewed students were not included in the overall item analysis reported earlier.

Table 24
Total Interview scores (-20 to +20), and test scores for students who took Versions A and B. Two interview scores are shown. The first interview score is based on the first revision of the CINS. Students were re-interviewed on questions 9 and 10 after a second set of revisions based on the data from the first interview.

Student	Total Interview score	Test score	Version
1	12/11.5	9	A
2	3/1.75	4	A
3	11.25/--	8	A
4	0/1.5	8	A
5	6.75/3.75	7	A
6	-3.5/-4.5	5	A
7	-5.75/-7	5	B
8	2.75/1.75	7	B
9	13.25/13.25	8	B
10	7/5.5	5	B
11	-12.5/-13	3	B
12	-1.5/-1.75	2	B

--indicates student was unavailable for re-interview.

Scatter plots of the interview scores versus the test scores before and after revisions of question 9 and 10 were performed on the data (Figure 7). Pearson's correlation coefficients were calculated on the first interview score versus the test score and the second interview score and the test score. Pearson's coefficient r, was 0.713 between the first interview score and the test score. Pearson's coefficient r, was 0.707 between the second interview score and test score. Both r values were above the critical values of 0.553 and 0.576 with degrees of freedom of 11 and 10, respectively, and a *p*-value of 0.05. The calculated two-tailed *p*-values were 0.009 and 0.0149, well below the criteria p-value of 0.05. The scatter plot shows that test scores were either true positives, i.e. students scored well on both the interviews and the test, or true negatives, i.e. students scored low on both the interviews and the test. There were no false positives, i.e. students who scored high on the CINS (7-

10) but low on the interview (-5 to -15), nor were there any false negatives, i.e. students who scored low on the CINS (0-3) but high on the interview (5-15).

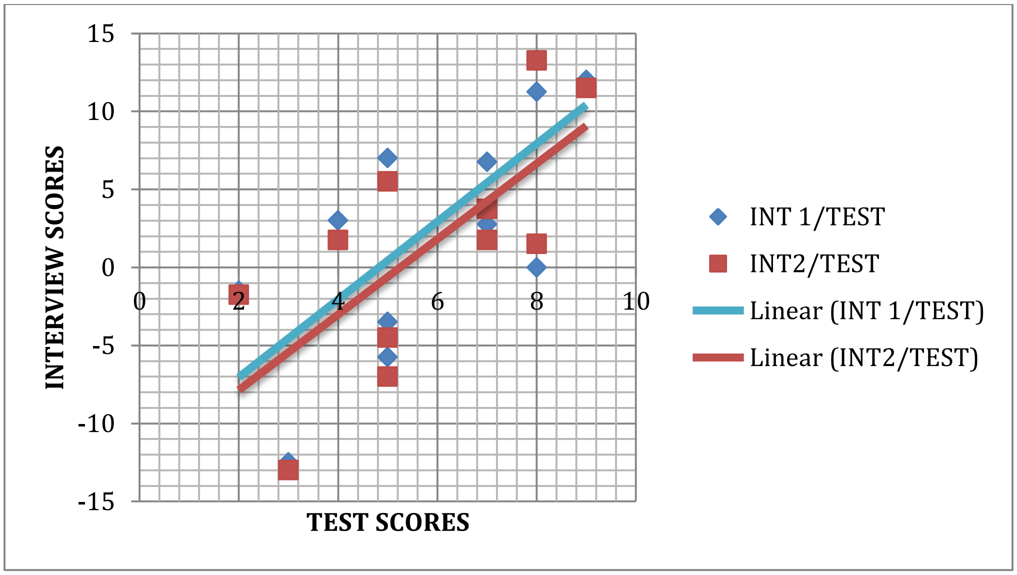


Figure 7
Scatter plot of the test score versus the first interview scores and the second interview scores for twelve interviewees.

Interview excerpts. Although there was improvement in overall student performance on the new CINS versus the original CINS, especially with regards to Version A, some questions were still difficult for students even though Chi-squared analysis indicated that students were no longer guessing on items as they had on the original version. The fact that students missed certain questions despite revisions to the old test strongly suggests that they missed these items based on firmly held alternative conceptions, and the questions do not need further revision. Two pieces of evidence support this idea. First, most students did not have trouble understanding the questions or the wording of the questions during the interviews (see tables 26 and 27). Second, bi-dimensional analysis shows clearly that on most of the questions, students had an elaborate answer, even if it was incorrect.

Analysis of the interview data on these missed questions sheds light on what students actually thought and why they answered incorrectly.

Questions where students held onto alternative conceptions

The concept of random and the alternative choices of "need" and "environment". At least two items on Versions A and B of the CINS were consistently missed by students, i.e. items 6 and 8, suggesting that the issue was not the wording of the tests, but rather, students were holding onto alternative conceptions. Question 6 dealt with the concept of random mutation and the question and the answer choices, as written in the different versions are as follows:

Question 6 on Version A:

6. How did the different types of beaks first appear in the birds?

a. The changes in the finches' beak size and shape happened because of their need to be able to eat different kinds of food to survive.
b. Changes in the size and shape of the beaks of the finches happened at random.
c. The changes in the beaks of the birds happened because the environment caused helpful changes in the genes.
d. The beaks of the birds changed a little bit in size and shape during each bird's life, with some getting larger and some getting smaller.

Question 6 on Version B:

6. Where did the variation in body size of the three species most likely come from?

a. The lizards needed to change in order to survive, so new helpful traits formed.
b. Random changes in the genes help to create new traits.
c. The environment of the island caused changes in the traits of the lizards.
d. The lizards wanted to become different in size, so helpful new traits slowly appeared in the population.

Going into the interviews we knew that question 6 was going to be interesting. Question 6 uses the word *random*, which raised the sentence to a 16^{th} grade level. However, we were reluctant to change this word because students at this age are taught probability in math and also encounter the word as part of the curriculum. It is also difficult to capture the essence of the word *random* using other means. We decided to see how students did on the interviews. Not surprising, students missed question 6 on both of the new versions of the CINS. Interestingly, it appears that students did not miss the item because they did not understand the word random, but rather, they did not understand the concept of how something could be generated randomly. A closer look at the interviews supports this view.

Student four, who took Version A, didn't think a random event could occur: "Uh the random, it probably wouldn't happen, well it's most likely not to happen…the environment can't really change the genes, only the parents can…" Student eleven who took Version B interpreted random events in the answer as occurring in the here and now: "…random changes, that's, that's not anything cause it also seems like one day he wakes up and he has different traits…" Student twelve who took Version B of the test just didn't know how the random event would take place, period: "Because I don't know how they would get a random gene change…"

Interestingly, students who took Version C also missed this concept, but students who took Version D did not. A closer look at question nine on Version D (the corresponding question to Versions A and B) reveals that the answers are much more detailed, possibly cluing in students to the right answer. The word random is mixed in with such words as genetic, combination of genes, mating and variation. In other words, the correct answer in this case is heavily laden with scientific wording as shown below in answer c.

> 9. According to the theory of natural selection, where did the variation in body size of the three species most likely come from?
>
> a. The lizards needed to change in order to survive, so beneficial new traits developed.
> b. The lizards wanted to become different in size, so beneficial new traits gradually appeared in the population.
> c. Random genetic changes and new combinations of genes produced through mating both produce variation.
> d. The island environment caused genetic changes in the lizards.

It is worth noting that several students stated that some answers on Versions A and B of the CINS just did not have enough detail. Student two had this to say:

> The changes in beaks of the finches happened at random, that is quite true,' but it's not specific enough and um, the reason why they have changes of the beaks of the finches happen at random is because probably of their genes.

In this case, the student knew that random changes were responsible, but didn't think the answer had enough detail, so she didn't choose that answer.

Item analysis and the interview data clearly show students avoided the answer choice that contained the word random. But what did students choose instead? Students gravitated towards answers that included the word "need" or the word "environment". Student 4 chose answer A, a "need" answer, and explained how need could cause a change in an organism: "I think it's A…cause when they eat other types of foods, their beaks will be able to adapt and change to what kind of food it is". In other words, the beaks change when they needed to change. Student five also chose answer A, and explained her choice this way: "Because as the time goes they would start to adapt to their habitat like how the mice did when it got darker and like in some of the places". Student five was referring to a video the students watched on mice in the Arizona desert that exhibited natural selection in real time. The mice were different colors depending on the terrain they lived on.

Student seven took Version B and looked to C, the“environmental”answer:

> Um I chose C because um there for A, ‘helpful traits,’ it’s not possible for a trait to be formed. They have to be inherited for most of their physical traits and for B um it’s kind of not possible for random changes in the genes to happen and for D um as I was saying, random changes or changes in the genes, it’s not possible to have it happen even if the lizards wanted to become different in size. New traits would not be able to appear unless it was inherited.

In the above example, student seven understood that organisms could not cause a trait to change on their own, but did not understand the concept of random mutations, and thus was left with environment as a viable avenue of change. Student one didn’t think random changes occurred out of the blue and also gravitated towards the environment as a cause: “I didn’t choose B (the correct answer) because it doesn’t happen at random. It requires certain climate changes or lack of food and predators”.

The concept of “Survival of the Fittest” and the alternative conceptions of strong and fast. Question 8 was also a commonly missed question on both tests. The focus of question 8 was on students’ understanding of the concept *differential survival* or the phrase “survival of the fittest”. The question as written on both versions is as follows:

Version A:

8. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Which trait would someone who studies these fish think is the most important in deciding which fish are the "most fit"?

a. Large body size and able to swim quickly away from predators.
b. High number of young that live to reproductive age.
c. Excellent at being able to compete for food.
d. High number of matings with many different females.

Version B:

8. Fitness is a term often used by biologists to explain the success of certain organisms. Below are descriptions of four lizards. Which lizard might someone who studies these lizards think is the "most fit"?

	Lizard A	Lizard B	Lizard C	Lizard D
Body length	20 cm	12 cm	10 cm	15 cm
Offspring surviving to adulthood	19	28	22	26
Age at death	4 years	5 years	4 years	6 years
Comments	Lizard A is very healthy, strong, and clever	Lizard B has mated with many lizards	Lizard C is dark-colored and very quick	Lizard D has the largest territory of all the lizards

a. Lizard A b. Lizard B c. Lizard C d. Lizard D

Interviews clearly show students did not understand the phrase “survival of the fittest”. On Version A, these were the responses of several students. Student one did not equate the number of young with being fit, because to the student, fit meant strong, healthy, etc: “I didn’t choose B (the right answer) because that doesn’t, even if you are young you can survive, if most of your young can survive doesn’t mean you are fit yourself”. Student three also equated fit with strength: “I think it would be A, ‘large body size and able to swim away from predators’ because the um they would be the most fit because they could escape from predators…" Student five not only did not understand the term survival of the fittest, but thought fitness in the physical sense could change when needed:

> 'Kay, um I would choose A because if it were to adapt it could protect itself and be a lot more defensive. For B, um I'm not really sure because um what it means by, the whole like choice, and for C, um I think it kind of connects to A cause if it had a larger body it would be able to swim quickly. It would also be able to get food faster and for D, um I think mating might go down depending on how fit they are.

For Version B, these were the responses of several students. Student eight was swayed by the words healthy, strong and clever:

> I chose lizard A…because lizard A is 'healthy strong and clever,' that's like you can escape predators with that and you mate and their offspring can be like lizard A, I guess….Um, I didn't choose B because even if lizard B mated with a lot, with many lizards, he could lack like the ability to be strong or clever or fast like that…

Student nine (an EL student) was very analytical, using the data to help support her answer along with the words healthy, strong and clever. "I just don't think its B because even though it like it had 28 (offspring) it's not that long, it's like 20 cm., but and same with other two, but lizard A is 20 cm long and has 19 and it's very healthy, strong and clever". Student eleven, was torn between the right answer B and choice D. Choice D was attractive because of the physically fit factor:

> Um, probably lizard B, or D…because for me, lizard B because they, the offspring, there's more that was the most offspring that survived and it said that lizard B 'mated with many lizards,' um and it was 12 cm long and lizard D also had a lot of offspring surviving and it said lizard D had the 'largest territory' out of all the lizards so it had to be fit to fight off other lizards I guess.

The fact that questions 6 and 8 were missed by the majority of students suggests that they are diagnosing student conceptions correctly and that further revision to either of these questions is unwarranted.

Questions that may need further revision

Several questions may need further revision. Several reasons stood as possible areas of focus. Students may have been influenced by the stem of questions, the context or headers of the questions or the ambiguity of certain words such as "need".

The stem as an influence on student answer choice. Most students did well on questions 1 through 5 on both versions of the CINS, with the exception of question 3 on Version B. Analysis of interviewed students on question 3, Version B reveals that students may have been confused by the stem of the question and the use of the word variety. The stem says:

> 3. Guppies eat a variety of insects and plants. Which statement describes the availability of food for guppies in the stream?

While the correct answer choice was answer A, "Sometimes there is enough food, but at other times there is not enough food for all of the guppies", the majority of students chose answer B which says, "Since guppies can eat a variety of foods, there will always be enough food for all of the fish at all times". During the interviews, at least three students were influenced by the word "variety". One student who chose answer B, gave the following reason for his answer: "Because they always eat a variety of food and um some of the insects might reproduce and make more food for the guppies". Another student who chose answer B over the correct answer specifically stated that the word "variety" was affected his answer choice:

> Because it says that there is a variety of foods and the question said variety of insects and plants, so there'll be like different species that can still eat if one dies out.

A third student actually chose answer D over the correct answer, but again, cited the word "variety" in the stem as a consideration: "Well, cause they said there is a variety so I'm just thinking that there's gonna be plenty of food, so yeah".

In comparison to question three on Version B, question three on Version A's stem did not use the word "variety" eliminating a source of confusion:

3. Finches on the Galapagos Islands require food to eat and water to drink. Which statement is true about the birds and the available resources?

Interestingly, the majority of students who took the original version of the CINS, also missed this question, but it, too, contained the word variety:

3. Lizards eat a variety of insects and plants. Which statement describes the availability of food for lizards on the Canary Islands?

Context or headers as a determining factor in student answer choice.

Question 9 dealt with the concept *change in population* and the concept of percentage, and was one of the few questions on both versions that was practically identical in wording, with the exception of the word guppies on Version A changed to lizards on Version B. Students were interviewed twice on this question because significant changes were made to answer choice B after the first round of interviews. The wording shown to students in the first interview was as follows:

Version A:

9. What is the best way to describe the evolutionary changes that happen in the guppy population over time?

a. The traits of each guppy change little by little.
b. The number of guppies that have different traits changes even if the size of the population does not change.
c. Helpful behaviors that are learned by certain guppies are passed on to their offspring.
d. Mutations happen that meet the needs of the guppies as the environment changes.

Version B:

9. What is the best way to describe the evolutionary changes that happen in the lizard population over time?

a. The traits of each lizard change little by little.
b. The number of lizards that have different traits changes even if the size of the population does not change.
c. Helpful behaviors that are learned by certain lizards are passed on to their offspring.
d. Mutations happen that meet the needs of the lizards as the environment changes.

While question 9 was originally revised through OKAPI! to read at a seventh to eighth grade reading level, it was noted that many of the students did not score well on this item during the first interview; ten out of twelve chose the incorrect answer. The problem was trying to convey the idea of percentage to students, who perhaps did not know the word percentage. After the first interview it was felt that the wording was too confusing on answer choice B, as expressed by several students including student eleven who when asked whether or not they had problems with the wording said: "No, just number B was a little bit confusing, the number thing…". The question was changed, and the students were re-interviewed. The answer was reworded to read as follows:

b. Guppies (lizards) with certain traits become more common in the population.

Despite efforts to clarify the correct answer choice, the revision did not seem to help much. Of the students who were re-interviewed only two out of eleven changed their answers, and the new answers were not the correct answer, B. And in fact, question 9 was a conundrum. A small majority of students got question 9 correct on Version A, whereas students missed this question on Version B, even though the answers to question 9 on both tests were virtually the same. A closer look at the item analysis shows that students were torn between answers B and D on Version A and between answers A, B, and D on Version B. Perhaps one reason students did better on Version A than Version B is because question 9 on Version A was about guppies whereas question 9 on Version B was on lizards, and the students had studied about guppies. Thus, there was a difference in context. Another possible explanation for the difference is that students were using other test clues, such as headers and information from other questions to help formulate their answers. If students were using the headers, they would have noted that the guppies breed much faster than the lizards which would account for students choosing choice A, "The traits of each lizard change little by little" as opposed to the fast changes that would occur in the guppy population.

Student interpretation of the word "need" may be different than previously thought. Question 10 was another item that gave mixed results. Question 10 dealt with the concept *origin of species* and is shown below for both Versions A and B.

Version A:

10. What could cause populations of guppies having different spot colors and sizes to become different species in the different streams?

a. The guppies vary a lot from each other. Those who are able to avoid the bigger fish and still attract mates could reproduce more in the different streams.
b. All guppies are alike and there are not really many different species.
c. The guppies would need to attract mates, so they changed their spots in different ways and became different species.
d. Guppies in different streams could change their spots to less bright colors because they need to avoid predators.

Version B:

10. What could cause one species to change into three species over time?

a. Groups of lizards may have grown up on different islands, and more and more genetic changes may have happened in these lizard populations over time.
b. There may be small variations between the lizards, but all the lizards are mostly alike and all are members of a single species.
c. Different groups of lizards needed to adapt to the different islands, and each organism in the groups slowly changed to become a new lizard species.
d. Groups of lizards found different island environments, so the lizards needed to become new species with different traits in order to survive.

