THE INTERRELATIONSHIPS AMONG STUDENTS' CONCEPTIONS OF THE NATURE OF SCIENTIFIC KNOWLEDGE, INDUCTIVE REASONING, AND ACHIEVEMENT IN SCIENCE

JAMES ALLEN DUFFETT, B.Sc., B.Ed.

ISBN 0-315-73343-8

THE INTERRELATIONSHIPS AMONG STUDENTS' CONCEPTIONS OF THE NATURE OF SCIENTIFIC KNOWLEDGE, INDUCTIVE REASONING, AND ACHIEVEMENT IN SCIENCE

BY

© James Allen Duffett, B.Sc., B.Ed.

A Thesis submitted to the School of Graduate
Studies in partial fulfilment of the
requirements for the degree of
Master of Education

Faculty of Education
Memorial University of Newfoundland
February 1992

St. John's Newfoundland

ABSTRACT

Interrelationships among conceptions of the nature of scientific knowledge, inductive reasoning, and achievement in science were examined for a sample of 305 suburban high school students. Epistemological beliefs were measured using a questionnaire developed for the study on the basis of the philosophic literature and portions of philosophical models of existing instruments. The final version consisted of 56 items organized into seven subscales representing different dimensions of scientific knowledge. The Essay Test of Inductive Reasoning Strategies, Part A, developed by Norris and Ryan (1987a), was selected as the measure of inductive reasoning. The test requires subjects actively to employ inductive reasoning strategies. Students' achievement in science was measured by the final grade received in the past school year for general science, biology, chemistry, physics, and earth science.

Students' conceptions included beliefs that scientific knowledge: (a) represents real world phenomena, (b) is fallible, (c) is changeable, (d) is a product of the human imagination, (e) must be subjectable to empirical test, (f) is acquired slowly, and (g) should be questioned when reasonable to do so. Students' inductive reasoning was characterized by a superficial treatment of the reasoning tasks. Students tended not to (a) withhold judgement, (b) seek additional information, (c) suggest alternate conclusions, and (d) monitor their own progress.

Results of a path analysis indicated that students' conceptions of scientific knowledge exerted strong significant effects on achievement in general science, biology, chemistry, and physics. The effect of scientific knowledge conception on reasoning was significant for the biology group. Significant effects were also found for reasoning on general science and biology achievement. Reasoning was found to play a much smaller role in determining science achievement than did conceptions of scientific knowledge.

A factor analysis of the questionnaire subscales empirically divided the variable of scientific knowledge conception into four factored variables. The results revealed that the large direct effect for conception of scientific knowledge on reasoning and science achievement was due to beliefs that knowledge: (a) was a representation of real world phenomena, (b) must be testable, (c) is fallible, (d) is changeable, and (e) should be questioned when appropriate to do so.

ACKNOWLEDGEMENTS

The author wishes to extend a sincere thank-you to Dr. Stephen Norris for sound advice and knowledgeable guidance willingly offered during the writing of this thesis; to the Cabot Chapter of The Canadian College of Teachers for their financial support of this thesis; and to Gail, whose endless patience and encouragement kept the project moving toward its completion.

The Essay Test of Inductive Reasoning Strategies, Part A is reproduced with the permission of the authors.

This thesis is dedicated to the memory of Susanna, whose love and support have been the source of constant inspiration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: OVERVIEW

This study examines the relationships among high school students' conceptions of the nature of scientific knowledge, their inductive reasoning ability, and their achievement in science. A sample of 305 senior high school students were administered an epistemological questionnaire and The Essay Test of Inductive Reasoning Strategies, Part A (Norris & Ryan, 1987a), while their achievement in science was measured using final grades in general science, biology, chemistry, physics, and earth science. A path analysis was performed to determine the strength of causal relationships among the variables.

Students reason using their existing knowledge to derive conclusions and new knowledge. In addition, it has been clearly demonstrated (Rubba, 1976; Cotham & Smith, 1981; Aikenhead, 1987) that students' conceptions of scientific knowledge vary. It is plausible to infer, then, that knowledge viewed differently may be used differently in reasoning, affecting the quality of reasoning, the type of reasoning strategies used, and the conclusions drawn.

Isolating the role that epistemological beliefs have on students' reasoning could have a profound effect on curriculum design and instruction. Our present science curricula place very little emphasis on epistemological issues (Perkins & Simmons, 1988). If it can be shown that accurate conceptions of scientific knowledge lead to improved reasoning, and to improved science achievement, then the motive to develop science curricula with an epistemological focus would be stronger.

Statement of the Problem

The present study addresses five specific research questions:

1. What conceptions of scientific knowledge are held by high school students?
2. What inductive reasoning strategies are used by high school students?
3. What is the relationship between students' conception of scientific knowledge and their inductive reasoning ability?
4. What is the relationship between students' conception of scientific knowledge and their achievement in science?
5. What is the relationship between students' inductive reasoning ability and their achievement in science?

The major hypotheses are: (a) students who have a deeper understanding of the nature of scientific knowledge are better inductive reasoners; (b) students who have a deeper understanding of the nature of scientific knowledge will attain higher science achievement scores; and (c) students who are better inductive reasoners will attain higher science achievement scores. Justification for proposing these hypotheses is outlined in the following sections.

Educational Considerations

Important goals of science education include developing in students an adequate conception of the nature of scientific knowledge and proficient reasoning ability. There is considerable evidence to suggest that our present science curricula are incapable of accomplishing these goals. It is well documented that students have misconceived notions of the nature of scientific knowledge (Cooley & Klopfer, 1963; Aikenhead, 1973; MacKay, 1971; Rubba, Horner, & Smith, 1981; Welch, 1981; Aikenhead, 1987; Carey, Evans, Honda, Jay, & Unger, 1988). Likely contributors to the development of these misconceptions are current teaching practices and curriculum materials that misrepresent science and convey a stereotyped image of scientific knowledge production. The notions of once-and-for-all disproof, indubitable observation, precise definition, and perfect precision, among others, pervade our science texts in the guise of the normal practice of science. The step-by-step scientific method is a common sight in high school science texts, even though most philosophers of science agree that there is no such scientific method (Norris, in press). Hodson (1986b) states:

> The failure of modern science courses to achieve fully some of their declared goals in relation to children's

> understanding of the nature of science is due, in part, to a degree of confusion in the philosophical position underpinning many contemporary curricula and, in part, to the continuing failure to provide teachers with an adequate understanding of the basic issues in the philosophy of science and their importance in the design of learning experiences. (p. 222)

Also, numerous authors have reported lack of adequate reasoning skills in students. Perkins, Allen and Hafner (1983) characterized learners as "make sense epistemologists". Many learners analyse a problem situation only to the point where it makes superficial sense. They tend not to reflect on their thought processes nor to consider alternate ways of creating a solution. Schoenfeld (1985) describes examples where students perform meaningless calculations on a problem while paying no attention to whether their approach is justified, or even to whether progress is being made.

The present curriculum encourages such poor reasoning processes. Wasserman (1984) states that we succeed in developing students good at performing hundreds of school exercises that require single, correct answers. These same students have problems with tasks that call for imagination, for suggesting hypotheses, or for taking any cognitive or creative risks. The recitation method dominates most schools and it is the major approach in most teachers' repertoires.

With so little emphasis on the process of thinking, the students who succeed in the present system are not necessarily those who have the greatest understanding, but those who use their memory best. Students frequently get good grades in science with no more than a superficial understanding of its concepts and relationships, and therefore don't have the ability to use these in the real world.

Problem-solving as portrayed in school curricula is much too mechanistic. Teachers and textbooks normally (a) list rules, (b) provide examples and (c) have students apply the rules in a context where they are told which rules apply where. Even in novel situations, decisions that should require thought become automatic. Similar beliefs are echoed by many researchers such as Baron (1985), Freire (1974), Perkins (1985), Perkins and Simmons (1988), and Stice (1987).

If we are to accomplish our educational goals of developing more appropriate conceptions of scientific knowledge and competent reasoning ability, we must become familiar with those variables that influence the attainment of these goals. Our science curricula should address not only <u>what</u> is known by science, but should also address <u>how</u> science has come to arrive at such knowledge (Duschl, 1988). We must discover the relationships between views of what knowledge is and where it comes from, and how that knowledge is used in the reasoning process.

Rationale of the Study

In the fields of educational research and philosophy there is mounting support for the idea that higher order thinking skills are linked to epistemological beliefs. Despite the fact that there have been very few experimental studies in the area, there is considerable theoretical support from such authors as Posner, Strike, Hewson and Gertzog (1982), Perkins and Simmons (1988), and Newmann (1988).

In proposing their model of conceptual change, Posner and his colleagues emphasize epistemological beliefs as an influencing factor in the conceptual change process. They claim that the degree to which one analyses the nature of evidence, understands the importance of parsimony in a theory, and believes in the orderliness of nature, affects the way in which one accepts conceptual change.

Perkins and Simmons (1988) identify four "frames" or categories of student understanding. The _content_ frame deals with the facts and definitions of the subject matter. The _problem-solving_ frame contains problem solving strategies, beliefs about problem solving, and processes that help the learner stay organized during the problem solving process. The _inquiry_ frame is comprised of knowledge and attitudes necessary for challenging and extending the knowledge in a domain. Finally, the _epistemic_ frame "focuses on the general norms having to do with the grounding of the concepts and constraints in a domain" (p. 311). "Facts" are considered valid by the standards established in the epistemic frame.

In our science curricula, students have little opportunity to learn about the inquiry or epistemic frames or about the interrelationships between the different frames. Students mainly receive instruction in content. Perkins and Simmons state:

> Real understanding consists in a web of relationships that connect with content knowledge but also with knowledge in the problem-solving, epistemic, and/or inquiry frames. Failure to recognize this web of interrelationships leads to instruction that allows and even exacerbates the naive and ritual patterns of misunderstanding. (p. 323)

Newmann (1988) concurs that thoughtfulness involves attitudes, personality traits, and general values and beliefs about the nature of knowledge. He explains that normally there is resistance on the part of the learner to the development of higher-order thought processes. In order to generate student engagement in tasks requiring higher-order processes, this resistance must be addressed openly. He reports that several researchers have found that students prefer passive well-defined roles with simple, mechanistic answers and an absence of mental conflict (McNeil, 1986; Powell, Farrar, & Cohen, 1985; Willis, 1977). This "lower-order mindset" is characterized by beliefs that include the following: (a) most knowledge is certain, (b) knowledge is created by outside authorities, (c) knowledge is to be learned as quickly as possible, and (d) knowledge may seem counter-intuitive or mysterious with respect to one's experience, but should be believed anyway.

Schommer (1989) provides experimental evidence to show that students' beliefs about the nature of knowledge have distinct effects on their reading comprehension. She concludes that epistemological beliefs affect the critical interpretation of knowledge and thus, to some degree, determine how we draw conclusions from information. Specifically, poorly established epistemological beliefs in the areas of speed of knowledge acquisition and certainty of knowledge result in conclusions that are oversimplified or inappropriately absolute. Schommer agrees that epistemic concerns should be recognized explicitly. Raising students' consciousness about their own beliefs, and teaching them how these views influence their learning, has the potential for far-reaching effects in the field of education.

Theoretical Framework

Conceptions of Scientific Knowledge

People's conceptions of scientific knowledge are comprised of beliefs about the source and status of that knowledge. Statements such as "A scientific statement is true if most scientists believe it", or "Observation in science is influenced by opinion" can constitute part of a conception of scientific knowledge. There is no single universally accepted view of the nature of scientific knowledge among scientists and philosophers of science since there are many issues on which different positions are held. Several of these issues and the positions generally considered most accurate are described below:

(1) The fallibility of observation. The predominant view is that observations are based on inadequate sense experience. Observations are based upon fallible prior knowledge and must be interpreted in the light of current theoretical beliefs (Hanson, 1958; Hodson, 1986a; Norris, 1984).

(2) The role of human creativity and imagination. The view of Popper (1972) is that concepts and theories are products of creative minds.

(3) The role of inductive generalization. The predominant view is that inductive generalization is inadequate as a description of the process of scientific knowledge production (Popper, 1972).

(4) The nature of change in scientific knowledge. The consensus within the philosophic literature is that scientific knowledge is tentative. Concepts and theories change and develop over time while some are discarded (Conant, 1951; Kuhn, 1962; Schwab, 1960).

(5) The nature of the scientific method. The dominant view is that there is no single universally followed, step-by-step

"scientific method" (Black, 1954; Conant, 1955).

(6) The testability of scientific knowledge. Hempel (1966) states that scientific hypotheses must be subjectable to empirical test.

Other issues relating to the role and status of theoretical knowledge and the role of the community of scientists in scientific practice tend to be more contentious. Regardless, student conceptions or misconceptions of science must envelop these and other philosophical issues. The ideas have direct bearing on science curricula and are taught either explicitly, or, as some researchers have suggested, implicitly, through language and behaviour (Herron, 1977; Lederman, 1986b; Munby, 1973, 1976; Zeidler and Lederman, 1989). Instruments constructed to measure students' conceptions must model the diversity of views of these different epistemological issues.

Inductive Reasoning

There have been numerous definitions of inductive reasoning put forward in the past four decades. Guildford and Lacey (1947) defined it as the ability to see trends and relations. French, Ekstrom, and Price (1963) described induction as "the finding of general concepts that will fit sets of data; the forming and trying out of hypotheses." (p. 19). More recently, Colberg, Nester, and Cormier (1982) state that induction should be defined as a non-demonstrative argument in which the conclusion does not necessarily follow from the premises. Skyrms (1975) distinguishes between deduction and induction in the following way:

> An argument is deductively valid if and only if it is impossible that its conclusion is false while its premises are true. An argument is inductively strong if and only if it is improbable that its conclusion is false while its premises are true. (p. 7)

According to Ennis (1987) inductive reasoning is one component of critical thinking. He defines critical thinking as "reasonable reflective thinking that is focused on deciding what to believe or do" (p. 10), and he has proposed a comprehensive model for the concept. In his model, thinking starts from a basis of information. Reasoning is carried out on the information leading to some decision or conclusion. The reasoning process, or inference, is of three main types: deduction, induction, or value judgement. Inductive inferences are inferences that generalize information or inferences that explain information. Since conclusions derived from reasoning are inferred from information, the process of inference has a fundamental role in critical thinking (Norris, 1988b). Ennis's conception of critical thinking and the place of inductive reasoning in it has been adopted as part of the conceptual framework for this study.

Methods

The instrument used to measure students' epistemological beliefs is a questionnaire that was developed for the study. The questionnaire consists of 56 Likert-scale items compiled in seven subscales. Each of the subscales is organized around a specific dimension of scientific knowledge. These dimensions are: (a) scientific truth, (b) the fallibility of scientific knowledge, (c) the changeability of scientific knowledge, (d) the role of creativity in scientific knowledge production, (e) the testable nature of scientific knowledge, (f) the speed of scientific knowledge acquisition, and (g) the role of authority in scientific knowledge acquisition. The subscales are designed so that opposite ends of each scale represent alternate views or conceptions of that particular dimension of scientific knowledge.

The Essay Test of Inductive Reasoning Strategies, Part A (Norris & Ryan, 1987a) has subjects imagine they are on an unknown planet where their task is to use various clues to explain how they would search for living creatures. The test requires subjects to actively employ inductive reasoning strategies. Their reasoning is then measured against an ideal model for the situation. Subjects are graded according to their use of the following strategies: (a) taking all relevant information into account; (b) seeking more information when it is appropriate; (c) generating alternative conclusions, explanations and hypotheses; (d) withholding judgement when appropriate; (e) monitoring the progress of one's reasoning; (f) handling complex problems in an organized manner; and (g) keeping focused to the main point. The essay test format allows the scorer more accurately to assess the process of reasoning since it is more apparent how conclusions are derived than on other types of pencil-and-paper tests.

Student achievement in science is measured using science grades from the previous academic year. Grades in five subject areas are used: general science, biology, chemistry, physics, and earth science.

Both the epistemological questionnaire and the inductive reasoning test were administered to a group of 305 high school students in grades 10-12. Pearson product-moment correlations are used to relate variables while the strength of causal relationships are examined using path analysis.

Summary

The present study examines the interrelationships among the variables of students' conceptions of scientific knowledge, inductive reasoning ability, and achievement in science. Failure of the present curriculum to produce adequate student conceptions of the nature of scientific knowledge and proficient reasoning ability provide motivation for the study.

The basic assumptions underlying the study are: (a) knowledge conceived differently will be used differently in the reasoning process, and (b) adequate conceptions of scientific knowledge will result in better reasoning and better achievement in science. An epistemological questionnaire developed for the study is used to measure conceptions of scientific knowledge, while the Essay Test of Inductive Reasoning Strategies, Part A (Norris & Ryan, 1987a) is used to measure inductive reasoning ability. Student science achievement is assessed using final grades from the previous school year.

CHAPTER II: REVIEW OF RELATED LITERATURE

In order to understand the relationship between students' conceptions of scientific knowledge and their reasoning ability, I have chosen to examine literature in three areas. The review starts with a discussion of research related to students' conceptions of scientific knowledge. Variables correlating with students' conceptions will be examined and instruments used to measure students' conceptions will be evaluated. Barriers to research in the area will also be discussed. The second part of the review deals with current research in the area of inductive reasoning ability. Variables correlating with this ability will be examined. Tests to measure inductive reasoning ability will be described and assessed. The third and final section of the review will describe recent research that has attempted to relate student epistemology and cognitive skills. Research in this area, however, is quite limited with only several studies being carried out, all in recent years.

Student Conceptions of the Nature of Scientific Knowledge

Developing an understanding of the nature of science and scientific knowledge has long been an important objective of science educators. In recent years both the National Science Teachers Association in the United States and the Science Council of Canada have expressed renewed interest in the educational goal of developing scientifically literate students with a greater understanding of the nature of science (NSTA, 1982; Science Council of Canada, 1984).

Appraisal of Student Conceptions

Students' conceptions have been measured on numerous occasions with students of various ages. Results indicate typically that students possess inadequate conceptions of the nature of science (Cooley & Klopfer, 1963; MacKay, 1971; Rubba et al., 1981; NSTA, 1982; Aikenhead, 1987; Carey et al., 1988).

One of the most comprehensive appraisals of students' conceptions was conducted by Welch as part of Project Synthesis (Welch, 1981). This project, funded by the National Science Foundation, was a joint effort of 23 researchers representing a wide variety of roles and perspectives within the science education community. The purpose was to portray the state of science education in the United States in the late 1970's (Kahl & Harms, 1981). Welch concluded that although there was some cognizance of the nature of scientific inquiry, there was a general lack of in-depth student understanding.

More recently, in an extensive study by Aikenhead (1987), over 10,000 students were asked to make a written reaction to a statement concerning a science-technology-society topic and to write a paragraph explaining the reasons for their response. The statement was one of 46 statements used on Views on Science-Technology-Society (Aikenhead, Fleming, & Ryan, 1987). Aikenhead found that students generally hold contradictory beliefs about scientific knowledge. For example, almost 100% of a subsample of 236 students believed scientific knowledge is tentative, but their reasons were varied. Most students (45%) viewed the tentativeness of scientific knowledge from a reconstructionist position. That is, new knowledge replaces old. Many students (20%) viewed scientific knowledge as tentative strictly in the cumulative sense. Scientific knowledge does not change but is added to. Others (20%) believed that changes to scientific knowledge were a result of technological advances. Aikenhead noted that students did not have uniform meanings for the frequently used terms "scientific fact" and "scientific method" nor did they seem to be aware of outside influences on scientific knowledge.

Carey et al. (1988) described the epistemological stance of 76 junior high students as believing that knowledge is acquired passively and is a faithful copy of the world. They believed the inquiry process is limited to observing nature rather than constructing explanations of the phenomena of nature.

Factors Influencing Development of Student Conceptions

With few exceptions, experimental research has not produced evidence that indicates the source of student conceptions. Generally researchers have been concerned with what the specific student conceptions are but not with how these conceptions have come to be adopted. The basic assumption, as first described by Robinson (1969), is that through the normal discourse of teaching, teachers' conceptions are adopted by their students. While seeming logical, this has not been demonstrated conclusively in research. If the assumption is true, there are two relevant questions that need to be answered: (a) What is the origin and degree of adequacy of teacher conceptions? and (b) How do teachers transmit their conceptions of scientific knowledge to their students?

Philosophy and teacher conceptions. In his seminal article, Robinson (1969) declared not only that teachers' conceptions of the nature of science are an important force in shaping their classroom behaviour, but he also asserted that the teacher training of the time did not provide the necessary philosophical background to develop a philosophy of science teaching consistent with the nature of scientific knowledge. During the last two decades, numerous authors, following Robinson's lead, have claimed that the relationship between philosophy of science and science education should be examined. Martin (1972) stated that the relevance of the philosophy of science to science education has been unexplored, but philosophy of science can help science educators in their educational practice. Ennis (1979) noted that problem areas in science education such as the nature of the scientific method and the part played by scientists' value judgements could profit from research in the philosophy of science. Abimbola (1983) recognized that important goals of the philosophy of science are directly related to the science curriculum, including how scientific knowledge is established and validated, and how it eventually changes form and meaning.

Other authors have concluded that teachers' conceptions of scientific knowledge are inadequate since they are grounded in inappropriate philosophical beliefs. A report published by the Association for Science Education in the United Kingdom (Association for Science Education, 1979), states that most science teachers are products of an education system that places emphasis on scientific knowledge while neglecting the history and philosophy of science. As a result, science teachers have a scant understanding of the nature of science itself.

Elkana (1970) claims that teachers' philosophical views lag contemporary views by as much as twenty to thirty years. Hodson (1988) agrees that teaching practices in the classroom are impeded by teachers who operate under principles of science that philosophers have long considered inadequate. Specifically, he criticizes the commonly held inductivist view that from simple unbiased observations students can inductively spawn scientific generalizations. He contends that scientists bring speculation and opinion to observation and that new knowledge must be anchored to the learners' prior knowledge. Thus, the assumption that unbiased observations lead automatically to infallible conceptual explanations is unfounded.

Duschl (1988) claims that school science is dominated by the authoritarian view that scientific knowledge is considered absolute and final. The source of this view, he believes, is rooted in the outdated philosophy of logical positivism developed during the first half of the twentieth century. Duschl maintains that a more accurate view of the nature of scientific knowledge must attend to humanistic and social issues in addition to the facts of science.

Two recently conducted ethnographic studies have provided insight into influences on teachers' conceptions of scientific knowledge. Gallagher (1991) carried out a study of 25 preservice secondary science teachers enrolled in his university methods course. He concluded that since there is no formal teacher education in history, philosophy, or sociology of science, teachers' knowledge is limited to the body of science. Science textbooks used in teacher training reinforce the role of teaching content knowledge to students, and many instructors act as presenters of factual knowledge. The result is that scientific training has not provided much understanding of the process by which scientific knowledge is formulated.

