COMPARATIVE STUDY OF AUTHENTIC SCIENTIFIC RESEARCH VERSUS GUIDED INQUIRY IN AFFECTING MIDDLE SCHOOL STUDENTS' ABILITIES TO KNOW AND DO GENETICS

A Thesis

by

JANE METTY SCALLON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2006

Major Subject: Curriculum and Instruction

COMPARATIVE STUDY OF AUTHENTIC SCIENTIFIC RESEARCH VERSUS GUIDED INQUIRY IN AFFECTING MIDDLE SCHOOL STUDENTS' ABILITIES TO KNOW AND DO GENETICS

A Thesis

by

JANE METTY SCALLON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Carol Stuessy
Committee Members,	Janie Schielack
	Lawrence Griffing
Head of Department,	Dennie Smith

May 2006

Major Subject: Curriculum and Instruction

ABSTRACT

Comparative Study of Authentic Scientific Research Versus Guided Inquiry in Affecting Middle School Students' Abilities to Know and Do Genetics.

(May 2006)

Jane Metty Scallon, B.S., Stephen F. Austin State University

Chair of Advisory Committee: Dr. Carol Stuessy

This exploratory mixed methods study addressed the types of gains students made when engaged in one of two forms of inquiry. Gains were measured on three levels: conceptual understanding, the process of scientific investigation, and use of practical reasoning skills. One hundred-thirty 8th grade students from a rural public school in East Texas participated in this study. Classes of students were randomly assigned to one of two treatment groups: guided inquiry or authentic student research learning. Non parametric statistical analysis and constant comparative qualitative analysis were used to triangulate pre-tests and post-tests, student journals, and student drawings to address the research questions. Findings support greater gains in conceptual understanding of domain specific content in a highly scaffolded guided inquiry. Further authentic scientific research learning was more effective for developing understanding of scientific investigation as a process and application of knowledge through practical reasoning skills.

DEDICATION

This thesis is dedicated to my family for the support, encouragement, and assistance they gave me throughout my studies. Special dedication goes to my sister Beth for her particular encouragement as well as her valuable assistance in editing and refining this thesis.   I dedicate this thesis to my family (my mom, brothers, and sisters) and to my sons, Daniel, Samuel, Simeon, and Nathan who provided a steady source of encouragement throughout my graduate studies.

ACKNOWLEDGEMENTS

I would like to acknowledge the exceptional dedication and work Dr. Carol Stuessy provided me as my chair. It was her excellent guidance and mentoring that made this study a success. I am honored and privileged to have been able to learn and work under her excellent leadership. I would like to acknowledge Dr. Larry Griffing and Dr. Alan Pepper for opening up their laboratory and assisting me in the technical aspects of the design of this study as well as the assistance in construction of the light boxes. Dr. Griffing's insights, advice and enthusiasm provided vision for me as I designed and implemented my research. Also, I would like to acknowledge Dane Bozeman for his help in the statistical analysis, his patience and instruction insured the accurate analysis and interpretation of my data. Additionally, I would like to acknowledge the contribution of both Toni Ivey and Susan Skidmore in helping with edits and providing ongoing emotional support.

Further, I would like to acknowledge Dr. Janie Schielack's support and the Information and Technology in Science (ITS) Center in funding me and providing me the opportunity to pursue my studies here at Texas A&M University. Also, acknowledgement for the funding of the technology used in this study goes to the Toshiba Americas Foundation.

TABLE OF CONTENTS

Page

CHAPTER

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

A Call for Inquiry

In 2000, approximately three-quarters of the eighth grade students in the United States lacked a conceptual understanding of the science information they received in school. Most students were not able to rise above rote factual recall to successfully perform scientific investigations, predict, interpret, or explain them. Furthermore, most students were unable to apply their understanding in new and real world applications. Findings from the 2000 National Assessment of Educational Progress (NAEP) confirmed that 74 % of eighth grade students performed below proficiency in science, having only a basic or below basic understanding of science concepts appropriate for the eighth grade, with no apparent statistical improvement in performance from 1996 to 2000 (National Center for Educational Statistics [NCES], 2004a). NAEP also confirmed that many science teachers had attempted the shift toward inquiry-based instruction to enhance conceptual understanding, promote practical reasoning, and aid students in conducting scientific investigations. However, science teachers in the United States have not yet accomplished this goal.

Research findings continue to support student inquiry as an effective method of enhancing students' conceptual understanding. There are an abundance of studies that

---

This thesis follows the style of Journal of Research in Science Teaching.

point to the benefits of student inquiry (Bonnstetter, 1998; Bruer, 1999; Bransford, Brown & Cocking, 2000; Chinn & Malhotra, 2002; Driver, Asoko, Leach, Mortimer, & Scott, 1994; Feldman & Minstrell, 2000; Polman, 2000). Based on this broad consensus of research promoting inquiry, the question that must be asked is, "Why don't the NAEP scores reflect improvement in students' understanding of science?"

Perhaps the answer lies in the type of inquiry in which K-12 students typically engage. In researching the literature, no studies were found that directly compared the impact of one inquiry method over another on student understanding. This apparent gap in the literature punctuated my need to investigate the effects different inquiry-based instructional methods have on students' achievement in science.

Inquiry is a generic term that has broad meaning. Inquiry is sometimes referred to as *scientific* inquiry, f*ull*-inquiry, and immersion (Duschl & Grandy, in preparation). The term *inquiry* has a wide variety of uses in the literature. It is described as a way of knowing, a way of becoming scientifically literate (National Research Council [NRC], 1996), a way of doing science, (American Association for the Advancement of Science [AAAS], 1989), a way of developing conceptual understanding (NRC, 1996) and a way of promoting scientific discourse. Inquiry is routinely described as an activity; an ability; a way of learning; a teaching strategy; an action; a process; as iterative cycle of experimentation, reflection and revision; a curriculum sequence; as probing; and as exploration. To add to the confusion, inquiry has been divided into levels of complexity by some and descriptive of only one event by others. Still others specify that the learning environment directs the inquiry (Duschl & Grandy, in preparation). The dissensus over how long an inquiry should last spans the continuum from one tight experience completed within one class period to months of ongoing investigations.

Lehrer, Schauble, & Petrosino (2001) describes inquiry as “more than just experimentation, involving complex forms of argument deeply embedded within domain-specific practices of modeling, representation and material manipulation of the world” (p. 251). Inquiry has been defined as probing of the natural world in search of explanations based on evidence, leading toward an understanding of the world (AAAS, 1989). Often inquiry is described as *learning with understanding* (AAAS, 1989; Bransford, Brown & Cocking, 2000; NRC, 1996). This phrase merges factual recall within a conceptual framework to emphasize the importance of understanding a phenomenon more completely than either is able to do independently. Bonnstetter (1998) defines inquiry by assigning five levels of inquiry, which span a continuum from very teacher directed to completely student directed. Finally, Etheredge & Rudnitsky (2003) stress the cyclical, dynamic nature of inquiry. They describe inquiry as a process, rather than a method of instruction. Inquiry as iterative with repeated cycles of reflection, revision, and experimentation further asserting that inquiry is not a step-by-step method.

This wide breadth of interpretation makes it essential to define how the term inquiry will be used within this study. For the purposes of this study, inquiry is a process in which the learner manipulates a system of variables to develop conceptual understanding of the system. Inquiry investigations range from traditional hands on to authentic scientific research. The inquiry model (Bonnstetter, 1998) used in this research defines levels of student inquiry, as follows:

1. *Traditional hands-on*, in which the teacher directs all aspects of the investigation (topic, question, materials, procedures/design, results/analysis and conclusion).
2. *Structured inquiry*, in which the teacher provides the topic, question, material, and design of the experiment. Together the teacher and student collect and analyze the data and the student derives the conclusion.
3. *Guided inquiry*, in which the teacher provides the topic, question to be investigated, and materials. Together the teacher and students design the experiment; and the students collect the data, analyze it, and forms their own conclusions.
4. *Student directed inquiry*, in which the teacher provides only the topic, and both the teacher and students develop the question. The students determine the materials to be used, and the procedures to follow. They collect and analyze the data and draw their own conclusions.
5. *Student research*, in which the teacher and student decide on the topic. The students then develop the question, decide which materials and procedures to use, collect the data, perform the analysis, and form the conclusion.

The levels of inquiry progress from completely teacher directed to largely student directed. Research findings support the assumption that most inquiry done in the classroom falls somewhere in the middle of this spectrum between structured inquiry and guided inquiry (Bonnstetter, 1998; Chinn & Malhotra, 2002).

The term “authentic” also occupies an interesting position in the discussion on inquiry. Is the inquiry authentic to the student? If it is, does that mean that the activity has meaning and relevance to the student? Or, is the classroom inquiry authentic to the scientist in the laboratory? If it is, does that mean that the procedures are the same as those which are done by a scientist? Or is the classroom activity authentic to the field of science? In that regard, does that mean that the classroom inquiry is aimed toward contributing to the knowledge base of what is known about the natural world? Edelson (1999) extends student research to involve scientists in the inquiry process. Authentic science research learning allows students to learn science through authentic methods typically employed in scientists’ laboratories. These methods have been carefully modified for the classroom. “Adapting the practices of science to classrooms can provide benefits of authenticity for science learning. Current theories hold that authentic learning activities are the key to developing understanding that will serve learners beyond the classroom” (Edelson, 1999, p. 5). Bonnstetter’s model of the levels of inquiry and Edelson’s model of classroom science research learning fit together as shown in Table 1.

Table 1

Combination of Bonnstetter and Edelson Model

Categories	Bonnstetter Model					Edelson Model
Inquiry	Traditional Hands-On	Structured Inquiry	Guided Inquiry	Student Directed Inquiry	Student Research	Authentic Scientific Research Learning
Topic Choice	Teacher	Teacher	Teacher	Teacher	Teacher/Student	Student
Question	Teacher	Teacher	Teacher	Teacher/Student	Teacher/Student	Student
Materials	Teacher	Teacher	Teacher	Student	Student	Student
Procedure & Design	Teacher	Teacher	Teacher/ Student	Student	Student	Student
Results & Analysis	Teacher	Teacher/ Student	Student	Student	Student	Student
Conclusion	Teacher/Student	Student	Student	Student	Student	Student
Instructional Focus	Curriculum Alignment	Curriculum Alignment	Curriculum Alignment	Curriculum Alignment & Student Interest	Student Interest	Scientific Knowledge
Outcome	Known Prior	Known Prior	Known Prior	Known Prior	Known Prior/Unknown	Unknown
Contribution	Rote Learning	Rote Learning	Collaborative Instruction	Collaborative Instruction	Student Interest	Contribute to the body of scientific knowledge
Educational Focus	Teaching	Teaching	Teaching/ Learning	Learning	Learning	Generation of new knowledge& Learning about how science is done

Stuessy & Scallon, 2006

The Argument

Debate exists as to whether students are able to engage in authentic scientific research learning (ASR). Those that argue that student research can not be done in schools say that, "authentic scientific inquiry is a complex activity, employing expensive equipment, elaborate procedures and theories, highly specialized expertise, and advanced techniques for data analysis and modeling. Schools lack the time and resources to reproduce such research tasks" (Chinn & Malhotra, 2002, p. 177). Those who believe students can and should be encouraged to engage in authentic scientific research cite the benefits they have observed (Feldman & Minstrell, 2000; Stuessy & Scallon, in press). Bonnstetter (1998) claims that authentic scientific research is inquiry's ultimate goal. In authentic scientific research the student simply needs support and guidance from the teacher. Bonnstetter (1998) goes on to point out that not all students will be able to engage in student research. Edelson (1998) provides examples of successful authentic scientific research learning within the classroom. It is anticipated that this research will help to identify and clarify the benefits a broad range of students gain from authentic scientific research.

Rationale

The rationale for developing conceptual understanding, practical reasoning, and an understanding of the scientific process lies in the need for people in all situations of life to be able to apply what they have learned. Those who really learn something are able to apply it in a new context. Learning must go beyond factual content that is disconnected from application and therefore of limited use. “Informational and societal changes require education to develop individuals with the knowledge, problem solving skills, cognitive processes, intellectual dispositions, and habits of mind necessary to engage in lifelong learning” (Costa, 2001).

The ability to take content (factual knowledge) and apply it to different situations transforms learning into knowledge that is useful in building new understanding. Conceptual understanding separates the expert from the novice (Bruer, 1993). Development of conceptual understanding is important because all people in our society from entry-level workers to college students are expected to have these types of skills when they leave high school (Bruer, 1993). These skills are the focus of the NAEP assessment, which serves as the framework on which the assessments in this study are built. The introduction to *Science for All Americans* (AAAS, 1989) states that,

> Education has no higher purpose than preparing people to lead personally fulfilling and responsible lives…..education in science… should help students to develop the understandings and habits of mind they need to become compassionate human beings able to think for themselves and to face life head on.

> It should equip them also to participate thoughtfully with fellow citizens in building and protecting a society that is open, decent, and vital… The most serious problems that humans now face are global…What the future holds in store for individual human beings, the nation, and the world depends largely on the wisdom with which humans use science and technology (p. xiv).

That only a few students achieve this level of understanding in school is no longer acceptable. High expectations from the labor force, secondary institutions of learning, and post-secondary institutions of learning dictate that all students should leave high school with practical reasoning skills. Instructional practices can incorporate strategies that encourage higher-order thinking among students. This study’s importance resides in its potential in contributing to our knowledge about the impact of different inquiry forms on students’ conceptual understanding of the material, on their understanding of the scientific process, and on their abilities to effectively engage in scientific investigation, and finally, their ability to practically reason.

Problem

Inquiry-based instruction most commonly found in the classroom has left students unable to conceptually understand the science they were taught. Students lack the skills to perform scientific investigations successfully, to accurately interpret their data or justify their results. They are unable to transfer their learning to new or real-world applications. In effect, they have left school ill-prepared for future work or life long learning.

Previous experiences in the classroom with authentic scientific research, as well as the experiences of other teachers using student research (Edelson, 1998; Feldman & Minstrell, 2000; Minstrell & van Zee, 2000), have led me to informally conclude that authentic scientific research can efficiently and effectively enhance the development of the desired outcomes called for by the National Research Council. Positive results of this research could encourage science educators and teachers to invest the time to engage their students in authentic scientific research in the classroom. In addition, these results will contribute to the body of knowledge about the effects of scientific inquiry on learning. Perhaps an answer to poor student performance on NAEP can be found in the nuances of the type of inquiry in which students are typically engaged.

Purpose

Do alternative forms of inquiry differently impact conceptual understanding, practical reasoning, and the understanding of scientific investigations? Does a particular type of inquiry best facilitate student development of conceptual understanding, which in turn leads to practical reasoning and changes in conceptual models? The purpose of this study was to compare a well scaffolded form of guided inquiry with authentic scientific research.

Research Questions

What is the value of authentic scientific research and guided inquiry in enhancing students' abilities to know and do genetic research?

1. How do these two forms of inquiry affect conceptual understanding?
2. How do these two forms of inquiry affect students' understanding of scientific investigations?
3. How do these two forms of inquiry affect practical reasoning ability?

Limitations

The process by which students were chosen for this study prevented these findings from being generalized to the larger population of eighth grade students. The students within each class were pre-determined by the school counselor who intentionally placed students in specific classes for a variety of reasons, few of which were random placement. Entire classes were randomly assigned to a treatment group. For this reason, results of this study are not generalizeable beyond the students in this study. These students serve as the population. A total of six classes of eighth grade students (N = 130) participated in this study. However, attrition was high resulting in low numbers of students for whom I had matched pre- and post-test data sets. Reasons for high attrition occurred primarily as a result of the time of year in which the study was done. These reasons are expanded on in Chapter V. The low numbers of matched pre- and post-tests constrained the availability of justifiable analysis techniques. I made the decision to use non-parametric methods to analyze differences between the two treatment groups.

Finally, I was constrained by district mandates directing what material was to be covered at specified times of the year and the length of time instruction could focus on that material. Tight controls over how the teacher allocates instructional time made it sufficiently difficult to conduct this study over the prolonged time required to complete the research. It was essential that I gained permission to deviate from the prescribed district syllabus. This limitation is becoming increasingly common as more districts take control away from the teacher, centralizing instruction around district goals rather than student needs. This tight control over time spent on specific topics at specific times during the year would likely discourage a teacher from such open-ended inquiry as ASR. Authentic scientific research by nature requires significant lengths of time to complete. Teachers may find that they have to move on before the student research is complete.

Deficiencies

As is often the case, questions raised in studies are not likely to be resolved in one intervention. It is far more likely that a combination of variables working together would be found to be responsible for changes in student understanding. Some of these combinations include individual teacher differences both in instructional preferences and pedagogical knowledge. Each teacher brings unique experiences, instructional strategies, and beliefs about teaching into the classroom. Brown (1992) states that variations in teachers' pedagogical abilities and understanding of the theoretical underpinnings of an intervention significantly impact student learning. It is quite possible that another teacher conducting this same study could find results that were different, caused by variables other than the inquiry intervention. This study does not address the influence of the teacher on the inquiry process.

Other important combinations that may influence conceptual understanding include the level of metacognition students possess and the degree of emphasis the teacher places on developing these skills within the student. Metacognitive skill development is also thought to play an integral part in students' development of conceptual understanding and practical reasoning skills. Furthermore, teaching metacognitive skills may enhance conceptual understanding and practical reasoning abilities among students. Schraw and Dennison, (1994) and Sperling et al., (2001) claim that perhaps more effective than the inquiry style alone is the degree of metacognitive processes that are taught. Inquiry style and instruction in metacognitive skills in partnership may not only enhance students' practical reasoning and conceptual understanding but maximize it (Sperling, 2001). Changes in students' metacognitive abilities were not addressed in this study.

Theories of expertise and the effect of experts over novice learners on conceptual understanding also were not addressed in this study. Development of expertise occurs over many years (Goldman, Petrosino & Cognition and Technology Group at Vanderbilt, 1999). Studies in expertise show that experts have better conceptual understanding of domain specific knowledge and they are able to transfer this knowledge to new situations (Bruer, 1993). The length of this study was not sufficient to allow for complete development of expertise. Conceptual understanding and practical reasoning skills almost certainly would not have had the time to fully develop in the time frame of this research. In this research, however, movement toward this type of learning could be evidenced. Therefore, this study has set the stage for a more in-depth, long-term study of the events that influence students' conceptual understanding, practical reasoning abilities, and their understanding of scientific investigation.

It is quite possible that the content of this research held little interest for some students. If this were the case and students did not find the content of the study meaningful, less cognitive engagement could have occurred on the part of the student with lower conceptual gains in all areas of measure. It would be a mischaracterization to conclude that the inquiry methods were ineffective when it may have been a lack of engagement or interest on the part of the student. Student interest in the content of the inquiry was not measured.

In conclusion, it was unrealistic for me to try to measure the impact of all of these events simultaneously in the scope of this exploratory investigation. Instead, this research sought to study one event and how the inquiry method influences student learning within a specific time frame and with one experienced teacher filling the role of teacher and researcher.

Definitions

Authentic Scientific Research Learning. Adaptation of the practice of science into the classroom consists of three essential components, the attitudes of scientists (uncertainty & commitment), tools and techniques that are shared across the community of scientists which allow them to pose and investigate questions, and social interactions such as the sharing of results in an atmosphere of cooperation and competition (Edelson, 1998).

Conceptual Understanding. The body of scientific knowledge consisting of essential scientific concepts which involve a variety of information, including: facts and events the student learns from both science instruction and experiences with the natural environment; and scientific concepts, principles, laws, and theories that scientists use to explain and predict observations of the natural world (O'Sullivan & Wiess, 1999, p. 5).

Guided Inquiry. This level of inquiry has the teacher deciding the question, materials, and procedures; the student is invited to contribute to the analysis and conclusions. Guided inquiry is focused on problems with known answers and content that aligns with curricular goals rather than student interest (Bonnstetter, 1998).

Inquiry. First, it refers to the abilities students should develop to be able to design and conduct scientific investigations and to the understanding they should gain about the nature of scientific inquiry. Second, it refers to the teaching and learning strategies that enable scientific concepts to be mastered through investigations (NRC, 2000 p.xv).

Practical Reasoning. Practical reasoning probes students’ abilities to use and apply science understanding in new, real world application (O’Sullivan & Wiess, 1999, p. 5).

Proficient. Solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter (NCES, 2004b).

Scientific Investigation. Scientific investigation includes the application of appropriate scientific knowledge, problem-solving skills, and thinking processes to the creation of new knowledge and understanding. The NAEP 1996 and 2000 science assessments probe students' abilities to:

- acquire new information;
- plan appropriate investigations;
- use scientific tools; and
- communicate their results to a variety of audiences (NCES, 2004a).

Scientific Literacy. Scientific literacy means that a person can ask, find, or determine answers to questions derived from curiosity about everyday experiences. Scientifically literate people have the ability to describe, explain, and predict natural phenomena; they can read with understanding articles about science, engage in social conversation about the validity of the conclusions. A scientifically literate person can identify scientific issues underlying national and local decisions, express positions that are scientifically and technologically informed, and they can pose and evaluate arguments based on evidence and to apply conclusions from such arguments appropriately (NRC, 1996).

Special-Needs Students. Students who have a physical or mental impairment, as defined by Americans with Disabilities Act, that substantially limits one or more major life activities, who have a record of such impairment or who are regarded as having such an impairment. (Concepts to Classroom, 2000)

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

> …what the subject matter comes to mean in the lives of learners still
> depends on the forms of participation available to them. Wenger, 1998

This study was designed to investigate what students learn as a result of participation in one of two forms of inquiry, guided inquiry (GI) or authentic scientific research learning (ASR). This review provides an in-depth look at the inquiry by answering four questions. First, who calls for inquiry and how is it defined? Second, what are the potential benefits and deficiencies of inquiry on student learning? Third, what do cognitive learning theories reveal about how learners gain knowledge that would advocate the use of inquiry? Finally, how are the effects of inquiry on learning best assessed? In short, what does the literature say about inquiry learning and its place in enhancing understanding of scientific facts and concepts? Further, this section reviews the literature on the understanding of scientific investigation as a process, and the ability to practically reason using new knowledge and finally how best to assess this learning.

Who Calls for Inquiry, and How Is It Defined?