The majority of students answered question 10 on Version A correctly, which was a guppy question. However, only 15.25% of students got this question correct on Version B. Unlike question 9, where the questions differed only in context, the questions and answers to question 10 on the two versions were different in wording, as well as context. A look at the interview data shows that all of the students interviewed on question 10 for Version A chose the correct answer, i.e. 6/6 chose A. This was for the first and second revision of the question. In contrast, 2/6 chose answer A on the first revision of B and 1/6 chose answer A on the second revision of

Version B. The majority of the interviewees, whether on the first revision or second revision of Version B chose answers C and D. These are the same two answers that were chosen by most students based on the item analysis presented earlier. Interestingly, both choices, C and D were "need" answers. However, it is unclear whether or not students thought that organisms "needed" to change themselves in the sense of "wanting" to change. Rather, evidence seems to suggest that students may have thought that organisms needed to change from a "second person" perspective, i.e. as an observer. The *students saw* that the organisms "needed" to change in order to survive. For instance, student seven clearly understood that organisms "couldn't change themselves", yet, chose answer choice C:

> Okay, um, I'm choosing C, no wait. I'm choosing C because for A, the 'genetic changes may have happened,' but they couldn't happen over time so they couldn't breed, because they have to keep on breeding in order to live, and then for B, there can be 'small variations between the lizards' and but it's not true that all the lizards are 'mostly alike' and it's not true they are all 'members of a single species' and for D, the lizards 'when they found the different island environments they did need to become a different species,' but they couldn't change themselves, besides natural selection, they couldn't just change some part for traits whenever they felt like it so…I chose C because it's true they needed to adapt, 'the different groups of lizards needed to adapt to the different islands' so the lizards on the different islands changed and then they became new lizard species from natural selection.

Student eleven chose D, another "need" answer, but clearly understood that organisms had to have the traits already present for the environment to act upon:

> It'd be D…cause that describes evolution basically, how, how they went somewhere and had to adapt, er, yeah, had to adapt with the traits they had and then those traits were passed down…

Although it might be argued that students chose need answers over the correct choice on Version B because they thought the organisms "wanted" to change, it is interesting to note that students were able to avoid those choices on Version A.

In addition, fifty percent of students who took Version D answered question ten correctly. Choice B on Version D read as follows:

> b. Groups of lizards may have been geographically isolated from other groups, and random genetic changes may have accumulated in these lizard populations over time**.**

Clearly other factors are a play with regards to this item. Student ten chose answer C, and surprisingly, did not mention the word need in his response.

> I got C…Um cause if the lizards are living in different places they'll end up adapting to the islands and they'll change, over time. They'll change into new lizard species.

Student ten, along with several other students, did state he had trouble with the phrase "so they could no longer breed":

> And (I did not choose) A because….if they live on different islands and they like change into a new species they won't be able to breed

Perhaps this was the phrase that "pushed" students toward the "need" answers. Another possible explanation for the discrepancy in the different versions of question 10 is again, context. Students were familiar with the guppy example, but not the lizard example. However, the fact that the majority of students who took Version D got this answer correct tends to rule out the latter possibility.

Conclusion

The CINS developed by Anderson, et al., was the first of its kind to be developed for the biological sciences and has served its target audience well, i.e. college students and their faculty. However, the CINS was of limited use at other grade levels in its original form due to its high reading level. The purpose of this study was to attempt to broaden the audience that could use the CINS determining the answer to the research question:

> Can CINS items be modified to make the test more equitable for middle school students without compromising the biology content?

Several lines of evidence including chi-square analysis to determine if students were guessing, bi-dimensional analysis to examine depth of student knowledge, student analysis of questions and words, and a comparison of interview performance versus CINS test results suggest that this goal was achieved. Chi-square analysis performed on the original and the new versions of the CINS indicate that revisions to the original CINS were successful in making the test more understandable to middle school students. Chi-square analysis of the original version of the CINS indicated a significant number of students were guessing on seven out of 20 questions (Table 15). Those questions, three, four, six, seven, eight, nine, and ten assessed the concept *origins of variation*, *variation inherited*, *differential survival*, and *changes in population*. In contrast, chi-square analyses of the new Versions A and B indicate there were not a significant number of students guessing on any of the newly revised questions. This analysis is of great importance because test users can be confident that students are choosing incorrect answers because they truly do not understand the concepts rather than because they don't understand the question, and then resort to guessing.

Like the Chi-square analysis, bi-dimensional analysis shows that most of the students gave a solid reason for their answers whether right or wrong. For Versions A and B, the majority of student responses were either compatible/elaborate answers or incompatible/elaborate answers. Very few responses were non-existent (Table 20 and 21). Thus, most times students had well thought out reasons for their answers, even if they were not in line with expert views.

During the interviews, students were asked if they had any trouble with the questions and answers. Nearly all (104) responses were in the negative versus eleven responses which were in the affirmative. Thus, the majority of responses by students indicated that they understood what was being asked.

During the interviews, students were also asked if they had any trouble with the words. Nearly all (107 responses) students responded "no" versus four responses that were "yes". Thus, again, the majority of responses indicate that students did not have much trouble understanding the words in the questions.

The overall implications of these results are that few students had trouble with the questions and wording of the test.   Rather, as indicated by the bi-dimensional data, students either knew the concept, or were purposely choosing alternative conceptions. The fact that they were not guessing, but actively choosing alternative conceptions, is supported by the chi-square data, which shows that the majority of students were not guessing on the answers to the questions. For example, question six on the concept of *origin of variation* was a commonly missed question. Students did not miss question six because they didn't understand the question or words like random. Instead, they didn't understand the concept of randomness. They chose alternative conceptions that involved organisms needing or wanting to change or they chose alternative conceptions where environment was the cause of the change.

In addition, the correlation graph of interview scores vs. test scores (Figure 7) indicates that the CINS-middle school version is a useful substitute for interviews.

There were no false positives and no false negatives.   Students who scored well in the interview, also scored well on the test and vice versa.

Although the middle school CINS is an improvement over the original CINS, there are still some issues that need clarification. Students performed better on Version A than they did on Version B. It is unclear whether or not students did better on Version A because of the context of the questions or because more revisions need to be done on some of the questions, in particular questions nine and ten. Version A dealt with guppies, and Version B dealt with lizards. Students had been instructed on guppy natural selection, whereas they had not on lizard natural selection. It is possible that the differences seen in results may be due to this factor. Pearson's correlation indicates that students saw more connections between questions on Version A versus Version B. Perhaps these connections helped students perform better on Version A, than on B.

Another area of concern is the use of the word "need". While the word want is self-explanatory, the use of the word "need" in some of the answers seemed ambiguous to some of the students. Some students interpreted the word need as something the organism thought it needed to do. However, other students looked at the answers from a third person point of view where, they, as observers, were making an analysis of the situation. From the observer's point of view, the organism "needed" to change in order to survive. Further research on the use of this word with the CINS is highly recommended.

There was not a significant difference in the means between the new and original versions of the test, even though data suggests students were guessing on

7/20 questions on the old version of the CINS. This finding is not surprising, given that there were only ten questions on each test, and the fact that the sample size used for the older version of the test, while statistically significant, was not large. Comparison of larger sample sizes might yield more light on the subject.

Given the results of this study, there is now a plausible concept inventory for middle school on the topic of natural selection that mirrors the college level test. The newly revised CINS for middle school is the first biological concept inventory of its kind to take into account lower grade levels. The implications are far reaching. By providing a concept inventory that mimics the college level test, but which is accessible to lower grades, the middle CINS can now serve as a starting point to begin data collection on the learning progression of students at the domain's earliest point of instruction. Questions such as "how do students' concepts change over time?" and "What are differences in alternative conceptions between middle school and older students" can begin to be answered with regards to the domain of natural selection. The middle school CINS also serves another purpose. It allows middle school teachers a way to gauge their own instruction and to become aware of the alternative conceptions present at this grade. It allows teachers to formulate better ways to develop instruction such that fewer alternative conceptions are propagated into the upper grades. Already evident from this study is the fact that students need more work on the concept of variation, namely the source of variation and the idea of randomness, as well as the concept of survival of fittest. We are optimistic that the CINS will continue to be a useful tool in science education given the broader audience, even as there is a shift to more open ended assessments with the roll out of the common core standards. The utility of a quick, accurate method of determining the level of understanding of a student on the naïve-expert continuum cannot be underestimated.

References

AAAS Science Assessment beta (2011), http://assessment.aaas.org/ (retrieved October 10th, 2011).

Abedi, J., Lord, C., & Hofstetter, C. (2004). Accomodations for english language learners: Implications for policy-based empirical research. *Review of Educational Research*, 74 (1), 1-28.

Abedi, J., Courtney, M. Mirocha, J., Leon, S., & Goldberg, J. (2005). Language accomodations for English language learners in large-scale assessments: Bilingual dictionaries and linguistic modification. *CSE Report*, 1-93.

Accountability Progress Reporting (2011), http://api.cde.ca.gov/Acnt2011/2011GrowthSch.aspx?allcds=37683386107056 (retrieved November 12, 2011).

Anderson, D.L., Fisher, K. M., & Norman, G.J. (2002). Development and evaluation of the Conceptual Inventory of Natural Selection. *Journal of Research in Science Teaching*, 39 (10), 952-978.

Atlas of Science Literacy, Volumes 1 and 2: Mapping K-12 science learning (2011), http://www.project2061.org/publications/atlas/ (retrieved October 10th 2011).

Bishop, B.A. & Anderson, C.W. (1986). Evolution by teaching natural selection: A teaching module. Occasional paper 91. Institute for Research on Teaching, Michigan State University, East Lansing, MI.

Brown, W. R. (1965), Science Textbook Selection and the Dale-Chall Formula. *School Science and Mathematics*, 65: 164–167.

Catley, K., Lehrer, R., & Reiser, B. (2005). Tracing a prospective learning progression for developing understanding of evolution. Paper Commissioned by the National Academies Committee on Test Design for K-12 Science Achievement.

Clough, E.E. & Driver, R. (1986). A study of the consistency in the use of students' conceptual frameworks across different task contexts. *Science Education* 70, 473–496.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.)

Corcoran, T., Mosher, F.A., & Rogat, A. (2009). *Learning progressions in science: An evidence based approach to reform.* NY: Center on Continuous Instructional Improvement, Teachers College—Columbia University.

Creswell, JW, & Plano Clark VL. Designing and Conducting Mixed Methods Research. 2nd ed. Thousand Oaks, CA: Sage Publications; 2011.

Cummins, C.L., Demastes, S.S., & Hafner, M.S. (1994). Evolution: Biology education's under-researched unifying theme. *Journal of Research in Science Teaching,* 31, 445–448.

D'Avanzo C. (2008). Biology concept inventories: overview, status, and next steps. *Bioscience,* 58*, 1079-1085.*

DeBoer, George E. "Assessment as a Tool for Improving Science Teaching and Learning." *NSTA Reports* Summer (2011). NSTA. Web.

Dehaene, S. (2009). Reading in the brain. New York: Penguin Group.

Demastes, S., Good, R.G., & Peebles, P. (1996). Patterns of conceptual change in evolution. *Journal of Research in Science Teaching*, 33 (4), 407-431.

Duit, R. (2004). Bibliography: Students' and Teachers' Conceptions and Science Education Database. University of Kiel, Kiel, Germany. http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html. (accessed 1 March 2005 by Tanner and Allen).

Dwyer, D. (2011). Using interview data from non-major biology students to improve the Conceptual Inventory of Natural Selection. Point Loma Nazarene University, unpublished thesis.

Eldredge, Niles. *Darwin: Discovering the Tree of Life*. New York: W.W. Norton &, 2005.

Endler, J.A. (1986). Natural selection in the wild. Princeton: Princeton University Press.

Good, R.G., Trowbridge, J.E., Demastes, S.S., Wandersee, J.H., Hafner, M.S., & Cummins, C.L. (1992, December). Toward a research base for evolution education: Report of a national conference. EDRS Conference Proceedings, ED 361 183, SE 053 585, Evolution Education Research Conference, Baton Rouge, LA.

Ghiselin, Michael T. *The Triumph of the Darwinian Method*. Berkeley: University of California, 1969.

Gould, Stephen J. (2009), Big thinkers on evolution, http://www.pbs.org/wgbh/nova/evolution/big-thinkers-evolution.html (retrieved October 12th, 2011).

Grant, P.R., Grant, B.R., & Petren, K. (2000). The allopatric phase of speciation: The sharpbeaked ground finch (Geospiza difficilis) on the Galapagos islands. *Biological Journal of the Linnean Society*, 69, 287–317.

Grant, P.R., Grant, B.R., & Petren, K. (2001). A population founded by a single pair of individuals: Establishment, expansion, and evolution. *Genetica*, 112–113, 359–382.

Hogan K., & Fisherkeller, J. (1996). Representing students' thinking about nutrient cycling in ecosystems: Bidimensional coding of a complex topic. *Journal of Research in Science Teaching*, 33(9), 941-970.

Howe, A.C. (1996). Development of science concepts within a Vygotskian framework. *Science Education*, 80 (1), 35-51.

Jensen, M. S., & Finley, F. N. (1996). Changes in students' understanding of evolution resulting from different curricular and instructional strategies. *Journal of Research in Science Teaching*, 33 (8), 879-900.

Jungwirth, E. (1975). Preconceived adaptation and inverted evolution. *Australian Science Teachers Journal*, 21, 95–100.

Klymkowsky MW, & Garvin-Doxas, K. 2008. Recognizing students' misconceptions through Ed's Tools and the Biology Concept Inventory. PloS Biology 6: e3. doi:10.1371/journal.pbio.0060003.

Lawson, A.E. (1994). Research on the acquisition of science knowledge: epistemological foundations of cognition. In: *Handbook of Research on Science Teaching and Learning*, ed. D. Gabel, New York: Simon & Schuster Macmillan, 177–210.

Lee, O., & Fradd, S.H. (1998). Science for all, including students from non-English-language backgrounds. *Educational Researcher*, 27 (4), pp. 12-21.

Liu, X. & Lesniak, K. M. (2005), Students' progression of understanding the matter concept from elementary to high school. *Science Education*, 89: 433–450.
Mayr, E. (1982). The growth of biological thought: Diversity, evolution and inheritance. Cambridge, MA: Harvard University Press.
McVee, Mary V, Dunsmore, K., & Gavelek, J. (2005). Schema Theory Revisited. *Review of Educational Research, 75* (4), pp. 531-566.
National Research Council. (1996). National science education standards. Washington, DC: National Academy Press.
National Research Council (NRC). (2001). Knowing What Students Know. Washington, DC: National Academies Press.
Nehm, R.H. (2006). Faith-based evolution education? *Bioscience*, 56, 638–639.
Nehm, R., & Schonfeld, I.S. (2008). Measuring knowledge of natural selection: A comparison of the C.I.N.S., an open-response instrument, and an oral interview. *Journal of Research in Science Teaching, 45*, 1131-1160.
OKAPI!: Readability Statistics & CBA Reading Probe Generator (n.d.), http://www.lefthandlogic.com/htmdocs/tools/okapi/okapi.php (accessed Aug. 28th, 2011).
Parker J, et al. 2008. Frameworks for reasoning and assessment in Mendelian genetics. Paper presented at Conceptual Assessment in Biology Conference II; 3–8 January, Asilomar, California. (25 September 2008; *http:// bioliteracy.net/CABS.html*)
Scharmann, L. & Harris, W. (1992). Teaching evolution: Understanding and applying the nature of science. *Journal of Research in Science Teaching*, 29, 375–388.
Science Framework for California Public Schools: Kindergarten through Grade Twelve, with New Criteria for Instructional Materials. Sacramento, CA: California Dept. of Education, 2004.
"Schema | Define Schema at Dictionary.com." *Dictionary.com | Find the Meanings and Definitions of Words at Dictionary.com*. Web. 30 Nov. 2011.
Siegal, M.A. (2007). Striving for equitable classroom assessments for linguistic minorites: Strategies for and effects of revising life science items. *Journal of Research in Science Teaching*, 44 (6), 864-881.
Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English language learners. *Educational Researcher*, 32, 3-13.
Tanner, K., Allen D. (2005). Approaches to Biology Teaching and Learning: Understanding the Wrong Answers—Teaching toward Conceptual Change. *Cell Biology Education*, 4, 112–117.
Thorpe, R.S. & Brown, R.P. (1989). Microgeographic variation of the colour pattern of Canary Island lizards. *Journal of the Linnean Society*, 38, 303–322.
Wandersee, J.H., Mintzes, J.J., & Novak, J.D. (1994). Research on alternative conceptions in science. In: *Handbook of Research on Science Teaching and Learning,* ed. D. Gabel, New York: Simon & Schuster Macmillan, 177–210.
Wenning, Carl J. "Dealing More Effectively with Alternative Conceptions in Science." *J. Phys. Tchr. Educ. Online* 5.1 (2008). Web.
Williams, K., Fisher, K., & Anderson, D. (2008). Using diagnostic test items to

assess conceptual understanding of basic biology ideas: A plan for programmatic assessment. Paper presented at the Conceptual Assessment in Biology Conference II; 3–8 January, Asilomar, California. (25 September 2008; *http://bioliteracy.net/CABS.html*).