In a study involving 13 preservice teachers, King (1991) found that most teachers believed that the history and philosophy of science should be taught in their preservice programs. The teachers interviewed did not know how they would teach topics on the nature of science and perceived they did not have enough knowledge to do so. King quotes one student:

> I learned science as a collection of facts, with no knowledge of how these facts came to be facts, or why these facts are considered facts. When I talk about teaching my students to think critically, I guess what I mean is that they have that historical and philosophical knowledge so that they understand and can appreciate the hows and whys. I wonder how I'll be able to teach this way, given my shallow knowledge of science. The only thing I feel I'm prepared to do now is to teach my students the facts I learned. (p. 139)

Despite this literature, there are authors who believe teachers' conceptions are changing and may not be as inadequate as some researchers believe. In a study of preservice teachers' understanding of the nature of science, Anderson, Harty, and Samuel (1986) compared the responses of a group of 24 preservice teachers on Kimball's <u>Nature of Science Scale</u> (NOSS) in 1969 to a group of 21 preservice teachers in 1984. The 1984 group responded closer to the model response provided by Kimball than did the 1969 group. The authors conclude that teachers entering the field today have a more accurate conception of the nature of science.

A closer look at this study reveals, however, that these results must be interpreted very cautiously. The sample sizes for both groups were very small. The groups were not equivalent in age, sex, or bachelor's degrees held. The 1984 group was older, better educated, and had a higher proportion of females. The grade point averages of the two groups were not compared. The model of the nature of science on which NOSS is based was developed over twenty years ago. The test is based on debatable notions such as curiosity, regardless of outcomes or applications, being the only driving force for the generation of scientific knowledge, and scientists' being uniquely open-minded and willing to change in the face of new evidence. It is doubtful that performing well on this test means one necessarily holds a contemporary philosophical view.

Lederman (1986a) and Lederman and Zeidler (1986) studied 18 high school biology teachers and their 409 students to determine student and teachers' conceptions of the nature of science and to compare their conceptions to an "adequate conception". The instrument used was the <u>Nature of Scientific Knowledge Scale</u> (Rubba, 1976). Their results showed that teachers scored higher than students on every subscale of the instrument and that students scored higher than the neutral response on each subscale. They concluded that teachers and students do have adequate conceptions of the nature of scientific knowledge. This conclusion contradicts most research in the area. Lederman and Zeidler point out that even though teachers may possess adequate conceptions, this does not necessarily mean that they will exhibit teaching behaviours that will increase students' understanding.

Teaching behaviour and practices. Over twenty years ago, Robinson (1969) focused attention on teacher practices necessary for teaching how scientific knowledge is constructed. He specifically noted the importance of teacher language and the role of the science laboratory in developing students' conceptions of scientific knowledge. Since that time there has been relatively little research with the goal of determining the effect of teaching behaviour and practices on student conceptions. In the last several years there have been a number of studies that have attempted to isolate teaching behaviour, curriculum materials, classroom variables, and teacher language as influences on student conceptions of the nature of science. These vary in research methodology and have provided somewhat inconsistent results. Several of the more ambitious projects have been carried out by Norman Lederman and his colleagues. Lederman's work has concentrated on the dynamics of teaching behaviour and classroom variables as influences on student conceptions. Other studies, taking an ethnographic approach, have focused on the effect of curriculum materials and teachers' instructional decisions on students' conceptions.

With a sample of 18 teachers and 409 high school students, Lederman (1986b) and Lederman and Druger (1985) attempted to relate specific classroom variables and teaching behaviours to changes in student conceptions. Their methodology consisted of five basic procedures. First, they administered the Nature of Scientific Knowledge Scale (NSKS) (Rubba, 1976) as a pretest at the beginning of the school semester. Second, throughout the semester, intensive qualitative classroom observations were made. Third, teachers and students were given a posttest of the NSKS. Fourth, classroom variables were derived by a systematic qualitative comparison of teachers and classes ranked as generally scoring high on the NSKS with those generally scoring low on the NSKS. The final procedure was a quantitative analysis of the classroom variables to determine if the variables statistically discriminated between the high scoring teachers and classes and the low scoring teachers and classes.

Results of the studies indicated that teachers possessed conceptions generally considered to be "more adequate" than those of the students. Teachers scoring highest on the NSKS were more supportive and dynamic than low scoring teachers. Also they questioned more, related content to students' personal lives, placed less emphasis on rote memory, and assigned less seatwork than teachers who scored low on the NSKS. Teachers' conceptions were not found to be significantly related to changes in students' conceptions.

Zeidler and Lederman (1989) used the same sample and a similar procedure to examine Munby's (1973, 1976) proposition that implicit messages hidden in teachers' language influence student conceptions of the nature of science. From qualitative classroom observations, the authors categorized teachers on the basis of their language as holding the Realist or Instrumentalist conception of the nature of science. Their results indicate that teachers' language reveals implicit conceptions of the nature of science that can be conveyed subsequently to students. Thus, the ordinary language teachers use can provide the context in which student conceptions are developed.

Lederman and O'Malley (1990) attempted to identify the sources of students' beliefs about the tentativeness of scientific knowledge. Students completed an open-ended questionnaire followed by an interview in which the subjects were asked to state the sources of their beliefs or elicit descriptions of those experiences that altered their beliefs. The researchers found that when students were asked about the sources of their beliefs or experiences that led to changes in their beliefs, they were unable to identify either. The authors concluded that this inability to identify sources indicates that an understanding of the nature of science may be learned implicitly. Further support for this idea is provided by Herron (1977) and Lederman (1986b).

In contrast to Lederman's position, Gallagher (1991) and King (1991) suggest that secondary science teachers' conceptions play a fundamental role in forming the image of science held by students. As well, Mitman, Mergendoller, Marchman, and Packer (1987) describe factors such as: (a) curriculum materials, (b) the nature of the teacher's instructional activities, and (c) the teacher's verbal and written instructions explicitly referring to the nature of science, as influencing student conceptions of the nature of science.

These authors contend that ethnographic procedures are required to isolate the effect teachers' beliefs have on the learning environment. Criticism that can be made of Lederman's approach include the high level of inference required in the design and employment of observational instruments and a non-recognition of the effect of the curriculum on teacher decisions and practices. Mitman et al. (1987) point out that it is especially difficult to design an instrument that can isolate students' perceptions of classroom events independent of their basic feeling of the class.

Attempting to isolate variables of teacher behaviour and practice and to relate them to student conceptions is both difficult and intriguing. It appears that the variables of teacher behaviour, classroom practice, curriculum, and language form a complex mix from which the learner may be influenced to varying degrees either explicitly or implicitly.

Available Instruments to Measure Students' Conceptions of Scientific Knowledge

To carry out the present study, an instrument to measure student conceptions of scientific knowledge had to be found or developed. An extensive search of the literature revealed only one instrument that dealt specifically with appraising student conceptions of scientific knowledge: The Nature of Scientific Knowledge Scale (NSKS) (Rubba, 1976). Several other instruments attempt to measure student conceptions of the general area of the nature of science.

What follows here is a description of those instruments examined as part of the selection process. The evaluation of conceptions of scientific knowledge is an essential component of all the instruments described. The instruments have for the most part been widely used and accepted in the research field.

Test on Understanding Science (TOUS). Perhaps the best known measure of conceptions of the nature of science was developed by Cooley and Klopfer (1961a). The TOUS is a four-alternative 60 item multiple-choice test. Items belong to one of three areas or subscales:

Subscale I. Understandings about the scientific enterprise (18 items).

Subscale II. The scientist (18 items).

Subscale III. Methods and aims of science (24 items).

Items from subscale III relate most to student conceptions of scientific knowledge. There are eight themes in this area for which items were developed. The themes are: (a) generalities about scientific methods, (b) tactics and strategies used in the process of science, (c) theories and models, (d) the aims of science, (e) accumulation and falsification, (f) controversies in science, (g) science and technology, and (h) unity and interdependence of the sciences.

The test was validated by an analysis of scientists at work and references to the history and philosophy of science literatures. Practitioners in the fields of science, science education, teaching, and history and philosophy of science were consulted regarding the content of items. One of the preliminary forms of TOUS was field tested. The overall Kuder-Richardson formula 20 reliability coefficient was found to be 0.76 (Cooley & Klopfer, 1961b).

Science Process Inventory (SPI). The Science Process Inventory, Form C is a 150 item, forced-choice, agree-disagree, scale developed by Welch (1966). The inventory is intended to appraise knowledge of the process of science possessed by secondary school students. Specifically, the inventory assesses students' "understanding of the methods and processes by which scientific knowledge evolves" (Welch and Pella, 1968, p.64). A more recent version of the SPI, Form D, contains 135 items with which subjects must agree or disagree (Welch, 1969).

The theoretical base for the inventory was a list of elements of the scientific process derived from the writings of Beveridge, Conant, Kemeny, Lachman, Nash, and Wilson. The list consisted of congruous elements which appeared in three or more of the six reference books. The elements were then presented to a panel of research scientists for validity judgement. The resulting outline became the basis for developing the instrument.

Content validity was provided by expert opinion on the items as well as an item analysis of the results of a field trial using 380 high school students. The reliability of the SPI, as provided by a sample of 1283 high school students, was found to be 0.79.

Wisconsin Inventory of Science Processes (WISP). A similar inventory to the SPI, the Wisconsin Inventory of Science Processes consists of 93 items that subjects are to describe as accurate, inaccurate, or not understood. The scale was constructed by the Scientific Literacy Research Center at Madison, Wisconsin (1967).

The content of WISP is very similar to the SPI and is based on a set of assumptions that include:

1. The universe and its natural phenomena are real and intelligible. Relationships are consistent and causal.
2. The products of science are amoral, repeatable, parsimonious, tentative and probabilistic.

There is no information provided on the validation of the scale. A reliability coefficient of 0.82 was reported for an unidentified sample of grade 12 students (Aikenhead, 1973).

Nature of Science Scale (NOSS). Kimball (1967) developed the Nature of Science Scale to determine whether or not science teachers have the same view of science as scientists. The scale consists of 29 position statements for which subjects may select one of three responses: agree, disagree, or undecided. A model response for each item was obtained by preparing a model of the nature of science based on the philosophical views of Conant and Bronowski. The model was composed of eight of the following assertions about characteristics of science:

1. The fundamental driving force in science is curiosity concerning the physical universe;

2. Science is a dynamic, on-going activity, rather than a static accumulation of information;
3. Science aims at ever-increasing comprehensiveness and simplifications using mathematics as a simple, precise method of stating relationships;
4. There is no one scientific method, but as many methods as there are practitioners;
5. The methods of science are better characterized by some value-type attributes such as dependence on sense experience, the need for reproducibility and insistence on operational definitions rather than by techniques;
6. A basic characteristic of science is a faith in the susceptibility of the physical universe to human ordering and understanding;
7. Science has a unique attribute of openness: both openness of mind and openness of the realm of investigation;
8. Tentativeness and uncertainty are characteristic of all science.

Content validity was checked by a panel of nine science educators who judged whether the items were related to the model. Kimball established construct validity by the test's ability to discriminate between college graduates who were science majors and those who were not. The split-half reliability of the test found to be 0.72.

Nature of Scientific Knowledge Scale (NSKS). This 48 item Likert-type scale was developed by Rubba (1976) to assess secondary students' understanding of the nature of scientific knowledge. The NSKS is based on the characterization of scientific knowledge developed by Showalter (1974). This conception describes scientific knowledge as (a) amoral, (b) tentative, (c) expressing creativity of scientists, (d) parsimonious, (e) testable, and (f) unified. The test consists of an eight item subscale on each of these categories.

Rubba and Anderson (1978) describe the exhaustive process by which the NSKS was developed and validated. To begin, the validity of the six-factor model was established by a panel of three philosophers of science. Items were written following guidelines for construction of Likert-type scales and the reading level was established using a sample of nine sixth grade students. Items were then submitted to a panel of science education experts for evaluation of form and content and the instrument was field tested with a group of 31 high-ability high school juniors. After the first field test, item statements were submitted to a panel of experts consisting of two philosophers of science, two science educators, two scientists, two high school teachers, and a psychometrician. These experts evaluated the content validity of the item statements as compared to A Model of the Nature of Scientific Knowledge (Rubba, 1976). Finally, the instrument was field tested again with a sample of 674 high school science students.

Evidence of construct validity was gathered by comparing the NSKS results of 40 philosophy of science students with the results of 125 university freshman with no background in philosophy. Results showed the philosophy of science students had higher mean scores on five of the six NSKS subscales as well as higher overall scores.

Test-retest reliability was established with groups of 52 general science and 35 advanced chemistry students. The Pearson product-moment correlations were 0.59 and 0.87, respectively (Rubba & Anderson, 1978).

<u>Conceptions of Scientific Theories Test (COST)</u>. Cotham and Smith (1981) criticized existing tests that measure understanding of certain aspects of the nature of science as being based on single interpretations of the nature of science. These tests make the assumption that their interpretation of the nature of science is the correct one. Since many aspects of the nature of science are controversial, this assumption has been subject to criticism (Lucas, 1975; Martin 1972).

In developing the <u>Conceptions of Scientific Theories Test</u>, the authors attempted to produce an instrument that was sensitive to alternative conceptions of particular aspects of scientific theories, and from which an understanding of the tentative and revisionary conception of the nature of science could be inferred. They assume, in taking this approach, that the tentative nature of scientific knowledge is one of its least controversial aspects.

COST is an attitude inventory consisting of 40 Likert-type items. It includes four subscales which are: (a) ontological implications of theories, (b) testing of theories, (c) generation of theories and (d) theory choice. Each subscale is organized around an issue that can be polarized or represented by two alternative views or conceptions.

Content validity was established by referring to the philosophic literature of Hempel, Kuhn, Martin, and Nagel. Evidence of construct validity was gathered by two approaches: (a) administering COST to education majors, college chemistry students, and college philosophy of science students to determine the ability of the instrument to discriminate between the contrasting groups; and (b) a statistical multi-trait multi-method procedure of Campbell and Fiske (1959) in which scores for the same trait (subscale construct) are correlated using different methods of measurement (different theoretical contexts). Reliability data of COST was not provided.

The results of the test administration showed that elementary education majors, more than chemistry and philosophy majors, believed scientific theories to be (a) tested conclusively and (b) selected by objective means. Also, elementary education majors, more than philosophy majors, believed scientific theories are generated by inductive means, that is, derived from observations rather, than invented to account for them.

Views on Science-Technology-Society (VOSTS). This instrument was developed for a study carried out by Aikenhead, Fleming, and Ryan (1987) in which over 10,000 Canadian high school students were sampled for their beliefs about science-technology-society topics. Though the instrument assesses the general relationships between science, technology, and society, a category of items specifically examines the characteristics and limitations of scientific knowledge.

In developing their instrument, the authors rejected the traditional standardized instruments which I have summarized previously, in favour of an alternate approach. Instead of employing a Likert-scale, VOSTS requires subjects to give a two-part response. First, students react to a statement by agreeing, disagreeing, or saying they "can't tell". If they agree or disagree, then they write an argumentative paragraph to support or reject the statement. Subjects reply to one of 46 different statements which constitute the VOSTS instrument.

VOSTS is in part derived from the Nature of Science Scale (Kimball, 1967), Science Process Inventory (Welch, 1966), and Test of the Social Aspects of Science (Korth, 1968) incorporating portions of their philosophical models. Selection of additional items on science-technology-society was influenced by the work of Gauld (1982) and Ziman (1980).

The original version of VOSTS was field tested with an unidentified sample. Validation procedures are not specified and test-retest reliability data is not provided. Interjudge reliability is reported as 0.84.

Perhaps it may be assumed that, since VOSTS has been derived from standardized instruments, its validity rests with the validity of the philosophical base of those instruments. This view is problematic, however, as the theoretical models on which these instruments are based are subject to some debate and the authors of VOSTS themselves dismiss the previous standardized instruments as clearly inadequate.

Limitations of Existing Instruments

Despite the wide acceptance of these instruments, they have fundamental problems. The discussion that follows mainly addresses problems related to test construction and administration. As well, general conceptual limitations of the entire research area are noted.

Test construction. The most profound criticism of these instruments is that the philosophical models on which they are based and the methods by which these models are derived are inappropriate. Lucas (1975) argues that test developers should evaluate critically how they have come to accept a particular model of science for a test they are creating. He notes that most of the authors of standardized instruments examined in a review by Aikenhead (1973) simply appeal to the philosophical literature for a consensus of opinion. These authors are relying on the assumption that consensus will necessarily produce the correct model. This, despite the fact that there is still considerable debate in the field as to the nature of scientific knowledge.

Different philosophical models advocated by different researchers will result in variation, and sometimes contradictions, in the expected responses of test items. A higher total test score simply means the subject subscribes to the same philosophical view as the test developer (Lucas, 1975). Due to the lack of agreement on what constitutes the nature of science, many tests purport to measure the general conceptions of the nature of science but focus on different aspects of the domain.

In a study analyzing several instruments, Doran, Guerin, and Cavalieri (1974) found that there were very poor correlations among three instruments tested. With a sample of 300 high school students they compared results on the Nature of Science Scale (Kimball, 1967), Science Support Scale (Schwirian, 1968), and Test on the Social Aspects of Science (Korth, 1968). They found no relationships among broad topics covered on the tests and very few relationships between pairs of subtests. They concluded that each instrument seems to be measuring a different domain within the nature of science.

Tests on the nature of science have often presented an unrealistic picture of the scientific enterprise as a value-free endeavour. Scientists are often characterized as open-minded and unbiased. Lucas (1975) quotes from Kimball's (1967) model:

> Science has the unique attribute of openness, both openness of mind, allowing for willingness to change opinion in the face of evidence, and openness of the realm of investigation, unlimited by such factors as religion, politics and geography. (p. 111-112)

Lucas refutes this position by citing several examples from the history of science where prominent scientists have been less than open-minded.

Some instruments that attempt to appraise conceptions of the nature of scientific knowledge may not only test epistemology, but also reading comprehension and vocabulary. The following statements from The Wisconsin Inventory of Science Processes demonstrate the need for high school students to have ample ability in this area:

1. Inductive logic is more likely to yield valid conclusions than is deductive logic.
2. Hypotheses in science seldom have their origin in "speculative ideas," "inspired guesses," or "intuitive hunches".

Without an adequate understanding of terms such as "inductive", "deductive", "hypotheses", "speculative", and "intuitive", student responses to these items would have little meaning.

Since important decisions are to be made on the basis of test scores from these instruments, potential users must be aware of their short-comings. Results from these tests must be used with prudence. Comparisons among different samples and different tests are difficult indeed.

Conceptual limitations. Despite the abundance of literature, there continues to be major barriers to research in the test development area. The most serious is the lack of standards for defining an adequate conception of scientific knowledge. If we look to the experts we will not find a consensus. Philosophers of science do not agree on the exact nature of science and there is no universally accepted model of the nature of scientific knowledge. Educational researchers develop tests based on their own conceptions which do not necessarily correlate with tests created by other researchers. With no widely accepted standards, researchers judge conceptions to be adequate or inadequate on the basis of test scores. Scores judged best match the particular conceptions of the test developer, while alternate, but perhaps valid, conceptions are judged poorly. Thus, through this fundamental discrepancy, the validity of the whole research area is called into question.

Student cognitive development and language. Another problem with research in the area is the lack of recognition by researchers of the level of cognitive development of students whose conceptions are being evaluated. It seems that the cognitive development of even young children is assumed to be advanced enough to be able to understand abstract philosophical concepts. Is it reasonable to assume that school children have an adequate conception of anything when compared to adult experts in the field?

In a study conducted by Rubba, Horner, and Smith (1981), the authors set out to determine the degree to which grade seven and eight students believed the misconceptions that (a) scientific research reveals incontrovertible, absolute truth, and (b) that scientific theories eventually mature into laws. Their sample was 102 students attending a science fair. The instrument used was a Likert-type questionnaire developed by the authors. No reliability or validity data was provided for the questionnaire, except to say that the questionnaire was validated as it was developed. The authors found the questionnaire failed to discriminate from the neutral responses to any great degree. They concluded that the subjects did not understand the nature of science well enough to appreciate the tentative nature of scientific knowledge, and that they did not understand that laws and theories are two distinct types of explanations. From the results of this study a reasonable question might be: Should we expect students as young as those in grade seven and eight to have an understanding of the abstract nature of scientific knowledge? Ironically, the subjects of this study were high ability, science oriented students. One can assume that average ability students would have an even poorer understanding of these sophisticated concepts.

It seems apparent from recent research that students' language may lead researchers to misrepresent their conceptions. Lederman and O'Malley (1990) used a novel methodology to determine beliefs about the tentativeness of scientific knowledge. A total of 69 subjects in grades 9 - 12 completed open-ended questionnaires that categorized their view of science as tentative or absolutist. Twenty of these students participated in follow-up videotaped interviews.

Results from the questionnaire indicated that students possessed an absolutist view of scientific knowledge. However, results from the interview showed that students' language had led researchers to make critical misinterpretations about the students' beliefs. For example, students used words like "prove" to mean "providing evidence for" instead of using the word in the absolute sense. The authors concluded that responses given by students using pencil and paper tests are subject to multiple interpretations and therefore the interactive data collection of the interview is recommended when determining student conceptions. As noted previously, Aikenhead (1987) also ascertained that students have quite different meanings for common scientific terms and phrases. Munby (1982) describes this situation in which researchers assume that their perception of an item corresponds to the students' as "the doctrine of immaculate perception". In light of these findings, researchers who must, due to practical limitations, use pencil and paper tests only, should ensure that tests are soundly and validly constructed, and results from these tests should be interpreted cautiously.

An interesting finding of Lederman and O'Malley (1990) was the fact that many students admitted having no prior thoughts about the issues raised on the questionnaire. Thus, it seems reasonable to believe that some students form their conceptions of scientific knowledge at the time they are completing the test or questionnaire. If this "test effect" is the case, the assessment procedure may be actually creating the students' views it is purporting to measure.

Due to these significant problems, there is the disturbing possibility that tests on the nature of scientific knowledge are less valid than originally thought. Thus, research findings should be accepted cautiously and with considerable scrutiny.

Selection Based on Existing Instruments

Of the instruments examined, the NSKS came closest to fulfilling the theoretical requirements of this study. Its specific focus on the nature of scientific knowledge and its rigorous development process are its assets. The problem with the NSKS is that it does not address several dimensions of scientific knowledge. Specifically, the nature of scientific truth, the fallibility of scientific knowledge, the role of authority in acquiring scientific knowledge, and the speed or rate at which the knowledge is acquired, are not part of Rubba's theoretical model. Since there is evidence that these dimensions affect the reasoning process, the decision was made to develop a questionnaire to include them. A detailed description of the development of the epistemological questionnaire will appear in chapter IV.

Implications of Limitations of Existing Instruments for Epistemological Questionnaire Development

The most significant implications for questionnaire development concern the assumptions about the adequacy of various conceptions of the nature of science. Due to a lack of consensus on what constitutes adequate conceptions, no philosophical model on which items are based can be considered the only correct and valid interpretation of the nature of scientific knowledge.

The present questionnaire does not attempt to define or determine adequate conceptions. A philosophical framework that allows expression of alternate conceptions within major dimensions of scientific knowledge will be provided as the basis of the questionnaire. The study takes a neutral stance with regard to which philosophical conceptions are correct.

Results of the present questionnaire will be interpreted with cognizance of (a) the level of cognitive sophistication of the subjects, and (b) the possibility of multiple interpretations of the questionnaire statements. It is assumed that high school students do not have a sophisticated understanding of the philosophical concepts related to the nature of scientific knowledge and that their understanding of philosophical terminology is limited. It is also assumed that multiple interpretations of items are likely, and as such, measures will be taken to limit them, including the completion of a pilot version of the questionnaire and the use of subjects' comments about the questionnaire statements and wording.