Broad support exists for students to go beyond factual recall to develop deep conceptual understandings of domain-specific concepts (American Association for the Advancement of Science [AAAS], 1989; Bransford, Brown & Cocking, 2000; Bruer, 1993; Donovan & Bransford, 2005; Duschl, & Gitomer, 1997; Driver, Asoko,Leach, Mortimer, & Scott, 1994; Minstrell & van Zee, 2000; National Research Council, [NRC], 1996; Polman, 2000; Wiggins & McTighe, 1998). The NRC stresses the importance of students being scientifically literate and advocates the fundamental role of inquiry in education to develop scientific literacy (NRC, 1996). The National Science Education Standards (NSES) explains scientific literacy as follows:

> Scientific literacy means that a person can ask, find, or determine answers to questions derived from curiosity about everyday experiences. It means that a person has the ability to describe, explain, and predict natural phenomena. Scientific literacy entails being able to read with understanding articles about science in the popular press and to engage in social conversation about the validity of the conclusions. Scientific literacy implies that a person can identify scientific issues underlying national and local decisions and express positions that are scientifically and technologically informed. A literate citizen should be able to evaluate the quality of scientific information on the basis of its source and the methods used to generate it. Scientific literacy also implies the capacity

> to pose and evaluate arguments based on evidence and to apply conclusions from such arguments appropriately (NRC, 1996 p. 22).

Scientific literacy is seen as an essential component for every citizen to function in a free society. Citizens should be able to evaluate claims and make informed decisions (AAAS, 1989; Bransford, Brown & Cocking, 2000; Bruer, 1993; Costa, 2001). Inquiry learning is consistently presented as the preferred instructional method by which students meet the challenges of learning at the levels called for in not only the Texas state standards, but also the national standards for science (Texas Education Agency [TEA], 2005).

AAAS (1989) defends the need for inquiry in the following statement, "By gaining lots of experience doing science, becoming more sophisticated in conducting investigations, and explaining their findings, students will accumulate a set of concrete experiences on which they can draw to reflect on the process" (p. 4). The inquiry process is instrumental in learning with understanding and competent performance (NRC, 1994; AAAS, 1989). Tables 2 and 3 provide a summary of what the NRC suggests should be emphasized as well as deemphasized in science instruction.

Inquiry instruction has broad support from the science community (Bonnstetter, 1998; Bransford, Brown & Cocking, 2000; Bruer, 1993; Brown, 1992; Edelson, 1998, Feldman, 2000; Wiggins & McTighe, 1998). Edelson (1999) says, "Participation in inquiry can provide students with the opportunity to achieve three inter-related learning objectives: the development of general inquiry abilities, the acquisition of specific

Table 2

Emphasis of the NRC National Science Standards

Emphasis of National Science Standards Less Emphasis on:	Emphasis of National Science Standards More Emphasis on:
Memorization of facts and information	Understanding of scientific concepts through inquiry.
Learning science disciplines in isolation	Learning science in the context of integrated inquiry & technology in the context which is meaningful.
Science facts as separate from the process	Integration of concepts and scientific process
Presenting scientific process as a set of procedures	Present scientific process as a dynamic process that seeks to investigate, and analyze
Superficially teaching many science topics	Deep understanding of fewer topics

Table modified from Chiappetta & Koballa, 2000)

Table 3

NRC National Standard Emphasis in Inquiry

Less Emphasis on Inquiry that:	More Emphasis on Inquiry that:
Cook book type labs that demonstrate a concept	Activities that require investigation and analyzing of findings
Lack of process skills	Emphasis on process skills
Investigations completed in one class session	Investigations occurring over a prolonged time
Getting the right answer	Using evidence to develop and revise thinking to provide an explanation
Individual work that is not shared	Discussion of ideas and results as a class emphasizing justification and explanation
Few Investigations done is classroom	Many investigations done to develop understanding, ability, knowledge of inquiry and science content.

Table modified from Chiappetta & Koballa, 2000)

investigation skills, and the understanding of science concepts and principles” (p. 393). Even with this support inquiry remains a loosely defined instructional style having varying degrees of student-centeredness (Bonnstetter, 1998; Etheredge & Rudnitsky, 2003).

What Are the Potential Benefits and Limitations of Inquiry on Student Learning?

Recent studies have revealed unexpected benefits of inquiry learning for students with special needs (Brown, 1994; Champion, Shaprio, & Brown, 1995; Zohar & Dori, 2003). A synthesis (Zohar & Dori, 2003) of four independent studies investigating the role of inquiry on low achieving students' responses to instruction reported positive findings in all four studies. Zohar and Dori's (2003) analysis of these four studies revealed that most teachers believed that only those students who were academically accelerated would benefit from instruction designed to develop higher-order thinking. Further, Zohar and Dori found that low-level students were not typically given the same opportunities to engage in higher-order thinking and that they were held to lower levels of expectation. Zohar and Dori also noted that even when teachers were explicitly instructed to engage all students equally in higher-order thinking tasks, differences in the treatment of students persisted. Zohar and Dori concluded that these studies were thus inherently biased against lower-level students as a result of teacher beliefs that higher-order skills were beyond the ability of low-level students. Even in light of this bias, results of all four studies showed significant gains in higher-order, practical reasoning performance in both low- and high-achievement groups. Significant also in this study was the clear pattern of improved performance for the low-achieving population of students. This goes counter to what most teachers believe, making a strong case for the benefits of inquiry-based instruction that places special emphasis on higher-order reasoning skills (Zohar & Dori, 2003). These studies support inquiry as an effective means of challenging all learners at whatever academic level they are, thus complying with No Child Left Behind (NCLB) mandates to provide students with the least restrictive learning environment.

Brown (1994) summarized the benefits of inquiry for all students including special-needs students in this statement, “Group cooperation, where everyone is trying to arrive at consensus concerning meaning, relevance, and importance, helps ensure that understanding occurs, even if some members of the group are not yet capable of full participation. Because thinking is externalized in the form of discussion; beginners can learn from the contributions of those more expert than they” (p. 7). This quote summarizes not only the benefits of cognitive apprenticeship (Collins, Brown & Newman, 1991) and distributed expertise (Brown, 1994) in inquiry for special-needs students, but also for all students of various developmental levels. Learners participate in inquiry to the level that they are capable as well as develop the cognitive skills to participate more fully in future inquiry experiences.

Donovan and Bransford (2005) identify three benefits students gain when they engage in inquiry.

1. Inquiry allows students to confront their own misconceptions under the guidance of a teacher who is aware of misconceptions that need to be corrected.
2. Inquiry allows students to develop an understanding of what it means to *do* science.

3. Inquiry coupled with metacognitive strategies enhances learning. Students improve their conceptual understanding by developing their metacognitive abilities through reflection on inquiry and monitoring and critiquing claims of their own and their peers (White & Frederiksen, 2000; Lin & Lehmann, 1999).

Two studies conducted by Kuhn, Schauble, & Garcia-Mila (1992) further reinforce the importance of conducting inquiry with special-needs students. Further they emphasize the beneficial impact of inquiry over a prolonged period of time. Additionally, these studies provided evidence that support the importance of allowing students to deal repeatedly with the same phenomenon. Kuhn and her associates found that long term, self-directed inquiry improved students' reasoning strategies. In their studies, students dealt repeatedly with the same phenomenon over a period of months, undergoing iterative cycles of reflection and revision. These cycles were found to enhance students' conceptual understanding. As a result of repeated exposure, students developed a rich conceptual understanding atypical of *cookbook*-type experiences, "one shot" inquiry lessons, or teacher demonstrations. These authors found that students underwent revisions to their conceptual theories and also improved their reasoning strategies.

Students do not gain an understanding of what it means to *do* science by memorizing a set of steps, commonly referred to as the scientific method (Bransford, Brown & Cocking, 2000; Brown, 1994; Bruer, 1993; Chinn & Malhotra, 2002; Driver, Asoko, Leach, Mortimer, & Scott, 1994). Inquiry allows students to learn new concepts and theories with understanding, to experience the processes of inquiry that are key elements of the culture of science, and finally, to reflect metacognitively on their own thinking and participation in scientific inquiry (Bransford, Brown & Cocking, 2000).

Benefits of Authentic Scientific Research Learning

Petrosino, Lehrer, & Shauble (2003) cite benefits of the authentic inquiry process in the following statement33

> …the use of tools and procedures, in the context of authentic inquiry, become devices that allow students to extend their everyday experiences of the world and help them organize data in ways that provide new insights into phenomena (p.143).

Edelson (1998) makes a compelling argument in defense of authentic scientific research learning, the most student-centered form of inquiry. His three arguments are:

1. Authentic scientific research places the student in the setting in which the knowledge applies. Edelson (1998) draws attention to the attitudes and social aspects of authentic inquiry that are likely to be lacking in other more guided forms of inquiry.
2. Scientists pursue answers to unanswered questions. This uncertainty constrains scientists to re-examine and repeat their investigations. This uncertainty is lacking in inquiry experiences with known answers. With these types of experiences, the cognitive processes scientists experience in

the authentic science of uncertainty are lacking. Learners who engage in traditional guided inquiry are believed to entirely miss this aspect of authentic science.

3. Scientists collaborate. They share their findings, discuss, and debate their findings with others in the community of scientists. Unless students are explicitly exposed to discussion or small group interactions that encourage them to defend their ideas and provide evidence critiqued by others, they too miss this essential element of what it means to do science.

Scientists pursue things that are personally meaningful to them. They seek answers to questions they want to answer. Activities that dictate what students are to investigate deny ownership of learning. This ownership, by its very nature, elicits a level of commitment to resolve the question.

Limitations of Less Student-Centered Forms of Inquiry

Though more and more classrooms across the United States are engaging in inquiry-based instruction in response to the NRC, *National Standards*, and AAAS, much of the theory that drives classroom instruction is outdated. Brown (1994) observed that “...the design of school practice is influenced by theories of development more typical of the 1950’s than the 1990’s” (p.10). When inquiry does make it into the classroom, the inquiry experiences are typically cookbook-type labs that bear little to no resemblance to the authenticity of how science is done in the real world. Chinn and Malholtra (2002) go so far as to claim that cookbook-type experiences actually embed an inaccurate epistemological view of what science is and how it is done.

Chinn & Malholtra (2002) draw interesting comparisons between what they call *simple inquiry tasks* and *authentic science inquiry*. These authors make a very strong argument for the cognitive and epistemological differences between the two inquiry strategies. The first argument is that simple inquiry tasks foster an inaccurate understanding of authentic science. The second argument is that simple inquiry tasks do not engage the learner in the cognitive processes that are employed in authentic science. Consequently, students engaged in simple inquiry tasks will not experience the same cognitive benefits as those who engage in authentic science inquiry.

In typical guided inquiries, the teacher orchestrates the question, which may or may not be of interest to the student. The teacher also provides experiments and hands-on activities that have known outcomes. Inquiry experiences with known and predictable outcomes circumvent almost entirely the collaborative aspect of the scientific process. They also deny students the cognitive benefits inherent in solving a problem with no known solution. In addition, these types of investigations do not provide a context from which students can develop an accurate understanding of what it means to *do real science*. Epistemologically, their experience with scientific investigation does not resemble how scientific investigations are really done. As a result, students develop an inaccurate conceptual model of what it means to do science.

Bonnstetter's (1998) model of inquiry explains how the roles of the teacher and the student change with the level at which the inquiry is experienced. Implications for the student and the teacher are different at every level. As instruction moves from teacher-centered to student-centered, the impact on the student's intellectual development increases. The focus of the classroom changes from a focus on teaching to a focus on learning (Bonnstetter, 1998; Bransford, Brown & Cocking, 2000). In more student-centered forms of inquiry the student moves from simple recitation and recall of what they know to a more "internal reconstruction of the new information" that allows them to apply their knowledge to new situations such as those occurring in problem solving (Bonnstetter, 1998). These differences in experiences and the potential impact on learner outcomes influenced the choice of guided inquiry and authentic scientific research learning as the two treatments in this study.

What Does Cognitive Learning Theory Reveal About How Students Learn?

Conceptual understanding and the ability to transfer new knowledge to novel situations are primary goals of education (Etheredge & Rudnitsky, 2003). Inquiry in the classroom can be effective in accomplishing these goals (Bransford, Brown & Cocking, 2000; Bruer, 1993; Driver, Asoko, Leach, Mortimer, & Scott, 1994; Etheredge & Rudnitsky, 2003; Feldman & Minstrell, 2000; Goldman, Petrosino & Cognition and Technology Group at Vanderbilt, 1999; Kuhn, Schauble, &Garcia-Mila, 1992; Polman, 2000). What then, is the link between conceptual understanding and inquiry?

Bruer (1993) presents a convincing argument that the link is readily explained by what recent cognitive learning theory has revealed about how students learn. Advances in research techniques allow researchers as never before to observe how learning occurs. Studies have shaped and continue to shape new beliefs about what classroom instructional strategies enhance learning (Bruer, 1993; Donovan & Bransford, 2005). Figure 1 summarizes the link between cognitive learning theory and inquiry in producing scientifically literate learners.

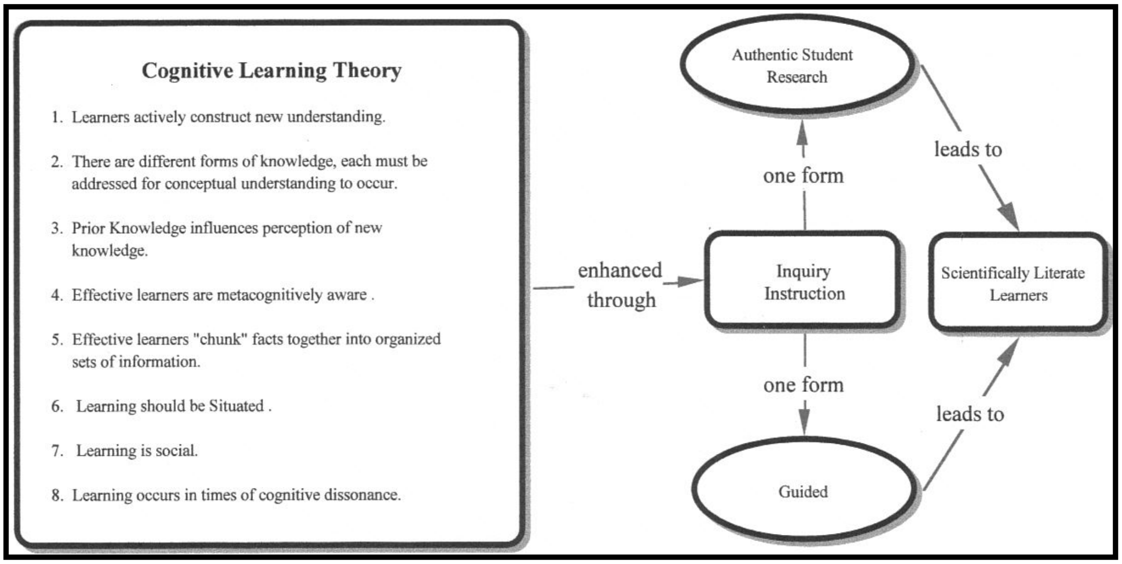

Figure 1 Concept Map. How Cognitive Theory Through Inquiry Can Lead to Scientific Literacy Among Learners.

Cognitive learning foundations on which the call for inquiry instruction rests and which are applicable to this study are as follows:

1. Learners actively construct new understanding. This understanding is built on the foundation of existing knowledge. Learners are not passive recipients

of knowledge (Bransford, Brown & Cocking, 2000; Bruer, 1993; Donovan & Bransford, 2005; Driver, Asoko, Leach, Mortimer, & Scott, 1994).

2. There are different forms of knowledge (Anderson & Krathwohl, 2001; Donovan & Bransford, 2005). Effective instruction incorporates the different forms of understanding as mutually important in creating conceptual understanding. One form should not be taught exclusive of the other forms (Bruer, 1993).
3. Conceptual constructs, prior knowledge, and prior experiences influence how learners interpret new knowledge (Bransford, Brown & Cocking, 2000; Brown, Collins, & Duguid, 1989; Bruer, 1993; Driver, Asoko, Leach, Mortimer, & Scott, 1994).
4. Learners need to be encouraged to develop habits of mind that allow them to assess their own learning. This process is referred to as *metacognition* (Bransford, Brown & Cocking, 2000; Bruer, 1993; Donovan & Bransford 2005; Goldman, Petrosino & Cognition and Technology Group at Vanderbilt, 1999; White & Frederiksen, 2000).
5. Effective learners “chunk” relevant facts together into organized sets of information rather than independent disconnected units. They readily see relationships within and between systems. They form schema or network structures that store and organize knowledge. These schemas help learners to interpret new knowledge as well as assist the learner in predicting (Bruer, 1993). Development of these schemas takes time and repeated exposure.

The ability to efficiently organize information separates the novice from the expert (Bransford, Brown & Cocking, 2000; Bruer, 1993; Donovan & Bransford, 2005).

6. Learners can more "flexibly" use new knowledge when it is acquired in a situation resembling the context in which the knowledge will be used (Bruer, 1999; Brown et al., 1989; Donovan & Bransford, 2005).
7. Learning is social. Learning occurs naturally within communities where diverse experiences, expertise and opinions are shared and critiqued. Learners need feedback and opportunities to reflect and revise their mental models as well as evaluate alternative conceptual models (Bransford, Brown & Cocking, 2000; Brown, Collins, & Duguid, 1989; Brown & Champione, 1996; Bruer, 1993; Driver, Asoko, Leach, Mortimer, & Scott, 1994; Etheredge & Rudnitsky, 2003; Wenger, 1998).
8. Learning occurs when we challenge the learner's current beliefs, referred to as *cognitive dissonance* (Bruer, 1993; Kuhn, Schauble, & Garcia-Mila, 1992).

Each of these eight theoretical statements is examined more closely in the following paragraphs.

Learners actively construct new understanding. Conceptual constructs, prior knowledge, and prior experiences influence how the learner interprets new knowledge. L*earning occurs when we challenge the learner's current beliefs.* Learners' preconceptions are often developed based on everyday experiences and are therefore difficult to change. These preconceptions are based on prior knowledge or beliefs and typically are incomplete, inaccurate conceptions (Donovan & Bransford, 2005). It is difficult, for example, for young children to accept that the world is a sphere when their everyday experience tells them they are on flat ground. This preconception is further reinforced when their own experience tells them that they can't stand on a curved surface, such as on a ball. These conceptions form the basis for our understanding of how the world works. Left unchallenged, learners will build constructs on these faulty foundations (Guzzetti, Snyder, & Glass, 1992). For example, when children are told that the Earth is round like a ball, studies have shown that they merge this new knowledge into their current schemas and construct a new understanding that the earth is round, but also flat forming a new belief that the Earth is shaped like a pancake, both round and flat (Kotulak,1996). Learner's preconceptions must be made visible to insure that new knowledge is not built upon a flawed foundation (Brown, Collins & Duguid, 1989).

Research reveals that when learners are exposed to new information that is in conflict with their current beliefs they often will memorize the information for purposes of passing a test, but when asked to apply this new knowledge, it becomes apparent that they have not merged this new knowledge into their schema. Instead, they will revert back to their previous conception using their original mental model to explain the phenomena (Donovan & Bransford, 2005). This explains in part how students are able to perform well on factual recall assessments yet fail miserably when asked to apply or explain their conceptions of the phenomena. Unless the underlying conceptual constructs are made visible and the misconceptions are challenged, learners are unlikely to develop coherent and accurate conceptual understandings of the information (Guzzetti, Snyder, & Glass, 1992).

Inquiry-based instruction promotes an environment in which the learners' current beliefs can not only be made visible through group discussion or written responses, but also challenged with the new evidence before them. Once this conflict occurs, the learner within the context of the inquiry experience can explore in depth those aspects that are in conflict with their current thinking. The intent is to allow the learner to resolve dissonance and form a new mental model that better explains the phenomena under investigation in the inquiry. Inquiry instruction allows learners to explore new concepts by immersing the learner in the phenomena and allowing them the freedom to investigate more thoroughly the specific aspects of the phenomena that are inconsistent with their current mental model. Production of new more complex schema is essential to conceptual understanding which leads to the ability to use this knowledge in novel situations.

There Are Different Forms of Knowledge. Cognitive theorists differ in how they define the different types of knowledge. Some cognitive learning theorists identify two forms of knowledge, factual and conceptual. Factual knowledge is readily observable. In *How People Learn*, Bransford, Brown & Cocking (2000) relate a story, *Fish is Fish*, to illustrate this point. In the story a frog tells the fish what the world outside the water is like. The fish is able to internalize the facts given, such as people wear clothes, they walk upright on two legs, and a cow has udders, four legs, and are black and white. These are examples of factual knowledge. Conceptual knowledge, however, is obscure and hard to uncover. Conceptual knowledge allows learners to evaluate new factual knowledge and either weave it into their existing mental constructs or to reject it.

Other theorists define four types of knowledge; factual, conceptual, procedural and metacognitive (Anderson & Krathwohl, 2001). Anderson defines factual knowledge as "the basic elements students must know to be acquainted with a discipline or solve problems in it" (p.29) and conceptual knowledge as "the interrelationships among the basic elements within a larger structure that enable them to function together" (p. 29); procedural knowledge as "how to do something, methods of inquiry, and criteria for using skills, algorithms, techniques and methods" (p.29), and metacognitive knowledge as "knowledge of cognition in general as well as awareness and knowledge of one's own cognition" (p. 29). Though theorists differ in how they define the forms of knowledge, there is agreement that knowledge in all its forms should be integrated into the learning process.

One form, factual knowledge, should not be taught to the exclusion of conceptual understanding. For the purposes of this discussion, the theory advanced by Bransford, Brown & Cocking (2000) of two types of knowledge will be used.

Bransford's fish story shows how factual knowledge is integrated with conceptual knowledge in the mind of the learner. The fish, having only a conception of what fish looks like, imagined cows as black and white fish with udders and four legs. The fish saw people as upright fish with clothes on. The fish took the new factual knowledge and fit it into its conceptual framework to form an understanding of what life was like outside the water. Merging factual knowledge and conceptual knowledge brings about competence and learning with understanding (Bransford, Brown & Cocking, 2000). Cognitive learning theory supports the belief that conceptual understanding must address the different forms of knowledge (Anderson & Krathwohl, 2001; Bransford, Brown & Cocking, 2000; Bruer, 1993).

Inquiry learning merges both factual knowledge and conceptual knowledge through exploration. Inquiry experiences allow the learner to take factual knowledge and use it in a context in which it applies giving it both context and relevance. The hallmark of effective inquiry is the natural ease with which the learner is able to merge facts with concepts to develop conceptual understanding. Donovan and Bransford (2005) state, "Competent performance is built on neither factual nor conceptual understanding alone; the concepts take on meaning in the knowledge-rich contexts in which they are applied"(p. 6).