Appendix A

Science Framework (California State Board of Education, 2004)
for grades first through ninth for the state of California.

Grade	Standard	Natural selection concept
First	**2. Plants and animals meet their needs in different ways. As a basis for understanding this concept:**	adaptation
	2. a. *Students know* different plants and animals inhabit different kinds of environments and have external features that help them thrive in different kinds of places.	
Second	**2. Plants and animals have predictable life cycles. As a basis for understanding this concept:**	reproduction, inherited traits and variation
	2. a. *Students know* that organisms reproduce offspring of their own kind and that the offspring resemble their parents and one another.	
	2. b. Students know many characteristics of an organism are inherited from the parents. Some characteristics are caused or influenced by the environment.	
	2.c. Students know there is variation among individuals of one kind within a population.	
Third	**3. Adaptations in physical structure or behavior may improve an organism's chance for survival. As a basis for understanding this concept:**	adaptation and how adaptations affect survival; variation
	3. a. *Students know* plants and animals have structures that serve different functions in growth, survival, and reproduction.	
	3. b. *Students know* examples of diverse life forms in different environments, such as oceans, deserts, tundra, forests, grasslands, and wetlands.	
	3. c. *Students know* living things cause changes in the environment in which they live: some of these changes	

	are detrimental to the organism or other organisms, and some are beneficial.	
	3. d. *Students know* when the environment changes, some plants and animals survive and reproduce; others die or move to new locations.	
Fourth	**3. Living organisms depend on one another and on their environment for survival. As a basis for understanding this concept:**	Survival, biotic and abiotic factors
	3. a. *Students know* ecosystems can be characterized by their living and nonliving components.	
	3. b. *Students know* that in any particular environment, some kinds of plants and animals survive well, some survive less well, and some cannot survive at all.	
Sixth	**5. d.** *Students know* different kinds of organisms may play similar ecological roles in similar biomes.	how the availability of resources affects survival
	5. e. *Students know* the number and types of organisms an ecosystem can support depends on the resources available and on abiotic factors, such as quantities of light and water, a range of temperatures, and soil composition.	
Seventh	**3. Biological evolution accounts for the diversity of species developed through gradual processes over many generations. As a basis for understanding this concept:**	Darwin and the lines of evidence that led him to develop his theory
	3. a. *Students know* both genetic variation and environmental factors are causes of evolution and diversity of organisms.	
	3. b. *Students know* the reasoning used by Charles Darwin in reaching his conclusion that natural selection is the mechanism of evolution.	
	3. c. *Students know* how independent lines of evidence from geology, fossils, and comparative anatomy provide the bases for the theory of evolution.	
	3. d. *Students know* how to construct a simple branching diagram to classify living groups of organisms by shared	

	derived characteristics and how to expand the diagram to include fossil organisms.	
	3. e. *Students know* that extinction of a species occurs when the environment changes and the adaptive characteristics of a species are insufficient for its survival.	
Ninth	**7. The frequency of an allele in a gene pool of a population depends on many factors and may be stable or unstable over time. As a basis for understanding this concept:**	students merge their understanding of genetics with the theory of Natural Selection
	7. a. *Students know* why natural selection acts on the phenotype rather than the genotype of an organism.	
	7. b. *Students know* why alleles that are lethal in a homozygous individual may be carried in a heterozygote and thus maintained in a gene pool.	
	7. c. *Students know* new mutations are constantly being generated in a gene pool.	
	7. d. *Students know* variation within a species increases the likelihood that at least some members of a species will survive under changed environmental conditions.	
	7. e.* *Students know* the conditions for Hardy-Weinberg equilibrium in a population and why these conditions are not likely to appear in nature.	
	7. f.* *Students know* how to solve the Hardy-Weinberg equation to predict the frequency of genotypes in a population, given the frequency of phenotypes.	
	8. Evolution is the result of genetic changes that occur in constantly changing environments. As a basis for understanding this concept:	
	8. a. *Students know* how natural selection determines the differential survival of groups of organisms.	
	8. b. *Students know* a great diversity of species increases the chance that at least some organisms survive major changes in the environment.	

	8. c. *Students know* the effects of genetic drift on the diversity of organisms in a population.	
	8. d. *Students know* reproductive or geographic isolation affects speciation.	
	8. e. *Students know* how to analyze fossil evidence with regard to biological diversity, episodic speciation, and mass extinction.	
	8. f.* *Students know* how to use comparative embryology, DNA or protein sequence comparisons, and other independent sources of data to create a branching diagram (cladogram) that shows probable evolutionary relationships.	
	8. g.* *Students know* how several independent molecular clocks, calibrated against each other and combined with evidence from the fossil record, can help to estimate how long ago various groups of organisms diverged evolutionarily from one other.	
	* standards that students should have the opportunity to learn vs. standards that students are expected to achieve.	

Appendix B

Results of readability as determined by OKAPI!. Headers, questions (Q) and answers of the original CINS (2011) run through OKAPI. Words that are underlined are considered difficult. OKAPI determines the reading level of the text as is shown in the right hand column.

Question	Original CINS 2011	Grade level
Header	*Scientists* have long believed that the 14 *species* of *finches* on the *Galapagos* Islands *evolved* from a single *species* of *finch* that *migrated* to the islands one to five million years ago.	13-15
	Recent *DNA* *analyses* *support* the *conclusion* that all of the *Galapagos* *finches* *evolved* from the *warbler* *finch.*	16
	Different *species* live on different islands.	7-8
	For *example,* the *medium* ground *finch* and the *cactus* *finch* live on one island.	13-15
	The large *cactus* *finch* *occupies* another island.	16
	One of the major changes in the *finches* is in their *beak* sizes and shapes as shown in this figure.	7-8
	Scientists have long believed that the 14 *species* of *finches* on the *Galapagos* Islands *evolved* from a single *species* of *finch* that *migrated* to the islands one to five million years ago. *Recent* *DNA* *analyses* *support* the *conclusion* that all of the *Galapagos* *finches* *evolved* from the *warbler* *finch.* Different *species* live on different islands. For *example,* the *medium* ground *finch* and the *cactus* *finch* live on one island. The large *cactus* *finch* *occupies* another island. One of the major changes in the *finches* is in their *beak* sizes and shapes as shown in this figure.	13-15
	Choose the one answer that best *reflects* how an *evolutionary* *biologist* would answer.	9-10
Q1	What would happen if a *breeding* pair of *finches* was placed on an island under ideal *conditions* with no *predators* and *unlimited* food so that all *individuals* *survived?*	11-12
1a	The *finch* *population* would stay small because birds only have enough babies to *replace* *themselves.*	11-12
1b	The *finch* *population* would double and then stay *relatively* stable.	11-12
1c	The *finch* *population* would *increase* *dramatically.*	16
1d	The *finch* *population* would grow slowly and then level off.	9-10
2	*Finches* on the *Galapagos* Islands *require* food to eat and water to drink. How does this fact *impact* the *population?*	11-12
2a	When food and water are *scarce,* some birds may be *unable* to *obtain* what they need to *survive.*	11-12
2b	When food and water are *limited,* the *finches* will find other food *sources,* so there is always enough.	9-10
2c	When food and water are *scarce,* the *finches* all eat and drink	9-10

	less so that all birds *survive.*	
2d	There is always plenty of food and water on the *Galapagos* Islands to meet the *finches'* needs.	7-8
3	Once a *population* of *finches* has lived on a *particular* island for many years, what will most likely happen to the *population?*	9-10
3a	The *population* *continues* to grow rapidly.	13-15
3b	The *population* remains *relatively* stable, with some *fluctuations.*	13-15
3c	The *population* *dramatically* *increases* and *decreases* each year.	16
3d	The *population* will *decrease* *steadily,* than *increase.*	16
4	What is the best way to describe the *evolutionary* changes that *occur* in a *finch* *population* over time?	11-12
4a	The *traits* of each *finch* *within* a *population* *gradually* change.	16
4b	The *percentages* of *finches* having different *traits* *within* a *population* change.	16
4c	*Successful* *behaviors* learned by *finches* are passed on to *offspring.*	
4d	*Mutations* *occur* to meet the needs of the *finches* as the *environment* changes.	13-15
5	Depending on their *beak* size and shape, some *finches* get *nectar* from flowers, some eat *grubs* from bark, some eat small seeds, and some eat large *nuts.* Which *statement* best describes the *interactions* among the *finches* and the food *supply?*	11-12
5a	Most of the *finches* on an island *cooperate* to find food and share what they find.	7-8
5b	Many of the *finches* on an island fight with one another and the *physically* strongest ones win.	7-8
5c	There is more than enough food to meet all the *finches'* needs so they *don't* need to *compete* for food.	7-8
5d	*Finches* *compete* *primarily* with closely *related* *finches* that eat the same kinds of food, and some may die from *lack* of food.	13-15
6	How did the different *beak* *types* first appear in the *Galapagos* *finches?*	13-15
6a	The changes in the *finches'* *beak* size and shape *occurred* because of their need to be able to eat different kinds of food to *survive.*	9-10
6b	Changes in the *finches'* *beaks* *occurred* *randomly,* and when there was a good match between *beak* *structure* and *available* food, those birds *usually* had more *offspring.`*	16
6c	The changes in the *finches'* *beaks* *occurred* because the *environment* caused the desired *genetic* changes.	13-15
6d	The *finches'* *beaks* changed a little bit in size and shape with each *successive* *generation,* some getting larger and some getting smaller.	9-10
7	What *type* of *variation* in the *finches'* *traits* is passed to the *offspring?*	16
7a	*Only Behaviors* that were learned during a *finch's* *lifetime.*	13-15

7b	Only *traits* that were *beneficial* during a *finch's* *lifetime*	16
7c	Only *traits* that were *genetically* *determined*	16
7d	Only *traits* that were *positively* *influenced* by the *environment* during a *finch's* *lifetime*	16
8	What caused *populations* of birds having different *beak* shapes and sizes to become *distinct* *species* *distributed* on the *various* islands?	13-15
8a	The *finches* were *quite* *varied,* and those whose *features* were best suited to the *available* food *supply* on each island *reproduced* most *successfully.*	16
8b	All *finches* are *essentially* alike and there are not really fourteen different *species.*	9-10
8c	Different foods are *available* on different islands and for that reason, *individual* *finches* on each island *gradually* *developed* the *beaks* they needed.	13-15
8d	Different lines of *finches* *developed* different *beak* *types* because they needed them in order to *obtain* the *available* food.	13-15
Header 2	*Guppies* are small fish found in streams in *Venezuela.* Male *guppies* are brightly colored, with black, red, blue and *iridescent* *(reflective)* spots. Males cannot be too brightly colored or they will be seen and *consumed* by *predators,* but if they are too plain, *females* will choose other males. *Natural* *selection* and *sexual* *selection* push in *opposite* directions. When a *guppy* *population* lives in a stream in the *absence* of *predators,* the *proportion* of males that is bright and *flashy* *increases* in the *population.* If a few *aggressive* *predators* are added to the same stream, the *proportion* of bright-colored males *decreases* *within* about five months (3-4 *generations).* The *effects* of *predators* on *guppy* *coloration* have been studied in *artificial* ponds with *mild,* *aggressive,* and no *predators,* and by *similar* *manipulations* of *natural* stream *environments.*	13-15
9	A *natural* *population* of *guppies* *consists* of hundreds of fish of a single *species.* Which *statement* best describes the *population* of *guppies?*	13-15
9a	The *guppies* share all of the same *characteristics* and are *identical* to each other.	9-10
9b	The *guppies* share all of the most important *characteristics* of the *species;* the small differences between them *don't* *affect* *survival.*	11-12
9c	The *guppies* are all *identical* on the inside, but have many differences in *appearance.*	9-10
9d	The *guppies* share most important *characteristics,* but also have differences that may *affect* *survival.*	11-12
10	*Fitness* is a term often used by *biologists* to explain the *evolutionary* *success* of certain *organisms.* Which *characteristics* would a *biologist* *consider* to be most important in *determining* which *guppies* were the "most fit" *according* to *Darwin's* *theory?*	13-15

10a	large body size and *ability* to swim quickly away from *predators*	16
10b	*excellent* *ability* to *compete* for food	16
10c	high number of *offspring* that *survived* to *reproductive* age	16
10d	high number of *matings* with many different *females*	16
11	*Assuming* ideal *conditions* with *abundant* food and space, and no *predators,* what would happen if a mating pair of *guppies* was placed in a large pond?	9-10
11a	The *guppy* *population* would grow slowly, as *guppies* would have only the number of *offspring* that are needed to *replenish* the *population.*	13-15
11b	The *guppy* *population* would grow slowly at first, then would grow rapidly, and thousands of *guppies* would fill the pond.	7-8
11c	The *guppy* *population* would never become very large, because only *organisms* such as insects and *bacteria* *reproduce* in that manner.	11-12
11d	The *guppy* *population* would *continue* to grow slowly over time.	11-12
12	Once a *population* of *guppies* has been *established* for a number of years in a pond with other *organisms* *including* *predators,* what will likely happen to the *population* if *conditions* remain *constant?*	13-15
12a	The *guppy* *population* will stay about the same size.	9-10
12b	The *guppy* *population* will *continue* to rapidly grow in size.	11-12
12c	The *guppy* *population* will *gradually* *decrease* until no more *guppies* are left.	16
12d	It is impossible to tell because *populations* do not follow *patterns.*	9-10
13	What is the best way to describe the *evolutionary* changes that *occur* in a *guppy* *population* over time?	11-12
13a	The *traits* of each *individual* *guppy* *within* a *population* *gradually* change.	16
13b	The *percentage* of *guppies* having different *traits* *within* a *population* change.	16
13c	*Successful* *behaviors* learned by certain *guppies* are passed on to *offspring.*	13-15
13d	*Mutations* *occur* to meet the needs of the *guppies* as the *environment* changes.	13-15
Header 3	The Canary Islands are seven islands just west of the *African continent.* The islands *gradually* became *colonized* with life: plants, lizards, birds, *etc.* Three different *species* of lizards found on the islands are *similar* to one *species* found on the *African continent.* Because of this, *scientists* *assume* that the lizards *traveled* from *Africa* to the Canary Islands by floating on tree trunks washed out to sea.	9-10
14	Lizards eat a *variety* of insects and plants. Which *statement* describes the *availability* of food for lizards on the Canary Islands?	7-8
14a	Finding food is not a *problem* since food is always in *abundant supply.*	9-10

14b	Since lizards can eat a *variety* of foods, there is likely to be enough food for all of the lizards at all times.	5-6
14c	Lizards can get by on very little food, so the food *supply* does not matter.	5-6
14d	It is likely that sometimes there is enough food, but at other times there is not enough food for all of the lizards.	4 and below
15	What do you think happens among the lizards of a certain *species* when the food *supply* is *limited?*	9-10
15a	The lizards will *cooperate* to find food and share what they find.	5-6
15b	The lizards fight for the *available* food and the stronger lizards kill the weaker ones.	5-6
15c	*Genetic* changes that would allow lizards to eat new food *sources* are more likely to *occur.*	9-10
15d	The lizards least *successful* in the *competition* for food are likely to die of *starvation* and *malnutrition.*	11-12
16	A *population* of lizards is made up of hundreds of *individuals.* Which *statement* describes how *similar* they are likely to be to other lizards in the *population?*	9-10
16a	All lizards are likely to be almost exactly the same.	4 and below
16b	All lizards are exactly the same on the inside, but *display* differences in their *external* *features.*	9-10
16c	All lizards share many *similarities,* yet likely have some *significant* differences in their *features.*	9-10
16d	All lizards are likely to be the same on the outside, but *display* differences in their *internal* *features.*	9-10
17	Which *statement* best describes how *traits* in lizards will be *inherited* by *offspring?*	13-15
17a	When parent lizards learn to catch *particular* insects, their *offspring* can *inherit* their *specific* insect-catching *skills.*	13-15
17b	When parent lizards *develop* stronger claws through repeated use in catching *prey,* their *offspring* can *inherit* their stronger-claw *trait.*	11-12
17c	When parent lizards' claws are *underdeveloped* because the *available* *prey* is easy to catch, their *offspring* can *inherit* their weakened claws.	11-12
17d	When a parent lizard is born with an extra claw on each limb, its *offspring* can *inherit* the extra claw.	7-8
18	*Fitness* is a *term* often used by *biologists* to explain the *evolutionary* *success* of certain *organisms.* *Below* are *descriptions* of four *fictional* *female* *lizards.* Which lizard might a biologist *consider* to be the "most *fit"* *according* to *Darwin's* *Theory?*	13-15
18a	*Lizard a*	
	Lizard b	
	Lizard c	
	Lizard d	

19	*According* to the *theory* of *natural* *selection,* where did the *variation* in body size of the three *species* most likely come from?	13-15
19a	The lizards needed to change in order to *survive,* so *beneficial* new *traits* *developed.*	11-12
19b	The lizards wanted to become different in size, so *beneficial* new *traits* *gradually* appeared in the *population.*	11-12
19c	*Random* *genetic* changes and new *combinations* of *genes* *produced* through mating both *produce* *variation.*	16
19d	The island *environment* caused *genetic* changes in the lizards.	9-10
20	What could cause one *species* to change into three *species* over time?	7-8
20a	Groups of lizards *encountered* different island *environments,* so the lizards needed to become new *species* with different *traits* in order to *survive.*	11-12
20b	Groups of lizards may have been *geographically* *isolated* from other groups, and *random* *genetic* changes may have *accumulated* in these lizard *populations* over time.	11-12
20c	There may be *minor* *variations,* but all lizards are *essentially* alike and all are members of a single *species.*	9-10
20d	In order to *survive,* different groups of lizards needed to *adapt* to the different islands, and so all *organisms* in each group *gradually* *evolved* to become a new lizard *species.*	11-12

Appendix C

Questions and answers of original (2011) and revised middle school Version A of the CINS. Original questions and answers were run through OKAPI to determine their current reading level. Questions and answers were then modified and rerun through OKAPI to lower the reading level of the questions and answers. Most of the new questions and answers now read at 7-8th grade reading level with a few exceptions.