Inductive Reasoning Ability

Inductive reasoning ability has been studied since the early stages of intelligence testing resulting in a rather extensive body of research. Ropo (1987) in his review describes induction as having been studied from two different perspectives. One major focus has been on concept formation including strategies and errors in hypothesis formation. Bruner, Goodnow, and Austin (1956) were among the first to begin this type of work. A second approach uses intelligence tasks such as classification, analogy, and series completion to analyse cognitive processes. This approach was introduced by Cronbach (1957) and later reinitiated by Glaser (1972). The first perspective most closely matches the methodology of this study.

Variables Correlating with Inductive Reasoning

Comparisons of the inductive reasoning skills of different age groups have generally found improved reasoning ability for older children. Sternberg and Rifkin (1979) discovered differences in solution times of analogy problems when they compared individuals 8, 10, 12 and 19 years old. Older children were able to manage the tasks more quickly than the younger children. They also found qualitative differences, with younger subjects using different processes than older subjects. Similar results were obtained by Sternberg and Nigro (1980) who, in addition, noted that process differences could be generalized to two separate performance levels; one characterizing the performance of 9 - 12 year-olds and the other characterizing 15 - 18 year olds. Goldman, Pellegrino, Parseghian, and Sallis (1982) reinforced these results when they found 10-year-old children were more accurate in inference and application processes than 8-year-olds.

Inductive reasoning differences in individuals of the same age have been largely studied by comparing skilled and less skilled reasoners (Ropo, 1987). Skilled reasoners process faster (Sternberg, 1977; Mulholland, Pellegrino, & Glaser, 1980) and are more accurate in executing the reasoning process (Alderton, Goldman, & Pelligrino, 1982; Whitely, 1980).

Pelligrino and Glaser (1982) found that differences in speed and accuracy of problem solving were due to individuals' memory management or knowledge base. They discovered several interrelated factors that seem to distinguish between good and poor reasoners. Poor reasoners lose information through poor memory management, have less knowledge of task restraints, and have lower-order conceptions of numbers and mathematical rules.

Gettys and Engelmann (1983) in their research on generating plans and hypotheses found that subjects typically fall short in their efforts to explore hypotheses fully or to explain thoroughly a situation. Subjects seem to overrate the extent to which they exhaust an issue and thus stop prematurely. Perkins et al. (1983) characterized these learners as "make-sense epistemologists". That is, they analyze a situation only to the point where it makes superficial sense. Bereiter and Scardamalia (1985) also addressed this phenomenon by concluding that students access only a fraction of their knowledge that relates to a specific topic.

Perkins (1985) carried out a study on informal reasoning ability involving 320 high school students, college students, non-students with a high school diploma, and non-students with a bachelor's degree. Even though the area of informal reasoning is somewhat broader than that of pure induction, the results raise some intriguing questions. Subjects participated in interviews in which they were presented with several issues about which they were to develop conclusions and provide arguments to support their conclusions. Through the interviews an estimate of the subjects' level of background knowledge was ascertained. After the interviews all subjects were given IQ tests.

Perkins concluded IQ was the most significant variable influencing informal reasoning ability. Prior knowledge was considered to be of minor significance. Contrary to the studies mentioned previously, Perkins found age to have no significant impact. When one cautiously combines the results of all these studies it appears that age is much more significant when dealing with younger subjects. As one might expect due to maturational factors, when subjects approach adulthood the differences in reasoning ability appear to be much less dramatic.

Perkins' most startling finding was that the number of years of education the subject completed was only borderline significant and seemed to have little effect on reasoning ability. If these results are valid, this study has widespread ramifications for the educational field. Should we continue to support a system of higher education that does little to make individuals better thinkers?

Tests of Critical Thinking and Inductive Reasoning

The design of the present study necessitates use of an instrument that accurately appraises inductive reasoning strategies. Several instruments, most containing induction as a component of critical thinking, were examined to determine their suitability for this study. What follows is a brief description of these instruments' characteristics and use. For an extensive analysis of most tests of critical thinking and their components, the reader should refer to Ennis and Norris (1989).

Watson-Glaser Critical Thinking Appraisal - Forms A and B. Developed by Watson and Glaser (1980), this 80 item test is perhaps the most widely used of all critical thinking tests. It contains sections on inference, assumption identification, deduction, interpretation, and argument evaluation. For the section on inference, the subject is to select one of five points on a true-false continuum. For the remaining sections, subjects are to select one of two choices: that an assumption was made or not made, that a conclusion follows or does not follow, or that an argument is strong or weak. Test reliability data include a test-retest estimate of 0.73 determined with a sample of 96 college students. Evidence of the test's validity was drawn from several unnamed studies.

New Jersey Test of Reasoning Skills. This test is a 50-item multiple choice test of reasoning and inquiry skills developed by Shipman (1983). The target range of the test is from grade 5 to college level. The test deals with topics including assumption identification, induction, classical syllogism, and the meaning of categorical statements. Reliability estimates of 0.84 and 0.91 are provided for samples of grade five and seven students. Information related to the test's validity is not provided.

Cornell Critical Thinking Test, Level X. This 71 item multiple-choice test was created by Ennis and Millman (1985). It is designed to be administered in a 50 minute period and includes sections on induction, credibility, observation, deduction, and assumption identification. The test is designed for use in elementary, junior high, high school, and beginning years of university. Reliability estimates using Spearman-Brown and Kuder-Richardson formulas range from 0.67 to 0.90. Content was validated by members of the Illinois Critical Thinking Project.

Cornell Critical Thinking Test, Level Z. This version of the Cornell test contains 52 items and has a target range from gifted high school students to students of college level. Its use of difficult language makes it inappropriate for use with elementary or junior high students.

Reliability estimates on the Level Z range from 0.50 to 0.77.

Ennis-Weir Critical Thinking Essay Test. This test was developed by Ennis and Weir (1985). It is a standardized essay test whose target range is grade 7 through college. Essays are graded according to criteria which include getting the point, recognizing reasons and assumptions, seeing other possibilities and explanations, and responding to overgeneralization, credibility problems, and the use of emotional language to persuade. The test is most appropriate for high school and college students. Inter-rater reliability with two samples of 27 and 28 students was recorded at 0.86 and 0.82. Content validity is demonstrated by a close match between the test problems and Ennis's conception of critical thinking.

Essay Test of Inductive Reasoning Strategies. This test was constructed by Norris and Ryan (1987a). Subjects must imagine they are searching for alien creatures on an fictitious planet. They must use clues that are available to explain how they would search for living creatures. Norris and Ryan have provided a model for ideal reasoning in the situation and student reasoning is measured against that ideal. In addition to assessing strategies, the test also attempts to address the very difficult concept of measuring student dispositions. Test development is not complete and as such inter-rater reliabilities and processes of validation have not been published.

Limitations of Existing Tests

In the past decade, many authors have addressed the shortcomings of present tests of reasoning and general critical thinking ability (McPeck, 1981; Ennis, 1984; Quellmalz, 1985; Whimbey, 1985; Norris, 1988a; Norris, 1988b; Ennis & Norris, 1989; Nickerson, 1989b; Norris, 1990). This section will focus on limitations of inductive reasoning tests and will be based primarily on the ideas of Ennis (1984), Ennis & Norris (1989), Norris (1986), and Norris (1988b). It will deal primarily with test format and problems of reliability, validity, and multiple interpretations of test items and instructions.

Machine-scorable vs. essay test format. There is considerable debate among test developers as to the merits of machine-scorable and essay type critical thinking tests. Early tests were strictly machine-scorable, but more recently there is an emphasis on the essay and interview format.

Tests that are machine-scorable include those that use multiple-choice questioning, Likert-type position statements, and true-false continuums. The major advantages of these types of tests is that they are much easier and economical to use. Ennis and Norris (1989) estimate that machine-scorable tests are economically superior by a factor of about 1000 when compared to a test such as the Ennis-Weir Critical Thinking Essay Test. Being quicker and easier to grade, machine-scorable tests have the added advantage of allowing larger sample sizes. To some degree machine-scorable tests can be graded more objectively than essay type tests. This is not to say that machine-scorable tests are totally objective. As Ennis and Norris (1989) point out, the choice of items on a machine-scorable test, the interpretation of the score, and the criterion set for mastery, are all at least as subjective as using an essay.

Ennis and Norris (1989) describe three problems related to the validity of machine-scorable tests. First, machine-scorable tests show only the final product of the reasoning. The process by which the answer is derived is really what the test developer is interested in, but it must be inferred that the appropriate answer choice means the correct reasoning process. But, for instance, it is not certain that the background assumptions the subject uses to reason through a problem are the same as the assumptions of the examiner. Thus the subject can arrive at the correct response for alternative reasons or at an incorrect response via appropriate reasoning processes.

Second, machine-scorable tests do not seem able to test the creative aspects of critical thinking. Frederiksen (1984) found low correlations between parallel open-ended and machine-scorable tests of hypothesis formulating.

The third problem is that machine-scorable tests do not effectively test critical thinking dispositions. Ennis and Norris (1989) define dispositions as what individuals have a tendency to do on account of their nature. They explain that individuals may have critical thinking abilities but may not be disposed to use them. An example of a critical thinking disposition might be the tendency to be open-minded. The disposition to act is difficult to evaluate effectively on any test, but particularly so on a machine-scorable test since the examiner sees only the product of the reasoning process. Ennis and Norris point out that this deficiency is particularly acute when dealing with poor responses. How can it be determined if a poor response is due to the student's lack of ability or lack of disposition to use the ability?

In addition, because of the test situation, subjects may perform the appropriate behaviour without normally being disposed to do so. Subjects may not normally be open-minded, for example, but may appear that way in a test situation if they perceive it to be the appropriate response.

Essay testing allows the examiner to see, at least to some degree, the reasoning process and the background assumptions made. The process is not fool-proof, however, as Ennis and Norris (1989) point out: "Essay tests only invite justification, but do not provide for further interaction when the student's justification itself is in need of clarification or justification" (p. 50). In the Essay Test of Inductive Reasoning Strategies, Norris (1990) notes that since students are not told what criteria will be used to grade their responses, they may not give reasons for what they write. Desired responses may be present or absent for reasons other than those inferred by the examiner. As an example, Norris points out that critical thinking dispositions may not be exhibited because the individual may not have appropriate subject-specific knowledge. Thus the examiner is still unable to clearly isolate the subject's dispositions.

The biggest drawback to using essay-type questioning, apart from its uneconomical qualities, is generally poor reliability. Since essay answers have to be interpreted by the examiner, a major concern is that different examiners are unable to score essay tests consistently. Generally, developers have attempted to develop standardized scoring formats to produce higher inter-rater correlations.

Reliability. Ennis (1984) provides an in-depth appraisal of the use of estimates of reliability for tests of reasoning and critical thinking. He points out that for practical considerations most test developers use measures of internal consistency (eg. the Kuder-Richardson formulas) as measures of reliability. He argues that items on a critical thinking test may not and perhaps should not be inter-correlated, since the ability to think critically may be heterogeneous. The degree to which critical thinking is homogeneous is the degree to which we should expect internal consistency. It is possible, as noted by Norris (1986), that a test could have a high test-retest reliability but a low measure of internal consistency. In addition, it is clear that measures of internal consistency can be manipulated by simply increasing or decreasing the number of test items. Larger numbers of items will increase the reliability estimates but not necessarily the test quality.

Validity. Establishing the validity of a test of inductive reasoning or critical thinking is very difficult indeed. Criterion-based validity depending on criteria already accepted as valid is impossible, since none exist (Ennis, 1984). Content validity is usually established by seeking a consensus among experts. Though this may be the best means available, universal agreement on such a complex topic as critical thinking or its components is very rare. In fact, one can argue that the truth does not necessarily reside with the majority.

Multiple interpretations. Ennis and Norris (1989) describe two problems especially pronounced in tests of inductive reasoning. First, induction testing depends on examinees having a set of background assumptions that are the same, or largely the same, as those of the test developer. However, a subject may base his or her response on assumptions different from those of the test developer and reasonably arrive at answers different from the response keyed correct. Second, subjects with different levels of sophistication may give different levels of support to a conclusion. In some cases, persons with a high level of sophistication may not accept conclusions that the test developer considers acceptable because they may feel they do not know enough to reach a definite conclusion. Thus, they have used correct processes to arrive at an answer they will be penalized for. Ennis and Norris (1989) describe several examples of both types of problems from the Watson-Glaser Critical Thinking Appraisal and the Cornell Critical Thinking Tests.

Other factors affecting test quality. Alternate subject interpretations of tests can also be a result of other factors. McPeck (1981) suggests the induction items on the Cornell Critical Thinking Test, Level Z are questions of reading comprehension. Clearly, if subjects do not understand terms or the meaning of a complete statement it will have a bearing on the student response.

Norris (1986) points out that subject interpretation of instructions will have a direct impact on the validity of a test. He specifically refers to the Ennis-Weir Critical Thinking Essay Test and notes that according to the test's instructions the subjects are to critique the thinking of the author of a passage. From a study completed by Norris and a colleague, it appeared that a large proportion of subjects responded to the content of the letter rather than the author's thinking processes. Students were not evaluating how someone was thinking but rather what someone was thinking, therefore calling into question the validity of the test as a whole.

Norris (1986) points out that critical thinking tests, for the most part, represent artificial situations which try to isolate single components of the critical thinking process. The degree to which critical thinking abilities in the complex real-world carry over to isolated manufactured situations is an interesting question.

Ennis and Norris (1989) note there are no clear ways to measure critical thinking dispositions or the willingness to use critical thinking abilities. A subject may respond to an item in the desired way when prompted in a test situation, but this is not a true indication of what the individual is normally disposed to do.

It is apparent that if educators are to base decisions on test scores from critical thinking tests they must do so very cautiously. They must recognize the particular limitations of test format and the limitations imposed by our present understanding of the concept of critical thinking.

Selection Based on Existing Tests

The Essay Test of Inductive Reasoning Strategies (Norris & Ryan, 1987a) was selected as the measure of reasoning skills for this study. Besides being the only test examined that tests inductive reasoning specifically, it has several features that are considered appropriate for this study. First, to complete the test, the subject is placed in a dynamic role and must actively employ inductive reasoning strategies. Second, the imaginary situation created in the test (searching for living creatures on planet Zed) requires little, if any, specialized background knowledge that will affect subjects' responses. Third, the essay type test allows the examiner to infer with greater accuracy the mental processes being used by the subject. A detailed discussion of this test will appear in the next chapter on methodology.

Epistemology and Thinking

As described in the previous chapter there is considerable theoretical support for the premise that a relationship exists between an individual's beliefs about the nature of knowledge, or "personal epistemology", and how that individual uses that knowledge. For example, Anderson (1984) points out that student epistemological beliefs may provide an explanation as to why some students integrate information and others do not. Spiro, Vispoel, Schmitz, Samarapungavan, and Boerger (1987) state that epistemology may explain why some students oversimplify information and others do not. Several authors have developed conceptual frameworks of human cognition containing epistemological components (Perkins & Simmons, 1988; Posner et al., 1982; Schoenfeld, 1983). Surprisingly, there is very little experimental support for these ideas. The section that follows will discuss the literature, both theoretical and experimental, that provides motivation for this study.

Theoretical Framework

As stated in the previous chapter, Posner et al. (1982) emphasize the role of epistemological beliefs in the conceptual change process. They propose that an individual's current concepts, or "conceptual ecology", will influence the development of a new concept. They describe conceptual change as a result of five primary determinants:

1. anomalies: ideas that fail are important for selecting their successors.
2. analogies and metaphor: these suggest new ideas and make them intelligible.
3. epistemological commitments:
    a. explanatory ideals - subject-matter specific views with norms for successful explanations in the field.
    b. general views about the character of knowledge - standards for successful knowledge such as parsimony and economy.

4. metaphysical beliefs and concepts:

   a. metaphysical beliefs about science - beliefs about the orderliness, symmetry and nonrandomness of the universe.

   b. metaphysical concepts of science - beliefs about the nature of the universe.

5. Other knowledge:

   a. knowledge in other fields

   b. competing concepts - selection of a new concept requires that it is an improvement on its competitors.

The belief that epistemological commitments directly influence the formation of new concepts lends theoretical support to the link between epistemology and reasoning. Standards for successful knowledge determine if information is believed, dismissed, or ignored. Within a discipline, norms for successful explanations influence how one will use inference, and determine if a conclusion drawn is justified. Also, metaphysical beliefs about science and the universe, described by the authors as an influencing factor on concept formation, affect one's personal epistemology. One's conception of scientific knowledge and scientific truth is encompassed by one's view of the universe. Metaphysical beliefs provide the context within which thought and reasoning take place.

Schoenfeld (1983) explores the way that belief systems shape people's behaviour as they solve problems. He has developed a framework for human problem-solving which includes three qualitatively different categories of knowledge and behaviour. They are:

1. resources - knowledge including facts and competencies possessed by the individual, that can be brought to bear on the problem at hand.
2. control - selection and implementation of resources.
3. belief systems - determinants of an individual's behaviour. These determinants can be conscious or unconscious and include beliefs about the individual, the environment, the topic and about the discipline.

Schoenfeld argues that the network of beliefs about oneself, the task, and about the nature of the discipline itself provides the context within which problems are solved and conclusions produced. In addition, understanding the context is essential to accurately interpret the conclusions drawn. He states:

> Beliefs about the very nature of facts and procedures will determine students' performance. The student who believes that mathematical knowledge must be remembered will be stymied when a particular object is forgotten, while another who believes that the procedure can be derived will act rather differently. (p. 350)

Perkins and Simmons (1988) offer strong support for the relationship between epistemology and reasoning. Their discussion of learning, as influenced by an interaction of different types of knowledge, provides an important theoretical base for this study. As outlined in the previous chapter, they propose four different categories or "frames" of student understanding. Briefly, they are: (a) the content frame, (b) the problem-solving frame, (c) the inquiry frame, and (d) the epistemic frame. The epistemic frame includes the norms by which knowledge is acquired and used. With regard to epistemology, their work parallels that of Posner et al. (1982) and Schoenfeld (1983). Epistemology provides the rules by which knowledge is used.

They claim it is an oversimplification that knowledge and understanding are acquired through information and practice. Deeper understanding is a result of an awareness of the relationships that connect content knowledge with knowledge in the problem-solving, inquiry, and epistemic frames. They urge that the educational community as a whole must shift the perspective of learning and understanding to encompass epistemic considerations.

Newmann (1988) explains that any theoretical framework of higher order thinking requires a component more fundamental than knowledge or skills. It requires what he calls dispositions of thoughtfulness. Dispositions have been mentioned previously and have been discussed at length by Norris (1985b) and Walsh and Paul (1987). They include such attitudes and personality traits as: (a) possessing a desire that claims be supported by reasons and that those reasons be scrutinized, (b) having a tendency to be reflective rather than blindly accepting the views of others, (c) having a curiosity to explore new questions, and (d) having the flexibility to seek alternate solutions.

Newmann states that dispositions also involve general values and beliefs about the nature of knowledge, such as beliefs that rationality is desirable and that knowledge is socially constructed, often indeterminate, and tentative. An awareness of these epistemological concerns is necessary for an understanding of higher order thought processes.

Though there are some specific differences in the conceptual frameworks discussed, there is extensive agreement that epistemology plays an important role in the cognitive process. In this context, it seems quite reasonable to expect a relationship between students' conceptions of scientific knowledge and their reasoning strategies.

Experimental Support

An extensive review of the literature revealed only a few studies that attempted to determine specifically students' epistemological beliefs. A single study was found that tried to isolate students' epistemological beliefs and to relate them to their comprehension. Schommer (1989) attempted to isolate different dimensions of students' personal epistemology and to relate these dimensions to their comprehension of a passage for which the students were to write a concluding paragraph. Since the literature search did not reveal any studies that attempted to isolate epistemological beliefs and relate them to inductive reasoning ability specifically, Schommer's work provides important practical support and will be discussed in detail.

The seminal work on students' personal epistemology was performed by William Perry in a study of college students (1968). Through a questionnaire and in-depth interviews Perry developed a scheme that included nine stages of epistemological development. He claimed that students move from a position of believing that knowledge is either right or wrong, certain, and handed down from authority, to the position characterized by an understanding that there are multiple possibilities for knowledge and that all knowledge is constructed.

Schommer (1989) reports that subsequent research based on Perry's work has been inconsistent and contradictory. She submits the problem is with Perry's unidimensional conception of epistemological beliefs. Her proposal is that epistemological beliefs can be subdivided into at least five more or less independent dimensions: (a) the structure of knowledge, (b) the certainty of knowledge, (c) the source of knowledge, (d) the control of knowledge, and (e) the speed of knowledge acquisition.

In her own study Schommer (1989) used a sample of 86 junior college students to test the effects of epistemology on comprehension. First, the subjects completed an epistemology questionnaire. Then they were asked to write the concluding paragraph to a passage. Subjects completed a ten-item mastery test on each passage and a confidence scale was administered. Subjects' prior knowledge was indicated by the number of classes they had taken in a courses relevant to the passages they had to read.

The results indicated that subjects who believed knowledge was acquired quickly tended to (a) make oversimplified conclusions, (b) perform poorly on the mastery test, and (c) overestimate their level of understanding as indicated on the confidence scale. Subjects who believed knowledge is certain rather than tentative tended to make what they believed to be certain conclusions.

The educational implications of Schommer's findings are quite significant. If appropriate epistemological beliefs can be learned through explicit instruction, then comprehension potentially can be improved. As Schommer points out, future research needs to test the generalizability of the effects of epistemological beliefs on different tasks, different domains, and different aspects of comprehension.

Summary and Conclusion

This chapter has attempted to assess the current status of research in the areas of conceptions of the nature of scientific knowledge and inductive reasoning. It is clear that there has been much research completed that is relevant to this study.

With regard to students' conceptions of scientific knowledge, the research generally indicates that students' conceptions are inadequate. The most convincing evidence has come from researchers such as Welch (1981) and Aikenhead (1987). Through the work of Robinson (1969), Lederman (1986b, 1986c), and others, teachers' conceptions and their effect on students' conceptions have been examined. These studies have produced somewhat inconclusive results. Many researchers believe the philosophical view portrayed in curricular materials is outdated and a source of student misconceptions regarding the nature of scientific knowledge (Duschl, 1988; Hodson, 1986a, 1986b, 1988; Norris, 1985a, in press).

Many instruments are available to test conceptions of the nature of science generally, but few instruments, with the exception of Rubba's (1976), are devoted specifically to testing the nature of scientific knowledge. Limitations of these instruments include: (a) lack of agreement by philosophers on a model of the nature of scientific knowledge, (b) basis in outdated philosophical models, (c) lack of a definition of an adequate conception of scientific knowledge, (d) lack of recognition of the level of cognitive development of young students and of different possible interpretations of language, and (e) poor construction containing difficult terminology. It was judged that none of the instruments examined would be appropriate for use in this study, so the development of an epistemological questionnaire was considered necessary.

With regard to inductive reasoning ability, the research indicates that these skills are also inadequate. Generally, it has been found that student reasoning is characterized by an unwillingness or inability to deeply analyse problems (Perkins et al., 1983; Schoenfeld, 1983, 1985).

Several tests of critical thinking exist that contain inductive reasoning as a component. These include ones by Ennis and Millman (1985), Ennis and Weir (1985), and Shipman (1983). As with tests of knowledge conceptions, tests of critical thinking and inductive reasoning have numerous limitations. These include: (a) conceptual problems of using measures of internal consistency as measures of reliability, (b) problems with determining validity, and (c) the problem of background assumptions and multiple interpretations. The Essay Test of Inductive

Reasoning Strategies, Part A was selected as the measure of reasoning for the study.