Effective Learners Use Metacognitive Strategies. Cognitive learning theory holds that the responsibility for learning lies ultimately with the learner (Bruer, 1993; Donovan & Bransford, 2005). Learners must develop the ability to take control or monitor their own learning, becoming metacognitively aware. In short, effective learners know how to learn. Metacognition includes being aware of when you understand something, seeing the need for clarification or feedback, and employing strategies effective in retaining information. Metacognition includes processes such as reflection, self explanation, feedback, self-monitoring ones' own comprehension, and employing strategies as needed to correct faltering comprehension such as re-reading a passage when the learner realizes a lack of comprehension (Donovan & Bransford, 2005). Palinscar and Brown

(1992) were able to show “strong effects” on learning with understanding through the process of reciprocal teaching, which employs metacognitive strategies such as self-monitoring, discussion, reflection and re-reading.

Thorndike identified feedback in a paper written in 1913 as a beneficial practice that provided support for learners as they engaged in self-assessment, aiding them in knowing what they know as well as refining their concepts. Teacher to student or student to student discussions afforded an opportunity to hear differing viewpoints, which were then used to inform and refine the learner’s conception (Bruer, 1993; Driver, Asoko, Leach, Mortimer, & Scott, 1994; Etheredge & Rudnitsky, 2003). Another component of feedback is allowing learners the opportunity to test their ideas through repeated experimentation to see if they work. In summary, learners need to be encouraged to develop habits of mind that allow them to assess their own learning (White & Frederiksen, 2000; Lin & Lehmann, 1999).

Inquiry learning as defined by Etheredge & Rudnitsky (2003) asks the learner, *How do you know what you know?* This question alone highlights the metacognitive benefits inherent in inquiry. Inquiry within the classroom typically involves small groups where student theories and explanations are presented, defended, critiqued, and modified. This process encourages learners to develop metacognitive abilities since they are made aware of their thinking in the process of the discourse.

Effective Learners “Chunk” Relevant Facts Together Into Organized Sets of Information Rather Than Independent Disconnected Units. Cognitive learning theory advances the idea that people organize knowledge into conceptual frameworks (Bruer, 1993; Donovan & Bransford, 2005). Donovan and Bransford (2005) cite studies with results that showed unsurprising results that experts had not only more factual knowledge than novice learners, but, experts were able to organize these facts into set of ideas or *chunks* that were more easily retrieved as related sets of information rather than as independent facts (Bransford, Brown & Cocking, 2000; Bruer, 1993; Donovan & Bransford, 2005). Novice learners see factual knowledge as individually distinct pieces of knowledge; they are not yet able to chunk facts together. Research on expertise emphasizes that effective learners organize knowledge in a connected, structured way so that the learner not only knows the factual knowledge but how that knowledge fits into the bigger conceptual picture. Another fundamental concept of expertise is that the process of becoming an expert takes a long time and can not be expected in an isolated experience. For learners to develop expertise and conceptual understanding, they need to be exposed to multiple representations as well as to immerse themselves in the phenomenon being studied. They must have multiple opportunities to manipulate, reflect and revise their work with the phenomenon (Etheredge & Rudnitsky, 2003; Goldman, Petrosino & Cognition and Technology Group at Vanderbilt, 1999; Kesidou & Roseman, 2002; Kozma, 2000; Driver, Asoko, Leach, Mortimer, & Scott, 1994).

During the inquiry process learners have multiple opportunities to manipulate variables with the system under study. With each iterative cycle the learner modifies their mental model refining their understanding to develop a more accurate conception through experimenting, reflecting, and revising their work. Inquiry typically requires more time in class than does didactic instruction. This additional time is essential and indeed sets inquiry above lecture because it allows the learner to begin to develop expertise.

Learners Can More Flexibly Use New Knowledge When it is Acquired in a Situation Resembling the Context in Which the Knowledge Will Be Used. Cognitive theory research supports this view of situated learning. The learners' ability to utilize knowledge is enhanced when the knowledge is acquired in an environment similar to the one in which the knowledge is likely to be used (Brown, Collins & Duguid, 1989). For example, a student is more likely to understand what the scientific process is when allowed to experience it in the context of scientific investigation as opposed to reading about it in a book or listening to someone explain it.

A way in which inquiry instruction embraces grounded knowledge of cognitive theory can best be summed up by comparing a student's understanding about small engine repair by reading about it in a manual or listening to a lecture as apposed to actually building a small engine under the guidance of an expert. Inquiry is ideal for learning to occur within the situation in which it is relevant. The student must construct the small engine requiring an understanding of how the parts work together. It is more likely that the student will develop a more accurate complete conceptual understanding of small engines through inquiry than through listening to a lecture, reading a manual, or even watching a demonstration on how to construct a small engine.

Learning Is a Social Endeavor (Goldman, Petrosino & Cognition and Technology Group at Vanderbilt, 1999; Brown, 1998; Bransford, Brown & Cocking, 2000; Donovan & Bransford, 2005; Bruer, 1993). Learners benefit from dialog with each other where ideas are explained, challenged and discussed. The diverse experiences other learners bring to the learning environment serve to challenge other learners to view alternative perspectives and alternative explanations. This creates an environment that promotes critical reflection and other metacognitive strategies shown to enhance conceptual understanding (Bransford, Brown & Cocking, 2000; Bruer, 1993; Driver, Asoko, Leach, Mortimer, & Scott, 1995; Etheredge & Rudnitsky, 2003; Lin & Lehmann, 1999; White & Frederiksen, 2000). Real world learning, such as the work done by scientists in their labs or businessmen in their profession is not done in isolation. Cognitive learning theory supports the belief that students do not learn best in isolation either.

Inquiry learning embraces the idea of small group interaction, fraught with discourse in which ideas are debated, investigated, modified, explained and, challenged by other learners. Contrary to didactic forms of instruction that force students to learn in isolation, inquiry requires social interaction as learners grapple with new ideas.

How Are the Effects of Inquiry on Learning Best Assessed?

A search to find existing assessments designed for inquiry learning that had been validated was unsuccessful. Studies dealing with assessment of genetics at the high school and middle school levels primarily assessed understanding of inheritance, allele combinations, meiosis and their impact on phenotypic expression. Many of these studies, which involved programs such as *GenScopes* and *BioLogica*, used performance-based assessments such as *New Worm* and *New Fly*, which were designed to go along with the

GenScopes project and measured transfer specific to that project (Buckley, Gobert, Kindfield, Horwitz, Tinker, Gerlitz, Wilensky, Dede, & Willett, 2004; Hickey, Kruger, Fredrick, Schafer, & Kindfield, 2002). It was hoped that these would yield validated pre-test and post-test questions that could be used in this study. Unfortunately, the aspects of genetics under investigation and the tested age group differed too much to be adapted to this study. These assessments, however, did serve to significantly inform the understanding of assessment design and implementation within the domain of genetics used in this study.

This section on assessment is divided into four sections. The first provides the rationale for using short, open-ended written response questions. The second concentrates on issues of validity and reliability. The third looks at the content focus of the assessments and finally what led to the assessment format and rubrics used in this study.

Why Short Answer Open-ended Response Questions?

Fellows (1994) described a study in which students were encouraged to write out their responses. This study investigated the impact on conceptual understanding of written responses verses multiple-choice responses. The study found that the students gained better conceptual understanding from recording their responses in written format. In addition, current theory on assessments demands that assessments do more than assess learning; they should serve as an instructional tool as well (Mertler, 2000; Messick, 1994). Messick (1994) refers to other studies done that used similarly open-ended assessment questions and found them successful with sixth grade students. These articles provided strong support as well as added validity for the use of an open-ended question format on the pre- and post-assessment.

Issues of Validity and Reliability

Literature of assessment validation and reliability informed the researcher's understanding of what validation is, the importance of validation, and what constitutes a valid assessment. Mertler (2000) defines validity as matching the test to the instructional method, assessments should cover content. They should be meaningful, fair, interpretative, as well as contain cognitive complexity. Mertler (2000) claims, "The more of these validation criteria that are considered in assessment development, the legitimacy of the instrument increases making its scientific foundations more credibility stronger" (p. 13). McMillian (1999) provided a chart which was found useful in further building support for the choice of open-ended questions as a valid method to assess the research questions (see Table 4). This chart provided a number match between the learning goal and the method of assessment based on guidelines set forth in the article for matching the assessment with the best method of assessment and objective of the assessment. On this chart, high numbers indicate better matches. This chart matches open-ended response questions with student-centered inquiry instruction (McMilllian, 1999).

Table 4
Assessment Method Adapted from - McMillian, (1999; p. 15)

Targets	Objective	Essay	Performance Based	Oral Question	Observation	Self Report
Knowledge	5	4	3	4	3	2
Reasoning	2	5	4	4	2	2
Skills	1	3	5	2	5	3
Products	1	1	5	2	4	4

McMillian (1999) echoes Mertler's view that for assessments to be valid they must match what is taught and be given after students have had adequate time to learn what is being assessed and be scored by a clearly defined rubric. They should be designed to avoid students' needing "pre-requisite" knowledge in order to be successful. These same criterions for validation of assessment design and the results obtained from the instrument were echoed in a paper presented by Wiggins & McTighe (1998) at the Center on Learning, Assessment and School Structures (CLASS) at Princeton. The basic premise of the paper was that credibility of the instrument rests on it measuring what it is supposed to do.

Messick (1994) also provided a rich source of information used to reinforce the concept of validity in assessment design and interpretation. "Validation must include evidential and consequential aspects" (p. 13). Evidential validity is concerned with the content of the domain being assessed. Consequential validity is concerned with the usefulness of the instrument in a variety of arenas such as, informing instruction and enhancing student learning. Together these papers enriched the understanding of what is involved in the development of both valid and reliable assessment questions.

Focus on Learning Targets

Clearly delineated learning targets that are age-appropriate and consistent with the overall goal of instruction are important to student learning (Dwyer, 1994; McMillian, 1999; Messick, 1994; Wiggins & McTighe (1998). The Texas Essential Knowledge and Skills (TEKS) are based on the National Science Standards and serve as the legal standard and statewide curriculum standards (Texas Education Agency [TEA], 1997) by which the teacher, curriculum director, and others in a position to direct curriculum make decisions about what will be taught in Texas classrooms. Compliance with state standards is measured annually in the spring of the school year through a standardized test called the Texas Assessment of Knowledge and Skills. The TEKS clearly delineated what the learning targets for this study were. The first five of the TEKS address scientific processes and served as justification for bringing scientific inquiry and experimentation into the classroom. These first five TEKS read as follows:

> (1) Scientific processes. The student conducts field and laboratory investigations using safe, environmentally appropriate, and ethical practices. The student is expected to:
>
> > (A) demonstrate safe practices during field and laboratory investigations;
> >
> > (B) make wise choices in the use and conservation of resources and the disposal or recycling of materials.

(2) Scientific processes. The student uses scientific inquiry methods during field and laboratory investigations. The student is expected to:

(A) plan and implement investigative procedures including asking questions, formulating testable hypotheses, and selecting and using equipment and technology;

(B) collect data by observing and measuring;

(C) organize, analyze, evaluate, make inferences, and predict trends from direct and indirect evidence;

(D) communicate valid conclusions; and

(E) construct graphs, tables, maps, and charts using tools including computers to organize, examine, and evaluate data.

(3) Scientific processes. The student uses critical thinking and scientific problem solving to make informed decisions. The student is expected to:

(A) analyze, review, and critique scientific explanations, including hypotheses and theories, as to their strengths and weaknesses using scientific evidence and information;

(B) draw inferences based on data related to promotional materials for products and services;

(C) represent the natural world using models and identify their limitations;

(D) evaluate the impact of research on scientific thought, society, and the environment; and

(E) connect Grade 8 science concepts with the history of science and contributions of scientists.

(4) Scientific processes. The student knows how to use a variety of tools and methods to conduct science inquiry. The student is expected to:

(A) collect, record, and analyze information using tools including beakers, petri dishes, meter sticks, graduated cylinders, weather instruments, hot plates, dissecting equipment, test tubes, safety goggles, spring scales, balances, microscopes, telescopes, thermometers, calculators, field equipment, computers, computer probes, water test kits, and timing devices; and

(B) extrapolate from collected information to make predictions.

(5) Scientific processes. The student knows that relationships exist between science and technology. The student is expected to:

(A) identify a design problem and propose a solution;

> (B) design and test a model to solve the problem; and
>
> (C) evaluate the model and make recommendations for improving the model (Texas Education Code, Subsection B: 112.24).

In addition, at the eighth grade level, TEKS 11 deals specifically with genetics. It reads as follows:

> The student knows that traits of species can change through generations and that the instructions for traits are contained in the genetic material of the organisms. The student is expected to: identify that change in environmental conditions can affect the survival of individuals and of species; make predictions about possible outcomes of various genetic combinations of inherited characteristics (Texas Education Agency, 2005)

This combination of TEKS provided the solid foundation for the justification to engage students in the domain of genetics through the process of scientific investigation.

Format, Framework and Rubric Design

The assessment instruments developed for this study were based on the framework of the United States Department of Education’s National Assessment of Educational Progress (NAEP) assessments. Bruer (1993) supports NAEP as “among the most useful indicators of student accomplishments” (p. 3). Very few questions directly pertained to the domain of genetic mutation in the released questions from the NAEP assessment. However, a clear pattern of open-ended question construction did emerge. This pattern was consistent throughout the NAEP assessment instrument.

The NAEP open-ended question design and accompanying rubric served as a foundation on which my assessment instrument was developed. The scoring rubrics accompanying the questions from the NCES (2004b) assessments provided the foundation for the rubric design used in this study. In addition, this framework also served as one form of instrument validation.

The use of the NAEP framework concurs with curriculum designers such as Wiggins & McTighe (1998), who state that validity is enhanced by the use of standards. "Assessment needs to be supported and judged in design by standards and peer review" (p. 18). The use of the NAEP and TAKS served as my standards. Peer review, and committee review, served to add to the validity of my instruments.

These literature pieces informed my careful design of the types of assessments to be used in this study.

Conclusion

My synthesis of the literature on inquiry and inquiry learning provided the basis for this study. The review provided proof that a "knowledge gap" indeed exists, that no one had investigated the differential effects of inquiry treatments on student learning.

The review also allowed me to build a strong theoretical foundation for the design of a study to investigate the effects of inquiry learning on student outcomes.

Inquiry learning merges factual knowledge and conceptual knowledge as well as knowledge of process (NCES, 2004b). Inquiry learning best occurs when learners attempt to seek answers to their own questions about some phenomenon of nature through scientific process that allow active exploration, modeling, representation, and manipulation of the world (Lehrer, Schauble, & Petrosino, 2001). Answering scientific questions with unknown answers increases the authenticity of the learning experience (Chinn & Malhotra, 2002; Edelson, 2003). Scientific understanding can be measured, using open-ended questions that require students to explain scientific facts and concepts, to explain the details of scientific investigation, and to use scientific reasoning to solve scientific problems (O’Sullivan & Weiss, 1999).

Classroom inquiry experiences allow learners to use factual knowledge in context. When learners engage in scientific inquiry that is similar to the ways in which scientists engage in their inquiries, learners learn and use factual knowledge; they learn about the processes by which scientists investigate natural phenomena; and they learn to use reasoning in solving problems. Donovan and Bransford (2005) state that “Competent performance is built on neither factual nor conceptual understanding alone; the concepts take on meaning in the knowledge-rich contexts in which they are applied” (p. 6).

Cognitive theory makes the connection between what students learn and how they learn it. Cognitive theory provided the foundation for my thinking about the creation of the ideal science learning environment, an environment that would allow my students to actively construct new understanding for themselves, to develop habits of mind that allow them to assess their own learning, to “chunk” relevant facts together into organized sets of information, to flexibly use new knowledge in new situations, and to learn in a social environment where new ideas were constructed, shared, and revised on the basis of feedback from peers and internal reflection.

These readings from the literature confirmed the personal relevance of my study. In my own professional context as an eighth grade science teacher, I wanted to explore the effects of design of instruction on students’ learning. I embraced the tenets of cognitive theory and reasoned that a design more like that of scientists, that is, more “authentic,” would provide richer learning opportunities for my students and that they would therefore learn more. I personally expected that more authentic scientific learning environments would lead to greater gains in my students’ scientific factual and conceptual knowledge, greater gains in their understanding of science as a process, and stronger, better developed reasoning abilities. More restricted inquiry learning environments, such as those in which the teacher guides and controls student learning, would provide fewer benefits to student learner.

Before this investigation, however, all I had were “gut feelings.” The following chapters explain how I designed a classroom-based study using my own eighth grade students to investigate the benefits of two types of inquiry learning environment, guided inquiry, and authentic scientific research learning, on student learning.

CHAPTER III

METHODOLOGY

Rationale for Mixed Methods Approach

Answers to three questions about the effects of two different forms of inquiry (Guided inquiry [GI], and authentic scientific research learning [ASR]) were the focus of this study.

1. How do these two forms of inquiry affect conceptual understanding?
2. How do these two forms of inquiry affect students' understanding of scientific investigations?
3. How do these two forms of inquiry affect practical reasoning ability?

To answer these questions a mixed methods approach with concurrent triangulation to confirm and cross-validate the data was used (Creswell, 2003). Creswell asserts that triangulation in a mixed method research approach will "result in well-validated and substantiated findings" (p. 217). Both quantitative and qualitative data were collected concurrently and integrated at the time of data analysis.

Mixed method methodology was selected for two reasons. First, the heart of the research was to explore the types of learning gains students achieved as a result of one of two different inquiry experiences. This question was best answered through qualitative questions that were then quantified. Second, mixed methods were ideal for analyzing investigations done within the messiness of the classroom (Brown, 1992; Chi, 1997). It was essential to "preserve the potential to capture unanticipated phenomena" (Maxwell, 1998, p. 75), a benefit of qualitative analysis. Emphasis on quantitative data alone could have significantly limited the scope of findings arising from the study, particularly in an exploratory study where hypotheses could not be generated a priori.

Further reasons for using mixed methods were best stated by Creswell (2003), "…all methods have limitations, researchers felt that bias inherent in any single method could neutralize or cancel the biases of other methods" (p. 15) and that "... results from one method can help develop or inform the other method" (p. 16). The choice of mixed methods was based on the anticipation that one method would inform the other and capture unanticipated outcomes. Feldman and Minstrell (2000) made the following point advocating the use of mixed methods, "The extent [to which] we can triangulate students' understanding from test results, discussions, laboratory activities, and written work, we establish the reliability of our findings" (p. 256). By using a mixed methods approach with congruent triangulation, I hoped to bring reliability to my results as well as provide rich, thick descriptions to my analysis (Geertz, 1973).

The remainder of this chapter will be devoted to the specific details of the study. They are presented in the following six sections: S*tudents and the Research Setting*, E*thical Considerations*, *Details of the Intervention*, *Data Collection and Assessment*, *Analysis* and, a concluding *Summary*.

Students and Classroom Research Setting

Six classes of eighth grade students (N=130) from a rural public middle school were randomly divided into two groups. These students were of mixed gender, ethnicity, interests, socio-economic status, and academic achievement levels. Particularly noteworthy about these classes were the extremely high numbers of children with special needs. The percentage of children identified as special-needs students in each class ranged from 24 to 50 percent. Class sizes ranged from 16 to 24 students. Each class met every day for a period of 45 minutes. There were six classes each day. These classes served as the population of this study.

The classroom was outfitted with four light box set-ups (see Figures 2, 3 and 4). Each of the light boxes had eight fluorescent bulbs, four of which were fluorescent *grow lights*; the other bulbs were regular white fluorescent bulbs. A 24-hour timer was attached at the outlet from which a power strip was connected. The technology available to the students included 15 digital cameras (see Figure 5), four web-cameras, one classroom computer, six computers in the library, and an LCD projector in the classroom, which was used in large-group viewing of student data sets and student presentations. Black construction paper was placed behind the plants. The paper served two purposes: first, it provided a strong contrast between the plant and the background, making morphological features of the plant more easily distinguishable and second, the

Figure 2 Light Box Set-ups with Several Student Data Sets Being Captured.

Figure 3 Light Box Set-up with Several Students Collecting Data in Their Journals.

Figure 4 Flats of *Arabidopsis* Plants.

Figure 5 Digital Camera Set-ups.

paper helped to focus the students' data collection on the plants under investigation rather than on all of the other plants around it. Each digital camera contained a data storage card on which the data was captured, powered by an A/C adaptor rather than the battery. A/C adapters were necessary because several of the data sets were captured over a 72-hour period. Obviously, the battery alone would be insufficient to power the camera over that length of time. Cameras were held in place using table-top tripods (see Figure 5).

The seeds, soil, plant flats, and dome lids used for this study were provided by research scientists from Texas A&M University. Most of the wild-type seeds were produced and supplied by the university. The mutant seeds were obtained by the university through an outside supplier. The light box design was that of Dr. Griffing and Dr. Pepper, research scientists from Texas A&M University, who assisted in constructing the first of the four light boxes and provided ongoing support throughout the project.

Prior to planting, seeds were soaked using an imbibing solution for 24 hours and kept refrigerated. Imbibed seeds were grown by one of two methods: in flats of soil, or on moistened paper towels in Petri dishes. Two methods of growing seeds were necessary for observing the germination process and the gross morphological development of the plants. Seeds were planted by using a pipette (see Figure 6), to place them on the surface of the soil, or onto moistened paper towels in Petri dishes. The seeds placed in flats of soil were not covered with additional soil (see Figure 7). Seeds placed in Petri dishes were sealed using a *Para-Film* wax sealer to prevent water loss. Students analyzed their data sets on library computers (see Figure 8 and 9). Before students looked at their data sets the data were downloaded to the classroom computer as a safety protocol. Once data had been analyzed and copied to the classroom computer, the data storage card was erased to make room for more data sets to be captured.

Each class was divided into small groups (see Figures 10 and 11) consisting of two to five members. Group composition varied throughout the study. Groups were formed by the students and were based on the research interests of the students.

Figure 6 Teacher Modeling Use of Pipette to Plant Seeds on Soil Surface.

Figure 7 Students Filling Trays with Soil.

Figure 8 Students Analyzing Time-lapse Video.

Figure 9 Student Groups Analyzing Data.

Figure 10   Student Groups Collecting Data.

Figure 11   Student Groups Collecting Data from Flats of Plants.