Original question number		grade level	New question Number and Version		grade level
Q1	What would happen if a <u>*breeding*</u> pair of <u>*finches*</u> was placed on an island under ideal <u>*conditions*</u> with no <u>*predators*</u> and <u>*unlimited*</u> food so that all <u>*individuals*</u> <u>*survived?*</u>	11-12	1A	What would happen if a <u>*breeding*</u> pair of <u>*finches*</u> was placed on an island with no <u>*predators*</u> and plenty of food so that all the birds lived?	7-8
1a	The <u>*finch*</u> <u>*population*</u> would stay small because birds only have enough babies to <u>*replace*</u> <u>*themselves.*</u>	11-12		The <u>*population*</u> of birds would stay small because <u>*finches*</u> only have enough babies to take their place when they die.	7-8
1b	The <u>*finch*</u> <u>*population*</u> would double and then stay <u>*relatively*</u> stable.	11-12		The <u>*population*</u> of <u>*finches*</u> would double and then stay about the same.	7-8
1c	The <u>*finch*</u> <u>*population*</u> would <u>*increase*</u> <u>*dramatically.*</u>	16		The <u>*population*</u> of birds would <u>*grow to a large number*</u> and would keep growing.	5-6
1d	The <u>*finch*</u> <u>*population*</u> would grow slowly and then level off.	9-10		The <u>*population*</u> of <u>*finches*</u> would grow slowly and then stay the same.	7-8
2	<u>*Finches*</u> on the <u>*Galapagos*</u> Islands <u>*require*</u> food to eat and water to drink. How does this fact <u>*impact*</u> the <u>*population?*</u>	11-12	*3A*	<u>*Finches*</u> on the <u>*Galapagos*</u> Islands require food to eat and water to drink. Which statement is true about the birds and the available resources?	11-12
2a	When food and water are <u>*scarce,*</u> some birds may be <u>*unable*</u> to <u>*obtain*</u> what they need to <u>*survive.*</u>	11-12		Sometimes there is enough food, but at other times there is not enough food for all of the finches.	5-6
2b	When food and water are <u>*limited,*</u> the <u>*finches*</u> will find other food <u>*sources,*</u> so there is always enough.	9-10		When food and water are <u>*limited,*</u> the <u>*finches*</u> will find other kinds of food so there is always enough.	7-8
2c	When food and water are <u>*scarce,*</u> the <u>*finches*</u> all eat and drink less so that all birds <u>*survive.*</u>	9-10		When food and water are <u>*limited,*</u> the <u>*finches*</u> all eat and drink less.	7-8
2d	There is always plenty of food and water on the <u>*Galapagos*</u> Islands to meet the <u>*finches'*</u> needs.	7-8		There is always plenty of food and water on the Galapagos Islands to meet the finches needs.	7-8
3	Once a <u>*population*</u> of <u>*finches*</u>	9-10	2A	A population of finches lives on an	7-8

	has lived on a *particular* island for many years, what will most likely happen to the *population?*			island for many years where there are predators. What will most likely happen to the population if everything remains the same?	
3a	The *population* *continues* to grow rapidly.	13-15		The *population* of birds grows rapidly.	7-8
3b	The *population* remains *relatively* stable, with some *fluctuations.*	13-15		The *population* remains stable, with few changes.	7-8
3c	The *population* *dramatically* *increases* and *decreases* each year.	16		The *population* gets larger and smaller each year.	7-8
3d	The *population* will *decrease* *steadily,* than *increase.*	16		The *population* will get smaller, then get larger.	7-8
4	What is the best way to describe the *evolutionary* changes that *occur* in a *finch* *population* over time?	11-12	9A	What is the best way to describe the *evolutionary* changes that happen in a *finch* *population* over time?	9-10
4a	The *traits* of each *finch* *within* a *population* *gradually* change.	16		The *traits* of each bird change little by little.	5-6
4b	The *percentages* of *finches* having different *traits* *within* a *population* change.	16		The number of birds that have different *traits* changes even if the size of the *population* does not change.	7-8
4c	*Successful* *behaviors* learned by *finches* are passed on to *offspring.*	16		Helpful *behaviors* that are learned by certain birds are passed on to the *offspring.*	7-8
4d	*Mutations* *occur* to meet the needs of the *finches* as the *environment* changes.	13-15		*Mutations* happen that meet the needs of the birds as the *environment* changes.	7-8
5	Depending on their *beak* size and shape, some *finches* get *nectar* from flowers, some eat *grubs* from bark, some eat small seeds, and some eat large *nuts.* Which *statement* best describes the *interactions* among the *finches* and the food *supply?*	11-12	4A	Depending on the size and shape of the *beak,* some *finches* get *nectar* from flowers, some eat insects in the bark, some eat small seeds, and some eat large *nuts.* Which sentence best describes how the *finches* will *interact* with each other over the food?	7-8
5a	Most of the *finches* on an island *cooperate* to find food and share what they find.	7-8		Most of the *finches* on an island *cooperate* to find food and share what they find so that they all live.	7-8
5b	Many of the *finches* on an island fight with one another and the *physically* strongest ones win.	7-8			
5c	There is more than enough food to meet all the *finches'* needs so they *don't* need to *compete* for food.	7-8			
5d	*Finches* *compete* *primarily* with closely *related* *finches* that eat the same kinds of food, and some may die from	13-15		*Finches* *compete* with other *finches* that eat the same kinds of food, and some die because there is less food. They do not get enough to	7-8

	lack of food.			live.	
6	How did the different *beak types* first appear in the *Galapagos finches?*	13-15	6A	How did the different *types* of *beaks* first appear in the birds?	7-8
6a	The changes in the *finches' beak* size and shape *occurred* because of their need to be able to eat different kinds of food to *survive.*	9-10		The changes in the *finches' beak* size and shape happened because of their need to be able to eat different kinds of food to *survive.*	7-8
6b	Changes in the *finches' beaks occurred randomly,* and when there was a good match between *beak structure* and *available* food, those birds *usually* had more *offspring.*	16		Changes in the *beaks* of the *finches* happened at *random.*	11-12
6c	The changes in the *finches' beaks occurred* because the *environment* caused the desired *genetic* changes.	13-15		The changes in the *beaks* of the birds happened because the *environment* caused helpful changes in the *genes.*	9-10*
6d6c	The *finches' beaks* changed a little bit in size and shape with each *successive generation,* some getting larger and some getting smaller.	9-10		The *beaks* of the birds changed a little bit in size and shape with each *generation,* some getting larger and some getting smaller.	7-8
7	What *type* of *variation* in the *finches' traits* is passed to the *offspring?*	16	7A	What kind of *variation* in the *traits* of the birds is passed on to their babies?	7-8
7a	*Only Behaviors* that were learned during a *finch's lifetime.*	13-1516		*Behaviors* that were learned during a bird's life.	7-8
7b	Only *traits* that were *beneficial* during a *finch's lifetime*	16		Only *traits* that were helpful during a bird's life.	5-6
7c	Only *traits* that were *genetically determined*	16		Only *traits* that were coded for by the *genes.*	9-10*
7d	Only *traits* that were *positively influenced* by the *environment* during a *finch's lifetime*	16		Only *traits* that were acted upon by the *environment* in a helpful way during a bird's life.	7-8
8	What caused *populations* of birds having different *beak* shapes and sizes to become *distinct species distributed* on *various* islands?	13-15	10A	What could cause populations of guppies having different spot colors and sizes to become different species in different streams?	9-10
8a	The *finches* were *quite varied,* and those whose *features* were best suited to the *available* food *supply* on each island *reproduced* most *successfully.*	16		The guppies *vary* a lot from each other. Those who are able to avoid the bigger fish and still attract a mate could reproduce more in the different streams.	9-10
8b	All *finches* are *essentially* alike and there are not really fourteen different *species.*	9-10		All *guppies* are alike and there are not really many different *species.*	7-8

8c	Different foods are *available* on different islands and for that reason, *individual finches* on each island *gradually developed* the *beaks* they needed.	13-15		The guppies would need to attract mates, so they changed their spots in different ways and became different species.	9-10
8d	Different lines of *finches developed* different *beak types* because they needed them in order to *obtain* the *available* food.	13-15		Guppies in different streams could change their spots to less bright colors because they need to avoid predators.	7-8
9	A *natural population* of *guppies consists* of hundreds of fish of a single *species.* Which *statement* best describes the *population* of *guppies?*	13-15	5A	A *population* of *guppies* has hundreds of fish of a single *species.* Which sentence best describes the group of fish?	7-8
9a	The *guppies* share all of the same *characteristics* and are *identical* to each other.	9-10		The fish share all the same *characteristics* and are *identical* to each other.	7-8
9b	The *guppies* share all of the most important *characteristics* of the *species;* the small differences between them *don't affect survival.*	11-12		The fish share all of the most important *characteristics* of the *species;* the small differences between them do not *affect* how long they live or reproduce.	7-8
9c	The *guppies* are all *identical* on the inside, but have many differences in *appearance.*	9-10		The fish are all *identical* on the inside, but have many differences in *appearance.*	7-8
9d	The *guppies* share most the important *characteristics,* but also have differences that may *affect survival.*	11-12		The fish share the most important *characteristics,* but also have differences that may *affect* how long they live and reproduce.	7-8
10	*Fitness* is a term often used by *biologists* to explain the *evolutionary success* of certain *organisms.* Which *characteristics* would a *biologist consider* to be most important in *determining* which *guppies* were the "most fit" *according* to *Darwin's theory?*	13-15	*8A*	*Fitness* is a term often used by *biologists* to explain the *evolutionary success* of certain *organisms.* Which *trait* would someone who studies these fish think is the most important in deciding which fish were the "most fit"?	9-10
10a	large body size and *ability* to swim quickly away from *predators*	16		large body size and are able to swim quickly away from *predators.*	5-6
10b	*excellent ability* to *compete* for food.	16		are *excellent* at being able to *compete* for food.	9-10
10c	high number of *offspring* that *survived* to *reproductive* age.	16		high number of young that live to *reproductive* age.	5-6
10d	high number of *matings* with many different *females*	16			

Appendix D

Questions and answers of original (2011) and revised middle school Version B of the CINS. Original questions and answers were run through OKAPI to determine their current reading level. Questions and answers were then modified and rerun through OKAPI to lower the reading level of the questions and answers. Most of the new questions and answers now read at 7-8th grade reading level with a few exceptions.

Original question number		Reading grade level	New question Number and Version		Reading grade level
11	*Assuming* ideal *conditions* with *abundant* food and space, and no *predators,* what would happen if a mating pair of *guppies* was placed in a large pond?	9-10	1B	If there were *abundant* food and space, and no *predators,* what would happen if a mating pair of *guppies* was placed in a large pond?	7-8
11a	The *guppy population* would grow slowly, as *guppies* would have only the number of *offspring* that are needed to *replenish* the *population.*	13-15		The fish *population* would grow slowly. The *guppies* would have only the number of babies that are needed to *replace* those that have died.	7-8
11b	The *guppy population* would grow slowly at first, then would grow rapidly, and thousands of *guppies* would fill the pond.	7-8			
11c	The *guppy population* would never become very large, because only *organisms* such as insects and *bacteria reproduce* in that manner.	11-12		The fish *population* would never become very large, because only *organisms* such as insects and *bacteria reproduce* in that manner.	9-10
11d	The *guppy population* would *continue* to grow slowly over time.	11-12		The fish *population* would keep growing slowly over time.	5-6
12	Once a *population* of *guppies* has been *established* for a number of years in a pond with other *organisms including predators,* what will likely happen to the *population* if *conditions* remain *constant?*	13-15	2B	A *population* of fish has been around for a number of years in a pond with other *organisms* and *predators.* What will likely happen to the *population* if *everything* remains the same?	7-8
12a	The *guppy population* will stay about the same size.	9-10		The fish *population* will stay about the same size.	5-6
12b	The *guppy population* will *continue* to rapidly grow in size.	11-12		The fish *population* will keep growing in size.	7-8

12c	The *guppy* *population* will *gradually* *decrease* until no more *guppies* are left.	16		The fish *population* will slowly get smaller until no more *guppies* are left.	7-8
12d	It is impossible to tell because *populations* do not follow *patterns.*	9-10			
13	What is the best way to describe the *evolutionary* changes that *occur* in a *guppy* *population* over time?	11-12	9B	What is the best way to describe the *evolutionary* changes that happen in the fish *population* over time?	7-8
13a	The *traits* of each *individual* *guppy* *within* a *population* *gradually* change.	16		The *traits* of each fish change little by little.	7-8
13b	The *percentage* of *guppies* having different *traits* *within* a *population* change.	16		The number of fish that have different *traits* changes even if the size of the *population* does not change..	7-8
13c	*Successful* *behaviors* learned by certain *guppies* are passed on to *offspring.*	13-15		Helpful *behaviors* that are learned by certain fish are passed on to their *offspring.*	7-8
13d	*Mutations* *occur* to meet the needs of the *guppies* as the *environment* changes.	13-15		*Mutations* happen that meet the needs of the fish as the *environment* changes.	7-8
14	Lizards eat a *variety* of insects and plants. Which *statement* describes the *availability* of food for lizards on the Canary Islands?	7-8	3B		
14a	Finding food is not a *problem* since food is always in *abundant* *supply.*	9-10		Finding food is not a *problem* since there is always plenty of food.	5-6
14b	Since lizards can eat a *variety* of foods, there is likely to be enough food for all of the lizards at all times.	5-6		Since lizards can eat a *variety* of foods, there will always be enough food for all of the lizards at all times.	5-6
14c	Lizards can get by on very little food, so the food *supply* does not matter.	5-6			
14d	It is likely that sometimes there is enough food, but at other times there is not enough food for all of the lizards.	4 and below			
15	What do you think happens among the lizards of a certain *species* when the food *supply* is *limited?*	9-10	4B	What do you think happens among the lizards of a certain *species* when the amount of food is low?	5-6
15a	The lizards will *cooperate* to find food and share	5-6			

	what they find.				
15b	The lizards fight for the *available* food and the stronger lizards kill the weaker ones.	5-6			
15c	*Genetic* changes that would allow lizards to eat new food *sources* are more likely to *occur.*	9-10		*Genetic* changes that would allow lizards to eat new *types* of food are more likely to happen.	7-8
15d	The lizards least *successful* in the *competition* for food are likely to die of *starvation* and *malnutrition.*	11-12		The lizards that can not *compete* for food well are likely to die from a *lack* of food or *nutrients.*	7-8
16	A *population* of lizards is made up of hundreds of *individuals.* Which *statement* describes how *similar* they are likely to be to other lizards in the *population?*	9-10	5B	A *population* of lizards is made up of hundreds of *individuals.* Which sentence describes how *similar* they are likely to be to other lizards in the *population?*	7-8
16a	All lizards are likely to be almost exactly the same.	4 and below			
16b	All lizards are exactly the same on the inside, but *display* differences in their *external* *features.*	9-10		All lizards are exactly the same on the inside, but show differences in their *external* *characteristics.*	7-8
16c	All lizards share many *similarities,* yet likely have some *significant* differences in their *features.*	9-10		All lizards share many *similarities,* yet but are likely to have some important differences in their *characteristics.*	7-8
16d	All lizards are likely to be the same on the outside, but *display* differences in their *internal* *features.*	9-10		All lizards are likely to be the same on the outside, but have differences in their *internal* *characteristics.*	7-8
17	Which *statement* best describes how *traits* in lizards will be *inherited* by *offspring?*	13-15	7B	Which sentence best describes how *traits* in lizards will be *inherited* by their young?	7-8
17a	When parent lizards learn to catch *particular* insects, their *offspring* can *inherit* their *specific* insect-catching *skills.*	13-15		When parent lizards learn to catch certain insects, their young can *inherit* the *ability* to catch those insects.	7-8
17b	When parent lizards *develop* stronger claws through repeated use in catching *prey,* their *offspring* can *inherit* their stronger-claw *trait.*	11-12		When parent lizards get stronger claws through repeated use in catching *prey,* their young can *inherit* their stronger-claw *trait.*	9-10
17c	When parent lizards' claws *underdeveloped* because the *available* *prey* is easy to catch, their *offspring*	11-12		When parent lizards' claws are weak because the *available* *prey* is easy to catch, their young can *inherit* their weakened claws.	7-8

	can *inherit* their weakened claws.				
17d	When a parent lizard is born with an extra claw on each limb, its *offspring* can *inherit* the extra claw.	7-8			
18	*Fitness* is a *term* often used by *biologists* to explain the *evolutionary success* of certain *organisms. Below* are *descriptions* of four *fictional female lizards.* Which lizard might a biologist *consider* to be the "most *fit"* *according* to *Darwin's Theory?*	13-15	*8B*	*Fitness* is a term often used by *biologists* to explain the *evolutionary success* of certain *organisms.* Below are *descriptions* of four lizards. Which lizard might someone who studies these lizards think is the "most fit".	7-8
18a				*Lizard a*	
				Lizard b	
				Lizard c	
				Lizard d	
19	*According* to the *theory* of *natural selection,* where did the *variation* in body size of the three *species* most likely come from?	13-15	6B	Where did the *variation* in body size of the three *species* most likely come from?	7-8
19a	The lizards needed to change in order to *survive,* s *beneficial* new *traits developed.*	11-12		The lizards needed to change in order to *survive,* so new helpful *traits* formed.	7-8
19b	The lizards wanted to become different in size, so *beneficial* new *traits gradually* appeared in the *population.*	11-12		The lizards wanted to become different in size, so helpful new *traits* slowly appeared in the *population.*	7-8
19c	*Random genetic* changes and new *combinations* of *genes produced* through mating both *produce variation.*	16		*Random changes in the genes help to create new traits.*	16
19d	The island *environment* caused *genetic* changes in the lizards.	9-10		The *environment* of the island caused changes in the *traits* of the lizards.	7-8
20	What could cause one *species* to change into three *species* over time?	7-8	10B		
20a	Groups of lizards *encountered* different island *environments,* so the lizards needed to become new *species* with different *traits* in order to *survive.*	11-12		Groups of lizards found different island *environments,* so the lizards needed to become new *species* with different *traits* in order to *survive.*	9-10
20b	Groups of lizards may have been *geographically*	11-12		Groups of lizards may have grown up on different islands,	7-8