There is a great deal of support for the theoretical relationship between epistemology and reasoning (Newmann, 1988; Perkins & Simmons, 1988; Posner et al., 1982; Schoenfeld, 1983). Basically, the relationship can be summarized by stating that the rules and norms that govern the acquisition of knowledge influence how that knowledge is used in the reasoning process. However, with the exception of work by Schommer (1989), there is relatively little experimental support for the relationship. It is the belief of this investigator that knowledge about student epistemology has been a neglected area of educational research and can provide valuable insight into the reasoning process.

CHAPTER III: METHODOLOGY

The Nature of the Causal Model

The theoretical model tested in this study is presented in Figure 3.1. The model depicts the causal relationships among conceptions of scientific knowledge, inductive reasoning, and science achievement. The model proposes how these three variables vary together, that is, how a change in the value of any one variable will affect the values of the other variables.

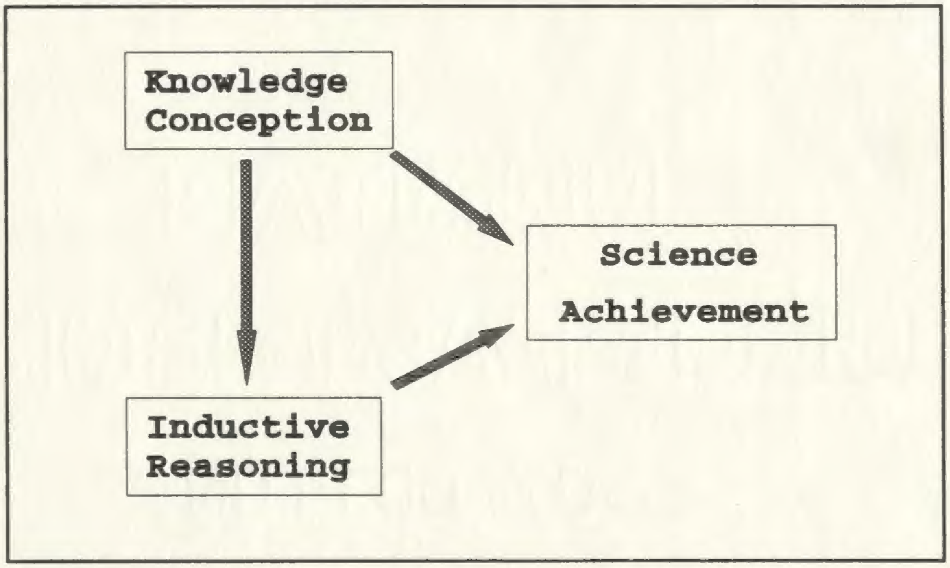

Figure 3.1. Theoretical model of causal relationships among conceptions of scientific knowledge, inductive reasoning, and achievement in science.

The following relationships are proposed: (a) that students' conceptions of scientific knowledge will directly affect their achievement in science, (b) that students' conceptions of scientific knowledge will directly affect their inductive reasoning, (c) that students' inductive reasoning will directly affect their achievement in science, and, consequently, (d) that students' conceptions of scientific knowledge will indirectly affect their achievement in science through changes in inductive reasoning.

The model is derived from the assumption of Newmann (1988), Perkins and Simmons (1988), Posner et al. (1982), and Schoenfeld (1983, 1985), that beliefs about the nature of knowledge will influence how that knowledge is used. Varying conceptions will cause differences in how one performs mental tasks, including inductive reasoning. Students who are better able to perform mental tasks should achieve better in science courses. The findings of Schommer (1989), that conceptions of the nature of knowledge influence comprehension, provide empirical support for the relationships proposed in the model.

<u>The Nature of Related Variables</u>

Within the context of the model, students' conception of scientific knowledge is an independent or predetermined variable. Thus, the total variation in conceptions of scientific knowledge is assumed to be caused by variables outside the model. This type of variable is called "exogenous" (Land, 1969). The remaining variables, inductive reasoning, and achievement in science, are taken as dependent or "endogenous" (Wright, 1934; Land, 1969).

The variation of the endogenous variables is assumed to be determined by a combination of the prior exogenous or endogenous variables and a "residual" variable. In other words, the variation in science achievement is assumed to be determined by inductive reasoning and conception of scientific knowledge, as well as by influences from outside the proposed model. The residual variable is assumed not to be correlated with the set of variables immediately determining the endogenous variable under consideration and it accounts for the variance of the endogenous variable not explained by the prior measured variables (Land, 1969).

Assumptions of the Model

In proposing the causal model, I make three basic assumptions. First, all relationships are linear. Changes in one variable always occur as a linear function of changes in other variables. That is, the equations used to describe the relationships are those in which the value of one variable is defined simply in terms of a linear combination of the values of the other variables (Heise, 1969).

Second, it is assumed the model is recursive. That is, the system of variables contains no reciprocal causation or feedback loops. The direction of effects are assumed to be away from the exogenous variable: conceptions affect achievement, but achievement does not affect conceptions, either directly, or through reasoning.

Third, the variables are assumed to be continuously acting. As a result, the instruments that measure them must do so on an interval scale. Imposing dichotomies or treating the variables as ordinal would be inappropriate since these interpretations imply that the variable behaves in an all-or-nothing fashion (Land, 1969).

Statistical Analysis

The first two research questions about students' conceptions of the nature of scientific knowledge and their inductive reasoning strategies will be answered using descriptive statistics. The final three research questions, as well as the hypotheses proposed, attempt to identify causal interrelationships between the three variables of conceptions of scientific knowledge, inductive reasoning, and science achievement. These questions will be analyzed by two correlational methods. First, the Pearson product-moment correlations for all variables will be determined. Second, the evidence for the strength of causal relationships will be examined using path analysis.

Path Analysis

Path analysis is a method of testing the validity of a theory about causal relationships among three or more variables that have been studied using a correlational research design (Borg & Gall, 1983). The process is a method of decomposing and interpreting linear relationships among a set of variables assuming that (a) the causal ordering among the variables is known, and (b) the relationships among the variables are causally closed (Kim, 1975).

The method of path analysis was originally developed by Wright (1921, 1934, 1954, 1960a, 1960b) in a series of general essays on data analysis in the field of genetics. Wright (1921) describes the purpose of a path analysis as:

> a method of measuring the direct influence along each separate path in such a system and thus finding the degree to which variation of a given effect is determined by each particular cause. The method depends on the combination of knowledge of the degree of correlation among the variables in a system with such knowledge as may be possessed of the causal relations. In cases in which the causal relations are uncertain, the method can be used to find the logical consequences of any particular hypothesis in regard to them. (p. 557)

Borg and Gall (1983) describe three fundamental steps to complete a path analysis. The first step is to formulate a theory that causally links the variables of interest. Then measures for the variables must be selected. The final step is to compute the path coefficients to show the strength of the relationship between each of the pairs of variables that are causally linked. Path coefficients generated from the analysis are typically standardized regression coefficients indicating the direct effect of one variable on another. Wright (1934) describes path coefficients as:

> the fraction of the standard deviation of the endogenous variable for which the designated variable is directly responsible in the sense of the fraction which would be found if this factor varies to the same

> extent as in the observed data while all other variables are constant. (p.162)

Thus, the path coefficients between an independent and dependent variable indicates the expected changes in the dependent variable when the independent variable is actually changed by one standard deviation and all other independent variables are held constant.

A major advantage of a path analysis is that it allows identification of direct and indirect effects of the independent and intervening variables upon dependent variables. In the causal model presented in this study, the indirect effect is that part of the total effect of conception of scientific knowledge on science achievement that works through the variable of inductive reasoning.

Measures

Epistemological Questionnaire

The questionnaire to measure students conceptions of scientific knowledge consists of 56 Likert-type items compiled in the following seven subscales: (a) scientific truth, (b) the fallibility of scientific knowledge, (c) the changeability of scientific knowledge, (d) the role of creativity in scientific knowledge production, (e) the testability of scientific knowledge, (f) the speed of scientific knowledge acquisition, and (g) the role of authority in scientific knowledge production. The questionnaire was designed so that extremes of each subscale represent alternate conceptions. Since the questionnaire was constructed specifically for the present study, its development is discussed separately and in detail in the next chapter.

Essay Test of Inductive Reasoning Strategies

In this test, subjects are to imagine they are on an imaginary planet "Zed" for four days. Subjects are told they must search planet Zed for living creatures. They are asked to read about the things that happen each day, and to think about what these events mean for their search. Subjects are then asked to write what they are thinking about the things that happened on that and previous days; and write what they plan to do on their search because those things happened.

The events of each day are presented in sequence and subjects are not to go on to the next day until they have discussed what the events of the current day mean for their search and what they plan to do as a result of the information provided on the current day. Day 1 has been completed as an example for the examinees by the authors of the test.

Theoretical framework. The test was designed within the context of Ennis's (1987) conception of critical thinking. Three types of inference are central to this conception: induction, deduction, and value judgement. Inductive inferences may be of two types: those that generalize information or those that explain information.

The Essay Test of Inductive Reasoning Strategies, Part A focuses on inductive explanation.

The test manual describes seven strategies used in inductive reasoning: (a) recognizing relevant information, (b) seeking information when appropriate, (c) producing alternate conclusions or explanations, (d) withholding judgement, (e) monitoring progress, (f) being organized, and (g) staying focused. Table 3.1 provides a descriptive summary of the authors' comments (Norris & Ryan, 1987b, p.4) on the importance of each strategy to the concept of inductive reasoning. The order in which the strategies are presented is arbitrary, since different tasks require different strategies used in different sequences.

Scoring. A detailed explanation of all scoring procedures, their rationale, and sample calculations of strategy, depth, and breadth scores (to be described subsequently) can be found in the test manual (Norris & Ryan, 1987b). The authors have provided scoring guides that incorporate models of ideal reasoning for each of the test days. They note that the scoring guide will provide considerable direction to the scorer but there is still a need for individual judgement and understanding of the reasoning required in the test situations.

Table 3.1

A Description of Inductive Reasoning Strategies

Strategy	Description
1	All information is taken into account. Since information is the basis for reasoning, if all information is not taken into account, then conclusions may be reached that would have been ruled out by information that was ignored, or conclusions that might be suggested by the ignored information will fail to be reached.
2	More information is sought when appropriate. This strategy is a supplement to strategy 1 when the information given in a problem situation is insufficient.
3	Alternative conclusions, explanations, hypotheses, and plans are generated. Induction does not force unalterable conclusions. An individual who reasons well will realize this and consider alternatives. This strategy involves the creative element of inductive reasoning as alternatives are not always obvious.
4	Judgement is withheld where appropriate. While one considers alternatives, judgement must be withheld. This follows from strategy 3.
5	The likelihood of progress is monitored. Progress must be evaluated as one moves towards a desired goal. An individual must know how close he or she is to a solution or when the solution is at hand. Without this assessment, there is no way to guide the inductive reasoning process or to bring it to a close.
6	Complex problems are handled in a clear and orderly fashion. (See strategy 7.)
7	The main point is kept in mind. The authors describe strategies 6 and 7 as implicit within the whole process of inductive reasoning. These strategies "permeate good inductive reasoning and are displayed in the reasoner's tying every action to the purpose at hand and never losing sight of the goal".

Each scoring guide consists of a chart comprised of a series of columns and rows. The scoring chart of Day 2 is presented in Table 3.2. The events of Day 2 can be found in the test presented in Appendix B.

Each row of the chart represents a different inductive reasoning strategy, either strategy 1, strategy 3, or strategy 2 and 3 together. Each column represents what the authors call a "data package". Each data package centres around a piece of information from each particular day. These pieces of information appear at the top of each column. Sometimes they are designated as "need to be explained"; other times they are designated as "relevant". These are the pieces of information judged most important for the examinees to consider.

Below the piece of information at the top of each data package are the appropriate kinds of actions for the examinees to perform with regard to the information. Also indicated to the right of each statement of action is the score to be awarded when the examinee takes the action indicated. Below this are spaces for recording the total score for each column as well as the percent of the possible total.

In addition, there are four questions concerning different inductive reasoning strategies that apply to the subjects' answers as a whole. Points are awarded as indicated for use of these respective strategies.

Table 3.2

Scoring Chart for Day 2 of the Essay Test of Inductive Reasoning Strategies.

Strategy	Data Package 1		Data Package 2		Data Package 3	
1	River and Valley (relevant)	2	Photographs (relevant)	2	Eggshell-like objects (need to be explained)	2
3					Creatures	5
					Inorganic matter	5
					Other	5
2+3	Return to explore	5	Examine for creatures	5	Try to discover constituents	5
	Check in water	5	Examine for clues of creatures	5	Explore for whole eggshells	5
	Other	5	Other	5	Other	5
Column total						
Possible total		17		17		32
Column %						

1. (Strategy 4) -- To what extent are alternatives other than living creatures considered? (Range: 0-5) ____

2. (Strategy 5) -- Is the likelihood of progress monitored? (Range: 0-5) ____

3. (Strategy 6) -- Is the response clear and orderly? (Range: 0-5) ____

4. (Strategy 7) -- Is the focus good? (Range: 0-5) ____

From these raw scores three further scores are calculated: (a) strategy scores, (b) depth of consideration scores, and (c) breadth of consideration scores.

The strategy score gives as a percentage how much a strategy was used compared to how much it could have been used in the ideal response. From Table 3.2 it can be observed that on Day 2 it is possible to score six points for strategy 1, two points for each data package. If an examinee recognizes that the photographs and eggshell-like objects are relevant, but does not take into account the river and valley, he or she will score four of a possible six points. Thus the score for strategy 1 on Day 2 is .67. To determine an overall test score for strategy 1, the strategy 1 scores of all three days are averaged. The overall total score of the test is the average score for all strategies.

The depth of consideration score is an indication of the amount of action taken in response to relevant pieces of data or data in need of explanation. The authors describe depth of consideration as a continuum from a minimum of just referring to a piece of information to a maximum of providing possible explanations for it and recognizing the need for additional information to evaluate possible explanations. Depth of consideration is indicated by how far down a data package examinees proceed when considering the information heading that data package. The depth score for a data package is the fraction of the total possible score received for that data package. For example, in Table 3.2, it can be seen that an examinee can score a maximum of 17 points in data package 1. If that examinee recognizes that the river and valley are relevant, and suggests to check in the water for living creatures, he or she will receive 7 of the possible 17 points, or a .41 depth of consideration score for data package 1. This score is calculated for each data package and averaged to produce the depth of consideration score for Day 2. The overall depth of consideration score is the average of all depth scores from all data packages for all days.

The authors describe ideal breadth of consideration as dividing one's effort proportionately among each data package. This is achieved when effort is distributed according to the proportion of the total possible score that a data package can contribute. In other words, if a particular data package can contribute 10% to the total possible score, an examinee's response in that data package should constitute 10% of his or her total effort. The authors call this the ideal proportion of effort.

By referring to Table 3.2, it can be observed that data package 1 can contribute 17 out of a total possible 66 points, or about .26 of the possible score for Day 2. This is the ideal proportion of effort for data package 1. Assuming an examinee scored 7 points on data package 1, and a total of 41 points for all Day 2 data packages, the actual proportion of effort on data package 1 would be .17. Thus, the examinees proportion of effort undershot the ideal by .09.

To determine the breadth of consideration score, deviations, in absolute values, from the ideal proportion of effort are calculated for each data package and the average deviation is determined. Widest breadth of consideration would result in no deviations from the ideal, and the breadth of consideration score would be zero. The larger the score the smaller the breadth of consideration.

To aid the task of scoring the essay test, the authors developed a computer program that (a) allowed the examiner to record personal data on the examinee, (b) allowed the examiner to record the scores of all data packages and strategy questions, (c) averaged all strategy scores, (d) calculated the overall test percentage and (e) calculated the average depth and breadth of consideration scores. The program displays each scoring chart as it appears in the test manual, with the user able to enter scoring values directly into the scoring charts. The overall effect of the program is to greatly improve the consistency of the grading process.

Achievement in Science

Achievement in science was assessed using students' final science grades from the previous school year. The major advantages of using this data as a measure of science achievement was that it was obtainable from existing records and was measured over the entire school year using numerous assessment instruments and procedures. One would expect the result to be a reasonably valid indication of achievement. Despite this, there are several concerns with using past grades as a measure of achievement.

First, past grades are a measure of past achievement. The study correlates current conceptions and abilities with achievement of the school year past. The danger is that influencing experiences could have taken place between the time the science achievement grades were assigned and the time the students completed the instruments for the study. I am assuming that any such influences were small, since the time delay is relatively short. The main study was conducted during the month of November while the final grade for the previous year was awarded in June. Thus, the time delay is in the order of four months.

Second, there was no way to control for different grading standards from teacher to teacher or from course to course. It is conceivable that two individuals in the study have equal grades but do not have equal abilities in science. In response to this limitation it was decided to use only grades from those science courses designed for students of average or higher academic ability. The standards for these courses are likely to be more consistent since at the end of high school these students must pass provincially set examinations. Thus, in this study, science achievement refers the final grades assigned in the following subjects: (a) general science, (b) biology, (c) chemistry, (d) physics, (e) and earth science.

Sample

The sample consisted of 305 students from a public suburban senior high school in an eastern Canadian province. The school was a regional one with students from a wide geographic area attending. The total school population was approximately 800 in grades 10 - 12 with a teaching staff of 43. Students were of varied socio-economic status and were from various cultural backgrounds, but were predominantly English Canadian.

The sample was comprised of 165 females and 140 males. The subjects were distributed over the three high school grades with 91 grade ten, 116 grade eleven and 98 grade twelve students. Only subjects whose parents or guardians had previously completed a consent form were allowed to participate. Although the total number of students who completed at least one of the instruments was 346, due to absenteeism on either day and to students who exercised their right not to take part in either portion of the study, the actual number of students to complete both instruments was 305. Table 3.3 provides a detailed summary of subjects' sex and grade.

Table 3.3

Number of Subjects by Sex and Grade Completing Both Instruments

	Grade			
	10	11	12	Total
Male	40	52	48	140
Female	51	64	50	165
Total	91	116	98	305

A total of 17 classes, including seven grade 10, four grade 11, and six grade 12 classes, were involved in the study. Enrolment in different courses was open to students of different grades, provided core and pre-requisite conditions were met. For example, grade 12 students could be found in the grade 10 biology and grade 10 physics classes. However, there were no grade 10 students in the grade 11 and 12 courses.

Three cooperating teachers were involved in the study. They held science and education degrees at the Bachelor's level and all had majored in biology. One teacher held an education degree at the Master's level. All were teaching at the school for a minimum of six years and had seven, eleven, and twenty-two years of teaching experience. All three teachers were male.

Administration of Instruments

Instruments were administered during regular 40 minute class periods under the supervision of the cooperating teacher. The epistemological questionnaire was administered first followed by the inductive reasoning test in the same class period the very next day. In that class period on the third day any students who were absent from either of the previous two periods were administered the instrument that they had not completed. Students who had completed both instruments were assigned seatwork in their subject area.

At the beginning of each session, subjects were again informed of the purpose of the study, the first time being when the original consent letter was sent to their parents. The instructions were then read aloud by the cooperating teacher before the subjects began. Subjects were not given extra time and all questionnaires and tests were collected at the end of the class period.

CHAPTER IV: QUESTIONNAIRE DEVELOPMENT

Philosophical Dimensions of the Questionnaire

This section attempts to establish the construct validity of the dimensions of scientific knowledge logically derived from the literature. Proposing a specific philosophical model on which to base the questionnaire is inappropriate since, as noted previously in chapter II, there is no consensus on what that model should be. Instead, the approach taken in this study is to derive from the literature various dimensions of scientific knowledge that (a) are considered essential to one's conception of scientific knowledge, and (b) are believed to affect the use of knowledge in the reasoning process. Thus, the author does not argue for or against any particular conception of scientific knowledge.

Dimensions were developed from: (a) portions of the philosophical models of existing instruments (COST, NOSS, NSKS, SPI, & WISP), (b) selected philosophical literature (Conant, 1951; Hanson, 1958; Hardwig, 1985; Hempel, 1966; Hodson, 1986a; Kuhn, 1962; Norris, 1984; & Popper, 1972), and (c) psychological literature that deals with epistemology and understanding (Schoenfeld, 1983, 1985; Schommer, 1989). The contribution of these sources and the descriptions of the dimensions will be presented in the subsequent sections. Table 4.1 provides a summary of the seven dimensions derived from the synthesis of the literature.

Table 4.1

A Summary of the Questionnaire Dimensions

Dimension	Summary
Scientific truth	The relationship between scientific statements and the nature of the universe, its phenomena and relationships (Kimball, 1967; Scientific Literacy Research Center, 1967).
Fallibility	The role of observation and the influence of current theoretical beliefs (Hanson, 1958; Hodson, 1986a; Norris, 1984).
Changeability	The development of scientific knowledge over time (Conant, 1951; Kimball, 1967; Kuhn, 1962; Rubba, 1976; Schwab, 1960; Scientific Literacy Research Center, 1967).
Creativity	The role of creativity in accounting for natural phenomena (Popper, 1972; Rubba, 1976).
Testability	The question of whether scientific knowledge must be reproducible (Cotham & Smith, 1981; Hempel, 1966; Scientific Literacy Research Center, 1967).
Speed of acquisition	The rate at which scientific knowledge is developed (Schoenfeld, 1983, 1985).
Role of authority	The role of questioning in the acceptance of scientific knowledge (Hardwig, 1985).

Items of the questionnaire were based on these seven dimensions, thus forming seven subscales. Two conceptual alternatives for each subscale were developed and items constructed relating to each. Thus, each subscale was polarized, with one set of item statements written representing one pole, and another set of statements written to reflect the alternate pole. By representing conflicting views, the subscales provided subjects with an opportunity to hold varied conceptions.

Since the questionnaire format chosen was a five point Likert-type rating scale (to be described subsequently), subscale average scores below a value of 3 represent beliefs leaning towards one pole of a dimension of scientific knowledge, while average subscale scores of greater than a value of 3 represent beliefs leaning towards the alternate pole. Table 4.2 provides an overview of the dimension poles and sample statements from the questionnaire worded consistently with the alternate conceptions. Greater detail on scoring is provided in a subsequent section of this chapter.

In most cases, there is a consensus in the literature about which view is more accurate. In some cases, though, the issue is not as apparent. For example, within the fallibility subscale, the conception that scientific knowledge can be doubted and is sometimes created in error is much more prevalent. If one examines the history of

Table 4.2

Summary of Questionnaire Dimensions, Dimension Poles, and Sample Statements Worded Consistently with Alternate Conceptions

Dimension	Dimension Poles	Sample Statement
Scientific truth	Correspondence	A scientific statement is true when it describes the world the way it really is.
	Coherence/ consensus	Statements about which scientists agree are true.
Fallibility	Fallible	We accept statements of scientific knowledge even though they contain error.
	Certain	Scientists do not make errors in their conclusions if they follow the scientific method.
Changeability	Changeable	Today's scientific theories, laws, and concepts may have to be changed in the face of new evidence.
	Fixed	After scientists think they have found the solution to a problem, they feel the problem has been solved once and for all.
Creativity	Created	Scientific knowledge is the product of human imagination.
	Found	Scientific theories are discovered, not created by scientists.
Testability	Testable	Test results that can be repeated by other scientists are required for the acceptance of scientific knowledge.
	A priori	Scientific knowledge need not be capable of being tested by experiments.
Speed of acquisition	Slow	The solutions to scientific problems are developed gradually over time.
	Fast	If you are ever going to be able to understand a scientific statement, it will make sense to you the first time you hear it.
Role of authority	Reasonable	Knowing when to seek an expert opinion is important when trying to solve a scientific problem.
	Authoritative	There is no purpose to questioning accepted scientific theory.

science, the fallibility of scientific knowledge can be demonstrated unquestionably. On the other hand, with regard to the truth subscale, neither the conception that scientific knowledge is true when it accurately represents natural phenomena nor when it enjoys a consensus among the leading scientists of the day, can be clearly demonstrated and is thus subject to great debate among philosophers. The sections that follow describe the seven dimensions in more detail.