Ethical Considerations

To comply with Institutional Review Board (IRB) requirements that students not be coerced or threatened, data collection occurred anonymously. All consent forms were printed in both Spanish and English and were distributed to the students and their parents. Forms were distributed and collected by a collaborating teacher (Mrs. Ruth Brooks). Mrs. Brooks held all forms for the duration of the study. As the teacher/researcher, I had no knowledge of which students participated in the study and which only participated as a part of the regular class work. To further ensure that students did not feel coerced in anyway, students completed the study as a non-graded activity. In addition, none of the pre-tests, post-tests, or student journals was analyzed until after the entire study was complete and student grades for the grading period had been submitted.

Mrs. Brooks collected all tests and student journals. She removed students' names and replaced them with a code that identified the data by class period and designated a number for each student within the class. It was essential that I knew which class period data came from, as different classes had different treatments. Any students who had not given consent for their work to be part of this study were separated from the rest of the data and kept by Mrs. Brooks. At the end of the school year after grades had been submitted to the administration, I received the data. Students therefore were assured that their consent or refusal would have no impact on their class grade.

Credibility of the Research. One might question the credibility of research done by a classroom teacher. To address this concern, I would like to summarize my qualifications to conduct valid classroom research. My participation in the Information Technology in Science (ITS) Center's long-term professional development program prepared me to conduct these interventions, both in theory and in practice. I was immersed in the inquiry process, read significant numbers of scholarly publications that informed my understanding of inquiry, pedagogy, and educational research methodology. In addition to these courses, I spent a significant amount of time with the science researchers in their laboratory learning about not only their research with the *Arabidopsis* plant but also how science is done in the real world, in other words how to *do* scientific research. These experiences uniquely qualified me to engage my students in different forms of inquiry and to guide them in developing a correct epistemological view of the scientific process. Furthermore, during my second summer with the ITS center, I was exposed to educational research design and was guided in the development of this study. This study design, although my idea was enhanced and critiqued by a team of researchers in science education, science, and educational psychology.

As part of my course work at Texas A&M, I completed several classes in research methodologies and techniques. Much of my research was informed by a course in educational research in which Creswell's book, *Quantitative, Qualitative and Mixed Methods,* was used (Creswell, 2003). I also completed a class that explored design experiments and considerations for conducting research within the classroom. Finally, because I was the teacher of the students in this study, I had the unique advantage of being able to detect changes in my students that likely would have gone unnoticed by another researcher not acquainted with these students.

The Pilot Study. Prior to this study, the intervention and the assessment instruments that were developed for this study were pilot tested. Analyses of these pilot studies were thorough and reviewed by me and discussed with the members of my research team. Results of the pilot study led to changes in the research design and in the revision of assessment instruments.

The Intervention

Six classes were divided into two groups. Though selection of the students within a class was out of my control, I was able to randomly assign entire classes to a treatment. Students in my first, fourth, and seventh period classes were assigned to the ASR treatment. Students in my second, third, and eighth period classes were assigned to the GI treatment. Both groups were similar in the total number of students involved in each treatment and in the distribution of special-needs students.

Both groups experienced three Phases in the study. Both groups received the same treatments in Phase I and II of the study and differed in Phase III of the study. Group A (the GI group) experienced one Phase of direct instruction and traditional hands-on experiences, followed by two Phases of highly scaffolded guided inquiry experiences. Group B (the ASR group) experienced one Phase of direct instruction and traditional hands-on activities, followed by one Phase of highly scaffolded guided inquiry, and a final Phase of authentic scientific research. Tables 5 and 6 summarize the research plans for groups A and B, respectively.

Intervention. Phase I - Direct Instruction. Prior to Phase I, both groups A and B took a pre-test consisting of eight open-ended response questions over a five day period. Students also made drawings of the *Arabidopsis* plant in their journals while observing this plant (see Figures 12 and 13). During Phase I, both groups participated in direct instruction in which basic content knowledge was delivered. As part of this Phase, students engaged in well scaffolded, teacher-led discussions. These discussions were intended to model and encourage practical reasoning.

In addition, this Phase encouraged students to apply genetic concepts they had learned and observed. Students also received instruction in use of digital still photography and time-lapse photography. They used this technology in a traditional hands-on investigation in which a control *Arabidopsis* plant (the *wild-type*) was

Table 5

Research Plan for Group A (Guided Inquiry Group)

Pre-Study	Phase I Activity & Assessment	Phase II Activity & Assessment	Phase III Activity & Assessment	End of Study Assessment
Pre-Test	Direct Instruction	Guided Inquiry	Guided Inquiry	Post–Test
	Guided Inquiry	Student Journals	Student Journals	Drawing
Drawing	Student journals	Open-ended questions	Open-ended questions	Open-ended questions
Week 1	Weeks 2 & 3	Weeks 4 & 5	Weeks 6, 7 & 8	Week 9

Table 6

Research Plan for Group B (Authentic Student Research Group)

Pre-Study	Phase I Activity & Assessment	Phase II Activity & Assessment	Phase III Activity & Assessment	End of Study Assessment
Pre-Test	Direct Instruction	Guided Inquiry	Student Research	Post-Test
	Guided Inquiry	Student Journals	Student Journals	Drawing
Drawing	Student Journals	Open-ended questions	Open-ended questions	Open-ended questions
Week 1	Weeks 2 & 3	Weeks 4 & 5	Weeks 6, 7 & 8	Week 9

Figure 12 Student Recording Observations in His Journal I.

Figure 13 Student Recording Observations in His Journal II.

compared with a mutated form of the *Arabidopsis* plant (see Figure 14), referred to as a *variant.* This Phase was used to scaffold the scientific process and prepare students to engage in guided inquiry. As additional assessments during Phase I, students periodically answered one or two open-ended response questions in their science journals. These questions were typically done at the end of class. These questions were designed to make students' thinking and learning both visible and measurable. During the traditional hands-on investigation occurring in Phase I, students observed differences between the control plant and the variant (see Figure 15). This variant, also called the *det-mut,* caused the plant to behave as though it were constantly in light. As a result, the variant was dwarfed in size. The students reasonably concluded that since the det-mut plant was so much smaller than the wild-type *Arabidopsis* and both varieties were planted at the same time, the function of the gene was to control the growth of the plant. The classroom teacher/researcher conducted *think-aloud* protocols to model observation techniques fundamental to scientific research. In addition to observing the plants, students were encouraged to record observations, as well as develop hypotheses based on their observations. Data recording and researchable questions were modeled by the teacher/researcher for both groups.

In a second activity of Phase I, students as groups, with my guidance, were asked to hypothesize what they believed the physical appearance of each seed would be based on their previous observations of the plants and their hypothesis that the gene in question controlled growth. Would they expect the dwarf mutated variant of *Arabidopsis* to have smaller seeds than the larger wild-type plant? Using images captured by digital

Figure 14 Sample of Wild-type *Arabidopsis* and Det-mut Variant.

Figure 15 Group of Students Evaluating Data Using Time-lapse Cameras.

photography and enlarging these images, students observed the size of the seeds for both the wild-type and the det-mut plants to confirm or reject our hypothesis. Students observed that the det-mut seed was actually larger than the wild-type seed. As a result of this observation most students modified their hypothesis to further specify that the altered gene affected the *rate* of growth rather than growth in general.

As a final activity in Phase I, students were asked to consider our refined hypothesis and predict whether the det-mut or the wild-type plant would germinate first. The idea behind this activity was to have students consider that if the gene played a role in the rate of growth, was it plausible that the wild-type would germinate faster than the det-mut? This became our hypothesis. Once again, the students found that their predictions though logical were not accurate. The det-mut germinated before the wild-type plant. These two experiments prepared the students to question their original hypothesis thus prepared them to entertain alternative hypotheses. To guide the students toward the actual function of the det-mut gene, it was essential that they be brought face to face with discrepancies in their thinking; this dissonance was the foundation on which students would be able to move forward and entertain the alternative hypotheses presented in Phase II.

Intervention. Phase II - Guided Inquiry. Both treatment groups, A and B, engaged in an experiment to determine if or how the wild-type plant would be affected by environmental influences, specifically changes in the amount of light to which the plants were exposed. Student were presented with wild-type plants and asked what they thought they would see. Some of which had been grown in continuous light, while others remained in a controlled lighting environment consisting of sixteen hours of day and 8 hours of darkness.

The focus of this activity was two fold. First, by showing students plants that were genetically the same, with only the environmental conditions altered (amount of light), I hoped that students would discover environmental effects on phenotype at a deeper level than just understanding that if plants are not given water they would die. The second focus was to capitalize on the student dissonance created in Phase I. As a group students were asked to generate an alternative hypothesis. Students could not agree on one single hypothesis, so I made the decision to allow the students to conduct exploratory research without first generating a hypothesis. I saw this as an excellent opportunity to convey to my students that not all research follows a prescribed *scientific method.* To accomplish this, students worked in small groups to compare morphological differences between the wild-type plants grown in continuous light and those that remained in controlled lighting.

Results and supporting evidence from group observations were discussed in a large group led by the classroom teacher/researcher. Class members were encouraged to critically evaluate the claims made by other groups. Group discourse was done to mimic the social nature of science as scientists routinely present their claims, support them with evidence, and respond to criticisms raised by other scientists regarding their work. Students noted that the wild-type plants were shorter and had more trichomes than the wild-type plants grown in controlled lighting. Students were next asked to compare the wild-type plants grown in continuous light with the det-mut plants. Students observed that the wild-type plants exhibited characteristics similar to that which they observed in the det-mut plants. Based on the finding of these two activities, we further modified our original hypothesis to state that the gene had some function in light sensing and somehow this impacted the way the plants grew.

The next activity done in Phase II was done with both groups, designed to prepare the students for independent research, a skill the ASR group would use in Phase III. An essential element of ASR is that students work with a phenomenon in which the answer is unknown. This next activity exposed students in both groups to *Arabidopsis* plants with genetic alteration on a gene of unknown function. Students in both groups were provided seed sets that contained a control and a plant with an altered gene (function unknown). With my guidance, seeds were planted in two different mediums. Some seeds were planted in soil, and some were put in Petri dishes similar to procedures used in Phase I with the wild-type and det-mut plants. As in Phase I, students captured the germination process with time-lapse photography by placing seeds on moistened paper towels and sealing them in clear Petri dishes. Students also captured the growth process of the plants emerging from the soil. Both data sets were analyzed in large-group. From both of these experiments, students observed growth patterns exhibited by the plants to try to uncover the function of the unknown gene. As with Phase I, small groups analyzed their data, presented their claims, and responded to questions from other students in the class. By the end of Phase II all content information had been presented to both groups at least once.

Intervention. Phase III – Overview. In the third Phase of the research, the type of inquiry instruction differed significantly in the two groups. Until this point, both groups had engaged in a well-scaffolded, guided inquiry format of instruction intended to prepare students for a more independent inquiry learning experience. Phase II instruction for group A mirrored exemplary guided inquiry and represented the best practice of guided inquiry. This guided inquiry experience continued into Phase III for group A only. All students had engaged in experiences designed to help them understand the process of scientific investigation as well as the use of time-lapse technology. In addition, they had been exposed to *think-aloud* protocols and group discussions designed to help develop practical reasoning skills. *Think-aloud* protocols and group discussions provided an observable model of practical reasoning in the classroom.

In Phase III, Group B pursued authentic student research (ASR) learning. In this study ASR learning differed significantly from guided inquiry in three ways:

1. Students directed their own learning.
2. Students experienced first hand how scientists *do* science, they generated their own questions, designed their own experiments, analyzed (see Figures 16 and 17) their own data, and made their own knowledge claims using technology similar to that used by researchers in their laboratories. They also had to provide evidence to justify their claims.
3. Data sets generated from the research were made available on an internet website for other scientists to view and use contributing to the body of scientific knowledge.

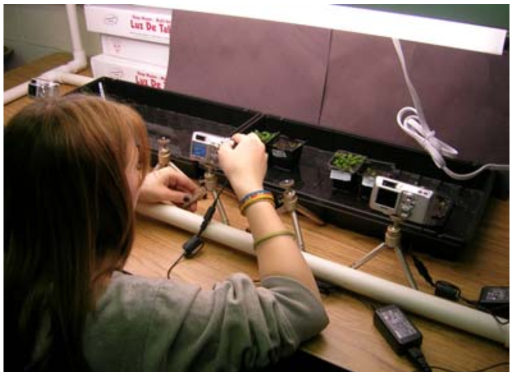

Figure 16 Student Setting Up a Time-lapse Sequence.

Figure 17 Students Making Observations of Emerging *Arabidopsis* Plants.

Unique to ASR learning is that scientific questions are generated by the student researchers that have no known answers. Students worked collaboratively as colleagues in answering their own questions. All contributions were relevant, which made the intervention meaningful, valid, and significant. The ASR intervention provided real-world application situated within the context of real science.

Intervention. Phase III - Guided Inquiry, Group A. Phase III of the study, Group A engaged in a final GI experience. In order to further guide students toward the correct function of the gene mutation in the det-mut and keeping in mind what students currently believed or hypothesized the function of the gene to be, I designed a GI experience in which the wild-type *Arabidopsis* plants and the det-mut plants were placed in total darkness for a period of 10 days. Time-lapse cameras were used to capture plant movement and growth of both types of plants in such a way that side-by-side comparisons could be made. Figure 15 shows a similar setup. The emphasis of this segment in the GI sequence was to continue to redirect students' thinking by highlighting the effects of light or lack of light on *Arabidopsis.* Data were compared in large-group discussions which I led. Students observed changes in the circadian rhythm of the wild-type similar to what they had observed in the det-mut. The det-mut exhibits very little circadian rhythm. Students noted that the *whirling* pattern of plant growth diminished the longer the wild-type plants were left in total darkness. Students observed that the det-mut plants appeared to have no reaction to changes in lighting. As in the preceding Phase, students recorded data in their journals, made observations, and answered open-ended questions about what they had observed and learned.

Intervention. Phase III - Authentic Student Research, Group B. Group B students determined what they wanted to investigate in Phase III. Students with similar questions were paired with each other. Groups were kept to no more than five members. Each small group worked collaboratively to refine their research question, determine their materials, design their experiment, collect and analyze their data, and prepare to present their conclusions in large-group class discussion. In addition, small groups communicated with the university scientist by e-mail and posted significant videos to a website set up by the scientist. In Phase III with Group B students, my role as the classroom teacher/researcher was that of a coach.

Instruments and Data Sources

Table 7 provides an overview of the assessment instruments and data sources used in this study. The table provides details of the data collected, by whom the data were recorded, and how this data was analyzed. Each of the assessment instruments is further explained in this section.

Researcher's Journal. The classroom teacher/researcher recorded observations of student interactions. In addition, changes in students' thinking which were verbalized through comments were then recorded in the journal. Data was typically updated at the end of each class. This journal served to primarily document what took place during each session of instruction.

Students' Journals. Students periodically responded to open-ended questions at the close of the class period. Students were asked to record responses to questions like, *What did you learn today?* or *How do scientists do science in the real world*? These journals also contained students' experimental data, observations, and conclusions. Students' responses to open-ended questions, data, observations, and conclusions provided insights into students' learning as they progressed through the Phases of their work with *Arabidopsis*.

Arabidopsis Drawing. Initial drawings of the *Arabidopsis* plant provided a baseline of students' observational skills and the level of detail students chose to represent in recording data. These drawings were compared to final drawings done at the conclusion of the study to evidence these changes. Drawing skills were emphasized because of the

Table 7

Assessment Instruments. Data Sources and Types of Analyses

Data Source- Rationale	Participating Group	Type of Analysis
Researcher's Journal (Maintain record from researcher perspective & log of occurrences)	Teacher/Researcher	Qualitative Quantitative
Student Journal (Scientific data, and reflections) (Record student learning as it occurs)	Student	Qualitative
Pre-Test – open ended (Baseline measure of pre-existing understanding)	Student	Quantitative (NAEP, 2004a)
Initial *Arabidopsis* Drawing (Baseline representative measure of student mental model of plant structures	Student	Qualitative Quantitative
Post –Test – open ended (Measure changes in students understanding of content & scientific practices)	Student	Quantitative (NAEP, 2004a)
Final *Arabidopsis* Drawing (Measure changes in representations of mental model of plant structures)	Student	Qualitative

importance of observation skills, accurate recording of data is essential to the scientific process.

Pre-Tests and Post-Tests. Appendix A displays a copy of the instrument used for both the pre-test and post-test that was developed to measure knowledge gains. The same instrument was used for both treatment groups and was administered at the beginning of the treatment sequence and again at the end of the sequence. Administration times of pre-test and post-test occurred approximately two months apart. In my deliberations for the design of the study, I felt that the time of the treatment was sufficient to allow the use of the same test for both pre-test and post-test administrations. I felt it was unnecessary for me to alter the tests in any way or to develop two independent tests to assure that the pre-test did not in some way skew the results of the post-test administrations. Both tests were administered over a one-week period to combat test exhaustion that might have occurred in a test where students are required to write their responses to eight separate questions in one class period.

Construction of questions followed a format similar to the National Assessment of Educational Progress (NAEP) framework. (See O'Sullivan & Weiss, 1999). The questions were designed, pilot tested, and revised to measure three domains of student learning: conceptual understanding of factual information about the role of genes and the environment on the phenotype (Questions 1-5), understanding of scientific investigation as a process (Questions 6 and 7), and practical reasoning skills (Question 8).

Scoring rubrics were used to score students' understanding in each of these domains and appear in Appendix B. The NAEP framework also guided the design of the rubrics. The rubric for each question consisted of 4 choices for evaluating the response. A score of zero was awarded when students made no attempt to answer the question. A score of "1" was awarded when the response was poor, unintelligible or wrong. A score of "2" was awarded when students' responses demonstrated partial understanding. A score of "3" was awarded when a student demonstrated complete understanding. Complete understanding was awarded when a response included all aspects of understanding in the domain.

I established content validity for the instrument in two ways. First, I aligned the questions with the state standards for eighth grade science, the *Texas Essential*

Knowledge and Skills. Second, I asked the members of my research team to review the questions for clarity, open-endedness, and thoroughness in the light of the *Arabidopsis* sequences I had planned for my students. Three educational and scientific researchers on the research team made suggestions for the improvement of questions before and after the administration of the pilot test. Reliability of the scoring rubric was established by using percent agreement between two reviewers. Eight randomly selected tests (with eight questions on each) were scored independently using the rubric. Both reviewers agreed on the scores for 56 of the 64 responses, resulting in an inter-rater reliability of 88 percent.

Analysis

Both qualitative and quantitative methods were used to analyze the data collected before, during, and after the completion of the two treatment groups, as follows:

Pre-tests and post-tests measuring students' understanding of science. Data sets were compared using quantitative techniques for matched student responses for each question on the pre- and post-tests. Differences in these responses, which had been evaluated by rubric, were determined using matched sets of question responses only. Small numbers of matched sets required non-parametric methods to test for statistical significance between pre-test and post-test responses for each treatment group, separately.

SPSS was used for all quantitative analyses. Wilcoxon signed-ranks tests were calculated for each question to reveal the distributions of positive, negative, and neutral gains for each question and to assess the statistical significance of the differences between pre-tests and post-tests. Confidence intervals were calculated and error box plots constructed to display means and standard deviations at the 95% level of confidence for each of the eight questions. Confidence intervals for each pre-test and post-test question in each treatment group, which had different numbers of matched pairs, were displayed next to each other for ease in visual comparison as no statistical tests were possible to determine whether the gains in understanding were indeed statistically significant.

Students' responses on Question 8 were also analyzed qualitatively by constant comparative methods to reveal differences in the depth and elaboration of students' responses and in their specific use of examples.

Arabidopsis Drawings. Data for which there were matched sets of *Arabidopsis* drawings before and after the treatments were also compared. The drawings were compared qualitatively for differences in accuracy and attention to detail.

Students' Journals. Students' responses to two journal entries were chosen for content analysis by constant comparison. Assignments of responses to categories emerging from the content analysis were used to tally responses and calculate frequency of occurrence in each of the treatment groups.

Synthesis

Research questions were used to synthesize the results of both quantitative and qualitative analyses. To answer Research Question 1, for example, results of the Wilcoxon signed-ranks tests and error box plots were used to describe the effects of the two forms of inquiry on conceptual understanding. The qualitative analyses, however, provided other types of data which could be used to support or dispute the quantitative findings. The actual mixing of the data to answer the research questions is explained in Chapter V.

Summary

The mixed methods employed in this study were chosen to collect both quantitative and qualitative data in order to provide a “tradeoff between breadth and depth” (Frechtling & Sharp, 1997, p. 3). Mixing methods is recommended by the National Science Foundation as being critical in telling the “important parts of a story” (Frechtling & Sharp, 1997, p. 1) that might otherwise be missed by employing only one technique or the other. In the light of the messy classroom context and the exploratory nature of this study, I felt completely comfortable in mixing methods. I left the paradigm wars behind for those more philosophically inclined. With practical outcomes in mind, I chose pragmatic methods to guide the design of this study. I approached the problem from the perspective of a classroom teacher wanting to know more about the types of inquiry options open to an eighth grade science teacher aspiring to make learning relevant and meaningful to her students.

This study combined both quantitative and qualitative analyses and methods to determine the gains students achieved in their conceptual understanding about the domain of genetics. Pre- and post-tests were administered by the classroom teacher/researcher. Students also recorded pre- and post-drawings of *Arabidopsis* plants in their journals. Student journals were used not only for data collection but also for recording responses to open-ended questions asked periodically throughout the course of the study. The teacher/researcher’s journal was updated at the end of each class period and was kept as a record of the study. Data from the students and the teacher/researcher were triangulated to test congruence in the findings and reveal insights not apparent in pre- and post-measures.

A three-phase approach was used to prepare students unaccustomed to inquiry learning for a more independent *minds-on* learning environment. The first two phases were heavily scaffolded by the teacher/researcher for three reasons:

1. To make visible the metacognitive processes used in scientific investigation, essential to practical reasoning (Bruer, 1999).
2. To make visible the iterative and social nature inherent in authentic scientific investigation. These processes include critical questioning, reflection, revision, evaluation, discussion, data recording, and justifications.

3. To emphasize the need for claims to be backed by evidence as well as the importance of repeated experimentation to ensure validity and enhance credibility of the findings.

The final phase, Phase III, differed between groups A & B in the level of student centeredness. One group continued with heavily scaffolded guided inquiry while the other group engaged in student-centered scientific research in which the students determined and directed all aspects of their investigation. For the first group the teacher/researcher guided the inquiry. In the second group the teacher/researcher provided assistance only when asked; serving more as a coach or assistant than a *teacher*.

Finally, the assessment instruments used in this study were evaluated using a NAEP framework. NAEP rubrics guided the design of evaluation rubrics for this study and were further evaluated by experienced science education researchers. They were piloted and revised to ensure both the reliability and validity.