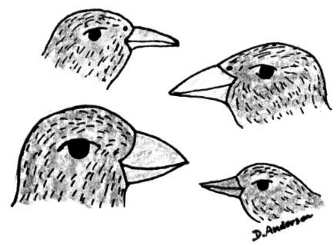
	isolated from other groups, and *random genetic* changes may have *accumulated* in these lizard *populations* over time.			and more and more *genetic* changes may have happened in these lizard *populations* over time.	
20c	There may be *minor variations,* but all lizards are *essentially* alike and all are members of a single *species.*	9-10		There may be small *variations* between the lizards, but all the lizards are mostly alike and all are members of a single *species.*	7-8
20d	In order to *survive,* different groups of lizards needed to *adapt* to the different islands, and so all *organisms* in each group *gradually evolved* to become a new lizard *species.*	11-12		Different groups of lizards needed to *adapt* to the different islands, and each *organism* in the groups slowly changed to become a new lizard *species.*	7-8

Appendix E

1st CINS revision, Version A, for middle school Questions 1-10

Highlighting indicates words or phrases were changed in the final Version of test following data gathered during interviews.

Introduction to Galapagos finch



- Studied on the Galapagos Islands by many scientists
- Generally accepted that all 14 species of finches on the Islands evolved from a single species
- New research suggests that they are closely related to tanagers.
- Original finches most likely migrated to the islands one to five million years ago
- Species found on the different islands have unique beak sizes and shapes

---

1. What would happen if a breeding pair of finches was placed on an island with no predators and plenty of food so that all the birds lived?

a. The population of birds would stay small because finches only have enough babies to take their place when they die.
b. The population of finches would double and then stay about the same.
c. The population of birds would grow to a large number and would keep growing.
d. The population of finches would grow slowly and then stay the same.

2. A population of finches lives on an island for many years where there are predators. What will most likely happen to the population if everything remains the same?

a. The population of birds will grow rapidly.
b. The population will remain stable, with few changes.
c. The population will get larger and smaller each year.
d. The population will get smaller, then get larger.

3. Finches on the Galapagos Islands require food to eat and water to drink. Which statement is true about the birds and the available resources.

a. Sometimes there is enough food, but at other times there is not enough food for all of the finches.
b. When food and water are limited, the finches will find other kinds of food so there is always enough.
c. When food and water are limited, the finches all eat and drink less.
d. There is always plenty of food and water on the Galapagos Islands to meet the finches' needs.

4. Depending on the size and shape of the beak, some finches get nectar from flowers, some eat insects in the bark, some eat small seeds, and some eat large nuts. Which sentence best describes how the finches will interact with each other over the food?

    a. Many of the finches on an island cooperate to find food and share what they find so that they all live.
    b. Many of the finches on an island fight with one another and the physically strongest ones win.
    c. There is more than enough food to meet all the finches' needs so they don't need to compete for food.
    d. Finches compete with other finches that eat the same kinds of food, and some die because they do not get enough to live.

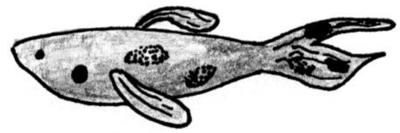
5. A population of finches has hundreds of birds of a single species. Which sentence best describes the group of birds?

    a. The birds share all the same characteristics and are identical to each other.
    b. The birds share all of the most important characteristics of the species; the small differences between them do not affect how long they live or reproduce.
    c. The birds are all identical on the inside, but have many differences in appearance.
    d. The birds share the most important characteristics, but also have differences that may affect how long they live or reproduce.

6. How did the different types of beaks first appear in the birds?

    a. The changes in the finches' beak size and shape happened because of their need to be able to eat different kinds of food to survive.
    b. Changes in the beaks of the finches happened at random.
    c. The changes in the beaks of the birds happened because the environment caused helpful changes in the genes.
    d. The beaks of the birds changed a little bit in size and shape with each generation, some getting larger and some getting smaller.

Introduction to Venezuelan guppies



- Small, colorful fish found in streams in Venezuela.
- Scientists have studied guppies in both natural streams and in lab experiments
- Male have black, red, blue and reflective spots.
- Natural selection and sexual selection push in opposite directions.
- Brightly colored males are easily seen and eaten by predators
- However, females tend to choose more brightly colored males
- In a stream with no predators, the proportion of males that is bright and flashy increases in the population.
- If predators are added, the proportion of bright-colored males decreases within about five months (3-4 generations).

7. What kind of variation in the traits of the guppies is passed on to their babies?

    a. Behaviors that were learned during a guppy's life.
    b. Only traits that were helpful during a guppy's life.
    c. Only traits that were coded for by the guppy's genes.
    d. Only traits that were acted upon by the environment in a helpful way during a guppy's life.

8. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Which trait would someone who studies these fish think is the most important in deciding which fish were the "most fit"?

    a. large body size and able to swim quickly away from predators.
    b. high number of young that live to reproductive age.
    c. excellent at being able to compete for food.
    d. high number of matings with many different females.

9. What is the best way to describe the evolutionary changes that happen in the guppy population over time?

    a. The traits of each guppy change little by little.
    b. The number of guppies that have different traits changes even if the size of the population does not change.
    c. Helpful behaviors that are learned by certain guppies are passed on to their offspring.
    d. Mutations happen that meet the needs of the guppies as the environment changes.

10. What could cause populations of guppies having different spot colors and sizes to become different species in the different streams?

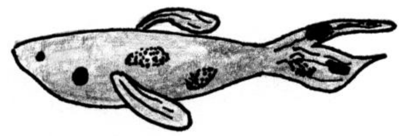
    a. The guppies vary a lot from each other. Those who are able to avoid the bigger fish and still attract mates could reproduce more in the different streams.
    b. All guppies are alike and there are not really many different species.
    c . The guppies would need to attract mates, so they changed their spots in different ways and became different species.
    d. Guppies in different streams could change their spots to less bright colors because they need to avoid predators.

Appendix F

1st CINS revision, Version B, for middle school questions 11-20

Highlighting indicates words or phrases were changed in the final Version of test following data gathered during interviews.

Introduction to Venezuelan guppies



- Small, colorful fish found in streams in Venezuela.
- Scientists have studied guppies in both natural streams and in lab experiments
- Male have black, red, blue and reflective spots.
- Natural selection and sexual selection push in opposite directions.
- Brightly colored males are easily seen and eaten by predators
- However, females tend to choose more brightly colored males
- In a stream with no predators, the proportion of males that is bright and flashy increases in the population.
- If predators are added, the proportion of bright-colored males decreases within about five months (3-4 generations).

---

1. If there were abundant food and space, and no predators, what would happen if a mating pair of guppies was placed in a large pond?

    a. The fish population would grow slowly. The guppies would have only the number of babies that are needed to replace those that have died.
    b. The fish population would never become very large, because only organisms such as insects and bacteria reproduce in that manner.
    c. The guppy population would grow slowly at first, then would grow to a large number, and thousands of guppies would fill the pond.
    d. The fish population would keep growing slowly over time.

2. A population of fish lives for a number of years in a pond with other organisms and predators. What will likely happen to the population if everything remains the same?

    a. The fish population will keep growing in size.
    b. The fish population will stay about the same size.
    c. The fish population will slowly get smaller until no more guppies are left.
    d. It is impossible to tell because populations do not follow patterns

3. Guppies eat a variety of insects and plants. Which statement describes the availability of food for guppies in the stream?

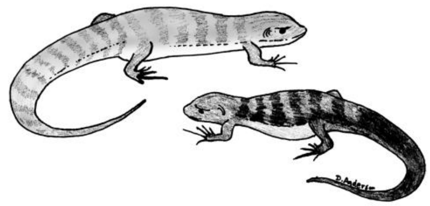
    a. Sometimes there is enough food, but at other times there is not enough food for all of the guppies.
    b. Since guppies can eat a variety of foods, there will always be enough food for all of the fish at all times.
    c. Guppies can get by on very little food, so the food supply does not matter.
    d. Finding food is not a problem since there is always plenty of food.

4. What do you think happens among the guppies living together when the amount of food is low?

    a. The guppies cooperate to find food and share what they find.
    b. The guppies fight for the available food and the stronger guppies kill the weaker ones.
    c. Genetic changes that would allow guppies to eat new types of food are more likely to happen.
    d. The guppies that cannot compete for food well are likely to die from a lack of food or nutrients.

Canary Island Lizards

- The Canary Islands are seven islands just west of the African continent.
- The islands gradually became colonized with life: plants, lizards, birds, etc.
- Three different species of lizards found on the islands
- These three species are similar to one species found on the African continent
- Scientists think that the lizards traveled from Africa to the Canary Islands by floating on tree trunks washed out to sea.



---

5. A population of lizards is made up of hundreds of individuals. Which sentence describes how similar they are likely to be to other lizards in the population?

    a. All lizards are likely to be almost exactly the same.
    b. All lizards are likely to be the same on the outside, but have differences in their internal characteristics.
    c. All lizards are exactly the same on the inside, but show differences in their external characteristics.
    d. All lizards share many similarities, but are likely to have some important differences in their characteristics.

6. Where did the variation in body size of the three species most likely come from?

    a. The lizards needed to change in order to survive, so new helpful traits formed.
    b. Random changes in the genes help to create new traits.
    c. The environment of the island caused changes in the traits of the lizards.
    d. The lizards wanted to become different in size, so helpful new traits slowly appeared in the population.

7. Which sentence best describes how traits in lizards will be inherited by their young?

    a. When parent lizards learn to catch certain insects, their young can inherit the ability to catch those insects.
    b. When parent lizards get stronger claws through repeated use in catching prey, their young can inherit their stronger-claw trait.
    c. When a parent lizard is born with an extra claw on each limb, its offspring can inherit the extra claw.
    d. When parent lizards' claws are weak because the available prey is easy to catch, their young can inherit their weakened claws.

8. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Below are descriptions of four lizards. Which lizard might someone who studies these lizards think is the "most fit"?

	Lizard A	Lizard B	Lizard C	Lizard D
Body length	20 cm	12 cm	10 cm	15 cm
Offspring surviving to adulthood	19	28	22	26
Age at death	4 years	5 years	4 years	6 years
Comments	Lizard A is very healthy, strong, and clever	Lizard B has mated with many lizards	Lizard C is dark-colored and very quick	Lizard D has the largest territory of all the lizards

a. Lizard A
b. Lizard B
c. Lizard C
d. Lizard D

9. What is the best way to describe the evolutionary changes that happen in the lizard population over time?

a. The traits of each lizard change little by little.
b. The number of lizards that have different traits changes even if the size of the population does not change.
c. Helpful behaviors that are learned by certain lizards are passed on to their offspring.
d. Mutations happen that meet the needs of the lizards as the environment changes.

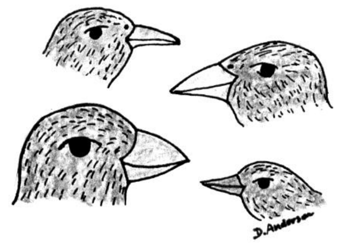
10. What could cause one species to change into three species over time?

a. Groups of lizards may have grown up on different islands, and more and more genetic changes may have happened in these lizard populations over time.
b. There may be small variations between the lizards, but all the lizards are mostly alike and all are members of a single species.
c. Different groups of lizards needed to adapt to the different islands, and each organism in the groups slowly changed to become a new lizard species.
d. Groups of lizards found different island environments, so the lizards needed to become new species with different traits in order to survive.

Appendix G: Final Version
Conceptual Inventory of Natural Selection
Middle School Version A

6th Edition – 2012 DRAFT

Your answers will test your understanding of the Theory of Natural Selection.
Please choose the answer that best shows how a **biologist** would answer each question.



Introduction to Galapagos finches

- These birds were studied on the Galapagos Islands by many scientists
- The original finches most likely came to the islands one to five million years ago
- 14 species of finches on the Islands evolved from a single species
- Species found on the islands have different beak sizes and shapes
- New research suggests that they are closely related to tanagers

---

1. What would happen if a breeding pair of finches was placed on an island with no predators and plenty of food so that all the birds lived?

    a. The population of birds would stay small because finches only have enough babies to take their place when they die.
    b. The population of finches would double and then stay about the same.
    c. The population of birds would grow to a large number and would keep growing.
    d. The population of finches would grow slowly and then stay the same.

2. A population of finches lives on an island for many years where there are predators. What will most likely happen to the population if everything on the island stays the same?

    a. The population of birds will grow rapidly each year.
    b. The population will remain stable, with few changes each year.
    c. The population will get larger and smaller each year.
    d. The population will get smaller, then get larger each year.

3. Finches on the Galapagos Islands require food to eat and water to drink. Which statement is true about the birds and the available resources?

    a. Sometimes there is enough food, but at other times there is not enough food for all of the finches.
    b. When food and water are limited, the finches will find other kinds of food so there is always enough.
    c. When food and water are limited, the finches all eat and drink less.
    d. There is always plenty of food and water on the Galapagos Islands to meet the finches' needs.

4. Depending on the size and shape of the beak, some finches get nectar from flowers, some eat insects in the bark, some eat small seeds, and some eat large nuts. Which sentence best describes how the finches will interact with each other over the food?

    a. Many of the finches on an island cooperate to find food and share what they find osthat they all live.
    b. Many of the finches on an island fight with one another and the physically strongest ones win.
    c. There is more than enough food to meet all the finches' needs so they don't need to compete for food.
    d. Finches compete with other finches that eat the same kinds of food, and some die because they do not get enough to live.

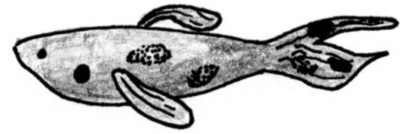
5. A population of finches has hundreds of birds of a single species. Which sentence best describes the group of birds?

    a. The birds share all the same characteristics and are identical to each other.
    b. The birds share all of the most important characteristics of the species; the small differences between them do not affect how long they live or reproduce.
    c. The birds are all identical on the inside, but have many differences in appearance.
    d. The birds share the most important characteristics, but also have differences that may affect how long they live or reproduce.

6. How did the different types of beaks first appear in the birds?

    a. The changes in the finches' beak size and shape happened because of their need to be able to eat different kinds of food to survive.
    b. Changes in the size and shape of the beaks of the finches happened at random.
    c. The changes in the beaks of the birds happened because the environment caused helpful changes in the genes.
    d. The beaks of the birds changed a little bit in size and shape during each bird's life, with some getting larger and some getting smaller.

Introduction to South American guppies



- These are small, colorful fish found in streams in Venezuela.
- Scientists have studied guppies in both natural streams and in lab experiments
- Males have black, red, blue and reflective spots.
- Brightly colored males are easily seen and eaten by predators
- However, females tend to choose more brightly colored males
- In a stream with no predators, the number of males that are bright and flashy increases in the population.
- If predators are added, the number of brightly-colored males gets smaller within about five months (3-4 generations).

7. What kind of variation in the traits of the guppies is passed on to their babies?

   a. Behaviors that were learned during a guppy's life.
   b. Only traits that were helpful during a guppy's life.
   c. Only traits that were coded for by the guppy's genes.
   d. Only traits that were affected by the environment in a helpful way during a guppy's life.

8. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Which trait would someone who studies these fish think is the most important in deciding which fish are the "most fit"?

   a. Large body size and able to swim quickly away from predators.
   b. High number of young that live to reproductive age.
   c. Excellent at being able to compete for food.
   d. High number of matings with many different females.

9. What is the best way to describe the evolutionary changes that happen in the guppy population over time?

   a. The traits of each guppy in the population change little by little.
   b. Guppies with certain traits become more common in the population.
   c. Helpful behaviors learned by certain guppies are passed on to their offspring and become more common.
   d. Mutations happen in the guppy population to meet the needs of the fish as the environment changes.