Scientific Truth

Of the instruments examined that measure conceptions of the nature of science (and scientific knowledge), none dealt specifically with scientific truth. Some instruments, such as TOUS, include truth under the umbrella of "Aims of Science". That perspective is characterized by the search for scientific truth as a motivating force for the practice of science, not the perspective of emphasizing conceptual differences in the nature of truth.

There are several possible reasons for the omission of scientific truth in the philosophical model of scientific knowledge. First, there is a lack of consensus among philosophers of science as to the nature of truth. Any philosophical model that includes a particular view of truth will be open to a degree of criticism. Second, the concept of truth is sophisticated and abstract. Students of high school age do not necessarily have well-formed, thoughtful views on the nature of truth. If subjects do not know what they believe, then accurately assessing their views is particularly difficult, because some of their views may germinate from the prompts of the questionnaire.

Scientific knowledge can be viewed as true from three perspectives. First, a scientific statement may be true because it corresponds to the natural world. Our knowledge is an approximation of reality. This is a traditional positivist view of scientific truth. Second, a statement may be considered true because it is consistent with our existing set of what we believe to be true statements. Finally, a statement may be considered true if it enjoys a consensus of belief among the members of the scientific community. These three viewpoints can be respectively labelled the "correspondence", "coherence", and "consensus" conceptions of scientific truth. To polarize the truth category of the questionnaire the coherence and consensus views were chosen to represent one extreme, while the correspondence view was chosen to represent the other.

The position taken in this study is that a perspective on scientific truth is fundamental to understanding scientific knowledge, and thus should not be omitted from the questionnaire. It is reasonable to assume that individuals who view knowledge as a reflection of reality will draw different conclusions during reasoning than those individuals who believe knowledge to be a consensus of opinion. Since these conceptions are still the subject of debate among philosophers, the decision as to which view is the correct one is still unresolved. Initially, the coherence and consensus view of truth was selected as the view to be scored highest. Results of the pilot study (to be described in a subsequent section) revealed that scoring the truth subscale high when responses matched this view of truth led to a subscale score that correlated negatively with the remaining subscales. Thus, the decision was made to score highest the statements agreeing with the correspondence view.

Fallibility

Many of the instruments examined in chapter II focused on the fallibility of scientific knowledge. The items in this category attempt to determine whether students believe scientific knowledge is certain or contains error and should be doubted. Believing that knowledge should be doubted should affect one's view of the conclusions drawn from reasoning. The consensus in the philosophy of science community is that scientific knowledge is fallible.

Changeability

This dimension of the questionnaire attempts to determine if students believe scientific knowledge is changeable or fixed. The history of science provides many examples of scientific knowledge that has changed either through an evolutionary or revolutionary process. Several instruments examined, including Aikenhead's, Fleming's, and Ryan's (1987), Cotham's and Smith's (1981), Kimball's (1967), and Rubba's (1976), focused on the changeability or tentativeness of scientific knowledge. It is nearly universally believed in the field of philosophy of science, that scientific knowledge is changeable.

Schommer (1989) provides an experimental link between this conception of knowledge and reasoning. She found that subjects who believed that knowledge was certain rather than tentative believed their conclusions to be also certain.

Creativity

The prevalent view of this dimension is that scientific knowledge is a product of the human mind that attempts to explain observed natural phenomena. A less sophisticated view is to believe that production of scientific knowledge simply involves uncovering what already exists. Creativity as a dimension of scientific knowledge is also emphasized in several philosophical models, for example Showalter's (1974), and Rubba's (1976) NSKS.

Testability

The role of testability of scientific knowledge is still the subject of debate among philosophers of science. The more modern view is that the production of scientific knowledge rests on principles derived from reason and may be independent of experience. This a priori view has stemmed from the writings of Kuhn (1962). The traditional positivist view states that scientific knowledge must be subjectable to empirical test. For ideas to become part of the body of scientific knowledge, consistency of test results is a requirement. The validity of scientific knowledge is established through repeated testing against accepted observations (Rubba & Anderson, 1978). This conception of scientific knowledge is a component of most instruments assessing the nature of scientific knowledge.

Speed of Knowledge Acquisition

This dimension of scientific knowledge can be derived from the work of Schoenfeld (1983, 1985). In his research with high school students' geometry proofs he found that certain students believed in quick, all-or-none learning. That is, they would spend a short time working on a problem, and, if they didn't find a solution in a few minutes, they assumed they would never find one.

This dimension was not found in any of the philosophical models of scientific knowledge on which testing instruments were based. Schommer (1989) incorporates the category into her model of general knowledge and found that it affected subjects' comprehension. She found that subjects who believed knowledge was acquired quickly tended to make oversimplified conclusions, performed poorly on a mastery test, and overestimated their level of understanding of a passage.

Role of Authority

The source of knowledge and the role of authority in producing that knowledge is a dimension that can be derived from the work of Perry (1968) who found that many students enter college with the belief that knowledge is simply handed down from authority. Other models that have incorporated the role of authority in knowledge production and acquisition are Welch's and Pella's SPI (1968) and Schommer's (1989).

One view of the role of authority is that knowledge is derived from reason rather than handed down from authority, and that questioning authority is always appropriate. The dimension used in this study however, is a slight variant of that theme. While one pole is the view that knowledge should be accepted blindly, the other pole is the view that reasonable questioning is desirable. Reasonable questioning means that sometimes it is appropriate to accept knowledge handed down from authority. Sometimes the most sensible course of action is to seek expert opinion.

This idea is what Hardwig (1985) describes as "epistemic dependence". He puts forth three arguments to support his contention that an appeal to intellectual authority constitutes justification for believing and knowing. First, he argues that one can have a good reason for believing a proposition if one has a good reason to believe that others have good reasons to believe it.

Secondly, he explains that since the layperson is the epistemic inferior to the expert, it may be more rational for the layperson not to rely on his or her own judgement in a matter that falls within the expert's area of knowledge. Thirdly, he maintains that the expert-layperson relationship is essential to the scientific and scholarly pursuit of knowledge. If all authority is questioned, then progress will not be made.

Questionnaire Items

Format

The final version of the questionnaire is presented in Appendix A. The format of the items is that of a bipolar rating scale (Andrich & Masters, 1988). A statement about scientific knowledge is presented and subjects are to indicate whether they strongly agree, agree, are undecided, disagree, or strongly disagree. There are eight statements assigned to each of the subscales, four with model "agree" responses and four with model "disagree" responses. The sequence of items within the questionnaire was randomly assigned.

With two extremes and a neutral response, this method of questioning was considered appropriate to isolate alternate conceptions. The form is economical for test construction, presentation, and administration. The procedure for completing this type of instrument was expected to be easily understood by high school students.

Having five points on the rating scale was considered large enough to take advantage of the subjects' capability to discriminate. Unlike a multiple-choice or an agree-disagree format, this particular format gives the subject an undecided or an "I don't know" option. A criticism of several existing instruments is that subjects are forced to respond to a statement as it is written with no provision for an undecided response (SPI and WISP).

Though follow-up interviews to check subjects' interpretations of the item statements would be desirable, they were not possible due to practical limitations and time constraints. It is believed that the careful preparation of the questionnaire through the development process and pilot study has reduced the possibility of multiple interpretations.

Construction

Questionnaire items were derived from the philosophical literature cited earlier in the chapter. As well, existing instruments were examined and many items adapted. The instruments having the greatest influence on the construction of the questionnaire were NSKS, WISP, SPI, Carey (1988) and Schommer (1989).

Item writing. Items were written by subscale following the guidelines suggested by Babbie (1973). Table 4.3 provides a general overview of these guidelines and measures taken to follow them.

Table 4.3

Summary of Guidelines Followed for Constructing Questionnaire Items

Guideline	Measure taken
Make items clear.	The objective is for all respondents to interpret the statement the way intended by the researcher. An attempt was made to make the statements as specific as possible. Vague and ambiguous words were avoided.
Avoid double-barrelled questions.	Statements were limited to a single idea or concept.
Respondents must be competent to answer.	An answer option of "undecided" was provided for subjects who were unable to make selections with any degree of confidence.
Simple items are best.	Long and complicated items were avoided.
Avoid biased items or terms.	An attempt was made to eliminate any clues of responses scored highest.

In addition to the precautions taken during item writing, measures were taken to improve the overall layout and organization of the questionnaire. These included: (a) checking grammar, spelling, and punctuation, (b) using an easily readable font (courier 10), (c) ensuring good duplication quality, (d) including instructions that were easy to understand along with an example of a completed item, and (e) avoiding cluttering within the questionnaire.

Readability. Phillips (1989) describes readability of text as its legibility, ease of reading, and ease of understanding. A considerable challenge when developing the instrument was to present the sophisticated philosophical concepts of the nature of knowledge in a readable form for high school students.

A conscious attempt was made to keep the length of the items short, to avoid using difficult vocabulary, and to present the statements in a concise, non-ambiguous manner. The questionnaire was designed such that it could be completed easily by high school students in a 40 minute class period. Its duplication was of high quality. Items were completely legible. The number of items was kept low since the statements require reflection. Existing instruments which contain as many as 150 items (SPI) may not allow the subject adequate time to consider his or her beliefs.

Several options were explored to assess the readability of the item statements. The idea of applying traditional readability formulae such as Fry's Readability Graph (Fry, 1968), which uses the number of syllables in the words used and sentence length as its criteria for measuring readability, was dismissed. These readability indexes are not sensitive to the context within which the text is used and do not take into account the background knowledge of the reader. As a result, their validity is called into question. Fry's Readability Graph is characterized by the following anomalies: (a) mono-syllabic words that are unfamiliar will improve readability; and (b) familiar multi-syllabic words will decrease readability.

Two sources were selected for a readability estimate: (a) student information from the pilot study, and (b) the judgement of high school science teachers. During the administration of the pilot questionnaire, students were instructed to write comments for items they did not understand or items that they thought were worded poorly. These comments were used to revise the questionnaire. The questionnaire was also examined by a panel of two high school science teachers. The reading level was estimated by them to be appropriate for high school students.

Scoring

The questionnaire was scored using the procedure of assigning the values of 5 to 1 to the subject responses for each item. The value of 5 was assigned when the subject strongly agreed with the response considered most prevalent in the literature. A value of 1 was assigned when the subject strongly disagreed with this response. Half the items for each dimension were written negatively and were therefore reverse scored. The subscale scores are simply the average of the individual item scores in that particular scale. The overall questionnaire score is the average of all the subscale scores and represents the extent to which subjects lean toward the poles scored highest. Table 4.4

provides a key displaying items worded consistently with specific conceptions within each subscale.

Table 4.4

Key Displaying Items Worded Consistently With Specific Conceptions Within Each Subscale

Dimension Poles	Items Representing
Truth	
Correspondence*	12, 38, 47, 51
Coherence/Consensus	24, 35, 39, 56
Fallibility	
Fallible*	2, 13, 27, 32
Certain	1, 14, 22, 41
Changeability	
Changeable*	15, 45, 52, 55
Fixed	3, 21, 31, 42
Creativity	
Creative*	5, 28, 33, 49
Found	4, 11, 20, 46
Testability	
Testable*	18, 34, 37, 43
A priori	6, 10, 16, 23
Speed of acquisition	
Slow*	7, 19, 48, 54
Fast	8, 26, 29, 50
Role of authority	
Reasonable*	9, 40, 44, 53
Authoritative	17, 25, 30, 36

* Conception scored highest

Pilot project

The pilot project was conducted in the same school as the main study. It involved 105 mixed ability students from two grade 10 and two grade 12 biology classes with one cooperating teacher. The students completed a preliminary form of the questionnaire.

The questionnaire was administered during a single period for each of the classes involved. Subjects were informed of the purpose of the pilot study, the instructions were read aloud by the cooperating teacher, and the subjects were instructed to write comments for any of the items they perceived to be poorly constructed. Results of the pilot study are presented in the next chapter.

CHAPTER V: RESULTS AND DISCUSSION

Discussion is organized into three major sections. The first deals with the presentation and discussion of descriptive statistics relating to the instruments used. The second addresses the first two research questions by describing the status of students' conceptions, reasoning, and science achievement. The final section addresses the final three research questions and deals with the correlational results including the path analysis of the general causal model. Also, the results of a factor analysis of the questionnaire, and of a path analysis based on the factored variables, are presented and discussed.

Descriptive Results: Measures

Pilot Questionnaire Results

The results of the pilot project indicated that several of the items were difficult for high school students to comprehend. Several items were lengthy and/or poorly worded. The analysis revealed that some items did not correlate with the questionnaire as a whole. The alpha reliability for the preliminary form of the questionnaire was 0.74. Table 5.1 provides a summary of the descriptive data gathered from the pilot questionnaire.

Average responses on all subscales, with the exception of those on the truth subscale, leaned towards the pole scored highest. The fallibility, changeability, speed of acquisition, and role of authority subscales had mean responses approximately representing the "agree" response on the five point rating scale.

Table 5.1

Per-item Means, Standard Deviations, and Sample Sizes by Subscale for the Pilot Questionnaire

Subscale	n	Mean	SD
Truth	96	2.85	.324
Fallibility	95	3.69	.353
Changeability	94	3.80	.402
Creativity	91	3.35	.372
Testability	90	3.78	.372
Speed of acquisition	100	3.99	.426
Role of authority	94	3.73	.303
Total		3.60	.366

A surprising and interesting finding was that the truth subscale, for which the response scored highest represented the humanistic coherence and consensus view, correlated negatively with all other subscales. From these initial results, students' conceptions of scientific knowledge are shown to be represented by beliefs that knowledge: (a) is a reflection of the real world; (b) is fallible; (c) is changeable; (d) is the product of the human imagination; (e) must be testable; (f) is acquired slowly; and (g) should be reasonably questioned.

Questionnaire Revisions

The revised questionnaire was shortened from 67 to 56 items. Those items containing difficult vocabulary, that were imprecise or lengthy, were reworded or deleted. Items that failed to correlate with the general questionnaire were omitted. On the basis of the empirical evidence from the pilot, the decision was made to change the response scored highest for the truth category from the coherence/consensus view to the correspondence view. This change ensured that all subscales correlated positively with each other.

The alpha reliability of the revised questionnaire was measured at 0.73, almost exactly the same as the pilot version. Even though the changes made would have improved the reliability, it is clear that there were counteracting influences as well. First, the overall number of items was reduced significantly, thereby reducing reliability. Also, several items were deleted that were considered difficult, ambiguous, or redundant, but were highly correlated with the general scale. Thus, their deletion had a negative effect of the instrument's overall reliability.

Essay Test Inter-rater Reliability

Since the Essay Test of Inductive Reasoning Strategies is in the early stages of its development, a measure of inter-rater reliability was necessary. The scorers were the author of this study and a cooperating teacher. The teacher taught high school science at a rural school. He held science and education degrees at the Bachelor's level. The school in which he taught was an all-grade school with a high school population of approximately 150.

A sample of 30 randomly selected essay tests was used for the inter-rater reliability measure. The cooperating teacher was not one of those involved in the administration of the instruments and had no other connection with the study. He did not receive coaching or advice on the scoring of the tests other than being given a copy of the test manual. The inter-rater reliability was measured as a Pearson product-moment correlation of 0.81.

Descriptive Results: Conceptions, Reasoning, and Science Achievement

The discussion in this section provides a general view of the descriptive findings related to conceptions of scientific knowledge, reasoning strategies, and achievement in different science courses. Subscale scores of the epistemological questionnaire are examined to determine specific conceptions while inductive reasoning strategies scores give insight into the types of reasoning processes used on the essay test. General comparisons are made between the scores of each of the subscales of the questionnaire and between the inductive reasoning strategy scores.

Students' Conceptions of Scientific Knowledge

Results of the questionnaire analysis are presented in Table 5.2. The average item score on the total questionnaire was 3.67 on the 5 point Likert-type scale. Though there is no standard to represent adequacy of student conception, the score of 3.67 is considerably above the "neutral" position, and is clearly in agreement with the questionnaire responses scored highest.

Table 5.2

Per-item Means, Standard Deviations, and Sample Sizes by Subscale for the Epistemological Questionnaire

Category	n	Mean	SD
Truth	331	3.36	0.37
Fallibility	331	3.56	0.42
Changeability	329	3.91	0.45
Creativity	334	3.43	0.38
Testability	333	3.79	0.38
Speed of acquisition	334	3.87	0.44
Role of authority	334	3.77	0.38
Total Score	304	3.67	0.23

The highest average score (3.91) is the category of changeability. It seems that high school students recognize that scientific knowledge can and will change. In addition, the high scores indicate that high school students' conceptions of knowledge include beliefs that: (a) knowledge is acquired slowly, (b) knowledge must be able to be tested, and (c) sometimes it is appropriate to question authority while at other times seeking expert opinion is a logical course of action. They believe, though less strongly, that scientific knowledge contains error and that it is the product of creative human minds. It appears there is some belief in the naive conception that scientific knowledge is simply "discovered". The lowest average score (3.36) was from the truth category. Students generally believe that scientific knowledge represents natural phenomena, that is, the correspondence view.

Students' Inductive Reasoning Ability

Results of the Essay Test of Inductive Reasoning Strategies are presented in Table 5.3. The mean total score of the test was .311. This result, while apparently very low, is to be considered with caution. It is not known what good reasoners of high school age should score on this test. As with conceptions of knowledge, there are no absolute standards. There is no absolute measure of good reasoning or even adequate reasoning. The essay test itself is in the process of being validated and has not been correlated with other reasoning tasks. All that can be said with confidence is that compared to the model scoring guide provided by the authors, students scored an average of .311 of what was possible for them to score. Taking these considerations into mind, this result seems to support previous findings that the level of reasoning ability shown by students is inadequate (Perkins, Allen, & Hafner, 1983; Schoenfeld, 1985).

Table 5.3

Means, Standard Deviations, and Sample Sizes for the Essay Test of Inductive Reasoning Strategies, Part A

#	Strategy	n	Mean	SD
1	Considering all information	311	.516	.183
3	Alternate explanations	311	.188	.107
2+3	Seek more information to generate explanations	311	.162	.084
4	Withholding judgement	310	.057	.163
5	Monitoring progress	311	.171	.241
6	Keeping process clear and orderly	311	.557	.214
7	Keeping focused	311	.793	.235
-	Total Score	311	.311	.162
-	Average depth	311	.182	.077
-	Average breadth deviation	311	.217	.066

Subjects scored an average of .217 on breadth. The overall breadth score is the average of breadth scores derived from measuring the actual proportion of effort for each data package compared to the ideal proportion of effort for each data package. This score is unlike the others in that lower values indicate better reasoning. Values closest to zero deviate the least from the ideal and indicate greater breadth of reasoning. The result indicates that subjects deviated an average of 21% from the ideal proportion of effort. While noteworthy, this deviation does not seem excessive. It appears that subjects actual proportion of effort is reasonably close to the ideal. Scores obtained from the data packages were generally proportionate to their total scores. Again, lack of standards for adequacy make this result difficult to interpret.

Subjects scored an average of .182 on depth. The overall depth score is the average of scores derived from measuring the number of explanations and alternate conclusions given <u>within</u> a data package as compared to the maximum number provided by the authors in the scoring key. It seems that the subjects only gave superficial treatment to the relevant points that needed explanation. The subjects provided less than 20% of the depth that was provided in the model of reasoning provided by the authors. This supports the previous research of Bereiter and Scardamalia (1985), Gettys and Englemann (1983), and Perkins, Allen and Hafner (1983) who respectively state: students only access a small portion of their knowledge when solving a problem, students fail to fully explore problem situations, and students analyse problem situations superficially.

When a comparison is made between the strategy scores, it can be observed that differences are quite dramatic. Subjects scored the highest (.793) on strategy 7, keeping focused to the main point. Subjects scored the lowest (.057) on strategy 4, withholding judgement. An interesting possibility is that the high score of strategy 7 may have contributed to the low score of strategy 4. It seems that during the test session subjects kept focused to the task at hand, the search for living creatures, but failed to consider the possibility the living creatures might not exist. It may be that subjects' perception of the task was to search for creatures <u>that were there</u>. If subjects perceived from the outset, for whatever reason, that creatures were present, there would be no reason to withhold judgement. It is true that as the days passed in the essay test simulation the accumulation of clues seemed to indicate an increasing likelihood that creatures were present.

Both the strategy 1 score (.516) and the strategy 6 score (.557) indicate these processes were handled less competently than strategy 7. With strategy 1 the subjects seemed to be able to recognize about half of the points of information that are relevant or need explanation. The strategy 6 score indicates that the problem was handled in a clear and orderly fashion about half of the time.

The remaining three strategies, providing alternate conclusions and explanations (.188), seeking more information when it's appropriate (.162), and monitoring progress (.171), all scored extremely low. This reasoning seems to be characterized by an inflexibility that forces the individual to follow a set path that may or may not lead to a solution. What's more, subjects seem not to be interested in determining if in fact they are on the right path at all. This supports Schoenfeld's (1985) contention that problem-solvers often pay little attention to the approach used to solve a problem, and often fail to consider whether the approach is justified or even whether progress is being made.

Students' Past Achievement in Science

Even though there is not a research question that specifically addresses differences in achievement between different science courses, science achievement is an important variable in the web of relationships being examined. Thus, it is relevant to note that different science courses offered in the school in which the study was carried out have a difference in average grade covering a range of more than 10%. Table 5.4 displays the comparative average grades in the science courses offered.

Table 5.4

Average Science Course Achievement

Course	Average Mark (%)
General Science	73.9
Biology	66.2
Chemistry	66.7
Physics	65.5
Earth Science	63.2

Test of the Causal Model

Since achievement in each science course is a different dependent variable, a separate correlational matrix was prepared for each. As well, a path analysis had to be completed separately for each measure of science achievement, and again, separately, for each measure of science achievement in a factored causal model. Results will be presented and discussed by science group, first for the correlational matrix, then for the general causal model, and finally for the factored causal model. A summary of all research questions and pertinent findings will be found at the end of the chapter.

Relationships Among Variables

Pearson correlations for all variables are provided in Table 5.5. Strong correlations were found between conceptions of scientific knowledge and science achievement for all sciences except Earth Science, which had a very small sample size of 20. Due to the small sample, and the resulting unstable analysis, the earth science data will not be included in subsequent data presentation and discussion.

Overall, students who hold conceptions of scientific knowledge that lean towards dimension poles scored highest achieve better in general science, biology, chemistry, and physics. This suggests three possible explanations.

Table 5.5

Pearson Correlations Among Conceptions of Scientific Knowledge, Inductive Reasoning, and Science Achievement for each Science Subject.