With these instruments, intervention protocols, and analyses, data were collected to answer three research questions regarding the differential effects of two inquiry-based instructional methods on student outcomes.

CHAPTER IV

RESULTS AND DISCUSSION

Introduction

> Classroom life is synergistic: Aspects of it that are often treated independently, such as teacher training, curriculum selection, testing, and so forth actually form part of a systemic whole. Just as it is impossible to change one aspect of the system without creating perturbations in others, so too it is difficult to study any one aspect independently from the whole operating system. Thus, we are responsible for simultaneous change in the system, concerning the role of students and teachers, the type of curriculum, the place of technology, and so forth. These are all seen as inputs into the working whole (Brown, 1992 pp. 142-143).

Pre-test and post-tests were administered in sections over five days according to the schedule presented in Chapter III. Students were often absent on one or more of the days of test administration, resulting in very small numbers of students who completed all questions on pre-tests and post-tests. Matched pre-test and post-test questions for the guided inquiry (GI) group ranged from 21 to 24 students, while matched pre-test and post-test questions for the authentic student research (ASR) group ranged from 13 to 30 students. Table 8 summarizes the numbers of students by group for which matched pre-test and post-test answers were available. Non-parametric analyses were conducted on individual test questions for which there were "matched" answers on both the pre-test and the post-test.

Table 8
Matched Pre-test and Post-test Questions Available for Comparison and Analysis

Group	Class	Numbers of Matched Questions for Pre- and Post-Test Comparisons							
		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
GI	1	7	9	9	9	9	7	7	9
GI	2	14	14	14	14	12	14	14	15
GI	3	0	0	0	0	0	0	0	0
Total GI		21	23	23	23	21	21	21	24
ASR	4	13	13	13	0	0	13	13	11
ASR	5	0	21	21	19	19	0	0	21
ASR	6	0	0	0	0	0	0	0	0
Total ASR		13	34	34	19	19	13	13	32

Non-Parametric Methods for the Analysis of Pre-tests and Post-tests

Wilcoxon signed-ranks tests were used (as the nonparametric alternative to the paired t-test) to compare "before" and "after" measures for each question for each group. SPSS was used to (a) calculate the absolute difference between each pair; (b) rank the absolute differences from smallest to largest, employing tied ranks where appropriate; and (c) assign to each rank a "+" sign when the pre-test mean was less than the post-test mean, a

“-” when the pre-test mean was greater than the post-test mean, and a “tie” when the means were the same. Mean ranks were then calculated for the numbers of “+,” “-,” and “tie” responses. Z-ratios and their associated one-tail probabilities were calculated to determine whether the differences between pre-test and post-test measures were significant.

Descriptive statistics were used to summarize the distribution of data for each pre-test and post-test question. Data regarding each question were reported in the form of a plot with “error bars” on either side of the mean statistic. The size of the error bar was used to display an estimate of the limits at two standard deviations, or 95 percent, for the student response data for that question. Error box plots were compared for each group to visualize (a) differences in performance between pre-test and post-test measures for each question within each group, (b) differences in performance between the two groups on pre-test and post-test measures, and (c) differences in variances between the two groups on all measures.

Methods for the Analysis of Students’ Journal Responses

Students were asked to answer several open-ended questions in their journals periodically throughout the course of the study. Student responses to the following questions were coded, clustered, and categorized: *What have you learned from this project?* and *What have you learned about how scientists in the real world do research?* Both questions were intentionally designed to be vague to ensure that students were not cued toward a specific answer. The first question was asked at the conclusion of the inquiry for both groups; and the second question was asked approximately mid-way through Phase III. Numbers of each type of response were counted and percent occurrence was calculated. Results from this analysis were used to supplement information provided by the pre-test and post-test comparisons in order to more fully understand the differences in student learning between the two groups.

Methods for the Analysis of Student Drawing Samples

Student drawings were used to collect non-verbal data about students' understanding of plant features. Students in both groups were asked to draw their observations of an *Arabidopsis* plant before and at the conclusion of the study. Matched sets of students' drawings were visually examined for detail, accuracy, and methods of representation.

Results

The rest of this chapter is organized to present the results of the analyses, of the pre-test and post-test questions, students' journal entries, and their drawings of *Arabidopsis* plants. All analyses compared students' work by treatment group so that data from GI and ASR groups were pooled to reveal patterns in learning within each group. Results are reported in three sections. The first section contains pre-test and post-test analyses for each of the questions, including the results of the Wilcoxon sign-ranks tests and error box plots. Conceptual understanding questions are presented first (Questions 1-5), following by the two questions that addressed students' perceptions of scientific investigation (Questions 6 and 7), and finally, one question formatted as a scenario to reveal students' abilities to practically reason (Question 8). The second section of this chapter reports the frequencies of students' journal responses to the two questions selected to provide information about students' reflections on their learning. The third section describes the results of visual comparisons of students' drawings of an *Arabidopsis* plant before and after the study was completed.

Analysis of Pre-tests and Post-tests. A rubric following the National Assessment of Educational Progress (NAEP; O'Sullivan & Weiss, 1999) was constructed to evaluate student responses to each question on the pre-test and post-test. Appendix B includes rubrics for each of the eight questions. Each question followed a similar scoring pattern, which evaluated students' responses on a score from zero to three. A score of zero indicated "no response" to the question; "one" indicated responses that were poor, unintelligible, or wrong; "two" indicated partial understanding, and "three" indicated complete understanding. Inter-rater reliability was established at 88% agreement between two independent raters who used the rubric to score eight student responses on eight tests.

Means and standard deviations for each question were calculated in order to construct error box plots for each question by group; and the distribution of scores were rank ordered to test significance using the Wilcoxon signed-ranks test.

Question 1: Can the Environment Have Any Effect on a Plant's Genes?

This question was crafted to reveal what students understood about the relationship between genes and the environment, more specifically, to reveal whether the students understand that the environment can affect genes. Matched data sets available for this question numbered 13 from the ASR group and 21 from the GI group. The Wilcoxon signed-ranks test analysis revealed that students in the GI group did not show a statistically significant change in their conceptual understanding of the environmental effects on plant genes ($z = -4.06$, $p = .684$). Analysis of the ASR group revealed statistically significant gains for this question ($z = -2.021$, $p = .043$). Tables 9 and 10 summarize these analyses for Question One.

Error box plots using a 95% confidence interval supported these findings and further revealed reduced variance in the GI groups' pre-test to post-test responses (See Figure 18). There was no apparent reduction in variance for the ASR group. Error box plots also showed that students' scores in the GI group remained at unsatisfactory levels (at about 1.0) based on the rubric design. The ASR groups' mean scores increased from below unsatisfactory (near 0.4) on pre-tests to post-tests that approached partial understanding (at about 1.3) for this question. Mean test scores on pre-test and post-tests for the ASR group also indicated increased gains in conceptual understanding. Visual inspection of the box plots revealed that pre-test scores were lower for the ASR group with higher means of post-test scores for the ASR group, further indicating that ASR students made stronger gains than GI students in their conceptual understanding of

Question One. The ASR group showed a small overlap of pre-test scores and post-test scores, the GI group showed complete overlap, indicating that most ASR students improved their conceptual understanding while the GI students did not.

Overlaps shown in the box plots were clarified by Wilcoxon signed-ranks tests, which showed that of the 13 students in the ASR group, ten improved their scores, one scored the same, and two actually scored lower on the post-test than the pre-test (see Table 10). Of the 21 GI students, eight students scored higher, seven scored the same and six students scored lower on the post-test (see Table 9).

Table 9

GI Group Question 1. Can the Environment Have Any Effect on a Plant's Genes?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	6	7.75	46.50
	Positive Ranks	8	7.31	58.50
	Ties	7		
	Total	21		

Table 10

ASR Group Question 1. Can the Environment Have Any Effect on a Plant's Genes?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	2	7.00	14.00
	Positive Ranks	10	6.40	64.00
	Ties	1		
	Total	13		

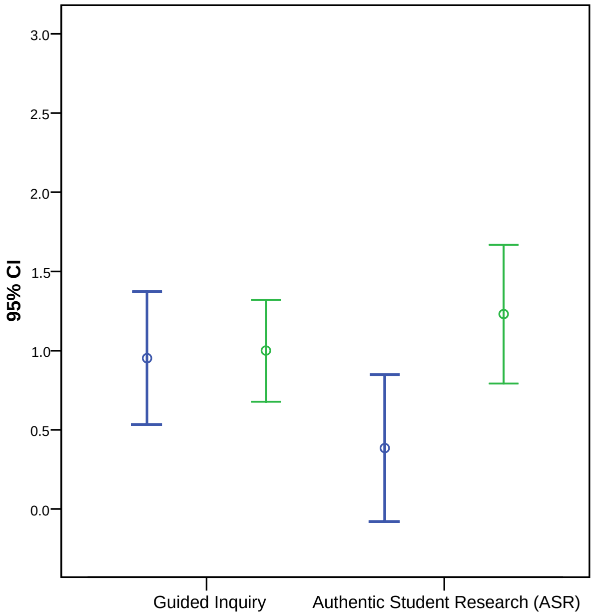

Figure 18 Error Box Plot Question 1. Can the Environment Have Any Effect on a Plant's Genes? Pre-test results precede post-test results for each group.

Question 2: What Do You Know About genes?

This question was intended to reveal what students understand about genes. The vagueness of the question was intentional to avoid cueing responses. Thirty-four matched data sets from the ASR group and 23 matched data sets from the GI group were used in the analyses of this question. The Wilcoxon signed-ranks test analysis revealed that both groups showed a statistically significant change in their conceptual understanding of what a gene is (ASR z = -1.66, p = .096 and GI z = -2.230, p = .026). Tables 11 and 12 and Figure 19 summarize the analyses for Question Two.

Table 11
GI Group Question 2. What Do You Know About Genes?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	2	4.50	9.00
	Positive Ranks	9	6.33	57.00
	Ties	12		
	Total	23		

Table 12
ASR Group Question 2. What Do You Know About Genes?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	5	6.50	32.50
	Positive Ranks	10	8.75	87.50
	Ties	19		
	Total	34		

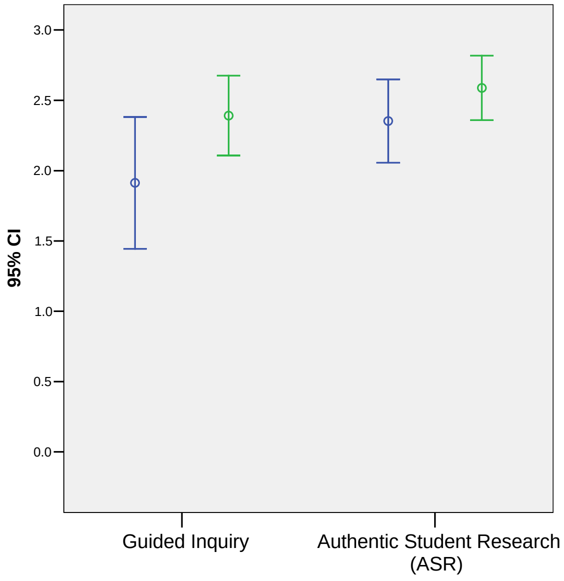

Figure. 19 Error Box Plot Question 2. What Do You Know About Genes?
Pre-test results precede post-test results for each group.

Error box plots using a 95% confidence interval supported these findings and further revealed reduced variance in both groups' pre-test to post-test responses (See Figure 19). Box plots further revealed a greater range of responses for the GI group (ranging from approximately 1.4 to 2.4) than the ASR group (ranging from approximately 1.9 to 2.5) on pre-tests. Error box plots also showed that students' mean scores in the GI group remained at levels of partial understanding (about 1.9 on pre-tests and near 2.3 on post-tests). The ASR groups' mean scores also remained at levels of partial understanding. Mean scores for the ASR group ranged from pre-test values near 2.3 to post-tests scores near 2.5. While responses for both groups achieved a level of partial understanding, the post-test scores for the ASR group are closer to approaching complete understanding. Mean test scores on pre-test and post-tests for the GI group indicated higher gains in conceptual understanding over the ASR group. Both groups showed considerable overlap in pre-test and post-test scores.

Overlaps shown in the box plots were clarified by Wilcoxon signed-ranks tests which showed that of the 34 students in the ASR group, ten students improved their scores, 19 showed no change, and five scored lower on the post-test. Of the 23 matched data sets for the GI group, nine students showed improvement, twelve showed no change, and two students scored lower on the post-test.

Question 3: What Do You Know About Mutations?

This question was intended to reveal what students understand about mutations. This question had matched data sets for 34 students from the ASR group and 23 matched sets from students from the GI group. The Wilcoxon signed-ranks test analysis showed that both groups had statistically significant differences in conceptual understanding of what a mutation is and how they occur (ASR z = –3.581, p = .000 and GI z = -2.183, p = .029). Tables 13 and 14 summarize the analyses for Question Three.

Error box plots using a 95% confidence interval supported these findings and further revealed reduced variance in the GI groups' pre-test to post-test responses (See Figure 20). There was no apparent reduction in variance for the ASR group. Error box plots also showed that students' scores in the GI group remained at unsatisfactory levels (at about 1.0). The ASR group mean scores increased from below unsatisfactory (near 0.4) on pre-tests to levels of partial understanding (at about 1.3) on post-tests.

Mean test scores on pre-test and post-tests for the ASR group also indicated increased gains in conceptual understanding over the GI group. Visual inspection of the box plots revealed that pre-test scores were lower for the ASR group than for the GI group and that the mean of post-test scores was higher for the ASR group. These findings further indicate that ASR students made stronger gains than GI students in their conceptual understanding of Question Three. The ASR group showed only a slight overlap, of pre-test scores and post-test scores while the GI group showed complete

Table 13
GI Group Question 3. What Do You Know About Mutations?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	3	4.00	12.00
	Positive Ranks	9	7.33	66.00
	Ties	11		
	Total	23		

Table 14
ASR Group Question 3. What Do You Know About Mutations?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	4	10.88	43.50
	Positive Ranks	23	14.54	334.50
	Ties	7		
	Total	34		

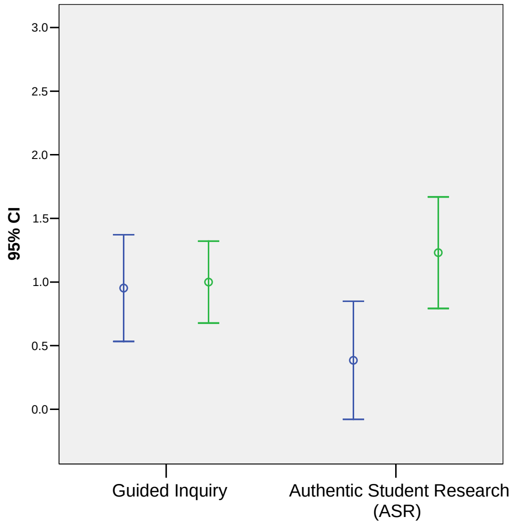

Figure 20 Error Box Plot Question 3. What Do You Know About Mutations?
Pre-test results precede post-test results for each group.

overlap indicating that more students in the ASR group improved their conceptual understanding than those in the GI group.

When taken together, overlaps shown in the box plots were clarified by Wilcoxon signed-ranks tests. Of the 34 matched data sets for the ASR group, 23 students showed improved scores, seven had no change, and four students scored lower on the post-test (see Table 14). In the GI group, nine students improved, 11 had no change and three students scored lower on post-tests (see Table 13).

Question 4: Why Do Plants Look the Way They Do?

This question was designed to uncover students' knowledge about the role of genes as well as environmental factors on phenotype. The vagueness of the question was intentional to prevent cueing in the response. For this question, matched data sets for the ASR group totaled 19; matched sets for the GI group totaled 23. Wilcoxon signed-ranks tests showed statistically significant differences in pre-test and post-test scores in both groups (ASR $z = -2.236$, $p = .025$ and GI $z = -3.082$, $p = .002$). Tables 15 and 16 summarize the analyses for Question Four.

Error box plots using a 95% confidence interval supported these findings and further revealed that the ASR group had a greater reduction in variance from pre-test to post-test than did the GI group (See Figure 21). Error box plots showed that students' mean

Table 15
GI Group Question 4. Why Do Plants Look the Way They Do?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	1	5.50	5.50
	Positive Ranks	13	7.65	99.50
	Ties	9		
	Total	23		

Table 16
ASR Group Question 4. Why Do Plants Look the Way They Do?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	3	6.50	19.50
	Positive Ranks	11	7.77	85.50
	Ties	5		
	Total	19		

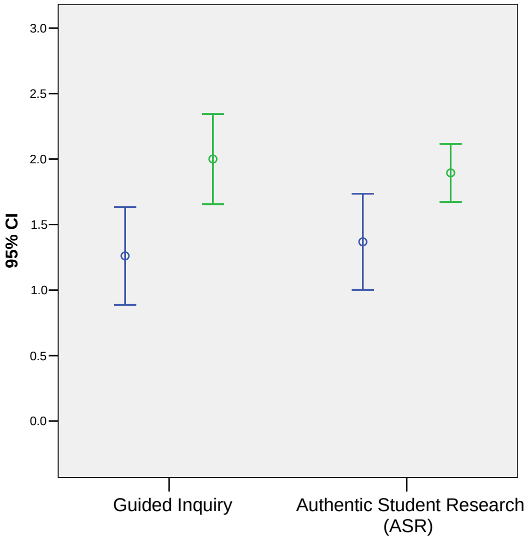

Figure 21 Error Box Plot Question 4. Why Do Plants Look the Way They Do?
Pre-test results precede post-test results for each group.

scores in the GI group improved from unsatisfactory levels (at about 1.2) to levels of partial understanding (at about 2.0). The ASR groups mean scores increased similarly from unsatisfactory levels on pre-tests (near 1.3) to post-tests approaching partial understanding (at about 1.9). Mean test scores on pre-test and post-tests for the GI group also indicated slightly increased gains in conceptual understanding over ASR group. Visual inspection of the box plots revealed that pre-test scores were similar for both the GI and ASR groups and that post-test scores for both groups showed only a small overlap from pre-test to post-test scores.

Overlaps shown in the box plots were clarified by Wilcoxon signed-ranks tests, which showed that of the 19 students in the ASR group, 11 improved their score, five scored the same, and three scored lower on the post-test (see Table 16). Of the 23 GI students, 13 students scored higher, nine scored the same, and one student scored lower on the post-test (see Table 15).

Question 5: Do Plants Move? Do They Have Patterns of Movement?

The purpose of this question was to reveal what students understood about plant movement. In other words, did students possess knowledge of patterns of plant movement beyond environmental influences such as air currents? The ASR group had 19 matched data sets for this question and the GI group had matched data sets totaling 21. Wilcoxon signed-ranks tests showed statistically significant gains in students understanding both groups (ASR z = -3.142, p = .001 and GI z = – 3.879, p = .000). Tables 17 and 18 and Figure 22 summarize the analyses for Question Five

Error box plots using a 95% confidence interval supported these findings and further revealed that the ASR group had a greater reduction in variance on post-tests. All students obtained partial understanding in the ASR group (See Figure 22). Error box plots also showed an improvement of students' mean scores in both groups from unsatisfactory levels (at about 1.1 and 1.2, respectively) to levels of partial understanding (at about 2.0 for both groups). Mean test scores on pre-test and post-tests for both groups revealed similar gains in conceptual understanding for this question.

Visual inspection of the box plots revealed that both the pre-test and post-test scores were similar for both groups. Similarly, both groups showed no overlap in scores from pre-test to post-test scores, indicating improvement in students' understanding of plant movement (See Figure 22).

Table 17
GI Group Question 5. Do Plants Move? Do They Have Patterns of Movement?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	0	.00	.00
	Positive Ranks	17	9.00	153.00
	Ties	4		
	Total	21		

Table 18
ASR Group Question 5. Do Plants Move? Do They Have Patterns of Movement?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	0	.00	.00
	Positive Ranks	13	7.00	91.00
	Ties	6		
	Total	19		

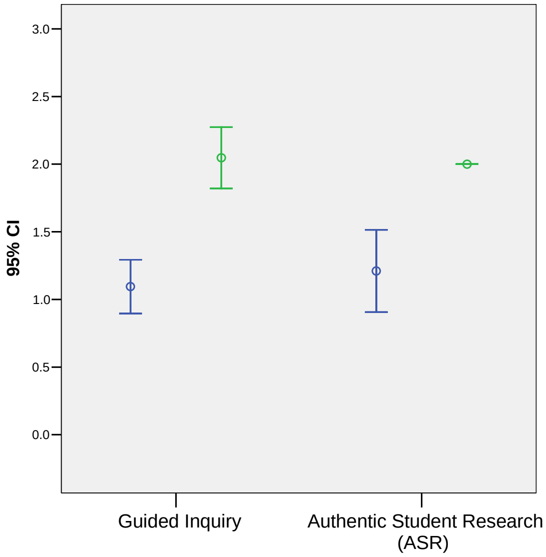

Figure 22 Error Box Plot Question 5. Do Plants Move? Do They Have Patterns of Movement?
Pre-test results precede post-test results for each group.

The variances indicated by box plots were clarified by Wilcoxon signed-ranks tests, which showed that of the 19 sets in the ASR group, 13 improved their understanding, six demonstrated no change, and no students scored lower on post-tests (see Table 18). Of the 21 matched data sets in the GI group, 17 students improved their scores, four showed no change and no student scored lower on post-tests (see Table 17).

The data does not allow an inference that one inquiry form allowed students to perform better than the other. Both treatments were effective in improving students' conceptual understanding of plant movement.

Question 6: Does an Experiment Have to Prove the Hypothesis Correct to Be a Good Experiment?

This question was designed with the intention of revealing what students knew about the relationship between the question (hypothesis) and the data collected. Students often associate experimental outcomes as failures if their hypotheses or questions are not supported by the data. They seldom see the benefit of unexpected outcomes. Measuring change in students' understanding of authentic scientific investigation as a process was seen as a critical indicator. For this question, matched data sets for the ASR group included 13 students; the GI group included matched sets for 21 students. Results of Wilcoxon signed-ranks test indicated that only the GI group had a statistically significant difference in conceptual understanding of the relationship between the hypothesis and data (ASR $z = -1.613$, $p = .107$ and GI $z = –2.236$, $p = .025$). Tables 19 and 20 summarize the analyses for Question Six.

Error box plots using a 95% confidence interval supported these findings and further revealed that the ASR group had a greater pre-test variance than the GI group. Both groups showed a reduction in variance in post-tests. Further analyses showed the GI group pre-test mean scores were below unsatisfactory (near 0.7), while the pre-test mean scores (about 1.2) were at the level of unsatisfactory for the ASR group (see Figure 23). Error box plots also showed that students' mean scores in both the GI and ASR groups improved to levels approaching partial understanding (about 1.2 and 1.7, respectively).