10. What could cause populations of guppies having different spot colors and sizes to become different species in the different streams?

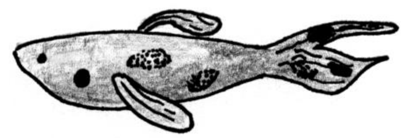
    a. The guppies that are able to avoid the bigger fish and still attract mates could reproduce more often. Over time the guppies in the different streams would not be able to breed with each other.
    b. All guppies are alike and there are not really many different species.
    c. The guppies that needed to attract mates could change their spots in different ways and become different species.
    d. Guppies that want to avoid predators in the different streams could change their spots so they are not so bright and in this way become different species.

Appendix H: Final Version
Conceptual Inventory of Natural Selection
Middle School Version B

6th Edition – 2012 DRAFT

Your answers will test your understanding of the Theory of Natural Selection.
Please choose the answer that best shows how a **biologist** would answer each question.

Introduction to South American guppies



- These are small, colorful fish found in streams in Venezuela.
- Scientists have studied guppies in both natural streams and in lab experiments
- Males have black, red, blue and reflective spots.
- Brightly colored males are easily seen and eaten by predators
- However, females tend to choose more brightly colored males
- In a stream with no predators, the number of males that are bright and flashy increases in the population.
- If predators are added, the number of brightly-colored males gets smaller within about five months (3-4 generations).

---

1. If there were abundant food and space, and no predators, what would happen if a mating pair of guppies was placed in a large pond?

   a. The fish population would grow slowly. The guppies would have only the number of babies that are needed to replace those that have died.
   b. The fish population would never become very large, because only organisms such as insects and bacteria reproduce in that manner.
   c. The guppy population would grow slowly at first, then would grow to a large number, and thousands of guppies would fill the pond.
   d. The fish population would keep growing slowly over time.

2. A population of fish lives for a number of years in a pond with other organisms and predators. What will likely happen to the population if everything in the pond remains the same?

   a. The fish population will keep growing in size.
   b. The fish population will stay about the same size.
   c. The fish population will slowly get smaller until no more guppies are left.
   d. It is impossible to tell because populations do not follow patterns.

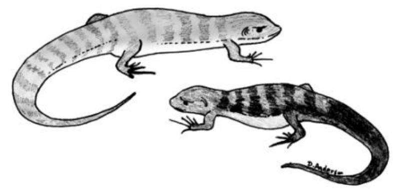
3. Guppies eat a variety of insects and plants. Which statement describes the availability of food for guppies in the stream?

   a. Sometimes there is enough food, but at other times there is not enough food for all of the guppies.
   b. Since guppies can eat a variety of foods, there will always be enough food for all of the fish at all times.
   c. Guppies can get by on very little food, so the food supply does not matter.
   d. Finding food is not a problem since there is always plenty of food.

4. What do you think happens among the guppies living together when the amount of food is low?

a. The guppies cooperate to find food and share what they find.
b. The guppies fight for the available food and the stronger guppies kill the weaker ones.
c. Genetic changes that would allow guppies to eat new types of food are more likely to happen.
d. The guppies that cannot compete for food well are likely to die from a lack of food or nutrients.

Introduction to Canary Island Lizards



- The Canary Islands are seven islands just west of the African continent
- The islands gradually became colonized with life: plants, lizards, birds, etc.
- Three different species of lizards are found on the islands
- These three species are similar to one species found on the African continent
- Scientists think that the lizards traveled from Africa to the Canary Islands by floating on tree trunks washed out to sea.

---

5. A population of lizards is made up of hundreds of individuals. Which sentence describes how similar they are to other lizards in the population?

a. All lizards are likely to be almost exactly the same.
b. All lizards are likely to be the same on the outside, but have differences in their internal characteristics.
c. All lizards are exactly the same on the inside, but show differences in their external characteristics.
d. All lizards share many similarities, but are likely to have some important differences in their characteristics.

6. Where did the variation in body size of the three species most likely come from?

a. The lizards needed to change in order to survive, so new helpful traits formed.
b. Random changes in the genes help to create new traits.
c. The environment of the island caused changes in the traits of the lizards.
d. The lizards wanted to become different in size, so helpful new traits slowly appeared in the population.

7. How are traits in lizards inherited by their young?

a. When a parent lizard learns to catch certain insects, its young can inherit the ability to catch those insects.
b. When a parent lizard get stronger claws through repeated use in catching prey, its young can inherit the stronger claw trait.
c. When a parent lizard is born with an extra claw on each limb, its offspring can inherit the extra claw.
d. When a parent lizard's claws are weak because the available prey is easy to catch, its young can inherit the weakened claws.

8. Fitness is a term often used by biologists to explain the success of certain organisms. Below are descriptions of four lizards. Which lizard might someone who studies these lizards think is the "most fit"?

	Lizard A	Lizard B	Lizard C	Lizard D
Body length	20 cm	12 cm	10 cm	15 cm
Offspring surviving to adulthood	19	28	22	26
Age at death	4 years	5 years	4 years	6 years
Comments	Lizard A is very healthy, strong, and clever	Lizard B has mated with many lizards	Lizard C is dark-colored and very quick	Lizard D has the largest territory of all the lizards

a. Lizard A b. Lizard B c. Lizard C d. Lizard D

9. What is the best way to describe the evolutionary changes that happen in the lizard population over time?

a. The traits of each lizard in the population change little by little.
b. Lizards with certain traits become more common in the population.
c. Helpful behaviors learned by certain lizards are passed on to their offspring and become more common.
d. Mutations happen in the lizard population to meet the needs of the lizards as the environment changes.

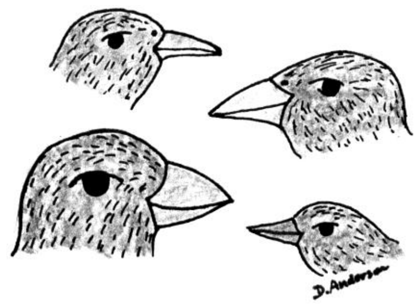
10. What could have caused one species to change into three species over time?

a. Groups of lizards lived on different islands. Many genetic changes may have happened in each of these populations over time so they could no longer breed.
b. There were small variations between the lizards, but all the lizards were mostly alike and all are members of a single species.
c. Different groups of lizards needed to adapt to the different islands. Lizards in each group slowly changed to become a new lizard species.
d. Groups of lizards found different island environments, so the lizards needed to become new species with different traits in order to survive.

Appendix I: Original 2011 CINS questions 1-10
Conceptual Inventory of Natural Selection
Version C-2011

Your answers will assess your understanding of the Theory of Natural Selection.
Please choose the answer that best reflects how a biologist would think about each question.

Galapagos finches



Scientists have long believed that the 14 species of finches on the Galapagos Islands evolved from a single species of finch that migrated to the islands one to five million years ago (Lack, 1940). Recent DNA analyses support the conclusion that all of the Galapagos finches evolved from the warbler finch (Grant, Grant & Petren, 2001; Petren, Grant & Grant, 1999). Different species live on different islands. For example, the medium ground finch and the cactus finch live on one island. The large cactus finch occupies another island. One of the major changes in the finches is in their beak sizes and shapes as shown in this figure.

1. What would happen if a breeding pair of finches was placed on an island under ideal conditions with no predators and unlimited food so that all individuals survived?
   a. The finch population would stay small because birds only have enough babies to replace themselves.
   b. The finch population would double and then stay relatively stable.
   c. The finch population would increase dramatically.
   d. The finch population would grow slowly and then level off.

2. Finches on the Galapagos Islands require food to eat and water to drink. How does this fact impact the population?
   a. When food and water are scarce, some birds may be unable to obtain what they need to survive.
   b. When food and water are limited, the finches will find other food sources, so there is always enough.
   c. When food and water are scarce, the finches all eat and drink less so that all birds survive.
   d. There is always plenty of food and water on the Galapagos Islands to meet the finches' needs.

3. Once a population of finches has lived on a particular island for many years, what will most likely happen to the population?
   a. The population continues to grow rapidly.
   b. The population remains relatively stable, with some fluctuations.
   c. The population dramatically increases and decreases each year.
   d. The population will decrease steadily.

4. What is the best way to describe the evolutionary changes that occur in a finch population over time?
   a. The traits of each finch within a population gradually change.
   b. The percentages of finches having different traits within a population change.
   c. Successful behaviors learned by finches are passed on to offspring.
   d. Mutations occur to meet the needs of the finches as the environment changes.

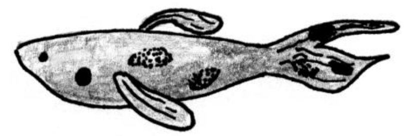
5. Depending on their beak size and shape, some finches get nectar from flowers, some eat grubs from bark, some eat small seeds, and some eat large nuts. Which statement best describes the interactions among the finches and the food supply?
   a. Most of the finches on an island cooperate to find food and share what they find.
   b. Many of the finches on an island fight with one another and the physically strongest ones win.
   c. There is more than enough food to meet all the finches' needs so they don't need to compete for food.
   d. Finches compete primarily with closely related finches that eat the same kinds of food**,** and some may die from lack of food.

6. How did the different beak types <u>first</u> appear in the Galapagos finches?
   a. The changes in the finches' beak size and shape occurred because of their need to be able to eat different kinds of food to survive.
   b. Changes in the finches' beaks occurred randomly, and when there was a good match between beak structure and available food, those birds usually had more offspring.
   c. The changes in the finches' beaks occurred because the environment caused the desired genetic changes.
   d. The finches' beaks changed a little bit in size and shape with each successive generation, some getting larger and some getting smaller.

7. What type of variation in the finches' traits is passed to the offspring?
   a. Only behaviors that were learned during a finch's lifetime
   b. Only traits that were beneficial during a finch's lifetime
   c. Only traits that were genetically determined
   d. Only traits that were positively influenced by the environment during a finch's lifetime

8. What caused populations of birds having different beak shapes and sizes to become distinct species distributed on the various islands?
   a. The finches were quite varied, and those whose features were best suited to the available food supply on each island reproduced most successfully.
   b. All finches are essentially alike and there are <u>not</u> really fourteen different species.
   c. Different foods are available on different islands and for that reason, individual finches on each island gradually developed the beaks they needed.
   d. Different lines of finches developed different beak types because they needed them in order to obtain the available food.

Venezuelan guppies



Guppies are small fish found in streams in Venezuela. Male guppies are brightly colored, with black, red, blue and iridescent (reflective) spots. Males cannot be too brightly colored or they will be seen and consumed by predators, but if they are too plain, females will choose other males. Natural selection and sexual selection push in opposite directions. When a guppy population lives in a stream in the absence of predators, the proportion of males that is bright and flashy increases in the population. If a few aggressive predators are added to the same stream, the proportion of bright-colored males decreases within about five months (3-4 generations). The effects of predators on guppy coloration have been studied in artificial ponds with mild, aggressive, and no predators, and by similar manipulations of natural stream environments (Endler, 1980).

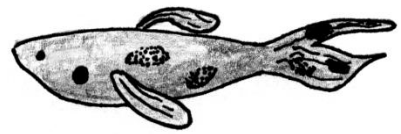
9. A natural population of guppies consists of hundreds of fish of a single species. Which statement best describes the population of guppies?
   a. The guppies share all of the same characteristics and are identical to each other.
   b. The guppies share all of the most important characteristics of the species; the small differences between them don't affect survival.
   c. The guppies are all identical on the inside, but have many differences in appearance.
   d. The guppies share most important characteristics, but also have differences that may affect survival.

10. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Which characteristics would a biologist consider to be most important in determining which guppies were the "most fit" according to Darwin's theory?
    a. large body size and ability to swim quickly away from predators
    b. excellent ability to compete for food
    c. high number of offspring that survived to reproductive age
    d. high number of matings with many different females

Appendix J: Original 2011 CINS questions 11-20
Conceptual Inventory of Natural Selection
Version D-2011

Your answers will assess your understanding of the Theory of Natural Selection. Please choose the answer that best reflects how a biologist would think about each question.

Venezuelan guppies



Guppies are small fish found in streams in Venezuela. Male guppies are brightly colored, with black, red, blue and iridescent (reflective) spots. Males cannot be too brightly colored or they will be seen and consumed by predators, but if they are too plain, females will choose other males. Natural selection and sexual selection push in opposite directions. When a guppy population lives in a stream in the absence of predators, the proportion of males that is bright and flashy increases in the population. If a few aggressive predators are added to the same stream, the proportion of bright-colored males decreases within about five months (3-4 generations). The effects of predators on guppy coloration have been studied in artificial ponds with mild, aggressive, and no predators, and by similar manipulations of natural stream environments (Endler, 1980).

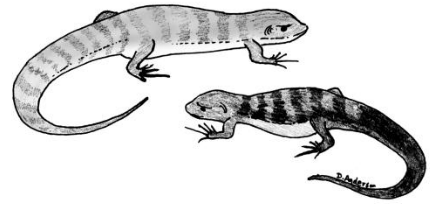
---

1. Assuming ideal conditions with abundant food and space, and no predators, what would happen if a mating pair of guppies was placed in a large pond?
    a. The guppy population would grow slowly, as guppies would have only the number of offspring that are needed to replenish the population.
    b. The guppy population would grow slowly at first, then would grow rapidly, and thousands of guppies would fill the pond.
    c. The guppy population would never become very large, because only organisms such as insects and bacteria reproduce in that manner.
    d. The guppy population would continue to grow slowly over time.

2. Once a population of guppies has been established for a number of years in a pond with other organisms including predators, what will likely happen to the population if conditions remain constant?
    a. The guppy population will stay about the same size.
    b. The guppy population will continue to rapidly grow in size.
    c. The guppy population will gradually decrease until no more guppies are left.
    d. It is impossible to tell because populations do not follow patterns.

3. What is the best way to describe the evolutionary changes that occur in a guppy population over time?
    a. The traits of each individual guppy within a population gradually change.
    b. The percentage of guppies having different traits within a population change.
    c. Successful behaviors learned by certain guppies are passed on to offspring.
    d. Mutations occur to meet the needs of the guppies as the environment changes.

Canary Island Lizards



The Canary Islands are seven islands just west of the African continent. The islands gradually became colonized with life: plants, lizards, birds, etc. Three different species of lizards found on the islands are similar to one species found on the African continent (Thorpe & Brown, 1989). Because of this, scientists assume that the lizards traveled from Africa to the Canary Islands by floating on tree trunks washed out to sea.

4. Lizards eat a variety of insects and plants. Which statement describes the availability of food for lizards on the Canary Islands?
   a. Finding food is not a problem since food is always in abundant supply.
   b. Since lizards can eat a variety of foods, there is likely to be enough food for all of the lizards at all times.
   c. Lizards can get by on very little food, so the food supply does not matter.
   d. It is likely that sometimes there is enough food, but at other times there is not enough food for all of the lizards.

5. What do you think happens among the lizards of a certain species when the food supply is limited?
   a. The lizards will cooperate to find food and share what they find.
   b. The lizards fight for the available food and the stronger lizards kill the weaker ones.
   c. Genetic changes that would allow lizards to eat new food sources are more likely to occur.
   d. The lizards least successful in the competition for food are likely to die of starvation and malnutrition.

6. A population of lizards is made up of hundreds of individuals. Which statement describes how similar they are likely to be to other lizards in the population?
   a. All lizards are likely to be almost exactly the same.
   b. All lizards are exactly the same on the inside, but display differences in their external features.
   c. All lizards share many similarities, yet likely have some significant differences in their features.
   d. All lizards are likely to be the same on the outside, but display differences in their internal features.

7. Which statement best describes how traits in lizards will be inherited by offspring?
   a. When parent lizards learn to catch particular insects, their offspring can inherit their specific insect-catching skills.
   b. When parent lizards develop stronger claws through repeated use in catching prey, their offspring can inherit their stronger-claw trait.
   c. When parent lizards’ claws are underdeveloped because the available prey is easy to catch, their offspring can inherit their weakened claws.
   d. When a parent lizard is born with an extra claw on each limb, its offspring can inherit the extra claw.

8. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Below are descriptions of four fictional female lizards. Which lizard might a biologist consider to be the "most fit" according to Darwin's Theory?

	Lizard A	Lizard B	Lizard C	Lizard D
Body length	20 cm	12 cm	10 cm	15 cm
Offspring surviving to adulthood	19	28	22	26
Age at death	4 years	5 years	4 years	6 years
Additional facts	Lizard A is very healthy, strong, and clever	Lizard B has mated with many lizards	Lizard C is dark-colored and very quick.	Lizard D has the largest territory of all the lizards.

a. Lizard A
b. Lizard B
c. Lizard C
d. Lizard D

9. According to the theory of natural selection, where did the variation in body size of the three species most likely come from?
a. The lizards needed to change in order to survive, so beneficial new traits developed.
b. The lizards wanted to become different in size, so beneficial new traits gradually appeared in the population.
c. Random genetic changes and new combinations of genes produced through mating both produce variation.
d. The island environment caused genetic changes in the lizards.

10. What could cause one species to change into three species over time?
a. Groups of lizards encountered different island environments, so the lizards needed to become new species with different traits in order to survive.
e. Groups of lizards may have been geographically isolated from other groups, and random genetic changes may have accumulated in these lizard populations over time.
c. There may be minor variations, but all lizards are essentially alike and all are members of a single species.
d. In order to survive, different groups of lizards needed to adapt to the different islands, and so all organisms in each group gradually evolved to become a new lizard species.