	Conception	Reasoning
Reasoning	.160	
General Science	.465	.358
Reasoning	.204	
Biology	.401	.266
Reasoning	.125	
Chemistry	.388	-0.065
Reasoning	.194	
Physics	.435	.224
Reasoning	.354	
Earth Science	-0.124	-0.020

The first is the possibility that past science grades and the epistemological questionnaire measure the same thing. The relationship could be influenced by evaluation in these courses that may, in some cases, measure these epistemological issues directly. Although this is a possibility, it is not likely to be a major influence. The science curricula offered in these subjects do not place a major emphasis on the nature of scientific knowledge. The different dimensions of the nature of scientific knowledge are not taught as part of the prescribed curriculum. Conversely, the epistemological questionnaire does not test science content.

The second explanation is that conceptions of scientific knowledge and achievement in science may be influenced together by a common variable not included in the causal model. It is likely that a variable, such as students' intelligence, affects their achievement in science. The effect of intelligence on conceptions of scientific knowledge, however, is unclear.

Third, the relationship may exist because students who have adequate conceptions of scientific knowledge also use that knowledge more effectively, and thus score higher science grades. This lends support to the hypothesis that students who have a deeper understanding of the nature of scientific knowledge will attain higher science scores.

Correlations between reasoning and achievement were somewhat lower, the highest being with general science. The correlations between reasoning and chemistry achievement and reasoning and earth science achievement were negative. As stated previously, the earth science data are, at best, unstable. The negative correlation between reasoning and chemistry, while not statistically significant, is nevertheless, interesting. Were this result sustained in further research, it would be difficult to explain why students who did better on the Essay Test of Inductive Reasoning Strategies performed more poorly in chemistry. One would expect the academic ability of the chemistry students to be similar to that of the physics students. Both are usually top academic students, generally of the highest ability.

The correlations between conception of scientific knowledge and reasoning ability are generally low, but consistently positive. The strongest relationship was shown for the biology and earth science groups. It appears that performing well on the epistemological questionnaire and on the reasoning test are related, but the order of the relationship is relatively low. This supports the hypothesis that students who have a deeper understanding of the nature of scientific knowledge are better reasoners.

Path Analysis of the General Causal Model

The standardized coefficients and R^2s for the general recursive model are reported in Table 5.6. The coefficients of determination (R^2s) indicate that significant portions of variance in science achievement can be predicted from conception of scientific knowledge and inductive reasoning. As much as 30% of the variance for general science achievement can be predicted by the two independent variables.

Table 5.6

Standardized Coefficients, R^2s, and Sample Sizes Among Conceptions of Scientific Knowledge, Inductive Reasoning, and Science Achievement for each Science Subject

Dependent Variables	n	Independent Variables		R^2
		Conception	Reasoning	
Reasoning	89	.160	-----	.026
General Science	89	.419***	.291**	.299
Reasoning	131	.204*	-----	.042
Biology	131	.362***	.192*	.196
Reasoning	78	.125	-----	.016
Chemistry	78	.403***	-0.115	.164
Reasoning	61	.194	-----	.038
Physics	61	.407**	.145	.210

* $p<.05$ ** $p<.01$ *** $p<.001$

It is observed that knowledge conception significantly affected general science, biology, chemistry, and physics achievement. All are powerful effects, three of them significant at the .001 level.

The effect of the intervening variable of reasoning on science achievement was generally weaker than that of knowledge conception, but in the case of general science and biology achievement the effect was still significant. With chemistry achievement, the intervening variable of reasoning had a weak, but negative, effect. This result is curious since conceptually is seems inconsistent that being able to reason better would result in poorer performance in a school subject. Further research is needed to determine if there are fundamental differences in the nature of course content, its presentation, or teaching methodologies between high school courses that might account for variation in reasoning. A repeat study, isolating the relationship between inductive reasoning and chemistry achievement would be desirable.

The effect of scientific knowledge conception on inductive reasoning ability was found to be significant in the case of the biology group. In the other sciences the effect was positive but less powerful. The path diagrams for each path analysis are presented in Figures 5.1 - 5.4.

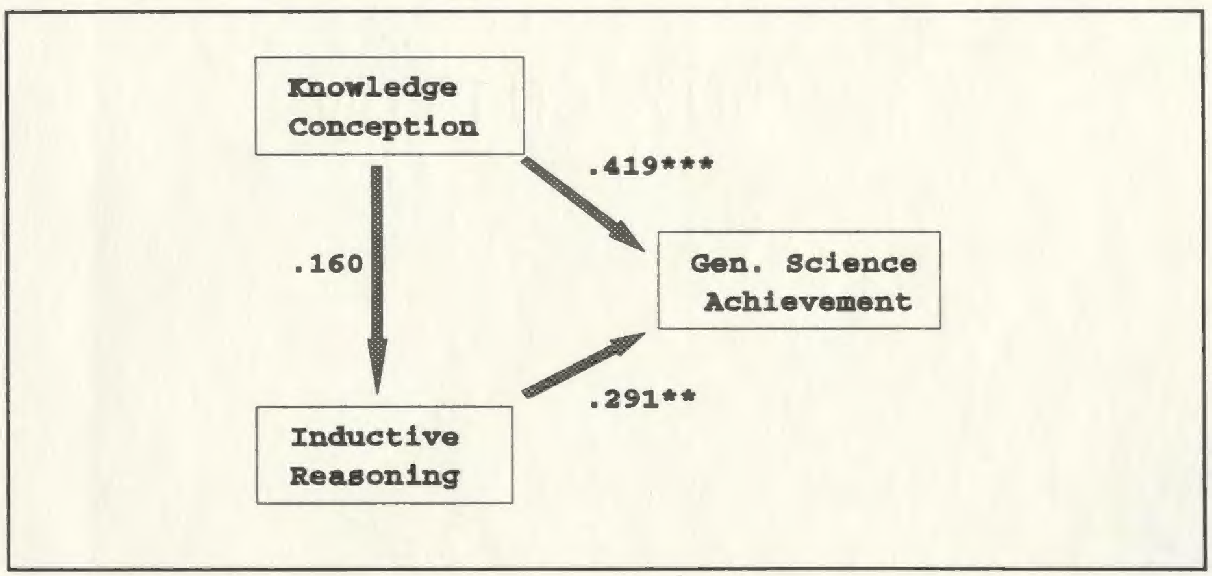

Figure 5.1. Path diagram of effects on general science.

(** p<.01; *** p<.001)

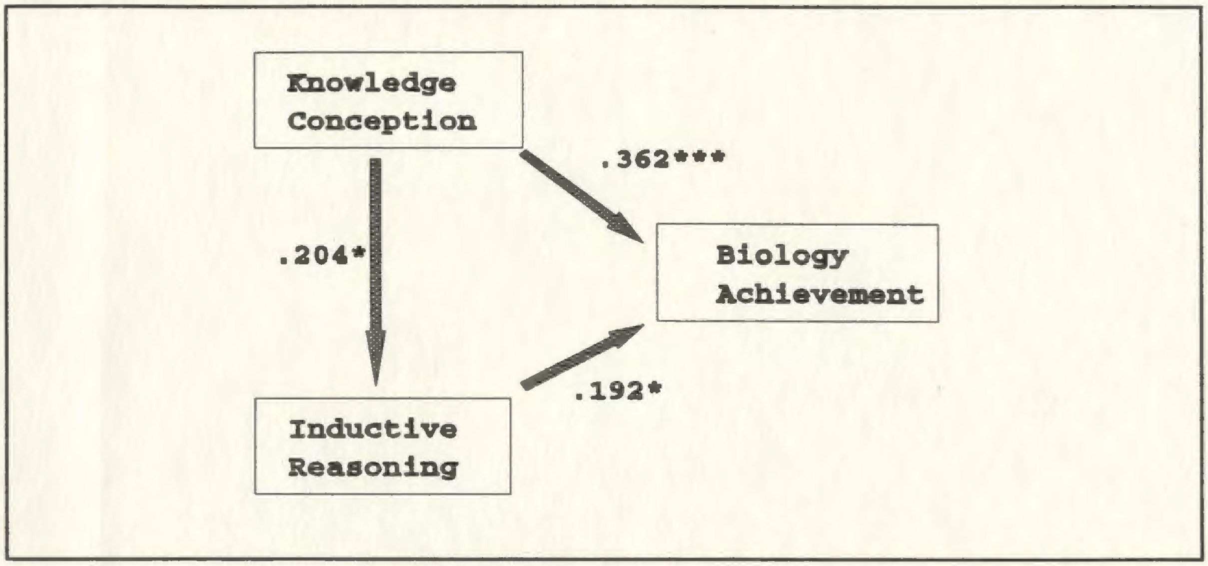

Figure 5.2. Path diagram of effects on biology.

(* p<.05; *** p<.001)

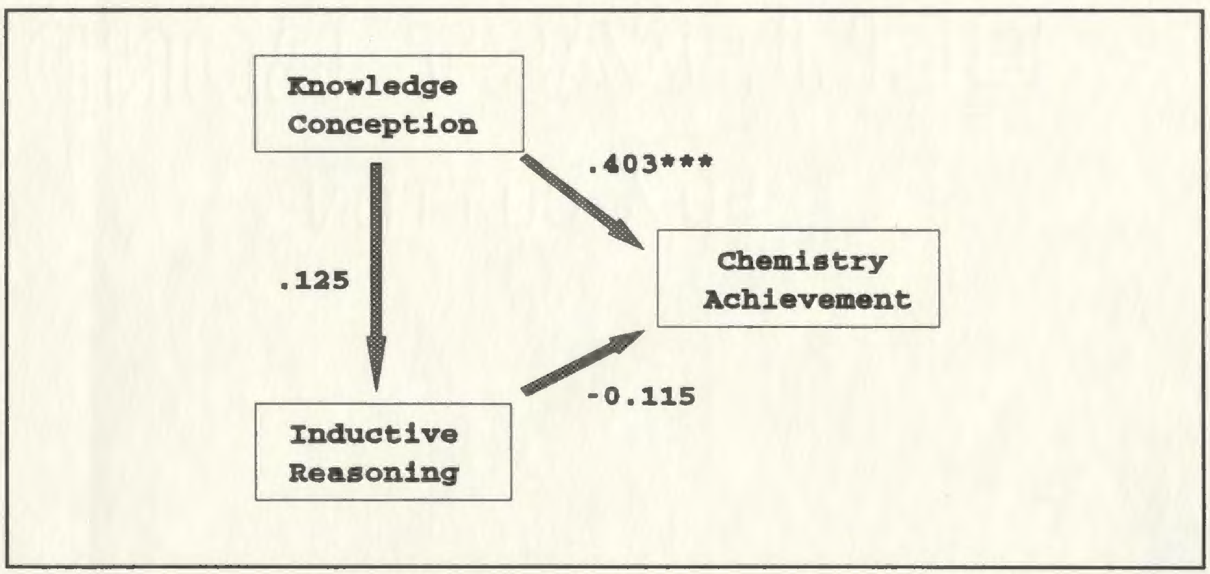

Figure 5.3. Path diagram of effects on chemistry.

(*** p<.001)

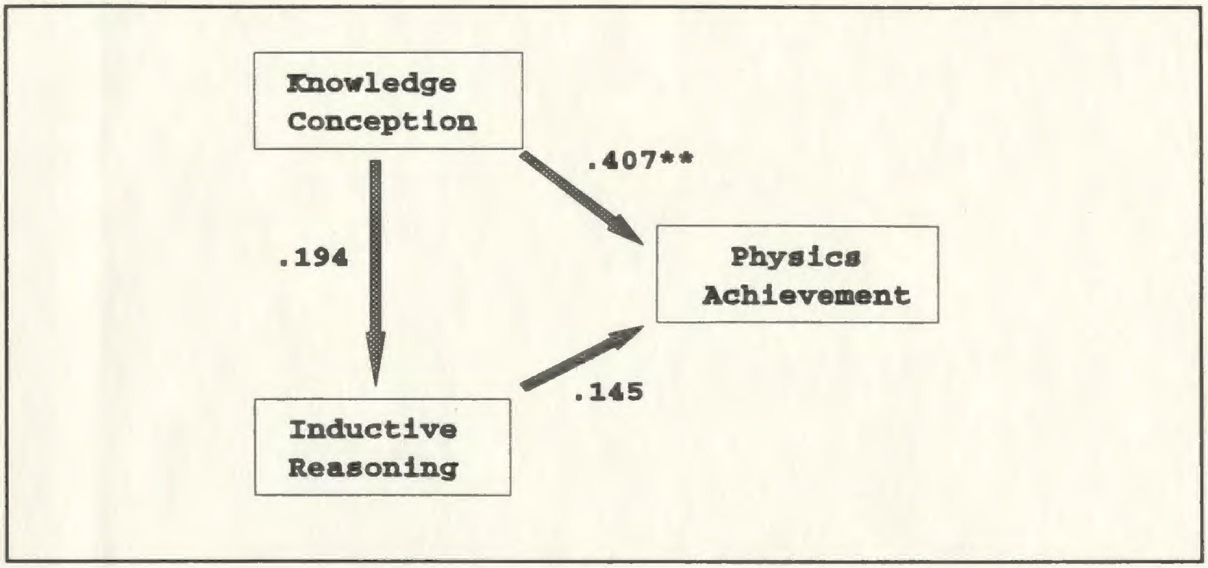

Figure 5.4. Path diagram of effects on physics.
(** p<.01)

In order to provide a closer examination of the effects of knowledge conception, the direct effects on reasoning and the direct and indirect effects on achievement are presented in Table 5.7. It can be seen that the mediating effects of reasoning are quite small. Direct effects of scientific knowledge conception are much more potent than the indirect effects. Even the negative indirect effect of reasoning on chemistry change the total effect of knowledge conception very little. This again indicates that reasoning plays a much smaller role in determining science achievement than does conception of scientific knowledge.

Table 5.7

Total, Direct, and Indirect Effects Among Conceptions of Scientific Knowledge, Inductive Reasoning, and Science Achievement for each Science Subject

Dependent Variables	Independent Variables	Direct	Indirect (Through Reasoning)	Total
Reasoning	Conception	.160	----	.160
G. Science	Conception	.419	.047	.466
	Reasoning	.291	----	.291
Reasoning	Conception	.204	----	.204
Biology	Conception	.362	.039	.401
	Reasoning	.192	----	.192
Reasoning	Conception	.125	----	.125
Chemistry	Conception	.403	-0.014	.389
	Reasoning	-0.115	----	-0.115
Reasoning	Conception	.194	----	.194
Physics	Conception	.407	.028	.435
	Reasoning	.145	----	.145

The comments of the anonymous reviewer of this thesis are relevant here. "To what extent is inductive reasoning an element of the academic activities and tasks that occur in our learning environments?" the reviewer asks. These results support the findings of several studies related to the current status of our instructional activities. They indicate that activities such as those involving inductive reasoning strategies are all but ignored in our present curriculum (Bloom, 1989; Duschl & Wright, 1989; Gallagher, 1991; Mitman et al., 1987; and Tobin, 1987). For example, Mitman et al. (1987) found that teachers rarely or never teach noncontent components of science in their presentations and academic work assignments, and that teachers' references to noncontent components were generally of poor quality. In addition, students perceive content as the prominent focus of teacher instruction.

Overall, the results of this study show there is limited support for the hypotheses that a deeper understanding of the nature of scientific knowledge will result in better reasoning. There is inconsistent support for the hypothesis that better reasoning will result in higher achievement in science. There is strong support for the hypothesis that a deeper understanding of the nature of scientific knowledge will result in higher science achievement.

Factored Dimensions of the Questionnaire

An exploratory factor analysis was completed to isolate specific components of the conception of scientific knowledge that influence inductive reasoning and science achievement. It was used to determine whether there was an underlying pattern of relationships existing among the seven questionnaire subscales. The factor analysis allowed the subscales to be reduced to a smaller set of factored epistemological dimensions to be used as variables in the statistical analysis. Thus, the significant effects of total knowledge conception on reasoning and achievement previously determined can be examined and analysed as the sum of the factored epistemological components.

The correlational matrix from which the variables were extracted is presented in Table 5.8. The factor analysis performed was a principal component analysis extracting four factors, with a varimax rotation. It was decided to extract four factors since the size of the eigenvalues after the fourth remained approximately constant.

Results of the Factor Analysis. Table 5.9 displays the rotated loadings from the factor analysis. Inspection of the factor 1 loadings indicate that truth and testability are highest correlated. The responses scored highest in these categories are a correspondence view of truth and the notion that knowledge must be empirically testable. Both these conceptions imply that knowledge reflects real, intelligible, and sensory perceptible phenomena. This is basically a Realist view of scientific knowledge and as a result, the factor 1 variable is labelled "realistic". That is, the knowledge represents phenomena that are able to be perceived clearly with the mind and senses.

Table 5.8

Correlational Matrix for Questionnaire Subscales

Subscale	1.	2.	3.	4.	5.	6.	7.
1. Truth	1.000						
2. Fallibility	.172	1.000					
3. Changeability	.200	.270	1.000				
4. Creativity	.044	.225	.091	1.000			
5. Testability	.314	.198	.264	.074	1.000		
6. Speed of acquisition	.191	.121	.256	.090	.218	1.000	
7. Role of authority	.312	.286	.308	.052	.304	.182	1.000

Table 5.9

Rotated Loading of the Factor Analysis

Rotated Loadings	1	2	3	4
Truth	.836	.027	.092	.053
Fallibility	.133	.473	-0.067	.635
Changeability	.044	-0.063	.359	.768
Creativity	.017	.938	.066	.002
Testability	.712	.048	.149	.210
Speed of acquisition	.182	.058	.927	.100
Role of authority	.468	-0.044	-0.085	.639
Variance explained by rotated components	1.1479	1.115	1.036	1.458
Percent of total variance explained	21.122	15.932	14.800	20.829

An examination of the factor 2 and factor 3 loadings reveal creativity and speed of acquisition factors. They will be labelled "creative" and "developed" to describe the responses that knowledge is a product of the human imagination, and knowledge is acquired slowly and developed over time, rather than acquired quickly.

The loadings of factor 4 reveal that fallibility, changeability, and role of authority are all highly correlated. As such they will be combined to be the fourth factored variable. The responses scored highest for these categories describe knowledge as (a) containing error, (b) being subject to a constant change process, and (c) being open to question depending on specific circumstances. These responses describe the human influence on the creation and interpretation of knowledge. To emphasize the error and instability inherent in knowledge, the factor 4 variable is labelled "humanistic".

Theoretical Model of Factored Epistemological Variables. Questionnaire sections that loaded on the same factors were added together creating four new factored variables. These were then used to model effects on reasoning and science achievement. The difference between the previously examined general model and the factored model is that the direct causal relationship between conception of scientific knowledge and science achievement, and between conception of scientific knowledge and reasoning, is subdivided into four different causal effects in the factored model. Thus, in the factored model, there are four direct effects of conceptions of scientific knowledge on achievement in science, and four indirect effects of conception of scientific knowledge on science achievement. There are four direct effects of conception of scientific knowledge on inductive reasoning. The theoretical model representing the factored epistemological categories is displayed in Figure 5.5.

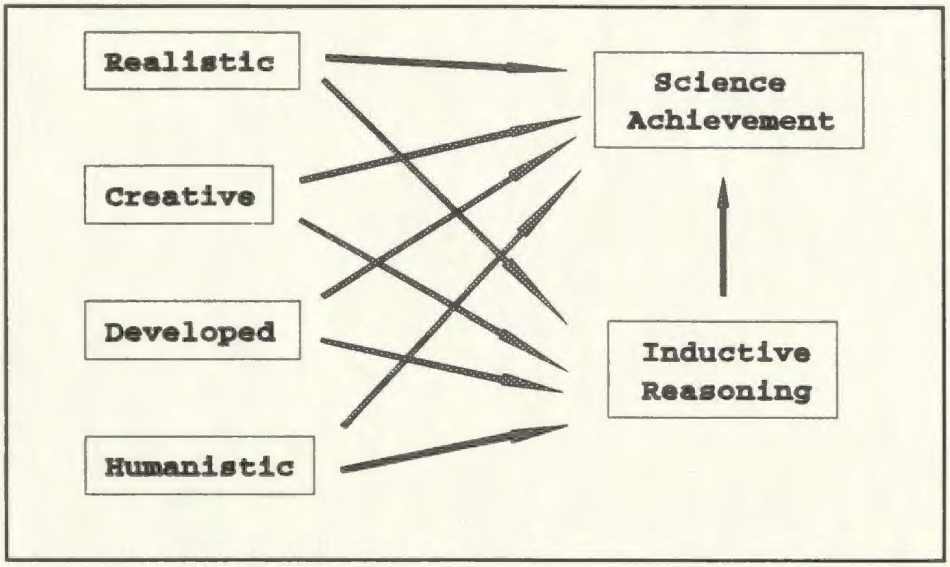

Figure 5.5. Theoretical model of causal relationships among factored epistemological categories, inductive reasoning, and achievement in science.

Path Analysis of the Factored Epistemological Model. The standardized coefficients and R^2s for the factored epistemological model are reported in table 5.10. The path diagrams for each analysis are presented in Figures 5.6 - 5.9 at the end of the section.

With regard to the dependent variable of science achievement, it is observed that realistic and humanistic factors have a significant effect on general science and biology achievement. This means that those general science and biology students who believe that knowledge (a) is an approximation of natural phenomena, (b) must be testable, and (c) influenced by human factors will achieve higher grades. The conception that knowledge is slowly

Table 5.10

Standardized Coefficients, R^2s, and Sample Sizes Among Factored Epistemological Variables, Inductive Reasoning, and Science Achievement for each Science Subject

Dependent Variables	n	Independent Variables					R^2
		Realistic	Creative	Developed	Humanistic	Reasoning	
Reasoning	89	-0.048	-0.094	-0.123	.327**	----	.095
G. Science	89	.241*	.066	-0.048	.303**	.263**	.325
Reasoning	131	.102	.017	-0.091	.221	----	.068
Biology	131	.175	.053	-0.052	.319***	.163*	.225
Reasoning	78	-0.081	-0.016	.188	.110	----	.041
Chemistry	78	.024	-0.042	.355**	.241	-0.156	.230
Reasoning	61	.220	-0.174	.117	.051	----	.088
Physics	61	.312*	.093	-0.060	.186	.142	.246

* $p<.05$ ** $p<.01$ *** $p<.001$

acquired or developed over a period of time significantly affected chemistry students only. The realistic factor only had a significant effect on physics achievement. This means that those students who believed in a correspondence view of truth, and that knowledge must be testable, achieved higher grades in this subject. Creative knowledge was not found to be significant on any variables.

With regard to reasoning as a dependent variable, it can be seen to be significantly affected only by the conception that knowledge is humanistic. Thus, for general science and biology students a belief that knowledge is fallible, changeable, and open to question results in better reasoning.

Overall, two patterns emerge from the data. First, the effects on general science, biology, and to some extent, physics achievement, seem to be parallel. They are affected significantly by conceptions that knowledge is realistic and humanistic, with the exception of physics which is only affected by the realistic conception of knowledge. They are negatively affected by the conception that knowledge is slowly acquired and developed over time. Chemistry achievement, alternately, is not affected by conceptions of realistic or humanistic knowledge, but is significantly affected by the conception of developed knowledge. Again, the exception for chemistry achievement is quite interesting. Chemistry is not perceptibly different from the other sciences with regard to the teaching of how knowledge is acquired.