Visual comparison of mean test scores on pre-test and post-tests for both groups revealed similar gains in conceptual understanding for this question. Box plots revealed that even though there were similar gains in both groups, the large variance and substantial overlap in ASR scores masked any significant difference for this group. Visual inspection of mean scores further revealed that the levels of conceptual understanding were higher for the ASR group. Post-test scores for the GI group were similar to the pre-test scores for the ASR group (1.3 and 1.2, respectively). Both groups showed considerable overlap from pre-test and post-test scores (See Figure 23).

Variances indicated by box plots were clarified in the Wilcoxon signed-ranks test analysis, which showed that of the 13 matched data sets for the ASR group, six students showed improved scores, five showed no change, and two scored lower on the post-tests (see Table 20). The 21 matched data sets from the GI group showed that 11 students improved, seven had no change, and three students scored lower on the post-test (see Table 19).

Table 19
GI Group Question 6. Does An Experiment Have to Prove the Hypothesis Correct to Be a Good Experiment?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	3	6.50	19.50
	Positive Ranks	11	7.77	85.50
	Ties	7		
	Total	21		

Table20
ASR Group Question 6. Does An Experiment Have to Prove the Hypothesis Correct to Be a Good Experiment?

		N	Mean Rank	Sum of Ranks
Pre – Post Differences	Negative Ranks	2	3.50	7.00
	Positive Ranks	6	4.83	29.00
	Ties	5		
	Total	13		

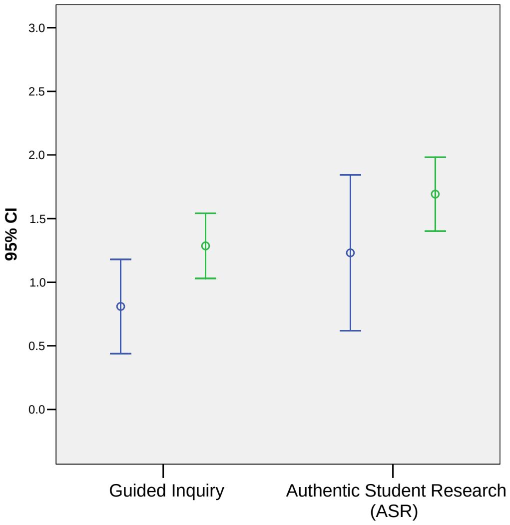

Figure 23 Error Box Plot Question 6. Does An Experiment Have to Prove the Hypothesis Correct to Be a Good Experiment?
Pre-test results precede post-test results for each group.

Question 7: How Do Scientists Go About Answering Scientific Questions?

This open-ended question was designed to reveal students' knowledge about the processes that scientists use to answer scientific questions. The wording of this question was intentionally general to prevent cuing responses in the hopes that students' misconceptions about the scientific process would be made visible. Matched data sets for the ASR group numbered 13 and 21 matched data sets for the GI group. Wilcoxon signed-ranks test showed statistically significant differences for both groups in their understanding (ASR z = -2.889, p = .004 and GI z = -3.291, p = .001). Tables 21 and 22 summarize the analyses for this question.

Error box plots using a 95% confidence interval supported these findings and further revealed that the GI group pre-test mean scores were below unsatisfactory (near 0.5) and that ASR pre-test mean scores at levels were also below unsatisfactory (about 0.3, see Figure 24). Error box plots showed that students' mean scores in both groups improved to levels approaching partial understanding (about 1.5 for both groups).

Visual comparison of mean test scores on pre-test and post-tests for both groups revealed slightly greater gains in conceptual understanding for the ASR group on this question. Box plots revealed no overlap in pre-test to post-tests scores for both groups, indicating that all students improved in their understanding of how scientists go about answering scientific questions.

Table 21

GI Group Question 7. How Do Scientists Go About Answering Scientific Questions?

		N	Mean Rank	Sum of Ranks
Pre to Post Difference	Negative Ranks	2	4.50	9.00
	Positive Ranks	15	9.60	144.00
	Ties	4		
	Total	21		

Table 22

ASR Group Question 7. How Do Scientists Go About Answering Scientific Questions?

		N	Mean Rank	Sum of Ranks
Pre to Post Difference	Negative Ranks	0	.00	.00
	Positive Ranks	10	5.50	55.00
	Ties	3		
	Total	13		

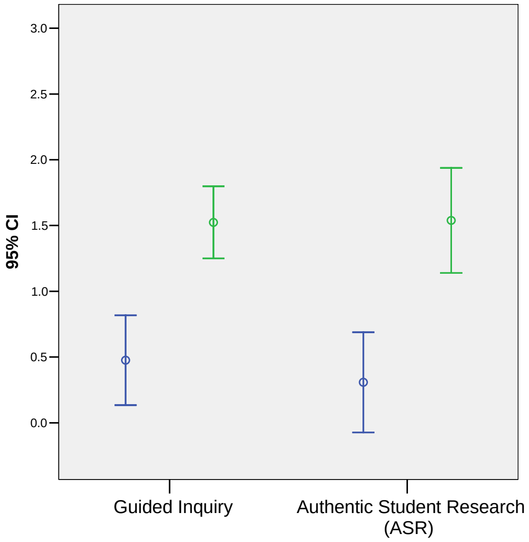

Figure 24 Error Box Plot Question 7. How Do Scientists Go About Answering Scientific Questions? Pre-test results precede post-test results for each group.

Variances shown in box plots were clarified in the analysis using the Wilcoxon signed-ranks test, which showed that of the 13 matched data sets for the ASR group, ten students improved their scores, three showed no change, and no students scored lower on the post-test (see Table 22). For the 21 matched data sets from the GI group, 15 students showed improved scores, four students scored the same, and two students scored lower on the post-test (see Table 21).

Question 8: Read the Story Below, and Then Answer the Question.

> A scientist has several pots of the same kind of plant in his office. They were spread all over the room to make the room pretty. He noticed that some of these plants didn't grow as tall as others, even though they were all planted at the same time and were the same type of plant. He wanted to know why some of his plants grew taller than others. He couldn't find any information to answer the question anywhere. How do you think the scientist would go about answering the question?

This question asked students to explain how a scientist would go about solving the problem specified in the stem of the question. The question was similar to what the students had been doing throughout the study, in that students were prompted to explain a design of an experiment to find out why some plants were doing well and others were not. The purpose of the scenario was to allow students to use practical reasoning to connect what they had learned about the effects of the environment and genes on plants with what they understood about scientific investigation as a process.

Matched data sets for this question consisted of 32 sets from the ASR group and 24 sets from the GI group. Both groups showed statistically significant gains in their use of practical reasoning skills (ASR z = -3.967, p = .000 and GI z = -3.667, p = .000). See Table 23 and 24.

Error box plots using a 95% confidence interval supported these findings and further revealed that all students in both groups made gains in their practical reasoning ability. Further analysis of error box plots showed the GI group pre-test mean scores were below unsatisfactory (near 0.7). The ASR group pre-tests mean scores were higher than the GI group but still at unsatisfactory levels (about 1.6). Figure 25 summarize the analyses for this question. Error box plots also showed student mean scores in both groups improved to levels approaching partial understanding (about 1.9 for both groups). Visual comparison of mean test scores on pre-test and post-tests for both groups revealed greater gains in practical reasoning for the GI group on this question. Box plots revealed no overlap in pre-test to post-test scores for both groups, indicating that all students improved scores for this question.

Variances indicated in box plots were clarified in the analysis using the Wilcoxon signed-ranks tests to reveal that of the 32 matched data sets for the ASR group, 22 students scored higher, eight students showed no change, and two students scored lower

Table 23
GI Group   Question 8.   Practical Reasoning Question

		N	Mean Rank	Sum of Ranks
Pre - Post Difference	Negative Ranks	2	6.50	13.00
	Positive Ranks	19	11.47	218.00
	Ties	3		
	Total	24		

Table 24
ASR Group   Question 8.   Practical Reasoning Question

		N	Mean Rank	Sum of Ranks
Pre – Post Difference	Negative Ranks	2	10.50	21.00
	Positive Ranks	22	12.68	279.00
	Ties	8		
	Total	32		

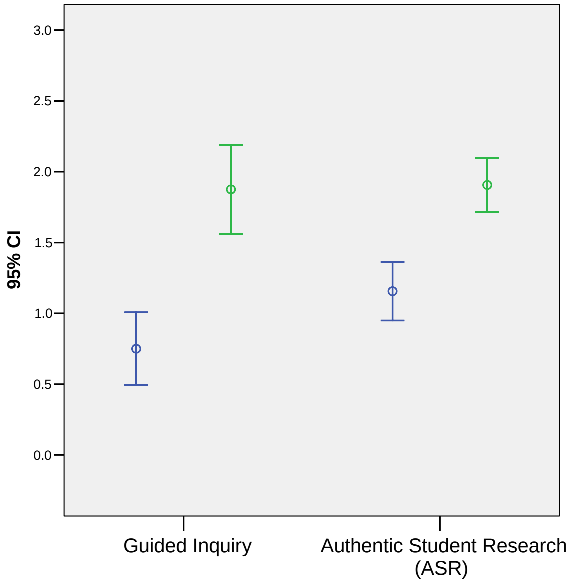

Figure 25   Error Box Plot Question 8.   Practical Reasoning Question.
Pre-test results precede post-test results for each group.

on post-tests than they did on pre-tests. Of the 24 matched data sets for the GI group, 19 students scored higher, three students showed no change in scores, and two students scored lower on post-tests. These data suggest that both treatments were similarly effective in getting students to flexibly apply their understanding of scientific investigation, but I felt that students' actual responses held more information that could be uncovered by an analysis of the content residing within the students' responses.

Content analysis of the responses to the question revealed that the ASR group provided nearly three times as many elaborations expanding on a point (see Table 25).

Table 25
Frequency Count of Responses to Practical Reasoning Question

	GI Post-Test (N = 38)		ASR Post-Test (N = 50)	
Student Responses	Counts	Percents	Counts	Percents
Elaboration, expanding a point	6	15.8%	22	44.0%
Elaboration, providing justification	1	2.6%	10	20.0%
Examples	2	5.3%	2	4.0%
Incorporate concepts from the study	8	21.1%	5	10.0%
No elaboration, examples, or incorporation of concepts	21	55.3%	11	22.0%
Total	38	100.0%	50	100.0%

Student responses included elaborations on hypotheses such as, "… then make your hypothesis about what you think…" or elaborations on experimental setup like, "He could set up different experiments such as putting some in controlled light, some in constant…" Elaborations that provided justification for points made were also made three times as often in the ASR group as they were in the GI group. A sample justifications from the ASR group was, "…put the other plants that don't get sunlight in the window…because plants deeper in the room don't get any sun." Content analysis further revealed that both groups referenced knowledge from direct experiences in the experiments such as time-lapse photography, controlling lighting, and genetic traits of the plants. There appeared to be no difference in the number of times examples were used between groups.

Frequency Counts from Students' Journals

Responses to the first question, "*What have you learned from this project?*" were clustered into four categories: scientific investigation, plant features, learning in general, and domain specific content. The total responses for each of the four main groups are shown below (see Table 26). Some students recorded more than one thing learned so the total number of GI and ASR group responses does not reflect the number of students, but the number of items recorded.

Categorization of this data revealed that the GI group recorded almost five times as many responses that addressed domain content knowledge (ASR = 6 and GI = 27). Further, no students from the ASR group recorded responses about environmental effects on plants, while 18 of the 73 GI responses address environmental impact on plants. Responses were high in both groups for the category of plant features (ASR = 34 and GI = 35). The ASR group recorded almost twice as many responses pertaining to scientific investigation (ASR = 27 and GI = 14) as well as aspects about learning in general, such as “learning is hard” and “I learned a lot” (ASR = 11 and GI = 5).

Table 26
Question: What I Learned from This Project

	GI (N=72)		ASR (N=78)	
Categories	N	Percent	N	Percent
Scientific Investigation				
How to make observations	6	8.3%	3	3.8%
How to take time-lapse	3	4.2%	9	11.5%
How to grow plants	0	0.0%	6	7.7%
Research takes a lot of hard work and thinking	0	0.0%	5	6.4%
How to conduct experiments	5	6.9%	4	5.1%
Category Total	14	19.4%	27	34.5%
Plant Features				
Hair-like structures (trichomes)	7	9.7%	10	12.8%
Circadian rhythm	3	4.2%	6	7.7%
How plants grow	25	34.7%	18	23.1%
Category Total	35	48.6%	34	43.6%
General Learning				
I learned a lot	5	6.9%	7	9.0%
This was a fun-interesting way to learn	0	0.0%	4	5.1%
Category Total	5	6.9%	11	14.1%
Domain Content				
Genes and mutations	9	12.5%	6	7.7%
Environmental effects on plants	9	12.5%	0	0.0%
Category Total	18	25.0%	6	7.7%
Grand Totals	72	99.9%	78	99.9%

Some general trends became apparent within the four categorizes of responses. The first category addressed responses tied to a new understanding of scientific investigation.

The GI group tended to record responses that dealt with the “how” of conducting research, such as “how” to make observations, “how” to take time-lapse videos and “how” to conduct experiments. The ASR group had similar responses but also included the culture of scientific investigation as shown in responses like, “research is hard work”, and “research makes you think”. The most common response for both groups was a new understanding for how plants grow (ASR = 18 and GI = 25).

Student responses to the second question, “What have you learned about how scientists in the real world do research?” showed that both groups displayed an accurate understanding of authentic science as done by scientists (See Table 27). Low numbers of responses were the result of attrition due to remediation pull outs, student absenteeism and school sponsored incentive trips.

Table 27
Question: What Have You Learned About How Scientists in the Real World Do Research?

	GI (N=15)		ASR (N = 23)	
Response	N	Percents	N	Percents
Scientists make lots of observations	1	6.7%	0	0.0%
Scientists make repeated trials	6	40.0%	8	34.8%
Scientists communicate and work with each other	1	6.7%	6	26.1%
Scientists sometimes get different results/ sometimes things go wrong	2	13.3%	2	8.7%
Research is hard work	4	26.7%	5	21.7%
Scientists don’t always know what they are looking for	1	6.7%	0	0.0%
Experimentation takes a long time	0	0.0%	2	8.7%
Total	15	100%	23	100.0%

Responses numbered 15 for the ASR group and 23 for the GI group. The most common response to this question for both groups was a new understanding that scientists repeat their experiments (ASR = 8 and GI = 6). The next most common response for both groups was that “research is hard work” (ASR = 5 and GI = 4). Interestingly, the ASR group recorded more responses that addressed communication with other scientists as part of what scientists do. Beyond this, both groups recorded similar responses (ASR = 6 and GI = 1).

Student Drawing Samples

This section provides examples of students’ initial and final drawings of the model plant, *Arabidopsis.* Students in both groups were asked to draw *Arabidopsis* while observing the plant before and at the conclusion of the study. Eight samples, two matched sets from each group, are provided below (Figures 26 - 33). Visual inspection of matched sets of students’ drawings revealed no distinguishable difference between the representations of *Arabidopsis* between groups.

Overall, students showed marked improvement in the level of detail and accuracy in their drawings. An interesting phenomenon found in both groups was that some students chose to attach a sample of *Arabidopsis* beside their drawing. This did not occur with any of the initial drawings.

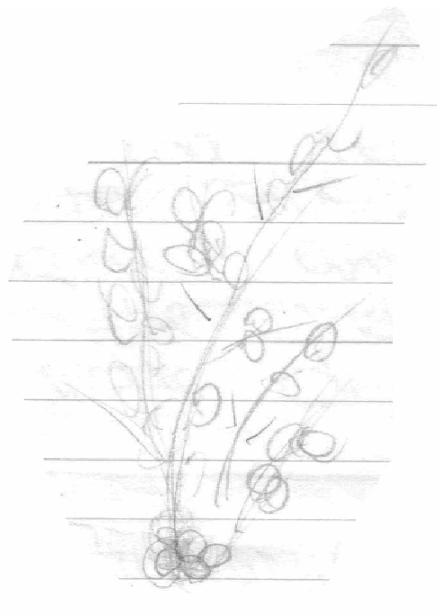

Figure 26 ASR Group Student Sample
Initial Drawing of *Arabidopsis* I

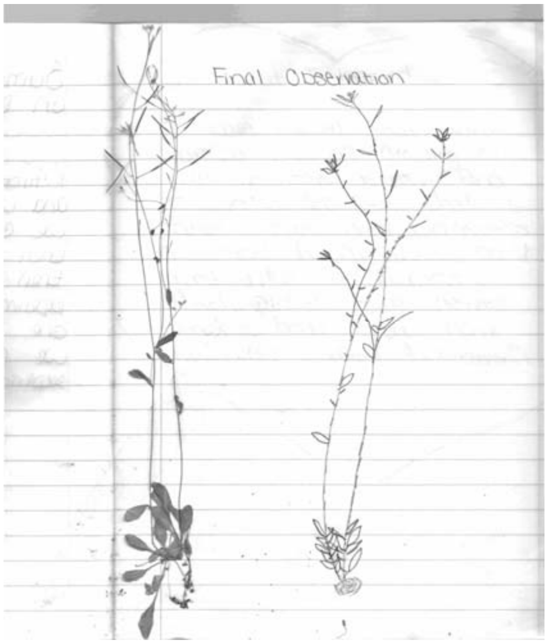

Figure 27 ASR Group Student Sample
Final Drawing of *Arabidopsis* I

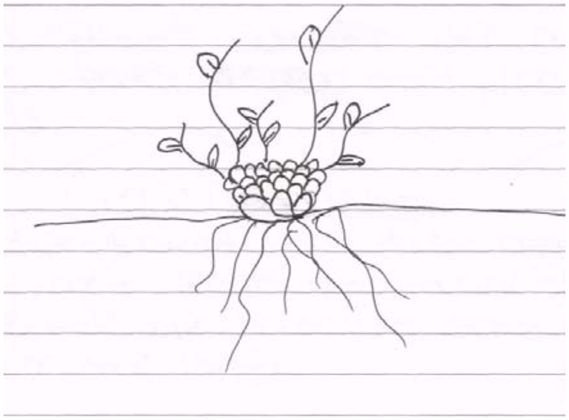

Figure 28 ASR Group Student Sample
Initial Drawing of *Arabidopsis* II

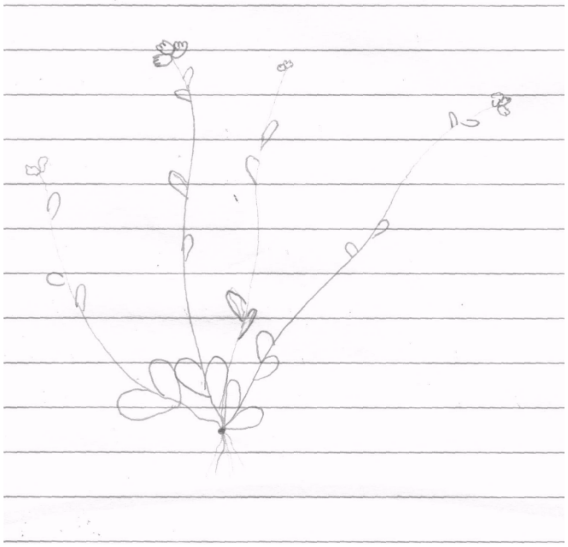

Figure 29 ASR Group Student Sample
Final Drawing of *Arabidopsis* II

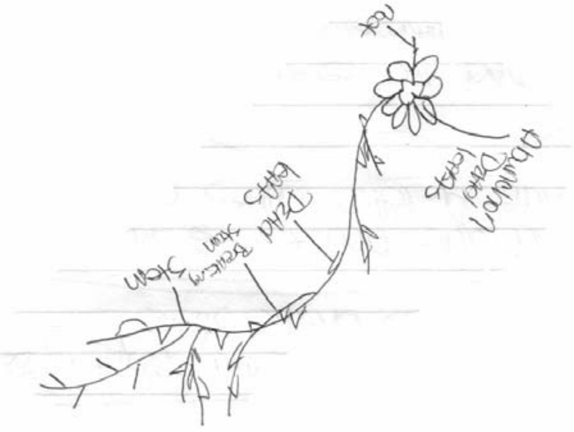

Figure 30 GI Group Student Sample
Initial Drawing of *Arabidopsis* I

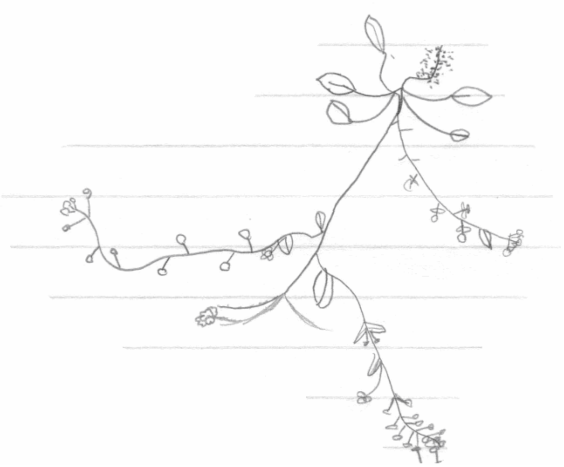

Figure 31 GI Group Student Sample
Final Drawing of *Arabidopsis* I

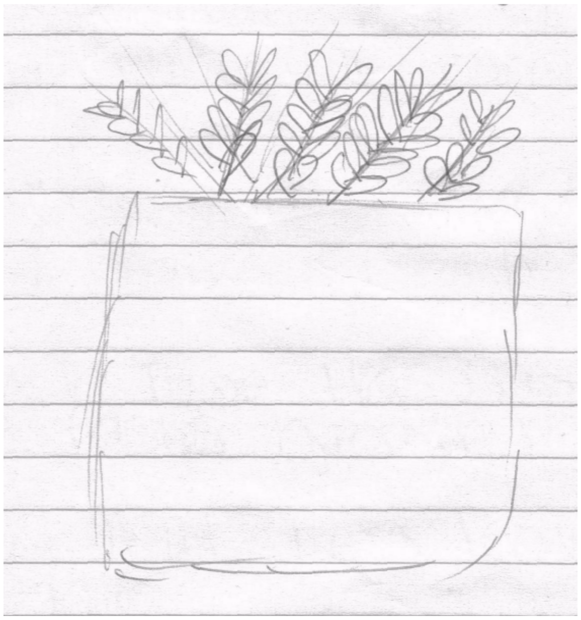

Figure 32 GI Group Student Sample
Initial Drawing of *Arabidopsis* II

Figure 33 GI Group Student Sample
Final Drawing of *Arabidopsis* II

Summary of Findings

Pre-test and post-test responses for the GI group showed statistically significant improvement in their conceptual understanding on four of the five questions addressing conceptual understanding of genes, the environment and their impact on the phenotype of the *Arabidopsis*. Statistically significant improvement was evidenced in all five of the questions addressing conceptual understanding for the ASR group (Questions 1 – 5, see Table 28). Responses to the questions on scientific investigations (Questions 6 and 7) revealed that only the GI group made statistically significant gains on both questions. The ASR group only showed gains in one of the two measures. Both groups showed statistically significant gains in response to the question using practical reasoning (Question 8).