Appendix K: Answer key for the CINS-Middle School Versions A and B, 2012
Relationship between Conceptual Inventory of Natural Selection (CINS-Middle School Version, 2012) and Darwin's Theory of Natural Selection
based on the logical development of the theory by Ernst Mayr (1982)

Fact or Inference*	**CINS-MS 2012 Version items**	**Answer key**
Fact #1: Biotic potential All species have such great potential fertility that their population size would increase exponentially if all individuals that are born would again reproduce successfully.	1, 11	C
Fact #2: Stable populations Except for minor annual fluctuations and occasional major fluctuations, populations normally display stability.	2, 12	B
Fact #3: Natural resources Natural resources are limited. In a stable environment, they remain relatively constant.	3, 13	A
Inference #1: Limited survival Since more individuals are produced than can be supported by the available resources, but population size remains stable, it means that there must be a fierce struggle for existence among the individuals of a population, resulting in the survival of only a part, often a very small part, of the progeny of each generation.	4, 14	D
Fact #4: Variation No two individuals are exactly the same; rather, every population shows enormous variability.	5, 15	D
Fact #5: Variation inherited Much of this variation is heritable.	7, 17	C
Inference #2: Differential survival Survival in the struggle for existence is <u>not random</u>, but depends in part on the hereditary constitution of the surviving individuals. This unequal survival constitutes a process of natural selection.	8, 18	B
Inference #3: Change in population Over the generations this process of natural selection will lead to a continuing gradual change of populations, that is, to evolution and to the production of new species.	9, 19 (change in population) 10, 20 (origin of species)	B A

*Direct quotes from Mayr's work

Additional concept on the CINS:

Concept	**CINS-MS Version 2012 items**	**Answer key**
Origin of variation: New variation appears randomly through mutation and sexual reproduction.	6, 16	B

Appendix L: IRB
Refinement of the Conceptual Inventory of Natural Selection (CINS) based on college, high school and middle school student feedback.

Dianne Anderson
PLNU Biology Department
October 31, 2011

POINT LOMA NAZARENE UNIVERSITY
Institutional Review Board (IRB)

SECTION C:
FULL REVIEW
(proposed research not eligible for Expedited Review)

1. Attach a succinct description of the proposed research, using numbered pages. Be sure to address each of the following issues, with particular attention to potential risks to participants and what will be done to minimize those risks.
1a. Briefly describe the study, giving its justification and rationale.

Natural selection is a central concept in biology, yet it can be difficult to understand for students at all levels. For this reason, a research-based diagnostic test using common alternative conceptions as distractors was developed, field tested, and published in 2002 (Anderson, Fisher, & Norman, 2002). Since its publication, the Conceptual Inventory of Natural Selection (CINS) has been used by science education researchers, as well as both high school biology teachers and biology professors, across the country. It is now seen as one of the models for the development of conceptual inventories on other science topics. However, it is clear that the items can be improved (Nehm & Schonfeld, 2008), so work began in 2010 to produce a revised version of the CINS (Dwyer, 2011). In particular, this new project will address the issue of determining if parallel items on this test do indeed accurately assess understanding of the same concept. In addition, feedback from middle school teachers who have used the CINS suggests that a middle school version of the CINS with a reduced reading level is needed, so this project will initiate that work.

In order to develop and field test the new versions of the CINS, this project will involve both collecting student responses to test questions as part of normal class activities in college, high school, and middle school classrooms, as well as conducting short interviews with college, high school, and middle school students to determine how they interpret question wording, and ultimately choose an answer on each item. This student feedback will be used to improve the test items before widespread field testing.

1b. Who are the subjects? How will you recruit them? How many will be used? Specifically, note whether subjects belong to a protected group as defined in 45 CFR 46.111(b) and see Subparts B and C there for further details. Click *Regulations* at http://www.hhs.gov/ohrp.

The participants will be students taking biology courses at the San Diego schools listed below. Data collection will be accomplished with the cooperation of both current and former PLNU graduate biology students who are teachers at these sites:

a. Point Loma Nazarene University
b. Helix High School
c. Thurgood Marshall Middle School
d. Challenger Middle School

Approximately 100 PLNU students in non-major biology classes will participate in the classroom testing. This testing will take place as a regular classroom activity. Approximately 20 PLNU students will also be interviewed. The interviews will not be part of a regular classroom activity.

Approximately 100 students from each high school and middle school site will participate in classroom testing. This testing will take place as a regular classroom activity. Approximately 20 students at each site will also be interviewed. The interviews will not be part of a regular classroom activity. Since nearly all high school students and all middle school students are under age 17, they are members of a protected group. As such, precautions will be taken to avoid coercion to participate in interviews by stressing that all participation is voluntary, that participation will not affect their grade in any way, and that participants can decide to stop participating at any time.

1c. What steps will you take to assure the participation is voluntary?

Please see Section 1k. Since the testing is part of normal class activities, all students will participate and it is not voluntary. The voluntary nature of interviews in the study is outlined in the script for each grade level.

1d. What will the subjects do? How will you interact with the subjects (e.g., describe any bodily invasive procedures)?

The participants will either take a pencil-and-paper 20-item multiple choice test, or will both take the test and participate in a 20-30 minute interview. There will not be any bodily invasive procedures done as part of this research study.

1e. Describe all the equipment you will use or with which the subject will interact.

A digital audio recorder will be used to record interviews.

1f. Attached questionnaires

Attachment A: Conceptual Inventory of Natural Selection, 2011 version
Attachment B: Interview Questions

1g. Note the estimated time duration of subject participation.

Students who are only taking the test will participate in the study for approximately 30 minutes, including a time for debriefing once all students have finished the test.

Students who both take the test and complete an interview will participate for a total of approximately 60 minutes, including time for debriefing at the conclusion of the interview.

1h. Will the subjects incur any expenses? If so, please explain.

There will be no expense to the participants as part of the in-class activities. The testing will not require extra time for the students since the testing will be part of a normal classroom activity.

The only expense will be the extra time for the interviewed students. Students completing an interview will be given a small gift card.

1i. Name the facilities other than PLNU where research will be conducted and provide copies of any letters of permission or support that you have obtained. If you have not obtained any letters, please explain.

a. Helix High School
b. Thurgood Marshall Middle School
c. Challenger Middle School

Letter of permission are currently being obtained. Research on related PLNU projects has been conducted at all of the sites within the past 5 years, and principals at all sites have been supportive in the past.

1j. List the foreseeable risk(s) to subjects, describe how you will minimize each risk, and why each risk is justifiable in light of benefits (either directly to the subject or indirectly to generalizable knowledge) to be gained by the research.

In terms of the testing, the risk to the student is no greater than any risk experienced on a normal school day since it will be part of normal classroom activities.

For students who participate in interviews, there is a slight risk of embarrassment in terms of a student's fear of saying something inaccurate. The interviewer will attempt to create a relaxed, conversational atmosphere, in which the student will be asked to "think aloud" as he/she works through multiple-choice items and explain why he/she would choose a particular answer and would not choose each of the others.

Students will be prompted just to "do what they can do", and that any information that they provide will be helpful in the study. They will be positively praised for having volunteered. Students will be reminded at the beginning of the interview that his/her participation is voluntary and is held in confidentiality. Any decision to stop participating will be respected and will not affect their grade in any way.

Based on previous experience with this type of data collection, most students enjoy sitting down one-on-one with someone to talk about these test items, and in some cases, students may actually develop a better understanding of the concepts just by going through the interview process.

All students will be assigned student numbers for confidentiality. All information will be kept confidential and secured, and all collected data will be destroyed after the completion of the data analysis by Dianne Anderson and/or her graduate students or research assistants.

1k. Document how informed consent will be gained. Include the exact words and method of delivery that will, prior to their agreement to participate, inform subjects of the nature of the study and of the extent of their involvement. Attach a copy of the consent form(s). This form will be examined closely.

Scripts to be used in college, high school and middle school classrooms follow…

PLNU students

The following script will be read to PLNU students participating in the testing:

> *My name is Dr. Dianne Anderson. I am conducting research related to improving student learning in biology. Specifically, my research team is focused on the development of valid and reliable test questions on various concepts in biology. You will be taking a short paper-and-pencil test on natural selection as part of your normal class activities today. After scoring, your name will be removed from your results, and your results will be combined will those of everyone else in your class for analysis.*
>
> *In addition to having students take the short tests, I also learn a great deal by interviewing students. This is where I would like your help, although your decision to participate is voluntary and will not affect your grade in the course. I am looking for several volunteers to complete a 20-30 minute audio-recorded interview about some of the test items in which you would explain why you chose particular answers and not others. Even if you do not understand the topic well, your feedback on the questions is still valuable for the research. Your responses will help us to improve the test questions, and you might even learn some biology in the process. Interviews will take place in a classroom or office nearby at a time that works well with your schedule, and you will be given a small gift card for your participation.*
>
> *Does anyone have a question?*
>
> *Each of you has the right to decide whether to participate or not. If you wish to volunteer to be interviewed, please read, complete and sign the form, then return it to me. Keep the second copy for your records.*

The college-level consent form (see Attachment C) will be used in PLNU classes. All forms will be collected, and any college students under age 17 will be eliminated from the study.

<u>High school students</u>

The following script will be read to high school students participating in the testing:

> *My name is Dr. Dianne Anderson. I am conducting research at Point Loma Nazarene University related to improving student learning in biology. Specifically, our research team is focused on the development of valid and reliable test questions related to the process of natural selection. You will be taking a short paper-and-pencil test on natural selection as part of your normal class activities today. After scoring, your name will be removed from your results, and your results will be combined will those of everyone else in your class for analysis. You will be given an informational letter to take home to your parent or guardian explaining this research.*
>
> *In addition to having students take the short test, I also learn a great deal by interviewing students. I am also looking for several volunteers to complete a 20-30 minute interview about some of the test items in which you would explain why you chose particular answers and not others. Your decision to participate is voluntary and will not affect your grade in the course. Even if you do not understand the topic well, your feedback on the questions is still valuable to our research. Your responses will help us to improve the test questions, and you might even learn some biology in the process. Interviews will take place in a classroom or office nearby at a time that works well with your schedule, and you will be given a small gift card for your participation.*
>
> *Does anyone have a question?*
>
> *Each of you has the right to decide whether to participate or not. If you are willing to volunteer for an interview, please read, complete, and sign the assent form, and return it to me if you wish to participate in the study. Keep the second copy for your records.*
>
> *In order to participate in an interview, you will also need your parent/guardian's consent. You will be given two copies of a consent form to take home. Please have your parent or guardian complete read, sign, complete and return one form, and keep the other copy for their records.*

The K-12 informational parent/guardian letter (see Attachment D), the high school level assent form (see Attachment E), and the K-12 parent/guardian consent form (see Attachment F), and will be used in middle school classes.

Middle school students

The following script will be read to middle school students participating in the testing:

> *My name is Dr. Dianne Anderson. I do research at Point Loma Nazarene University on improving student learning in science. Specifically, our research team is focused on the development of good test questions related to natural selection. You will be taking a test on natural selection as part of your normal class activities today. After scoring, your name will be removed from your results, and your results will be combined will those of everyone else in your class for analysis. You will also be given a letter to take home to your parent or guardian explaining this research.*
>
> *In addition to having students take the short tests, I also learn a lot by talking to students about some of the questions, so I would like to know if any of you would volunteer to talk to me for about 20 minutes. This is called an interview. I would ask why you would pick some answers and not others. I will also ask you to explain why. Even if you are not sure about the answers, just hearing your ideas and thoughts about the questions will be very helpful. It is hard for me to write down everything that students tell me, so I will record our voices to refer to later.*
>
> *You can help us to improve the test questions, and you might even learn some biology in the process. Your decision to participate is voluntary, and will not affect your grade in science. Interviews will take place in a classroom or office nearby at a time that works well with you and your parent/guardian, and you will be given a small gift card as a thank you gift.*
>
> *Does anyone have a question?*
>
> *While I read the form aloud, please follow along with your finger on the page. Raise your hand if you have questions.* (read form slowly while students follow along)
>
> *Each of you has the right to decide whether to participate or not. If you are willing to volunteer for an interview, please print your name, sign your name, and return it to me. Keep the second copy for your records.*
>
> *In order to do an interview, you will also need your parent/guardian's consent. You will be given two copies of a consent form to take home. Please have your parent or guardian complete*

read, sign, complete and return one form, and keep the other copy for his/her records.

The K-12 informational parent/guardian letter (see Attachment D), the middle school level assent form (see Attachment G), and the K-12 parent/guardian consent form (see Attachment F), and will be used in middle school classes.

1l. Explain how debriefing will be handled.

Debriefing related to classroom testing: once all students have answered the test questions, they will have a chance to ask the researcher any questions related to either the purpose of the study, or to their participation in the study.

Debriefing related to interviews: at the conclusion of each interview, the student will have a chance to ask the researcher any questions related to either the purpose of the study, or to their participation in the study.

1m. If copyrighted tests, scales, or inventories are to be used attach a copy of the approval letter.

None – test items are a revised version of the original test published by myself and coauthors K. Fisher and G. Norman (2002).

References

Anderson, A., Fisher, K., & Norman, G. (2002). Development and evaluation of the Conceptual Inventory of Natural Selection. *Journal of Research in Science Teaching*, 39 (10), 952-978.

Dwyer, D. (2011) Using Interview Data from Non-Major Biology Students To Improve the Conceptual Inventory of Natural Selection. Thesis, Point Loma Nazarene University, San Diego, CA.

Nehm, R., & Schonfeld, I.S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, and open-response instrument, and an oral interview. *Journal of Research in Science Teaching*, 45 (10), 1131-1160.

Attachment A

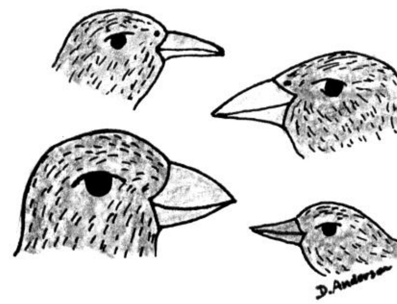
Conceptual Inventory of Natural Selection

5th edition – 2011 DRAFT

Your answers will assess your understanding of the Theory of Natural Selection.
Please choose the answer that best reflects how a biologist would think about each question.

Galapagos finches

Scientists have long believed that the 14 species of finches on the Galapagos Islands evolved from a single species of finch that migrated to the islands one to five million years ago (Lack, 1940). Recent DNA analyses support the conclusion that all of the Galapagos finches evolved from the warbler finch (Grant, Grant & Petren, 2001; Petren, Grant & Grant, 1999). Different species live on different islands. For example, the medium ground finch and the cactus finch live on one island. The large cactus finch occupies another island. One of the major changes in the finches is in their beak sizes and shapes as shown in this figure.



1. What would happen if a breeding pair of finches was placed on an island under ideal conditions with no predators and unlimited food so that all individuals survived?
   a. The finch population would stay small because birds only have enough babies to replace themselves.
   b. The finch population would double and then stay relatively stable.
   c. The finch population would increase dramatically.
   d. The finch population would grow slowly and then level off.

2. Finches on the Galapagos Islands require food to eat and water to drink. How does this fact impact the population?
   a. When food and water are scarce, some birds may be unable to obtain what they need to survive.
   b. When food and water are limited, the finches will find other food sources, so there is always enough.
   c. When food and water are scarce, the finches all eat and drink less so that all birds survive.
   d. There is always plenty of food and water on the Galapagos Islands to meet the finches' needs.

3. Once a population of finches has lived on a particular island for many years, what will most likely happen to the population?
   a. The population continues to grow rapidly.
   b. The population remains relatively stable, with some fluctuations.
   c. The population dramatically increases and decreases each year.
   d. The population will decrease steadily.

4. What is the best way to describe the evolutionary changes that occur in a finch population over time?
   a. The traits of each finch within a population gradually change.
   b. The percentages of finches having different traits within a population change.
   c. Successful behaviors learned by finches are passed on to offspring.
   d. Mutations occur to meet the needs of the finches as the environment changes.

5. Depending on their beak size and shape, some finches get nectar from flowers, some eat grubs from bark, some eat small seeds, and some eat large nuts. Which statement best describes the interactions among the finches and the food supply?
   a. Most of the finches on an island cooperate to find food and share what they find.
   b. Many of the finches on an island fight with one another and the physically strongest ones win.
   c. There is more than enough food to meet all the finches' needs so they don't need to compete for food.
   d. Finches compete primarily with closely related finches that eat the same kinds of food**,** and some may die from lack of food.

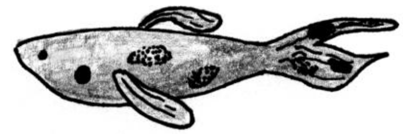
6. How did the different beak types first appear in the Galapagos finches?
   a. The changes in the finches' beak size and shape occurred because of their need to be able to eat different kinds of food to survive.
   b. Changes in the finches' beaks occurred randomly, and when there was a good match between beak structure and available food, those birds usually had more offspring.
   c. The changes in the finches' beaks occurred because the environment caused the desired genetic changes.
   d. The finches' beaks changed a little bit in size and shape with each successive generation, some getting larger and some getting smaller.