The second pattern is that conceptions that knowledge is creative and developed, for the most part, negatively affect achievement and reasoning. Thus, the large direct effect obtained for the general conception of knowledge on reasoning and achievement are due to the realistic and humanistic components of the total knowledge conception.

To allow a closer examination of the effects of the factored epistemological variables, the direct and indirect effects are presented in Table 5.11. Again it can be seen that the mediating effects of reasoning are quite small. The largest indirect effects were humanistic knowledge on general science achievement, humanistic knowledge on biology achievement, and realistic knowledge on physics achievement. Overall, the variables having the greatest effect on achievement are realistic and humanistic knowledge. The variable having the greatest effect on reasoning ability was humanistic knowledge.

Table 5.11

Total, Direct, and Indirect Effects Among Factored Epistemological Variables, Inductive Reasoning, and Science Achievement for each Science Subject

Dependent Variables	Independent Variables	Direct	Indirect Reasoning	Total
Reasoning	Realistic	-0.048	----	-0.048
	Creative	-0.094	----	-0.094
	Developed	-0.123	----	-0.123
	Humanistic	.327	----	.327
General Science	Realistic	.241	-0.013	.228
	Creative	.066	-0.025	.041
	Developed	-0.048	-0.032	-0.080
	Humanistic	.303	.086	.389
	Reasoning	.263	----	.263
Reasoning	Realistic	.102	----	.102
	Creative	.017	----	.017
	Developed	-0.091	----	-0.091
	Humanistic	.221	----	.221
Biology	Realistic	.175	.017	.192
	Creative	.053	.003	.056
	Developed	-0.052	-0.015	-0.067
	Humanistic	.319	.036	.355
	Reasoning	.163	----	.163
Reasoning	Realistic	-0.081	----	-0.081
	Creative	-0.016	----	-0.016
	Developed	.188	----	.188
	Humanistic	.110	----	.110
Chemistry	Realistic	.024	.013	.037
	Creative	-0.042	.002	-0.040
	Developed	.355	-0.029	.326
	Humanistic	.241	-0.017	.224
	Reasoning	-0.156	----	-0.156
Reasoning	Realistic	.220	----	.220
	Creative	-0.174	----	-0.174
	Developed	.117	----	.117
	Humanistic	.051	----	.051
Physics	Realistic	.312	.031	.343
	Creative	.093	-0.025	.068
	Developed	-0.600	.017	-0.043
	Humanistic	.186	.007	.193
	Reasoning	.142	----	.142

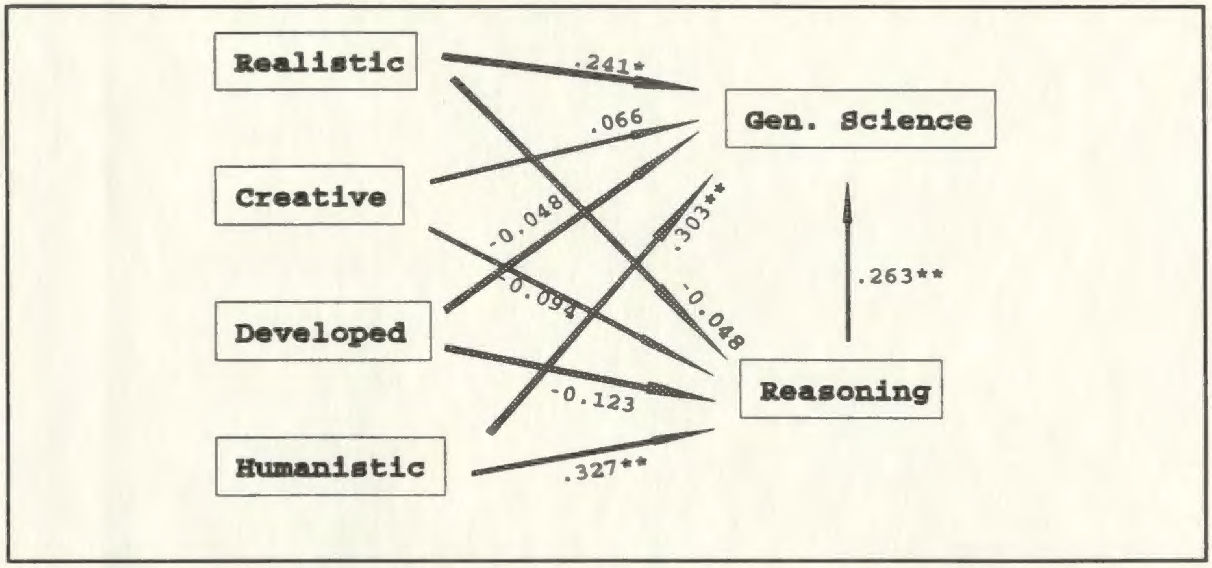

Figure 5.6. Path diagram of causal relationships among factored epistemological categories, reasoning, and general science achievement. (* p<.05; **p<.01)

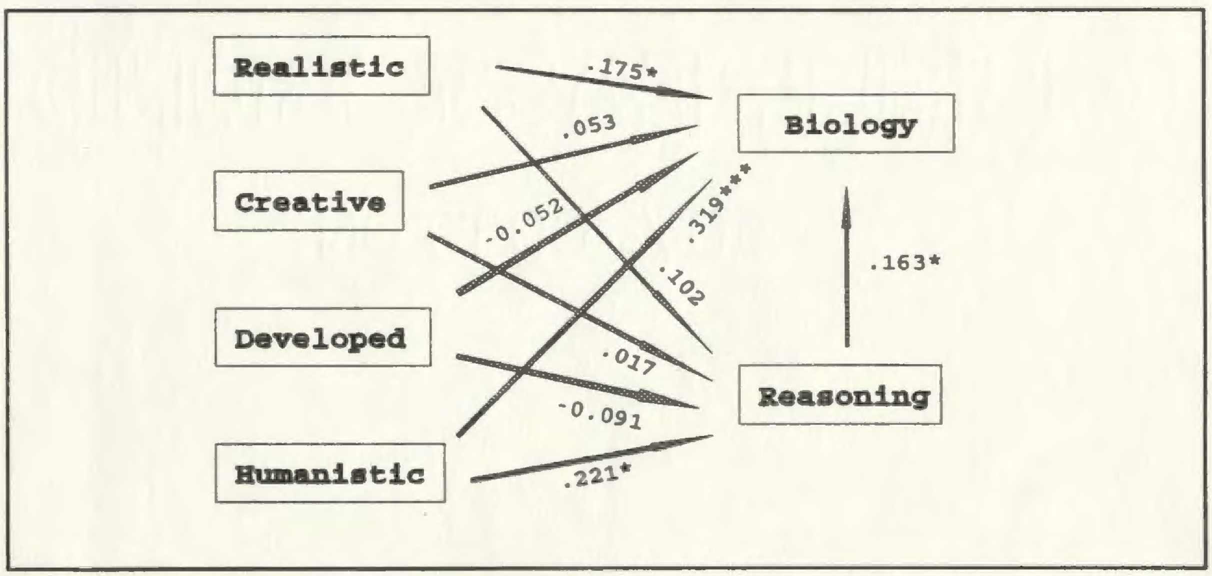

Figure 5.7. Path diagram of causal relationships among factored epistemological categories, reasoning, and biology achievement. (* p<.05; *** p<.001)

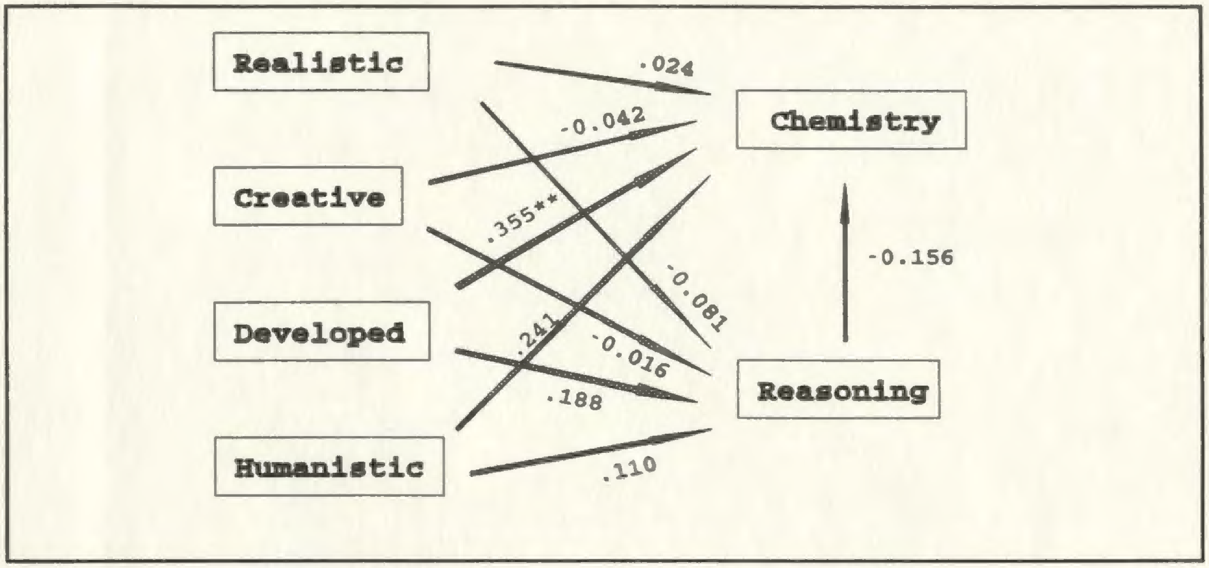

Figure 5.8. Path diagram of causal relationships among factored epistemological categories, reasoning, and chemistry achievement. (** p<.01)

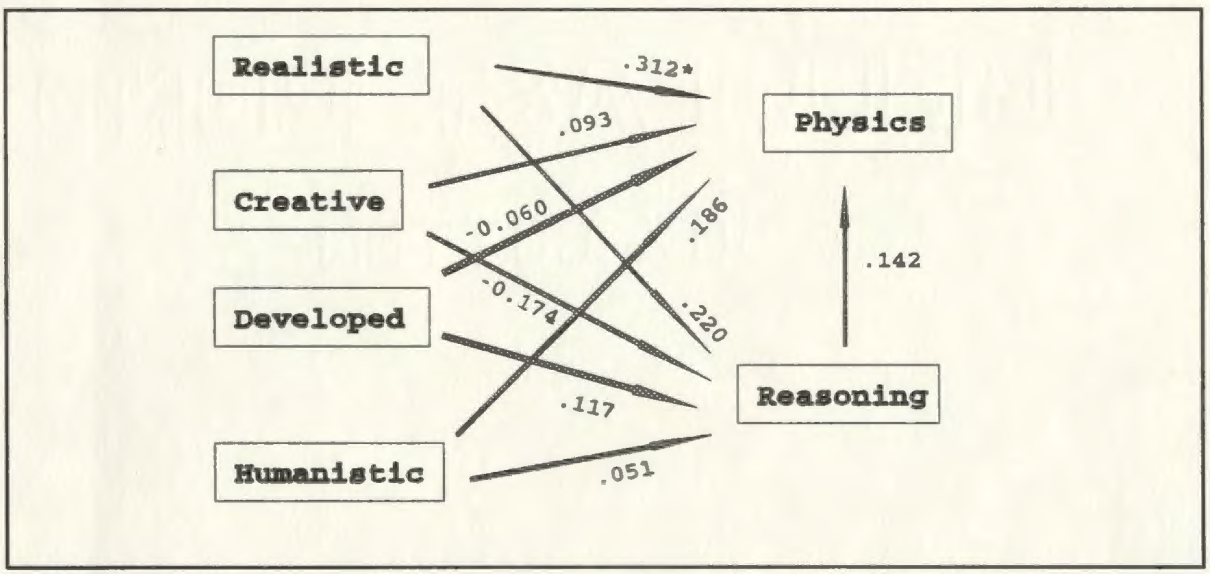

Figure 5.9. Path diagram of causal relationships among factored epistemological categories, reasoning, and physics achievement. (* p<.05)

Summary

Research Question #1: What conceptions of scientific knowledge are held by high school students?

Students' conceptions included beliefs that scientific knowledge: (a) represents real world phenomena, (b) is fallible, (c) is changeable, (d) is a product of the human imagination, (e) must be subjectable to empirical test, (f) is acquired slowly, and (g) should be questioned when reasonable to do so.

Research Question #2: What inductive reasoning strategies are used by high school students?

Students' strategy scores indicated that they were best able to keep focused on the reasoning task. They generally considered relevant information, and kept the reasoning process clear and orderly. Students scored poorly on strategies involving withholding judgement, seeking more information when necessary, producing alternate explanations, and monitoring their own progress while reasoning. A total average score of .311 on the Essay Test of Inductive Reasoning Strategies seems to indicate low overall reasoning ability.

Research Question #3: What is the relationship between students' conception of scientific knowledge and their inductive reasoning ability?

Students' conception of scientific knowledge had a positive, but limited, effect on inductive reasoning ability. Only for the biology group did knowledge conception exert a significant effect on inductive reasoning.

Research Question #4: What is the relationship between students' conception of scientific knowledge and their achievement in science?

The relationship between students' conception of scientific knowledge and their achievement in science was the strongest found. In all cases knowledge conception exerted a strong influence over achievement, with the effect on general science, biology, and chemistry significant at the .001 level. With physics, the relationship was significant at the .01 level.

Research Question #5: What is the relationship between students' inductive reasoning ability and their achievement in science?

The relationship between students' inductive reasoning ability and their achievement in science was found to be inconsistent. With general science and biology, inductive reasoning exerted a significant effect. With physics and chemistry, inductive reasoning did not have an effect.

Hypotheses

The first hypothesis, that students who have a deeper understanding of the nature of scientific knowledge are better reasoners, was supported only for the biology group. The second hypothesis, that students who have a deeper understanding of the nature of scientific knowledge will obtain higher achievement scores, was strongly supported for all groups. The third hypothesis, that students who are better inductive reasoners will attain higher science achievement scores, was supported for the general science and biology groups.

CHAPTER VI: SUMMARY AND DISCUSSION OF IMPLICATIONS

The study examined a causal model describing the relationships among conceptions of scientific knowledge, inductive reasoning, and science achievement. The model proposed the following relationships: (a) students' conceptions of scientific knowledge will influence their achievement in science, (b) students' conceptions of scientific knowledge will influence their inductive reasoning, (c) students' inductive reasoning will influence their achievement in science, and consequently, (d) students' conceptions of scientific knowledge will indirectly influence their achievement in science through changes in inductive reasoning. The basis for the model was the theoretical support, derived from the literature, for the assumption that beliefs about the nature of knowledge will influence how that knowledge is used.

Summary of Measures and Methodology

A review of the existing instruments assessing the nature of scientific knowledge revealed that none of those examined would be appropriate for use in this study. The Essay Test of Inductive Reasoning Strategies, Part A was selected as the measure of reasoning ability. Final science grades from the previous school year were used as the measure of science achievement.

Epistemological Questionnaire

Through an examination of the philosophic literature and existing instruments assessing conceptions of the nature of science, dimensions of the nature of scientific knowledge were isolated. These included: (a) scientific truth, (b) the fallibility of scientific knowledge, (c) the changeability of scientific knowledge, (d) the role of creativity in scientific knowledge production, (e) the testability of scientific knowledge, (f) the speed of scientific knowledge acquisition, and (g) the role of authority in scientific knowledge production.

The 67 item preliminary form was piloted with 105 students in a suburban regional high school. On the basis of the pilot data, the questionnaire was revised resulting in a final version of 56 items containing eight item statements pertaining to each dimension of scientific knowledge.

Essay Test of Inductive Reasoning Strategies

To complete the essay test, subjects had to imagine they were on an imaginary planet for four days with their task being to search the planet for living creatures. They were asked to read about the things that happen each day, and to think about what these events mean for their search. Subjects were then asked to write what they are thinking about the things that happened on that and previous days; and write what they plan to do on their search because those things happened. Subjects' reasoning was then measured against an ideal for the situation.

The theoretical basis for the test is Ennis's view of induction as a type of inference within the larger context of critical thinking. The test assesses seven strategies used in inductive reasoning: (a) recognizing relevant information, (b) seeking information when appropriate, (c) producing alternate conclusions or explanations, (d) withholding judgement, (e) monitoring progress, (f) being organized, and (g) staying focused.

Method

A sample of 305 mixed ability, high school students completed the epistemological questionnaire and the essay test. Descriptive assessments of students' conceptions and inductive reasoning ability were made and results on the epistemology questionnaire were correlated with overall performance on the essay test and with the students' past achievement in each science course. A path analysis was completed to determine the strength of the relationships proposed in the general causal model.

In order to isolate the causal effects of specific aspects of conceptions of scientific knowledge, an exploratory factor analysis was performed on the seven questionnaire subscales. Four factored epistemological variables were produced characterizing scientific knowledge as: (a) realistic, (b) creative, (c) developed, and (d) humanistic. A theoretical causal model incorporating the factored variables as conceptions of scientific knowledge was proposed and evaluated.

Summary of Findings

It was found that student responses on the epistemological questionnaire generally leaned towards the dimension poles scored highest. The highest subscale score was for the student belief that scientific knowledge was changeable; the lowest was for the student belief that scientific knowledge approximates real world phenomena. Students also held conceptions that scientific knowledge: (a) is fallible, (b) is a product of human creativity, (c) must be testable, (d) is acquired slowly, and (e) should be questioned where appropriate.

Depth scores from the inductive reasoning test revealed that subjects gave only superficial treatment to the relevant information provided. Poor strategy scores indicated that subjects tended not to: (a) seek more information when appropriate, (b) generate alternate explanations or conclusions, (c) monitor their progress, or (d) withhold judgement. In light of these scores and low total scores, student inductive reasoning was judged generally inadequate.

When all variables were correlated, strong relationships were found between conceptions of scientific knowledge and achievement in general science, biology, chemistry, and physics. The relationship between inductive reasoning and achievement in science was found to be inconsistent. Lower level correlations were found between conceptions of scientific knowledge and inductive reasoning ability.

Results of the path analysis for the general model showed significant effects of conceptions of scientific knowledge on achievement in all sciences. There was a smaller effect from reasoning, though the relationship was still significant for general science and biology achievement. The effect of conception of scientific knowledge on inductive reasoning was found to be significant only for the biology group. Direct effects of scientific knowledge conception were found to be much more powerful than the indirect effects through reasoning.

The path analysis for the factored epistemological model indicated that the large direct effect obtained for conception of scientific knowledge on achievement was generally due to the realistic and humanistic components of the total scientific knowledge conception. The direct effect of scientific knowledge conception on inductive reasoning was due to the humanistic component of the scientific knowledge conception.

Discussion of Implications

From the examination of the research literature and from the difficulties and concerns experienced when selecting and developing instruments for this study, it is apparent that there should be an emphasis placed on the testing and measurement of epistemology and reasoning. There is a need for more and better instruments to assess students conceptions of scientific knowledge. Tests should be developed that enable a portrayal of students' particular views of the nature of scientific knowledge, not of whether or not they hold the correct view. There is also a need for more and better tests of reasoning, including inductive reasoning, and there should be more studies to determine current levels of student reasoning ability. A consensus is required on what reasoning abilities students should hold.

The results of the present study support Schommer's (1989) contention that the conclusions students draw from knowledge are influenced by their interpretation of that knowledge. It was observed that general science and biology students who believed knowledge to be humanistic, that is, to be fallible, changeable, and open to reasonable questioning, scored significantly higher on the inductive reasoning test. Thus, holding these specific views of the nature of scientific knowledge modified how these students used that knowledge in the reasoning process.

An interesting finding relating to the status of the current science curriculum was the identification of those specific epistemological views that result in better science achievement. Since holding these views results in higher science scores, it is logical to assume that these views are endorsed by our present science curricula, either explicitly or implicitly. Specifically, general science and biology students who believed that scientific knowledge is realistic and humanistic, and physics students who believed that scientific knowledge is realistic, obtained significantly better achievement scores. Thus, holding the positivist view that scientific knowledge is a faithful copy of reality results in better grades in our present curriculum than believing that scientific knowledge results from consensus among members of the scientific community and from coherence with our existing set of beliefs.

This supports Duschl's (1983) view that science teachers' beliefs and curriculum materials are congruent with the conceptions of logical positivism. Science for these teachers typically consists of a body of knowledge arrived at by neutral, objective application of scientific method (King, 1991). Teachers and students are reinforced in this philosophy by curriculum materials in which authors tell a story of what we know. Textbooks are idealized and present scientific knowledge as "revealed truth" (Brickhouse, 1989; Gallagher, 1991).

Teachers' conceptions influence decisions that affect the learning environment. These include decisions related to laboratory instruction, use of demonstrations, word usage, instructional goals, and selection of curriculum materials (Brickhouse, 1991; Duschl, 1983). Clearly then, teachers holding conceptions characterizing logical positivism influence what and how students learn in these science courses, and thus influence student achievement. These results focus on the need to de-emphasize the outdated philosophy of logical positivism that dominates our high school curricula, and to place an emphasis on the humanistic and social issues related to the nature of science (Duschl, 1988).

Before curricula can be developed with a greater epistemological focus, three key questions must be addressed: (a) What epistemological beliefs should be taught? (b) How do we best teach epistemological beliefs? and (c) To what extent are the effects of epistemological beliefs on inductive reasoning and achievement generalizable to other tasks and domains?

Explicitly teaching a specific set of epistemological beliefs about scientific knowledge implies that the nature of scientific knowledge is known. A danger exists that students will perceive the particular philosophical view presented as the single correct view. Therefore, it is important that teachers and curriculum developers present the nature of scientific knowledge as open, dynamic, and subject to interpretation. Clearly, the most modern philosophical views should be taught, but students should also be made aware of alternate and older views. To demonstrate the dynamics of epistemological beliefs, topics should be presented in a historical context.

The level of cognitive development of the learners must be an important consideration in deciding what epistemological topics to teach. While some dimensions of knowledge, such as fallibility, may be easy for young children to grasp, other dimensions, such as the nature of scientific truth, are conceptually more difficult. Teachers and curriculum developers need to be aware that the views presented must be appropriate cognitively for the learners.

At present, it is not clear how best to teach epistemological beliefs to produce the greatest influences on reasoning and achievement. The consequences of different types of explicit instruction are unknown. Simply making students aware of their own epistemological beliefs may be enough to improve reasoning and achievement. Current learning situations that stress content to the exclusion of epistemology, forgo the potential benefits of improved student reasoning. Teachers should incorporate not only considerations for the learner, but also the nature of the subject matter when designing instructional tasks. Programs designed to teach more reasoned conceptions and to emphasize the relationships among epistemology, reasoning, and achievement, should produce more profound benefits.

The extent to which the effects of students' epistemology are generalizable to other tasks and domains determines the significance of epistemology to the field of education. Through this study and Schommer's (1989) it has been shown that epistemological beliefs affect inductive reasoning and comprehension of a text passage respectively. The present study has also shown that significant effects of knowledge conception can be found on achievement in each of four different school science subjects. Many studies are necessary to determine the effect of epistemology on other tasks, aspects of comprehension, forms of reasoning, and achievement in different subject areas. If individuals' personal epistemologies affect a wide range of mental abilities, then this potentially can have a major impact on learning. Isolating the role of epistemology in student thinking will potentially have significant effects on curriculum and instruction.