Analysis of journal responses revealed responses from the GI group tended to show new understanding of the logistics of the scientific process such as *how to* observe, *how to* take time-lapse pictures, and *how to* conduct experiments. Responses from the ASR group tended to show new understanding of both the logistics of scientific investigation, but also an understanding of some of the abstract aspects of investigation such as the thinking and difficulty inherent in the process. Responses from the GI group focused markedly on domain specific concepts of genes and the environmental effects. The ASR group responses focused on the learning processes and the process of scientific investigation. Responses from both groups addressed plant features, such as trichomes, circadian rhythms, and how plants grow.

Table 28

Summary of Pre-test and Post-test Results for GI and ASR Groups

Construct	Question	GI Pre-Post Differences	ASR Pre-Post Differences
Understanding of Genetics, Plant, and Environmental Concepts	1	-	+
	2	+	+
	3	+	+
	4	+	+
	5	+	+
Understanding of Scientific Investigation	6	+	-
	7	+	+
Practical Reasoning	8	+	+

+ = significant differences; – = no significant difference

Finally, initial and final drawings for both groups showed improvement in attention to detail as well as in improved accuracy of the drawings. Visual inspections did not support improvement in one group over the other. Improvement appeared to be consistent for both groups and all six classes. Conclusions inferred from these results are the focus of the next chapter. Table 29 summarizes the data analyses.

Table 29

Summary Comparison. Results on All Measures by Type of Inquiry Group

	Guided Inquiry	Authentic Scientific Research
Conceptual Understanding of Genetics		
Pre-test Post-test Questions 1-5	4 of 5 questions significantly different	5 of 5 questions significantly different
Gain Scores Question 1-5	Similar gains to ASR Questions 2,4,5	Greater gains on Questions 1 and 3
Journal Question “What have you learned?” – Domain Content Category	25% responses indicated content domain knowledge	7.7% responses indicated content domain knowledge
Journal Question “What have you learned?” --Plant Features Category	43.6% indicated plant features	48.6% indicated plant features
Arabidopsis pre- and post-Drawings	Indicated differences similar to ASR	Indicated differences similar to GI
Understanding of Scientific Investigation		
Pre-test Post-test Questions 6-7	2 of 2 questions	1 of 2 questions
Gain Scores	Similar gains for both groups	Similar gains for both groups
Journal Question 1, “What have you learned?” –Scientific Investigation Category	19.4% responses indicated information about scientific investigation	34.6% responses indicated information about scientific investigation
Journal Question 2, “What have you learned about scientists and how they do their work? --Communication and collaboration category	6% (1 out of 15 responses) indicated that “Scientists communicate and work with each other”	26% (6 out of 23 responses) indicated that “Scientists communicate and work with each other”
Journal Question 2, “What have you learned about scientists and how they do their work? -- Length of Time Category	0% indicated that “experimentation takes a long time”	9% (2 out of 23) indicated that “experimentation takes a long time”
Practical Reasoning		
Pre-test Post-test Question 8	Significant difference	Significant difference
Gain Scores	Higher gains evidenced by GI group	Lower gains evidenced by ASR group
Frequency Count for Responses	7 of 38 elaborations 10 of 38 concepts 21 of 38 no elaboration	32 of 50 elaborations 5 of 50 concepts 11 of 50 no elaboration

CHAPTER V

CONCLUSIONS

Introduction

> Quantitative and qualitative techniques provide a tradeoff between breadth and depth and between generalizability and targeting to specific (sometimes very limited) populations. ... Data collected through quantitative methods are often believed to yield more objective and accurate information because they were collected using standardized methods, can be replicated, and, unlike qualitative data, can be analyzed using sophisticated statistical techniques. In line with these arguments, traditional wisdom has held that qualitative methods are most suitable for formative evaluations, whereas summative evaluations require “hard” (quantitative) measures to judge the ultimate value of the project. This distinction is too simplistic. Both approaches may or may not satisfy the canons of scientific rigor. Quantitative researchers are becoming increasingly aware that some of their data may not be accurate and valid. ... On the other hand, qualitative researchers have developed better techniques for classifying and analyzing large bodies of descriptive data. It is also increasingly recognized that all data collection – quantitative and qualitative – operates within a cultural context and is affected to some extent by the perceptions and beliefs of investigators and data collectors (Frechtling & Sharp, 1997, pp. 3-4).

I chose mixed methods for the design of this research study for reasons not unlike those discussed in the National Science Foundation's *Handbook on Designing and Conducting Mixed Method Evaluations* (Frechtling & Sharp, 1997), which introduces a broader perspective to evaluators who have focused primarily in the past on quantitative techniques and who, as a result, "may miss important parts of a story" (p. 1). The exploratory nature of this study, the lack of other studies examining the differential effects of different forms of inquiry on student learning, and my inabilities to control aspects of the administration of my treatments, I believe, warranted the use of a broader perspective in the design of the study. My goals in the design of this investigation were to address not only questions of scientific rigor but also the broader need to collect data that might contribute to an important, otherwise absent, part of the story regarding the effects of two different forms of inquiry on student learning (Tashakkori & Teddie, 1998; Gall, Gall & Borg, 1999). The following sections in this introduction are meant to explain, summarize, and justify my pragmatic choice to use mixed methods rather than to philosophically remain bound to either quantitative or qualitative approaches to data collection, analysis, and synthesis.

Quantitative Techniques. My research questions regarding the advantages of inquiry instructional methods on student learning were basically quantitative in nature. I asked questions that required pre-test and post-test data to assess and compare the effects of two forms of inquiry-based instructional sequences on student learning. I asked three questions, which were informed by the framework of the nationally recognized National

Assessment for Educational Progress (NAEP; O’Sullivan & Weiss, 1999), which broadly assesses students’ understanding of science in three domains. The questions were:

1. How do these two forms of inquiry affect conceptual understanding?
2. How do these two forms of inquiry affect students’ understanding of scientific investigations?
3. How do these two forms of inquiry affect practical reasoning ability?

Three of my six classes were randomly assigned to the guided inquiry (GI) treatment; and three other classes were randomly assigned to the authentic scientific research learning (ASR) treatment. All classes had similar demographics in terms of students with special needs, ethnicity, and socioeconomic status. Quantitative data were analyzed appropriately using nonparametric statistical techniques to assess the statistical significance of students’ gains in understanding on eight matched pre-test and post-test questions. The content and design of these questions were revised after they were administered to a pilot group of students in the fall prior to the school year in which the study occurred. The content and form of the final questions were validated by two “experts” in educational research and one scientific researcher. A four-point rubric was constructed for assessing the quality of my students’ responses on each question, informed by conventions for developing rubrics for conceptual understanding that had been established by the NAEP (O’Sullivan & Weiss, 1999). Inter-rater reliability regarding the consistency of the rubric as a scoring device was established by comparing the scores of two independent raters on 64 pre-test and post-test responses (eight tests).

Confidence intervals and error box plots were constructed to display means and standard deviations at the 95% level of confidence. Error box plots for each question were displayed for ease in visual comparison between pre-test and post-tests for both treatments. Wilcoxon signed-ranks tests were also calculated for each question to reveal distributions of positive, negative, and neutral gains for each question and to assess the statistical significance of the differences between pre-test and post-tests.

Qualitative Techniques. Students' journal entries provided descriptive data about their learning that was not available from their pre-test and post-test answers. Students were required periodically at the end of class to reflect on their learning for the day by answering an open-ended question posed by the teacher/researcher. Students were also required to make drawings of an *Arabidopsis* plant before the inquiry instruction began and after the inquiry sequences were completed. These data were collected to fill in important parts of the story, to clarify and illustrate results of the pre-test and post-test data, and to test for complementarity between measures. Two journal entries were chosen for content analysis which allowed me to probe some of the underlying issues that revealed themselves in the students' responses to pre-test and post-test measures that were evaluated on a simple four-point scale.

Mixing the Techniques. While I chose predominantly quantitative questions to drive the research, I also allowed students' journal responses, drawings, and content analyses of students' responses on the pre-test and post-test questions to complete, enhance, reinforce, and question the results of the quantitative approaches taken to data analysis.

Conclusions

Research Question 1: How do these two forms of inquiry affect conceptual understanding? Guided inquiry students understood more about genetic concepts and environmental effects on phenotype than the authentic scientific research learning students. Quantitative analysis supported gains for ASR over GI, but qualitative data revealed that GI students were much more likely to record responses that included use of genetics and environmental concepts explored in the study. When results of both analyses were mixed, GI students made greater gains. Qualitative results did not support the statistical results. Quantitative data showed greater gains for the ASR group on two of the five pre-test and post-test questions and similar gains for both groups on the remaining questions. Quantitative analysis further revealed statistical significance in all five questions measuring conceptual understanding for ASR and statistical significance on four of the five questions for GI indicating slightly greater gains for the ASR group on these quantitative measures alone. However, when GI responses to the question, "*what have you learned?*" were qualitatively analyzed, a different picture emerged. GI responses focused on concepts from the study three times more often than the ASR group. Further the GI group recorded learning about plant features in 43.6 % of their responses. In addition, the GI group used concepts explored in the study twice as often as did the ASR group on the question using practical reasoning skills.

Scientific Investigation as a Process: How do these two forms of inquiry affect students' understanding of scientific investigations? ASR students understood more about scientific investigation as a process. Quantitative analysis supported similar gains for both groups, but an analysis of qualitative data revealed deeper understanding for the ASR students. Qualitative data did not support the statistical results. Quantitative data revealed statistically significant gains in the ASR group on one of the two pre-test and post-test questions. Closer inspection of error box plots revealed large variance on ASR pre-tests. Statistical significance was found on both questions for the GI group. Qualitative analysis of student journals revealed that 34.5% of ASR students' responses to the question, "*what have you learned from this project?*" addressed scientific investigation while only 19.4% of the GI students mentioned scientific investigation. Further, when asked, "*what have you learned about how scientists and how they do their work?*" 35% of ASR students' responses addressed collaborative communication and length of time involved in scientific investigation, while only 6% of GI students made these references.

Practical Reasoning Skills: *How do these two forms of inquiry affect practical reasoning ability?* ASR students made greater gains in demonstrating practical reasoning skills. Quantitative analysis supported similar gains for both ASR and GI, but qualitative data revealed that ASR students had a more developed sense of practical reasoning. ASR students were much more likely to provide elaborations and explanations to support their responses. Quantitative results from pre-test and post-test questions revealed similar gains. However, when additional statistical analyses were completed using error box plots, gains in the GI group over the ASR group became evident. GI pre-test mean scores were lower than ASR pre-test mean scores, while post-tests results were similar in both groups indicating greater gains for the GI group. Qualitative results did not support the statistical results. Quantitative analysis further revealed that when students were asked to use practical reasoning skills to answer a scientific question, 68% of ASR student used some form of elaboration or examples in their responses as compared to 24% of GI students. Only 22% of ASR students failed to provide some form of elaboration as compared to 55% of the GI students.

Lessons Learned

I designed interventions for both treatments with the best of intentions. I had gathered advice from members of my committee, including a researcher in education and a scientist. I had conducted a pilot test, although shorter in duration, in the fall of the previous year with a small number of eighth grade students. I made revisions on the basis of the pilot test, developed a full-blown, three-stage implementation plan, and began to organize my classroom to become a "teaching research laboratory" during the fall of the school year.

While I had planned to complete the data collection for the full study in the fall of the school year, I waited to conduct the study until I heard from two important agencies, one involved in funding the research and another involved in granting permission to conduct it. I had written a grant to the Toshiba Foundation in the fall of the year to buy technology, but notification of the award was not made until January. While the award allowed the opportunity for all of the students in my six classes to engage in one of two forms of classroom-based inquiry, the late award required me to wait until the early spring to set up my classroom with cameras and light boxes. Furthermore, I waited to receive university approval for my proposal to collect data and informed consent from eighth grade students and their parents. The approval from the Institutional Review Board (IRB) at Texas A&M University was also not awarded until the early spring. Implications for waiting until the spring of the school year will be discussed later in this chapter.

As with any study, there are things the researcher learns in the process of doing the study that serve to inform future studies. My study was no exception. In this section, I have identified four main concerns: (1) the impact of my graduate studies on the research, (2) the conflict I experienced in serving as both the teacher and the researcher, (3) the effects of delayed IRB approval and funding on the timing of the study, and (4) alterations in the research design.

First, the original design of the study was to compare a *typical* GI experience with an ASR learning experience. As I have progressed through my studies at Texas A&M, I became familiar with research-based learning theories that have transformed my teaching (Mundry, 2003). As I progressed through the GI instructional segments I found myself automatically incorporating these learning theories, in effect providing the best GI experience possible, as well as the best ASR experience. I found that I couldn’t go back, so to speak. I no longer was able to conduct the type of GI that more typically represents the level of GI done by teachers in most classrooms. I found myself spontaneously picking up on cues from my students that indicated they required more scaffolding or prompting, and I found myself constantly adjusting to meet the needs of my students. In effect, I offered a *model* GI experience for my students rather than a *typical* representation. Had I been able to conduct a GI that was more representative of the typical GI experiences, there may have been greater differences in all three measures of this study.

Second, I found the conflict between my responsibilities as a teacher to often be in conflict with my role as the researcher. It was very difficult to separate these two roles. The frustration of knowing what to do to assist my students and the constraints of the research that mandated no instructional variation was heavy indeed. I found myself, as did Wong (1995), coming face-to-face on a daily basis with conflicts between doing what I would intuitively do as a teacher (assist the student in learning) and acting the part of the researcher (assure validity in treatments). The constant conflict between what was best for the student and what was best for the research made me question if it was even beneficial to engage teachers as researchers. Perhaps a better model would have been a paired research model in which the researcher worked alongside the teacher.

Third, this study was riddled with situations beyond the control of the researcher that typified what Ann Brown (1992) referred to as the *messiness* of classroom research. Delays in funding and IRB approval pushed my implementation of the study from the fall into the spring. Unfortunately, as any classroom teacher can attest, spring is rife with student illnesses, incentive trips, field trips, practice tests, meetings for special education reviews, mandatory review sessions, spring break, and state-mandated testing.

As a result of these interruptions, my original population of 130 students yielded less than half the participants in both groups who were able to complete all parts the pre-test and post-test measures.

Fourth, I would do several things differently in future research to provide a richer description (Gertz, 1973) to further inform the research findings. A review of student journals revealed that students on several occasions ran out of time when recording their reflections or responding to open-ended guided response questions. My reflective journal also revealed that I did not allow sufficient time for students to fully reflect and respond in their journals. Furthermore, I did not look at any of my students' writings until the study was complete. I chose to do this in an effort to remain unbiased and in response to IRB mandates. I was therefore unaware of the types of responses some of my students were making. Student responses indicated students did not understand the questions. Had I reviewed the data during the study, I would have been able to modify the format of questions to encourage student responses that more clearly addressed the idea behind the question.

I also discovered (after the fact) that students were recording information in their journals in a way that was very hard to follow. If I had looked at journals during the course of the study, I could have detected the need to provide more guidance in the way the students structured their entries. A more organized structure for student journal responses would have aided me in finding and interpreting student responses. It might have been beneficial to have students respond to questions scaffolded like, *I used to believe*__________, and *now I believe* __________. I believe this would have provided a richer, more aware description of changes to student conceptual models as a result of a research intervention.

Finally, I myself gained a deeper understanding of the implications of missing data. As obvious as it is now, when I was in the midst of the study and wearing both the hat of a teacher and the researcher and bearing the responsibilities of both, I failed to see the implications of missing data on my study results. It wasn’t until I started the data analysis that I became aware that the missing data imposed serious limitations on my abilities to legitimately make claims about the effects of two treatments on student learning.

Recommendations for Further Study

Working within the messiness of the classroom environment has inherent inconsistencies that confound and cloud the most careful of research designs. This research proved to be just such a study, and yet it did yield indications of differences between two different inquiry-based instructional approaches. Additionally, this study highlighted unexpected benefits for special-needs students and implications of the role of teacher as researcher. The results revealed in this study provided the foundation on which the following recommendations for future study were made.

Time of Year. Interruptions occurred throughout this study, resulted in large part to the time of year in which the study was conducted. The spring of the year in a Texas classroom brings a wide assortment of deviations from the normal class schedule, none of which contributes to coherence in classroom instruction. I recommend that any similar study done over a sustained period of time be implemented earlier in the school year when students tend to be in class more and the educational focus is centered more on classroom instruction.

Despite the small numbers of students from which complete data was collected, I believe that there are trends in the data that warrant further study. This study revealed indications that gains in conceptual understanding of concepts occurred through inquiry-based instructional methods. Data from student journals and classroom observations further support this finding and suggested benefits in their understanding of the scientific investigation as a process, and increased in practical reasoning skills.

Special-Needs Students. Classroom teachers are typically skeptical of engaging special-needs students in any form of inquiry-based instruction believing that they will lose control of the classroom and that inquiry is beyond the capability of these students (Zohar & Dori, 2003). My findings concurred with Zohar and Dori (2003): Special-needs students gained substantial benefits from the challenges presented in both inquiry-based instructional methods. Inquiry-based instruction resulted in all students being actively engaged an obvious classroom management benefit (Stuessy & Scallon, in preparation). Repeated instances occurred where low-performing special-needs students performed at levels far beyond what I typically expected. These were students who typically did not actively engage in classroom discourse.

Teacher professional development is another implication of the apparent benefits of inquiry-based instruction for special-needs students. Benefits of inquiry-based instruction for special-needs students substantiated through repeated investigations could reform teachers' existing beliefs about the capabilities of special-needs students in participating in inquiry-based instruction. Repeated references in my reflective journal point to instances where special-needs students were highly motivated, participated with enthusiasm in student-centered inquiry, and excelled in the inquiry-based learning environment.

Data Collection Instruments & Rubric Design. Student records provided an incomplete picture of what I knew my students had learned but had no evidence to support. I listened to student discourse and I knew the richness of their thinking, however; students recorded enough to answer the question but did not write the full extent of what they really knew. The discrepancy between what students recorded and what I heard them say in group and classroom discussions was huge. My reliance on written evaluation instruments was insufficient to provide a complete picture of the impact of GI and ASR on student learning.

On several occasions, students did not have adequate time to finish entries in their journals. On more than one occasion, students would be in the middle of recording responses to open-ended questions when the bell rang, thus ending class. Obviously, there were lost opportunities to acquire insights from students' recording of questions such as, *What I learned today.* I recommend that in future studies the allotment of time be seriously and deliberately considered for students to reflect on what they have done and learned.

I further recommend the addition of personal interviews and audio taping of small-group discourse. These additions would provide a broader and richer collection of data for analysis. Follow-up interviews would allow researchers to probe students for more explicit answers to pre-test and post-test questions serving to triangulate the data. In addition, taping small-group discourse would allow researchers to glean from students' conversations a wealth of information that went untapped in this study.

Finally, I recommend that student journals be reviewed throughout the study as a formative evaluation of the matches between the intent of the researcher, questions, and student responses. Formative reviews would allow the researcher to make minor revisions in the way the questions were worded. It would also reveal if students were being given ample time to complete their responses in the allotted time.

Design of the Study. With respect to the actual design of the study, I now suspect that the three-phase approach was unnecessary. I suspect that if these students had engaged in ASR from the very beginning (as occurred in the pilot study) that there would have been more differences between student learning in the GI and ASR instructional models. I now believe that the three phases actually hindered the degree of participant engagement rather than providing a scaffold toward independent inquiry. I noted in my reflective journal that toward the end of the second phase of the study, both groups grew tired of the guided instruction and began to lose interest. During Phase III, the ASR group re-gained their enthusiasm, but the GI group did not. Separating the two interventions at the onset of the study would have better reflected the model of the pilot study, in which students engaged in guided instruction before engaging in ASR only long enough to make them comfortable and allow them to engage in the project. In design of this study, I believed this initial guided-inquiry would scaffold students toward a more independent student-centered, inquiry-based instructional model, providing an essential bridge toward ASR. However, I now believe I built in too many guided-inquiry experiences. I suspect that the progression from didactic to limited hands-on to guided-inquiry was too much scaffolding; in effect, stifling students who demonstrated readiness for the independent learning characteristic of the ASR experience.

Finally, I suggest that a paired research model be used; that is, a model that splits the research roles between the teacher and the researcher, rather than the teacher acting as both teacher and researcher. This study brought to light many of the conflicts associated with trying to fill two roles which often times have conflicting objectives (Wong, 1995). Specifically, the role of the teacher is to help students learn, while the role of the researcher is to understand a phenomenon. Where the teacher would alter instruction, events, and experiences to enhance student learning, the researcher would see these alterations as a threat to the validity of the study. The focus of the teacher is on the student; the focus of the researcher is on the study. Additionally, the burden of managing a study as well as doing the classroom teaching proved at times to be overwhelming. By using a paired research model, the researcher could offer guidance to the teacher as the study unfolds. Similarly, in planning the teacher could provide expertise to the researcher on the feasibility of the research design and alert the researcher to potential implementation pitfalls. In addition, because the teacher had worked with these students in a context outside the study, they are in a unique position to notice and identify differences in student performance, attitudes and actions specific to the study that would probably go unnoticed by a researcher unfamiliar with the students.

A Final Reflection on the Connections Between Theory and Practice.

Cognitive theory made the connections in my own thinking between what students learn and how they learn it. Cognitive theory also provided the foundation for my thinking about the creation of the ideal science learning environment. As a teacher, I was interested in knowing more about how to design a science learning environment that would allow my students to actively construct new understanding for themselves, to develop habits of mind that would allow them to assess their own learning, to "chunk" relevant facts together into organized sets of information, to flexibly use new knowledge in new situations, and to learn in a social environment where new ideas were constructed, shared, and revised on the basis of external feedback and internal reflection. Cognitive theory also led me to believe that a more open, *authentic* learning environment would lead to more gains in student learning. I focused on creating an open learning environment for my students in my role as a designer of instruction.