7. What type of variation in the finches' traits is passed to the offspring?
   a. Only behaviors that were learned during a finch's lifetime
   b. Only traits that were beneficial during a finch's lifetime
   c. Only traits that were genetically determined
   d. Only traits that were positively influenced by the environment during a finch's lifetime

8. What caused populations of birds having different beak shapes and sizes to become distinct species distributed on the various islands?
   a. The finches were quite varied, and those whose features were best suited to the available food supply on each island reproduced most successfully.
   b. All finches are essentially alike and there are not really fourteen different species.
   c. Different foods are available on different islands and for that reason, individual finches on each island gradually developed the beaks they needed.
   d. Different lines of finches developed different beak types because they needed them in order to obtain the available food.

—

Venezuelan guppies



Guppies are small fish found in streams in Venezuela. Male guppies are brightly colored, with black, red, blue and iridescent (reflective) spots. Males cannot be too brightly colored or they will be seen and consumed by predators, but if they are too plain, females will choose other males. Natural selection and sexual selection push in opposite directions. When a guppy population lives in a stream in the absence of predators, the proportion of males that is bright and flashy increases in the population. If a few aggressive predators are added to the same stream, the proportion of bright-colored males decreases within about five months (3-4 generations). The effects of predators on guppy coloration have been studied in artificial ponds with mild, aggressive, and no predators, and by similar manipulations of natural stream environments (Endler, 1980).

9. A natural population of guppies consists of hundreds of fish of a single species. Which statement best describes the population of guppies?
   a. The guppies share all of the same characteristics and are identical to each other.
   b. The guppies share all of the most important characteristics of the species; the small differences between them don't affect survival.
   c. The guppies are all identical on the inside, but have many differences in appearance.
   d. The guppies share most important characteristics, but also have differences that may affect survival.

10. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Which characteristics would a biologist consider to be most important in determining which guppies were the "most fit" according to Darwin's theory?
    a. large body size and ability to swim quickly away from predators
    b. excellent ability to compete for food
    c. high number of offspring that survived to reproductive age
    d. high number of matings with many different females

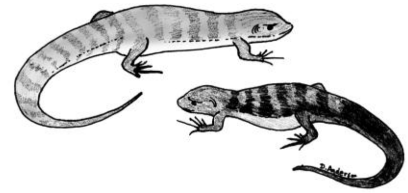
11. Assuming ideal conditions with abundant food and space, and no predators, what would happen if a mating pair of guppies was placed in a large pond?
    a. The guppy population would grow slowly, as guppies would have only the number of offspring that are needed to replenish the population.
    b. The guppy population would grow slowly at first, then would grow rapidly, and thousands of guppies would fill the pond.
    c. The guppy population would never become very large, because only organisms such as insects and bacteria reproduce in that manner.
    d. The guppy population would continue to grow slowly over time.

12. Once a population of guppies has been established for a number of years in a pond with other organisms including predators, what will likely happen to the population if conditions remain constant?
    a. The guppy population will stay about the same size.
    b. The guppy population will continue to rapidly grow in size.
    c. The guppy population will gradually decrease until no more guppies are left.
    d. It is impossible to tell because populations do not follow patterns.

13. What is the best way to describe the evolutionary changes that occur in a guppy population over time?
    a. The traits of each individual guppy within a population gradually change.
    b. The percentage of guppies having different traits within a population change.
    c. Successful behaviors learned by certain guppies are passed on to offspring.
    d. Mutations occur to meet the needs of the guppies as the environment changes.

Continue on the next page…

Canary Island Lizards



The Canary Islands are seven islands just west of the African continent. The islands gradually became colonized with life: plants, lizards, birds, etc. Three different species of lizards found on the islands are similar to one species found on the African continent (Thorpe & Brown, 1989). Because of this, scientists assume that the lizards traveled from Africa to the Canary Islands by floating on tree trunks washed out to sea.

14. Lizards eat a variety of insects and plants. Which statement describes the availability of food for lizards on the Canary Islands?
    a. Finding food is not a problem since food is always in abundant supply.
    b. Since lizards can eat a variety of foods, there is likely to be enough food for all of the lizards at all times.
    c. Lizards can get by on very little food, so the food supply does not matter.
    d. It is likely that sometimes there is enough food, but at other times there is not enough food for all of the lizards.

15. What do you think happens among the lizards of a certain species when the food supply is limited?
    a. The lizards will cooperate to find food and share what they find.
    b. The lizards fight for the available food and the stronger lizards kill the weaker ones.
    c. Genetic changes that would allow lizards to eat new food sources are more likely to occur.
    d. The lizards least successful in the competition for food are likely to die of starvation and malnutrition.

16. A population of lizards is made up of hundreds of individuals. Which statement describes how similar they are likely to be to other lizards in the population?
    a. All lizards are likely to be almost exactly the same.
    b. All lizards are exactly the same on the inside, but display differences in their external features.
    c. All lizards share many similarities, yet likely have some significant differences in their features.
    d. All lizards are likely to be the same on the outside, but display differences in their internal features.

17. Which statement best describes how traits in lizards will be inherited by offspring?
    a. When parent lizards learn to catch particular insects, their offspring can inherit their specific insect-catching skills.
    b. When parent lizards develop stronger claws through repeated use in catching prey, their offspring can inherit their stronger-claw trait.
    c. When parent lizards' claws are underdeveloped because the available prey is easy to catch, their offspring can inherit their weakened claws.
    d. When a parent lizard is born with an extra claw on each limb, its offspring can inherit the extra claw.

Continue on the next page…

18. Fitness is a term often used by biologists to explain the evolutionary success of certain organisms. Below are descriptions of four fictional female lizards. Which lizard might a biologist consider to be the "most fit" according to Darwin's Theory?

	Lizard A	Lizard B	Lizard C	Lizard D
Body length	20 cm	12 cm	10 cm	15 cm
Offspring surviving to adulthood	19	28	22	26
Age at death	4 years	5 years	4 years	6 years
Additional facts	Lizard A is very healthy, strong, and clever	Lizard B has mated with many lizards	Lizard C is dark-colore and very quick.	Lizard D has the largest territory of all the lizards.

a. Lizard A
b. Lizard B
c. Lizard C
d. Lizard D

19. According to the theory of natural selection, where did the variation in body size of the three species most likely come from?
a. The lizards needed to change in order to survive, so beneficial new traits developed.
b. The lizards wanted to become different in size, so beneficial new traits gradually appeared in the population.
c. Random genetic changes and new combinations of genes produced through mating both produce variation.
d. The island environment caused genetic changes in the lizards.

20. What could cause one species to change into three species over time?
a. Groups of lizards encountered different island environments, so the lizards needed to become new species with different traits in order to survive.
b. Groups of lizards may have been geographically isolated from other groups, and random genetic changes may have accumulated in these lizard populations over time.
e. There may be minor variations, but all lizards are essentially alike and all are members of a single species.
d. In order to survive, different groups of lizards needed to adapt to the different islands, and so all organisms in each group gradually evolved to become a new lizard species.

Attachment B: Interview questions

Questions about the CINS items:

1. Please read the question, then each answer choice. Then tell me which answer you think is correct, and why.
2. Now, please explain why you did not choose each of the other answers.

These questions will be used for each test questions. Students may be asked follow-up questions such as "can you tell me more?" or "why do you think that?"

Questions about student understanding of natural selection:

1. ''Please arrange these cards on the table so that the words that are closely related to each other are close together, and those that are unrelated to each other are far apart. If there are any terms that are unfamiliar to you or have nothing to do with natural selection, put them aside.''

adaptation	need
competition	offspring
fitness	population of rabbits
gene	random
individual rabbit	survival
mutation	variation
want	

   "Please explain why you put the cards in to this arrangement."

2. ''I am going to show you several pictures now. Please tell me whether each one is an example of one or more of the terms on the cards, then explain your answer.''

   Photos:
   1. Cheetah camouflaged in grasses
   2. Forest of pine trees
   3. Deer pulling the last leaves off branches during a snowstorm
   4. Hundreds of birds nesting on a rocky beach
   5. Herd of antelope with a couple of babies

Attachment C

Consent Form to Participate in a Research Interview
(College student version)

Introduction: The biology department at Point Loma Nazarene University is sponsoring this study for the purpose of improving a test that measures student understanding of natural selection. I understand that I am being invited to participate in the study. My participation is voluntary, and I have the option to withdraw at any time without penalty. My participation will not have any effect on my course grade. I understand that I must be at least 18 to participate.

Procedures: I understand that the proposed length of my participation in this study consists of one interview approximately 20-30 minutes long, including a few minutes for me to ask questions at the end. During the interview, I will read through test questions about natural selection and will explain which answers I would choose and why. The interview will be audio-recorded.

Risks: Although there are no risks currently known in relation to this study, I understand that there may be a slight risk that I will feel embarrassed if I do not know an answer.

Benefits: I may improve my understanding of the process of natural selection as a result of talking about the test questions during the interview. I will receive a small gift card at the end of the interview.

Confidentiality: I understand that my records will be held confidential to the extent permitted by law and that I will never be identified in any publication. I understand that a random number rather than my name will be used with the data. Any recordings will be transcribed, all identifying information will be removed, and then the recordings will be destroyed. I understand that my participation is voluntary and that I may refuse to participate or withdraw from the study at any time. Only signatures are needed for proof of consent; they will be separate from the other materials.

Debriefing: I understand that I have a right to have my questions about the study answered in sufficient detail for me to clearly understand the level of my participation, as well as the significance of the research. I understand that at the completion of the interview, I will have an opportunity to ask and have answered all questions related to my involvement in this study.

Receipt of consent form: I acknowledge having received two copies of the consent form, one to be returned to the researcher, and one for me to keep for my reference. I may call or e-mail the investigator in the study to discuss confidentially any questions about participation in the study.

Printed name: ______________________________ Age: ________

Signature: ________________________________ Date: ________
(I am at least 18 years or old.)

Investigator:
Dr. Dianne Anderson dianneanderson@pointloma.edu 619-849-2705

If you have any concerns about this study, you can also contact the Chair of PLNU's Institutional Review Board, Dr. Ross Oakes Mueller, whose job it is to make sure that you feel comfortable and safe when completing this study (619-849-2905 or RossOakesMueller@pointloma.edu).

Attachment D

Letter of Information for K-12 Parents/Guardians

Dear Parents or guardians,

As part of an on-going research project sponsored by the Point Loma Nazarene University biology department to improve learning in college and high school biology and middle school life science classes, your child's class will be participating in field trials of biology test questions. The biology concepts on the test are included in the California State Science Standards. All students in your child's class will be answering test questions used in the study as a part of normal classroom instruction under the supervision of their regular teacher.

The results from all of the students will be compiled, and your child's name will be removed from the data as soon as it is collected. The research project is focused on improving the test questions, and not on assessing your child's learning. The results will have no impact on your child's grade in the course.

Another aspect of the research is to ask for volunteers at each grade level to participate in one 20-30 minute interview in which the students will be asked to explain which answers they would choose on specific test questions, and to explain why they would make that choice. Interviewed students will receive a small gift card at the end of the interview. Interviews will be held on school grounds at a time that is convenient for both your child and you. This research is important in order to refine the wording of questions so that they are as clear as possible to students.

If your child has volunteered to participate in the interviews, he/she will bring home two copies of a consent form for you to sign if you are willing to allow your child to participate. Your consent is voluntary and can be withdrawn at any time. If you are willing to let your child participate, please read, complete and sign one copy of the consent form and retain the second copy for your records.

If you have any questions or concerns, please e-mail me at dianneanderson@pointloma.edu or call me at 619-849-2705. I will get back to you promptly.

If you have any concerns about this study, you can also contact the Chair of PLNU's Institutional Review Board, Dr. Ross Oakes Mueller, whose job it is to make sure that you and your child both feel comfortable and safe when completing this study (619-849-2905 or RossOakesMueller@pointloma.edu).

Sincerely,
Dr. Dianne Anderson
Point Loma Nazarene University Biology Department
San Diego, CA 92106

Attachment E

Assent Form to Participate in a Research Interview
(High School version)

Introduction: The biology department at Point Loma Nazarene University is sponsoring this study for the purpose of improving a test that measures student understanding of natural selection. I understand that I am being invited to participate in the study. My participation is voluntary, and I have the option to withdraw at any time without penalty. My participation will not have any effect on my course grade.

Procedures: I understand that the proposed length of my participation in this study consists of one interview approximately 20-30 minutes long, including a few minutes for me to ask questions at the end. During this time, I will read through test questions about natural selection and will explain which answers I would choose and why. The interview will be audio-recorded.

Risks: Although there are no risks currently known in relation to this study, I understand that there may be a slight risk that I will feel embarrassed if I do not know an answer.

Benefits: I may improve my understanding of the process of natural selection as a result of talking about the test questions during the interview. I will receive a small gift card at the end of the interview.

Confidentiality: I understand that my records will be held confidential to the extent permitted by law and that I will never be identified in any publication. I understand that a random number rather than my name will be used with the data. I understand that my participation is voluntary and that I may refuse to participate or withdraw from the study at any time. Only signatures are needed for proof of consent; they will be separate from the other materials. Also note that recordings will be transcribed, all identifying information removed, and then recordings will be destroyed.

If scores from a specific test will be linked with an interview, how will this be done, and how will identifying information be protected?

Debriefing: I understand that I have a right to have my questions about the study answered in sufficient detail for me to clearly understand the level of my participation, as well as the significance of the research. I understand that at the completion of the interview, I will have an opportunity to ask and have answered all questions related to my involvement in this study.

Receipt of assent form: I acknowledge having received two copies of the assent form, one to be returned to the researcher, and one for me to keep for my reference. I may call or e-mail the investigators in the study to discuss confidentially any questions about participation in the study.
In addition to the signed assent form, I understand that my parent or guardian will need to sign a consent form in order for me to participate in an interview.

Printed name: ______________________________ Age: ________

Signature: ______________________________ Date: ________

Investigator:
Dr. Dianne Anderson dianneanderson@pointloma.edu 619-849-2705

If you have any concerns about this study, you can also contact the Chair of PLNU's Institutional Review Board, Dr. Ross Oakes Mueller, whose job it is to make sure that you feel comfortable and safe when completing this study (619-849-2905 or RossOakesMueller@pointloma.edu).

Attachment F

Consent Form to Participate in a Research Interview
(K-12 parent/guardian version)

Introduction: The biology department at Point Loma Nazarene University is sponsoring a study for the purpose of improving a test on natural selection. I understand that my child has been invited to participate in the study as a volunteer, and he/she has the option to withdraw at any time without penalty. My child's participation will not have any effect on his/her course grade.

Procedures: I understand that the proposed length of my child's participation in this study consists of one audio-recorded interview 20-30 minutes long, including a few minutes for my child to ask questions at the end. During this time, my child will read through test questions about natural selection and will explain which answers he/she would choose and why.

Risks: Although there are no risks currently known in relation to this study, I understand that there may be a slight risk that my child will feel embarrassed if he/she does not know an answer.

Benefits: My child's understanding of the process of natural selection may improve as a result of talking about the test questions during the interview. My child will receive a small gift card at the end of the interview.

Confidentiality: I understand that my child's records will be held confidential to the extent permitted by law and that my child will never be identified in any publication. I understand that a random number rather than my child's name will be used with the data. Only signatures are needed for proof of consent; they will be separate from the other materials.

Debriefing: I understand that both I and my child have a right to have my questions about the study answered in sufficient detail for us to clearly understand the level of my child's participation, as well as the significance of the research. I understand that at the completion of the interview, my child will have an opportunity to ask and have answered all questions related to his/her involvement in this study.

Receipt of consent form: I acknowledge having received two copies of the consent form, one to be returned to the researcher, and one for me to keep for my reference. I may call or e-mail the investigator in the study to discuss confidentially any questions about participation in the study.

I give my consent for my son/daughter, ______________________________, to be a volunteer in the study.

(Printed student full name)

Parent/guardian's printed name: ________________________________

Parent/guardian's signature: ___________________________________Date: _________

Investigator:
Dr. Dianne Anderson dianneanderson@pointloma.edu 619-849-2705

If you have any concerns about this study, you can also contact the Chair of PLNU's Institutional Review Board, Dr. Ross Oakes Mueller, whose job it is to make sure that you and your child feel comfortable and safe when completing this study (619-849-2905 or RossOakesMueller@pointloma.edu).

Attachment G

Assent Form to Participate in a Research Interview
(Middle School version)

I have been informed by Dr. Anderson or a member of her team about the research study being done to improve science learning.

I understand that I have been invited to volunteer to do one 20-30 minute interview that will be audio-recorded. I will be able to ask any questions at the end of the interview.

I understand that my participation will not have any effect on my grade.

During the interview, I will be asked to explain my answer choices on science test questions.

I understand that my answers during the interview are helpful for the research, even if I don't understand the questions very well.

I understand that I might learn something about science by doing the interview.

I understand that my privacy will be respected, and that my name will never appear in a report based on the research because my results will be assigned a random number. I understand that I can decide to stop participating at any time.

I understand that I can ask questions about the research at any time.

I have received a copy of this assent form to keep.

I understand that my parent or guardian will also need to sign a consent form in order for me to participate in an interview.

Printed name: ____ Age: _________

Signature: ____ Date: _________

Investigator:
Dr. Dianne Anderson dianneanderson@pointloma.edu 619-849-2705

If you have any concerns about this study, you can also contact the Chair of PLNU's Institutional Review Board, Dr. Ross Oakes Mueller, whose job it is to make sure that you feel comfortable and safe when completing this study (619-849-2905 or RossOakesMueller@pointloma.edu).