I believe that the results of this investigation have connected my understanding of cognitive theory and practice in a new way. My choice of mixed methods in my classroom-based study contributed many important, otherwise absent, parts of the story. The lessons I learned were many, including that I needed to reserve more time for my students to reflect on their learning.

My Final Reflections Center on Student Reflection. Of course, I understand now that more time would have provided a better data source for analysis. However, I also have a new understanding about the role of student reflection in the study, which I never quite understood before. The results of this study indicated to me that the *design* of the environment is not the only important factor in creating the ideal learning environment. An unexamined factor in this particular study is the differences in the *role of the teacher* in these two environments, particularly in regard to the teacher's role in facilitating reflective practice in students.

The guided inquiry environment allowed many opportunities for me to scaffold my students' reflections on their learning, in large-group, on a daily basis. We constantly came together after small bursts of independent activity to review, reflect, and revise our practice. On the other hand, students in the authentic scientific learning classes did not have the benefits from large-group discussions in the same way. While small, independent ASR groups may have had opportunities to share and revise their ideas about the inquiries they were conducting on their own; the independent small-group structure did not provide as many opportunities for the more formal, guided exchanges between the teacher and students about their experiences. Multiple opportunities were provided for students in the GI group to communicate formally about new facts and concepts regarding mutations, environmental effects, genotypes, and phenotypes. By contrast, students in the small ASR groups talked among themselves without the benefit of more formalized, directed, teacher intervention.

My own personal reflections lead me to wonder about the role of those daily teacher-directed, scaffolded discussions in the development of students' understanding about the ways things in the natural world works. Perhaps these discussions provided the factual and conceptual benefits to students in the GI group demonstrated in their responses on post-tests and in their answers to journal questions. Without the thorough examination of the quality of student responses, as well as the results comparing pre-test and post-test scores in both groups, I do not think that I would be so willing to identify the possible connection between teacher-directed discussions and scientific knowledge gains.

Reflections From the Teacher's Perspective. The research results fill in only part of the story about the design of inquiry-based learning environments. From the professional teacher's perspective, I would like to fill in another part of the story, and that has to do with how I would change my practice as a result of the research I have just conducted.

In designing the "ideal" learning environment for my students, I would provide multiple opportunities for individuals, small groups, and entire teacher-led class discussions to share, reflect, and revise their ideas about what they are learning. I would encourage students to use factual and conceptual information in their discourse and writing, emphasizing the "big ideas" under girding the inquiry, just as I did on a daily basis with students in the GI group.

I would embed my students' hands-on learning, however, within an authentic scientific research learning approach. Students would ask their own questions, design their own experiments, and collect their own evidence for their own knowledge claims, just as I did with the students in the ASR group with one important difference. I would allow more time for students to come together in large groups. In these large groups, I would guide their sharing to reveal what small groups are learning independently. I would scaffold the discussion so that students could learn from one another about what is working, what they are thinking, and how they would revise not only their processes of doing science but also their thinking about science.

Overall, on a daily basis, I would spend much more time thinking about ways to make my students' thinking visible through not only what they say in class but what they write and draw. I would monitor and review student reflections daily to focus my decisions about the most important topics for the next day's large-group discussions. I would check specifically for clarity and accuracy in my students' use of important scientific ideas and note them for further discussion, and we would talk as a class about what we are learning about how the world works, how scientists do their work, and how effective solutions to real-world problems require a good conceptual understanding of both facts and process.

Finally, I would conduct other messy, classroom-based studies to see how well these recommendations hold together in other science learning contexts. I believe in time I will feel more comfortable in the teacher-as-researcher role (particularly without the constraints of the requirements for rigor applied to the completion of a master's thesis). I would like to engage in lesson studies, perhaps with other teachers, to "test" the effects of innovative inquiry-based lessons on my students' learning.

I am encouraged by the results I have gathered here and my own progress as a teacher-researcher. On the research side, I believe this study contributes new knowledge about the differential effects of guided inquiry and authentic scientific research learning on student learning. For me personally, however, I have experienced first-hand the benefits that research can have regarding better ways to increase my students' understanding about science. I have also experienced first-hand the benefits that research can have regarding the ways I can increase my own understanding about science teaching.

REFERENCES

American Association for the Advancement of Science(AAAS). (1989). Project 2061: Science for all Americans. Washington, DC: Author.

Anderson, L.W., & Krathwohl (Eds.). (2001). A Taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives. New York: Longman.

Bonnstetter, R.J. (1998). Inquiry: Learning from the past with an eye on the future. *Electronic Journal of Science Education,* 3. Retrieved October 27, 2002 from http://www.asd.k12.ak.us/Depts/Science/Elementary/inquiry.htm

Bransford, J. D., Brown, A.L., & Cocking, R.R. (Eds). (2000). How people learn: Brain, mind, experience and school. Washington, DC: National Academy Press.

Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. Journal of the Learning Sciences 2(2), 141-178.

Brown, A.L. (1994). The Advancement of learning. Educational Researcher, 23(8), 4-12.

Brown, A.L., & Champione, J.C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.) Innovations in learning: New environments for education (pp. 289-325). Mahwah, NJ: Erlbaum.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 32-41.

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

Buckley, B.C., Gobert, J.D., Kindfield, A.C.H., Horwitz, P., Tinker,R.F., Gerlitz,B.J., Wilensky, U., Dede, C., & Willett, J. (2004). Model-based teaching & learning with BioLogica™: What do they learn? How do they learn/How do we know? Journal of Science Education and Technology, 13(1), 23-41.

Champione, J.C., Shapiro, A.M., & Brown, A.L. (1995). In A. Shapiro, J. Lupart, & A. Marini (Eds.), Teaching for transfer: Fostering generalization in learning (pp. 35-68). Mahwah, NJ: Erlbaum.

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

Chiapetta, E., & Koballa, T.R., Jr. (2000) Science instruction in the middle and secondary schools (5th ed.). Columbus, OH: Merrill Prentice Hall.

Chinn, C., & Malhotra, B. (2002). Epistemologically authentic reasoning in schools: A theoretical framework for evaluating inquiry tasks. Science Education, 86, 175-218.

Collins, A., Brown, J.S., & Newman, S.E. (1991). Cognitive apprenticeship: Teaching the crafts of reading, writing and mathematics. In L.B. Resnick (Ed.), Knowing, learning, and instruction: Essays in honor of Robert Glaser. Hillsdale, NJ: Erlbaum.

Concept to Classroom. (2000). Workshop: Inquiry-based learning. Retrieved September 16, 2004, from http://www.thirteen.org/edonline/concept2class/inquiry/index.html

Costa, A. L. (Ed). (2001). Developing minds: A resource book for teaching thinking (3rd ed.). Alexandria, VA: Association for Supervision of Curriculum and Development.

Creswell, J.W. (2003). Research design: Qualitative, quantitative and mixed methods approaches (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Donovan, M.S., & Bransford, J.D. (2005). How students learn science in the classroom. Washington, DC: National Academies Press:

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott. P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23, 5-12.

Duschl, R.A., Gitomer, D.H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. Educational Assessment, 4(1), 37-73).

Duschl, R. A., & Grandy, R. E. (in preparation). Reconsidering the character and role of inquiry in school science: Framing the debates. In R. A. Duschl & R. E. Grandy (Eds.), TBA.

Dwyer, C. (1994). Criteria for performance-based teacher assessment: Validity, standard, and issues. Journal of Personal Evaluation in Education, 8, 135-150.

Edelson, D. (1998). Realising authentic science learning through the adaptation of scientific practice. In B. J. Fraser & K. G. Topin (Eds.), International handbook of science education (pp. 317-331). London: Kluwer.

Edelson, D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. Journal of the Learning Sciences, 8(3/4), 391-450.

Etheredge, S. & Rudnitsky, A. (2003). Introducing students to scientific inquiry: How do we know what we know? Boston: Allyn & Bacon.

Feldman, A., & Minstrell, J. (2000). Action research as a research methodology for the study of the teaching and learning of Science, Handbook of research design in mathematics and science, Erlbaum, NJ: Mahawh.

Fellows, N. (1994). A window into thinking: Using student writing to understand conceptual change in science learning. Journal of Research in Science Teaching, 31(9), 985-1001.

Frechtling, J., & Sharp, L. (Eds.). (1997). User-friendly handbook for mixed method evaluations. Washington, DC: National Science Foundation. Retrieved September 19, 2005 from http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAPU1.HTM

Gall, P.J., Gall, M.D., & Borg, R.W. (1999). Using research to improve educational practice. Applying Educational Research: A practical guide. Logman: New York.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz (Ed.) The interpretation of culture. (pp. 3-30). New York: Basic Books.

Goldman, S.R., Petrosino, A.J., and Cognition and Technology Group at Vanderbilt. (1999). Design principles for instruction in content domains: Lessons from research on expertise and learning. In F.T. Durso, R.S. Nickerson, R.W. Schvaneveldt, S.T. Dumais, D.S. Lindsay, & M.T.H. Chi (Eds.), Handbook of applied cognition (pp 568-627). New York: Wiley.

Guzzetti, B.J., Snyder, T.E., & Glass G.V. (1992). Promoting conceptual change in science: Can text be used effectively? Journal of Reading, 35(8), 642-649.

Hickey, D.T., Kruger, A.C., Fredrick, L.D., Schafer, N., & Kindfield, A.C.H. (2002). Balancing formative and summative science assessment practices: Year one of the GenScope assessment project. Paper presented at Annual meeting of American Educational Research Association, New Orleans, LA.

Kesidou, S., & Roseman, J. (2002). How well do middle school science programs measure up? Findings from project 2061 curriculum review. Journal of Research in Science Teaching, 39(6), 522-549.

Kotulak, R. (1996). Inside the brain: Revolutionary discoveries of how the brain works. Kansas City, MO: Andrews McNeal Publishing.

Kozma, R.B. (2000). The use of multiple representations and the social construction of understanding in chemistry. In M. J. Jacobcon & R. B. Kozma (Eds.), Innovations in science and mathematics education (11-46). Mahwah, NJ: Erlbaum.

Kuhn, D., Schauble, L., Garcia-Mila, M. (1992). Cross-domain development of scientific reasoning. Cognition and Learning, 9(4), 285-327.

Leher, R., Schauble, L., & Petrosino, A. J. (2001). Reconsidering the role of experiment in science education. In K. Crowley, C.D. Schunn, & T. Okada (Eds.), Designing for science: Implications from everyday, classroom, and professional settings (pp. 251-278). Mahwah, NJ: Lawrence Erlbaum Associates.

Lin, X.D., & Lehman, J. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. Journal of Research in Science Teaching, 36(7), 837-858.

Maxwell, J.A. (1998). Designing a qualitative study. In L.Bickman & D. S. Rog (Eds.), Handbook of applied social research methods (p. 69 -100). Thousands Oaks, CA: Sage.

McMillan, J.H. (1999). Establishing high quality classroom assessments. Metropolitan Educational Research Consortium, Virginia Commonwealth University (ERIC document reproduction service No. ED 429-146)

Mertler, C.A. (2000). Teacher-centered fallacies of classroom assessment validity an reliability. Mid-Western Educational Researcher, 13(4), 29-36.

Messick, S. (1994). The Interplay of evidence and consequence in validation of performance assessment. Educational Researcher, 23(2), 13-23.

Minstrell, J., & van Zee, E.H. (Eds.). (2000). Inquiring into inquiry learning and teaching in science. Washington, DC: American Association for the Advancement of Science.

Mundry, S. (2003). Honoring adult learners: Adult learning theories and implications for professional development. In J. Rhoton & P. Bowers (Eds.), Science teacher retention: Mentoring and renewal (pp.123-132). Arlington, VA: National Science Teachers Association Press.

Nation Center for Educational Statistics (NCES). (2004a). The NAEP questions tool v3.0. Retrieved September 15, 2004 from http://nces.ed.gov/nationsreportcard/itmrls/

National Center for Educational Statistics(NCES). (2004b). The NAEP science achievement levels. Retrieved September 15, 2003, from http://nces.ed.gov/nationsreportcard/science/achieve.asp

National Research Council (NRC). (1996). National science educational standards. Washington, DC: National Academy Press.

O'Sullivan, Y.C. & Weiss, R.A. (1999). Student work and teacher practices in science: A report on what students know and can do. NCES 1999-455 Washington DC: Education Publication Center.

Palincsar, A.S. & Brown, A.L. (1992). Fostering literacy learning in supportive contexts. Journal of Learning Disabilities, 25(4), 211-225.

Petrosino, A., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. Mathematical Thinking and Learning, 5(2 & 3), 131-156.

Polman, J.L. (2000). Designing project-based science: Connecting learners through guided inquiry. New York: Teachers College Press.

Schraw G., & Dennison, R.S. (1994). Assessing metacognitive awareness. Contemporary Educational Psychology, 19, 460-475.

Sperling R.A., Howard B.C., Miller L.A.., Murphy C. (2001). Measures of children's knowledge and regulation of cognition. Contemporary Educational Psychology 27, p51-79 online: http://www.idealibrary.com

Stuessy, C. S., & Scallon, J.M.(in press). The Learning research cycle: Bridging research and practice. In Press

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

Texas Education Agency(TEA). (2005). Texas essential knowledge and skills. Retrieved on September, 18, 2002 from http://www.tea.state.tx.us/teks/index.html

Thorndike, E.L. (1913). The psychology of learning. New York: Teachers College. U.S Congress (2001). No child left behind act 2001. Public Lawn 107-110. 107th Congress. Washington, DC: Government Printing Office.

Wenger, E., (1998). Community. Communities of practice: learning, meaning and identity. 72-85 New York: Cambridge University Press

White, B.C., and Frederiksen,J.R.,(2000). Technological tools and instructional approaches for making scientific inquiry accessible to all. In M.J. Jacobson and R.B Kosma (Eds.) Innovations in science and mathematics education 321-259 Mahwah, NJ: Lawrence Erlbaum Associates

Wiggins, G., & McTighe, J. (1998). Understanding by design. Alexandria, VA: Association for Supervision and Curriculum Development.

Wong, D.E., (1995). Challenges confronting the researcher/teacher: Conflicts of purpose and conduct. Educational Researcher, 24(3): 22-28.

Zohar, A., & Dori, Y. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? The Journal of the Learning Sciences, 12(2), 145-181.

APPENDIX A

Rubric for Pre-Post Test Questions

Why do plants look the way they do?

3 **Complete:** Both genes & environment impact phenotype. May also mention other phenomenon like chlorophyll, cell structure, etc.

2\. **Partial:** Mention either gene or environment impact phenotype. May also mention other phenomenon like chlorophyll, cell structure, function of the parts, etc.

1 **Unsatisfactory:** No mention of genes or environment as impacting phenotype only other phenomenon like chlorophyll, cell structure, function of parts, etc.

0 No response/ Illegible response

Do Plants Move?

3 **Complete:** Expressed understanding of rhythmic cycles like circadian rhythm, geotropism, etc. also, understanding that plants move as they grow,. Elongation, circular pattern seen in time lapse.

2 **Partial:** Express either rhythmic or patterns of growth or movement associated with growth.

1 **Unsatisfactory:** Expresses understanding that external impacts such as wind or stationary roots prevent movement.

0 No response/ Illegible response

Can environment have any affect on genes?

3 **Complete:** Environment affects both genes expression and an understanding that genes adapt to their environment.

2 **Partial:** Environment affects either gene expression or genes adapt to environment

1**.** **Unsatisfactory:** Yes – no explanation - Unrelated answers – answers do not address the question.

0 No response/ Illegible response

Does an experiment have to prove the hypothesis correct to be a good experiment?

3 **Complete:** Explanation expresses understanding of value of unexpected outcomes as well as hypothesis being unrelated to experimental validity.

2 **Partial:** Explanation expresses either a value of unexpected outcomes or understanding that hypothesis is unrelated to experimental validity.

1 **Unsatisfactory:** Yes– no explanation – Any yes answer. Statements expressing misconceptions.

0 No response/ Illegible response

Tell me everything you know about scientific method.

3. **Complete**: Expressed an understanding of scientific method as a dynamic process used to answer questions. There are multiple ways to experiment. Scientific method is NOT a sequential unalterable step by step method. Responses include an understanding of importance of accuracy, data collection and explanation.

2. **Partial**: Understanding of scientific method as a process, may mention steps but they don't imply regiment following of the steps.

1 **Unsatisfactory:** Step by step sequence that scientists use to experiment – or some other misconceptions expressed.

0 No response/ Illegible response.

What do you know about genes?

3. **Complete:** Expressed understanding that genes are inherited, the determine traits and they are instructional information.

2. **Partial:** Expressed one or two of the three, demonstrated some accurate understanding of the function of genes.

1. **Unsatisfactory:** Expressed only misconceptions – responses illegible.

0 No response or wrote something like "I don't know."

What do you know about mutations?

3 **Complete:** Mutations alter gene function or growth and mutations usually adversely affect an organism.

2 **Partial:** Expressed either understanding that mutations affect genes, or mutations adversely affect organisms, or some other response that demonstrates students has an accurate partial understanding of what a mutation is.

1 **Unsatisfactory:** Expressed only misconceptions/ illegible responses.

0 No response - student left blank or wrote something like "I don't know."

Practical Reasoning scenario (last question – students design an experiment)

3. **Complete:** Students demonstrate an understanding of the need for experimentation. Elements of a good experimental set-up are elaborated on, showing an understanding of the role of variables, data collection, observation.

2. **Partial:** Contains elements of experimental design, explaining what the scientists should do, but not linking it to experimentation.

1 **Unsatisfactory:** Response does not address the question / illegible responses.

0 No Response - left blank or wrote something like, "I don't know."

APPENDIX B

Howdy --- Here's your Pre-Test –

Name: ______________________________

Part 1

Class Period: _______ Date: ___________

In the space below, draw a plant with all of its parts. <u>Label as many parts as you can !! (You may draw several drawings if you wish)</u>

<u>Use any of these words in your explanations if you can use them correctly</u>.

Genotype
Inherit
Hypothesis
Conclusion
Phenotype
Traits
Data
Result
Adaptation
Genome
Observation
Experiment

What do you know about <u>genes</u>? Remember tell me everything you can !!!
(What are they, where are they, what do they do, etc)

____

____

____

____

____

What do you know about <u>mutations</u>? Remember tell me everything you can !!! (What are they, what causes them, what do they do, etc.)

____

____

____

____

____

____

____

Howdy --- Here's your Pre-Test –  Name: ______________________________

Part 2 Class Period: _______ Date: __________

Use any of these words in your explanations if you can use them correctly.

Genotype
Inherit
Hypothesis
Conclusion

Phenotype
Traits
Data
Result

Adaptation
Genome
Observation
Experiment

Why do plants look the way they do? Remember tell me everything you can !!

____

____

____

____

____

Can the way a plant looks change? __________ If so, HOW???

Remember tell me everything you can !!!

____

____

____

____

____

Use any of these words in your explanations if you can use them correctly.

Genotype
Inherit
Hypothesis
Conclusion
Phenotype
Traits
Data
Result
Adaptation
Genome
Observation
Experiment

How do plants grow? (What causes them to grow, what determines how they will grow?)

____

____

____

____

____

Do plants move? Do they have patterns of movement? Explain please.

____

____

____

____

____

____

____

____

Howdy --- Here's your Pre-Test –  Name: ___________________________

Part 3 Class Period: _____ Date: ____________

Use any of these words in your explanations if you can use them correctly.

Genotype	**Phenotype**	**Adaptation**
Inherit	**Traits**	**Genome**
Hypothesis	**Data**	**Observation**
Conclusion	**Result**	**Experiment**

How do scientist go about answering scientific questions?

____

____

____

____

____

What is model? Why are they used? For example what do you think is meant when people use terms like " model system " or model plant or "model of the solar system"?

____

____

____

____

____

Do plants respond to the environment? _____ Can the environment affect how a plant grows? _______ Explain your answer please. Give me as much detail as you can.

____

____

____

____

____

____

____

____

____

Does an experiment have to prove the hypothesis correct to be a good experiment? Explain your answer please.

____

____

____

____

____

____

____

____

____

Can the environment have any effect on a plant’s genes? Explain your answer please.

____

____

____

____

____

____

____

____

____

____

Howdy --- Here's your Pre-Test –  Name: ________________________
Part 4 Class Period: _______ Date: _______

Read the story below, and then answer the question.

A scientist has several pots of the same kind of plant in his office. They are spread all over the room to make the room pretty. He notices that some of these plants don't grow as tall as others, even though they were all planted at the same time and are the same type of plant. He wants to know why some of his plants growing taller than others. He can't find any information to answer the question anywhere.

How do you think the scientist would go about answering the question? Be as detailed as you can

____
____
____
____
____
____
____
____
____
____
____
____

VITA

Name: Jane Metty Scallon Graduate Research Assistant – ITS Center
PRISE Fellow - Policy Research/Science Education

Address: Texas A&M University Department of Teaching Learning & Culture
442 Harrington Tower 4232 TAMU College Station, TX. 77843

Email Address: scallonj2000@yahoo.com

Education: B.S., Biology, Stephen F. Austin State University, 1981

Certification: Gifted and Talented Certification
Environmental Science Endorsement (TNRC)
Secondary Science Composite Certification
Secondary Advanced Placement Biology Certification
Secondary Biology Certification

Publications: Stuessy, C.L., & Scallon, J. The learning research cycle: Bridging research and practice. *Journal of Science Teacher Education – accepted for publication May 2005*.
Scallon,J. Chapter VI in ITS publication: Information technology and today’s learners:Bridging the gap between research and teaching in science and education. In preparation
Stuessy, C.L. & Scallon.J What happens when eighth graders design their own scientific experiments? In preparation

Conferences: Scallon, J. Stuessy, C.L (2005) *What happens when eighth graders design their own scientific experiments?* Paper accepted for presentation at the annual meeting of School Science and Mathematics Association, Ft. Worth, November 2005.
Stuessy, C.L., Scallon, J. (2005) *The Learning Research Cycle: The teacher perspective* Paper accepted for presentation at annual meeting of the National Association for Educational Research, Montreal, April 2005.
Stuessy, C.L., Scallon, J. (2005) *The Learning Research Cycle: Authentic student research in the classroom* Paper accepted for presentation at the annual meeting of National Association of Research in Science Teaching, Dallas, April 2005.
Stuessy, C.L., Scallon, J. (2004) *The Learning Research Cycle: Bridging research and practice* Paper accepted for presentation at the annual meeting of School Science and Mathematics Association, Atlanta, October 2004.