REFERENCES

Abimbola, I. O. (1983). The relevance of the "new" philosophy of science for the science curriculum. School Science and Mathematics, 83, 181-193.

Aikenhead, G. S. (1973). The measurement of high school students' knowledge about science and scientists. Science Education, 57, 539-549.

Aikenhead, G. S. (1987). High-school graduates' beliefs about science-technology-society. III. Characteristics and limitations of scientific knowledge. Science Education, 71, 459-487.

Aikenhead, G. S., Fleming, R. W., & Ryan, A. G. (1987). High school graduates' beliefs about science-technology-society. I. Methods and issues in monitoring students' views. Science Education, 71, 145-161.

Alderton, D. C., Goldman, S. R. & Pelligrino, J. W. (1982). Multitask assessment of inductive reasoning skill. Paper presented at the annual meeting of the American Educational Research Association, New York.

Anderson, R. C. (1984). Some reflections on the acquisition of knowledge. Educational Researcher, 13(9), 5-10.

Anderson, H. O., Harty, H. & Samuel, K. V. (1986). Nature of science, 1969 and 1984: Perspectives of preservice secondary science teachers. School Science and Mathematics, 86, 43-50.

Andrich, D. & Masters, G. N. (1988). Rating scale analysis. In J. P. Keeves (Ed.), Educational research, methodology, and measurement: An international handbook (pp. 297-303). Elmsford, NY: Pergamon Press

Association for Science Education (1979). Alternatives for science education: A consultative document. Hatfield, MA: Association for Science Education.

Babbie, E. R. (1973). Survey research methods. Belmont, CA: Wadsworth.

Baron, J. (1985). Rationality and intelligence. New York: Cambridge University Press.

Bereiter, C. & Scardamalia, M. (1985). Cognitive coping strategies and the problem of inert knowledge. In S. S. Chipman, J. W. Segal, & R. Glazer (Eds.), Thinking and learning skills: Vol. 2 Current research and open questions (pp. 65-80). Hillsdale, NJ: Erlbaum.

Black, M. (1954). Problems of analysis. Ithaca, NY: Cornell University Press.

Bloom, J. W. (1989). Preservice elementary teachers' conceptions of science: Science, theories and evolution. International Journal of Science Education, 11, 401-415.

Borg, W. R. & Gall, M. D. (1983). Educational research: An introduction. New York: Longman.

Brickhouse, N. W. (1989). The teaching of the philosophy of science in secondary classrooms: Case studies of teachers' personal theories. International Journal of Science Education, 11, 437-439.

Bruner, J. S., Goodnow, J., & Austin, G. A. (1956). A study of thinking. New York: Wiley.

Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56 (2), 81-105.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1988, June). "An experiment is when you try it and see if it works": A study of junior high school students' understanding of the construction of scientific knowledge. (Rep. No. ETC-TR-88-19). Cambridge, MA: Educational Technology Center. (ERIC Document Reproduction Service No. ED 303 366)

Colberg, M., Nester, M. A., & Cormier, S. M. (1982). Inductive reasoning in psychometrics: A philosophical corrective. Intelligence, 6, 139-164.

Conant, J. B. (1951). On understanding science. New York: Mentor Books.

Conant, J. B. (1955). Robert Boyle's experiments in pneumatics. Harvard Case Histories in Experimental Science, 1, 3-63. Cambridge, MA: Harvard University Press.

Cooley, W. W. & Klopfer, L. E. (1961a). Test on Understanding Science. Princeton, NJ: Educational Testing Service.

Cooley, W. W. & Klopfer, L. E. (1961b). Test on Understanding Science: Test Manual. Princeton, NJ: Educational Testing Service.

Cooley, W. W. & Klopfer, L. E. (1963). The evaluation of specific educational innovations. Journal of Research in Science Teaching, 1, 73-80.

Cotham, J. C. & Smith, E. L. (1981). Development and validation of the Conceptions of Scientific Theories Test. Journal of Research in Science Teaching, 18, 387-396.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. American Psychologist, 12, 671-684.

Doran, R. L., Guerin, R. O., & Cavalieri, J. (1974). An analysis of several instruments measuring "nature of science" objectives. Science Education, 58, 321-329.

Duschl, R. A. (1983). Science teachers' beliefs about the nature of science and the selection, implementation, and development of instructional tasks: A case study. Dissertation Abstracts International, 45, 422-A.

Duschl, R. A. (1988). Abandoning the scientistic legacy of science education. Science Education, 72, 51-62.

Duschl, R. A. & Wright, E. (1989). A case study of high school teachers' decision making models for planning and teaching science. Journal of Research in Science Teaching, 26, 467-501.

Elkana, Y. (1970). Science, philosophy of science and science teaching. Education, Philosophy and Theory, 2, 15-35.

Ennis, R. H. (1979). Research in philosophy of science and science education. In P. Asquith & H. Kyburg (Eds.), Current research in philosophy of science (pp. 138-170). East Lansing, MI: Philosophy of Science Association.

Ennis, R. H. (1984). Problems in testing informal logic, critical thinking, reasoning ability. Informal Logic, 6(1), 3-9.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), Teaching thinking skills (pp. 9-26). New York: Freeman.

Ennis, R. H. & Millman, J. (1985). Cornell Critical Thinking Tests Level X and Z. Pacific Grove, CA: Midwest.

Ennis, R. H. & Norris, S. P. (1989). Critical thinking evaluation: Status, issues, needs. In S. Legge & J. Algina (Eds.), Cognitive assessment of language and math outcomes (pp. 1-42). Norwood, NJ: Ablex.

Ennis, R. H. & Weir, E. (1985). The Ennis-Weir Critical Thinking Essay Test. Pacific Grove, CA: Midwest Publications.

Frederiksen, N. (1984). The real test bias. American Psychologist, 39, 193-202.

Freire, P. (1974). Pedagogy of the oppressed. New York: Seabury Press.

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). Manual for kit of reference tests for cognitive factors. Princton, NJ: Educational Testing Service.

Fry, E. (1968). A readability formula that saves time. Journal of Reading, 11, 513-516.

Gauld, C. (1982). The scientific attitude and science education: A critical appraisal. Science Education, 66, 109-121.

Gallagher, J. J. (1991). Prospective and practicing secondary school science teachers' knowledge and beliefs about the philosophy of science. Science Education, 75, 121-133.

Gettys, C. F. & Engelmann, P. D. (1983). Ability and expertise in act generation. Norman, OK: University of Oklahoma, Decision Processes Labratory.

Guilford, J. P. & Lacey, J. I. (1947). Printed classification tests. AAF Aviation Psychology Research Program Reports: Report 5. Washington, DC: G.P.O.

Glaser, R. (1972). Individuals and learning: The new aptitudes. Educational Researcher, 1, 5-12.

Goldman, S. R., Pellegrino, J. W., Parseghian, P. E. & Sallis, R. (1982). Development and individual differences in verbal analogical reasoning. Child Development, 53, 550-559.

Hanson, N. R. (1958). Patterns of discovery. Cambridge: Cambridge University Press.

Hardwig, J. (1985). Epistemic dependence. The Journal of Philosophy, 82, 335-349.

Heise, D. R. (1969). Problems in path analysis and causal inference. In E. F. Borgatta (Ed.), Sociological Methodology (pp. 38-73). San Francisco: Jossey-Bass Inc.

Hempel, C. G. (1966). Philosophy of natural science. Englewood Cliffs, NJ: Prentice-Hall.

Herron, J. D. (1977). Implicit curriculum - Where values are really taught. The Science Teacher, 44(3), 30-31.

Hodson, D. (1986a). The nature of scientific observation. The School Science Review, 68(242), 17-29.

Hodson, D. (1986b). Philosophy of science and science education. Journal of Philosophy of Education, 20, 215-224.

Hodson, D. (1988). Toward a philosophically more valid science curriculum. Science Education, 72, 19-40.

Kahl, S. & Harms, N. C. (1981). Project Synthesis: Purpose, organization and procedures. In N. C. Harms & R. E. Yager (Eds.), What research says to the science teacher: Volume 3 (pp. 5-11). Washington, DC: National Science Teachers Association.

Kim, J. (1975). Special topics in general linear models. In N. Nie, C. Hull, J. Jenkins, K. Steinbrenner, & D. Bent (Eds.), SPSS: Statistical package for the social sciences (pp. 383-397). New York: McGraw-Hill.

Kimball, M. E. (1967). Understanding the nature of science: A comparison of scientists and science teachers. Journal of Research in Science Teaching, 5, 110-120.

King, B. B. (1991). Beginning teachers' knowledge of and attitudes toward history and philosophy of science. Science Education, 75, 135-141.

Korth, W. W. (1968). The use of the history of science to promote student understanding of the social aspects of science. Unpublished doctoral dissertation, Stanford University.

Kuhn, T. (1962). The structure of scientific revolutions. Chicago: The University of Chicago Press.

Land, K. C. (1969). Principles of path analysis. In E. F. Borgatta (Ed.), Sociological Methodology (pp. 3-37). San Francisco: Jossey-Bass Inc.

Lederman, N. G. (1986a). Students' and teachers' understanding of the nature of science: A reassessment. School Science and Mathematics, 86, 91-99.

Lederman, N. G. (1986b). Relating teacher behaviour and classroom climate to changes in students' conceptions of the nature of science. Science Education, 70, 3-19.

Lederman, N. G. (1986c, April). Science teachers' conceptions of the nature of science: Do they really influence teaching behaviour? Paper presented at the annual meeting of National Association for Research in Science Teaching, San Francisco, CA.

Lederman, N. G. & Druger, M. (1985). Classroom factors related to changes in students' conceptions of the nature of science. Journal of Research in Science Teaching, 22, 649-662.

Lederman, N. G. & O'Malley, M. (1990). Students' perceptions of tentativeness in science: Development, use, and sources of change. Science Education, 74, 225-239.

Lederman, N. G. & Zeidler, D. L. (1986). Science teachers'conceptions of the nature of science: Do they really influence teaching behaviour? Science Education, 71, 721-734.

Lucas, A. M. (1975). Hidden assumptions in measures of "knowledge about science and scientists". Science Education, 59, 481-485.

Mackay, L. D. (1971). Development of understanding about the nature of science. Journal of Research in Science Teaching, 8, 57-66.

Martin, M. (1972). Concepts of science education. Glenview, IL: Scott Foresman & Co.

McNeil, L. M. (1986). Contradictions of control: School structure and school knowledge. New York: Routledge & Kegan Paul.

McPeck, J. E. (1981). Critical thinking and education. Oxford: Martin Robinson.

Mitman, A. L., Mergendoller, J. R., Marchman, V. A., & Packer, M. J. (1987). Instruction addressing the components of scientific literacy and its relation to student outcomes. American Educational Research Journal, 24, 611-633.

Mulholland, T. M., Pelligrino, J. W., & Glaser, R. (1980). Components of geometric analogy solution. Cognitive Psychology, 12, 252-284.

Munby, H. (1973). The provisions made for selected intellectual consequences by science teaching: Derivation and application of an analytical scheme. Unpublished doctoral dissertation, University of Toronto.

Munby, H. (1976). Some implications of language in science education. Science Education, 60, 115-124.

Munby, H. (1982). The place of teachers' beliefs in research on teacher thinking and decision making, and an alternative methodology. Instructional Science, 11, 201-225.

National Science Teachers Association (1982). Science-Technology-Society: Science education for the 1980's. Washington DC: National Science Teachers Association.

Newmann, F.M. (1988). Higher order thinking in the teaching of social studies: Connections between theory and practice. In Perkins, D., Segal, J., & Voss, J. (Eds.), Informal reasoning and education. Hillsdale, NJ: Erlbaum.

Nickerson, R. S. (1989). New directions in educational assessment. Educational Researcher, 18(9), 3-7.

Norris, S. P. (1984). Defining observational competence. Science Education, 68, 129-142.

Norris, S. P. (1985a). The philosophical basis of observation in science and science education. Journal of Research in Science Teaching, 22, 817-833.

Norris, S. P. (1985b). Synthesis of research on critical thinking. Educational Leadership, 42(8), 40-45.

Norris, S. P. (1986). Evaluating critical thinking ability. The History and Social Science Teacher, 21, 135-146.

Norris, S. P. (1988a). Controlling for background beliefs when developing multiple-choice critical thinking tests. Educational Measurement Issues and Practice, 7(3), 5-11.

Norris, S. P. (1988b). Research needed on critical thinking. Canadian Journal of Education, 13(1), 125-137.

Norris, S. P. (1989). Can we test validity for critical thinking? Educational Researcher, 18(9), 21-26.

Norris, S. P. (in press). Practical reasoning in the production of scientific knowledge. In R. Duschl & R. Hamilton (Eds.), Philosophy of science, cognitive psychology, and educational theory and practice. New York: SUNY Press.

Norris, S. P. & Ryan, J. (1987a). Essay Test of Inductive Reasoning Strategies: Part A. St. John's, Newfoundland: Memorial University of Newfoundland, Institute for Educational Research and Development.

Norris, S. P. & Ryan, J. (1987b). Essay Test of Inductive Reasoning Strategies: Manual. St. John's, Newfoundland: Memorial University of Newfoundland, Institute for Educational Research and Development.

Pelligrino, J. W. & Glaser, R. (1982). Analyzing aptitudes for learning. In R. Glaser (Ed.), Advances in instructional psychology. Vol. 2 (pp. 269-345). Hillsdale, NJ: Erlbaum.

Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. Journal of Educational Psychology, 77, 562-571.

Perkins, D. N., Allen, R., & Hafner, J. (1983). Difficulties in everyday reasoning. In W. Maxwell (Ed.), Thinking: The expanding frontier (pp. 177-189). Philadelphia, PA: Franklin Institute Press.

Perkins, D. N. & Simmons, R. (1988). Patterns of misunderstanding: An integrative model for science, math, and programming. Review of Educational Research, 58, 303-326.

Perry, W. G. (1968). Patterns of development in thought and values of students in a liberal arts college: A validation of a scheme. Cambridge, MA: Harvard University, Bureau of Study Counsel. (ERIC Document Reproduction Service No. ED 024315)

Phillips, L. M. (1989). Developing and validating assessments of inference ability in reading comprehension. (Tech. Rep. No. 452). Champaign, IL: University of Illinois at Urbana-Champaign, Center for the Study of Reading.

Popper, K. R. (1972). Objective knowledge. Oxford: Oxford University Press.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accomodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66, 211- 227.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). The shopping mall high school: Winners and losers in the educational marketplace. Boston: Houghton Mifflin.

Quellmalz, E. S. (1985). Needed: Better methods for testing higher-order thinking skills. Educational Leadership, 43(2), 29-35.

Robinson, J. T. (1969). Philosophy of science: Implications for teacher education. Journal of Research in Science Teaching, 6, 99-104.

Ropo, E. (1987). Skills for learning: A review of studies on inductive reasoning. Scandinavian Journal of Educational Research, 31(1), 31-39.

Rubba, P. A. (1976). Nature of Scientific Knowledge Scale. Bloomington, IN: Indiana University, School of Education.

Rubba, P. A. & Anderson, H. O., (1978). Development of an instrument to assess secondary school students' understanding of the nature of scientific knowledge. Science Education, 62, 449-458.

Rubba, P. A., Horner, J. K., & Smith, J. M. (1981). A study of two misconceptions about the nature of science among junior high school students. School Science and Mathematics, 81, 221-226.

Schoenfeld, A. H. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. Cognitive Science, 7, 329-363.

Schoenfeld, A. H. (1985). Mathematical problem solving. New York: Academic Press.

Schommer, M. (1989). Students' beliefs about the nature of knowledge: What are they and how do they affect comprehension? (Tech. Rep. No. 484). Champaign, IL: University of Illinois at Urbana-Champaign, Center for the study of reading.

Schwab, J. J. (1960). Enquiry, the science teacher, and the educator. The Science Teacher, 27, 6-11.

Schwirian, P. (1968). On measuring attitudes toward science. Science Education, 52, 172-179.

Science Council of Canada (1984). Report 36: Science for every student. Hull, Quebec: Canadian Government Publishing Center.

Scientific Literacy Research Center (1967). Wisconsin Inventory of Science Processes. Madison, WI: The Regents of the University of Wisconsin.

Shipman, V. (1983). New Jersey test of reasoning skills. Upper Montclair, NJ: Institute for the Advancement for Philosophy for Children.

Showalter, V. M. (1974). What is unified science education? Program objectives and scientific literacy. Prism II, 2, 3-4.

Skyrms, B. (1975). Choice and chance: An introduction to inductive logic. Encino, CA: Dickenson Publishing Company.

Spiro, R. J., Vispoel, W. L., Schmitz, J., Samarapungavan, A., & Boerger, A. (1987). Knowledge acquisition for application: Cognitive flexibility and transfer in complex content domains. In B. C. Britton (Ed.), Executive control processes (pp. 177-200). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Erlbaum.

Sternberg, R. J. & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. Child Development, 51, 27-38.

Sternberg, R. J. & Rifkin, B. (1979). The development of analogical reasoning process. Journal of Experimental Child Psychology, 27, 195-232.

Stice, J. (1987). Learning how to think: Being earnest is important, but it's not enough. In J. Stice (Ed.), Developing critical thinking and problem-solving abilities: New directions for teaching and learning 30 (pp. 93-100). San Francisco: Jossey-Bass.

Tobin, K. (1987). What happens in high school science classrooms? Journal of Curriculum Studies, 19, 549-560.

Walsh, D. & Paul, R. W. (1987). The goal of critical thinking: From educational ideal to educational reality. Washington, DC: American Federation of Teachers.

Wasserman, S. (1984). Promoting thinking in your classroom II: Inconsistencies between means and ends. Childhood Education, 60, 229-233.

Watson, G. & Glaser, E. (1980). Watson-Glaser Critical Thinking Appraisal, Forms A and B. Cleveland, OH: The Psychological Corporation.

Welch, W. W. (1966). The development of an instrument for inventorying knowledge of the processes of science. Unpublished doctoral dissertation, University of Wisconsin.

Welch, W. W. (1969). Welch Science Process Inventory: Form D. Minneapolis, MN: Author.

Welch, W. W. (1981). Inquiry in school science. In N. C. Harms and R. R. Yager (Eds.), What research says to the science teacher: Volume 3 (53-72). Washington, DC: National Science Teachers Association.

Welch, W. W. & Pella, M. O. (1968). The development of an instrument for inventorying knowledge of the processes of science. Journal of Research in Science Teaching, 5, 64-68.

Whimbey, A. (1985). You don't need a special "reasoning" test to implement and evaluate reasoning training. Educational Leadership, 43(2), 37-39.

Whitely, S. E. (1980). Modelling aptitude test validity from cognitive components. Journal of Educational Psychology, 72, 750-769.

Willis, P. E. (1977). Learning to labour. Lexington, MA: D. C. Heath.

Wright, S. (1921). Correlation and causation. Journal of Agricultural Research, 20, 557-585.

Wright, S. (1934). The method of path coefficients. Annals of Mathematical Statistics, 5(September), 161-215.

Wright, S. (1954). The interpretation of multivariate systems. In O. Kempthorne (Ed.), Statistics and Mathematics in Biology (pp. 11-33). Ames: Iowa State College Press.

Wright, S. (1960a). Path coefficients and path regressions: Alternative or complimentary concepts? Biometrics, 16(June), 189-202.

Wright, S. (1960b). The treatment of reciprocal interaction, with or without lag, in path analysis. Biometrics, 16(September), 423-445.

Zeidler, D. L. & Lederman, N. G. (1989). The effect of teachers' language on students' conceptions of the nature of science. Journal of Research in Science Teaching, 26, 771-783.

Ziman, J. (1980). Teaching and learning about science and society. Cambridge: Cambridge University Press.

APPENDICES

APPENDIX A

Final Form of the Epistemological Questionnaire

SCIENTIFIC KNOWLEDGE

QUESTIONNAIRE

Identification (name): ____________________________

School: ____

Grade level: _____________ Male: _____ Female: _____

James A. Duffett
Department of Curriculum and Instruction
Memorial University
of Newfoundland
1990

STUDENT QUESTIONNAIRE

In this questionnaire, you are asked to read statements about scientific knowledge, and to tell how much you agree or disagree with each.

This is not a test, and there are no right or wrong answers. Nobody will be told what answers you pick. The purpose is to find out how students in general think.

Simply indicate how you actually feel about each statement.

Example

Read the statement below. On the right hand side of the page, there are five responses from which to choose.

The responses are :

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

Circle the one that best expresses how you feel about the statement.

	STATEMENT	HOW YOU ACTUALLY FEEL				
1.	Scientists always follow the scientific method.	SA	(A)	U	D	SD

Some items may appear to be the same as others. Be sure to answer every item.

Key

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

	STATEMENTS	HOW YOU ACTUALLY FEEL				
1.	Scientists do not make errors in their conclusions if they follow the scientific method.	SA	A	U	D	SD
2.	When there is some evidence against a scientific theory, scientists may still accept the theory.	SA	A	U	D	SD
3.	After scientists think they have found the solution to a problem, they feel the problem has been solved once and for all.	SA	A	U	D	SD
4.	Different scientists observing the same thing will draw the same conclusions.	SA	A	U	D	SD
5.	A scientific theory is similar to a work of art in that they both express creativity.	SA	A	U	D	SD
6.	Scientific knowledge need not be capable of being tested by experiments.	SA	A	U	D	SD
7.	The solution to a complicated scientific problem is not often obvious.	SA	A	U	D	SD

Key

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

	STATEMENTS	HOW YOU ACTUALLY FEEL				
8.	If a person tries long and hard to understand a scientific problem, they will most likely end up being confused.	SA	A	U	D	SD
9.	By reporting their findings, scientists influence the beliefs of other scientists.	SA	A	U	D	SD
10.	When evidence is discovered indicating that a theory is incorrect, the theory is abandoned.	SA	A	U	D	SD
11.	Scientists rarely produce hypotheses on the basis of hunches or guesses.	SA	A	U	D	SD
12.	Scientific statements are descriptions of the world as it really is.	SA	A	U	D	SD
13.	We accept statements as scientific knowledge even though they contain error.	SA	A	U	D	SD
14.	The truth of scientific knowledge should not be questioned.	SA	A	U	D	SD
15.	Scientists often change their opinion in light of new evidence.	SA	A	U	D	SD

Key

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

	STATEMENTS	HOW YOU ACTUALLY FEEL				
16.	When a scientist produces a set of experimental results, it is not important that other scientists are able to reproduce those same results.	SA	A	U	D	SD
17.	Scientists rely primarily on the published results of other scientists for their conclusions.	SA	A	U	D	SD
18.	Scientists make progress by forming hypotheses and testing them.	SA	A	U	D	SD
19.	The longer you work at a scientific problem the more likely you are to find a solution.	SA	A	U	D	SD
20.	An idea is scientific if and only if it is the result of a systematic process of logical thought.	SA	A	U	D	SD
21.	A scientific theory gives the final answers to scientific questions.	SA	A	U	D	SD
22.	Scientific observations are not influenced by scientists' feelings.	SA	A	U	D	SD

Key

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

	STATEMENTS	HOW YOU ACTUALLY FEEL				
23.	A scientific theory need not correctly predict future events.	SA	A	U	D	SD
24.	Statements about which scientists agree are true.	SA	A	U	D	SD
25.	There is no purpose to questioning accepted scientific theory.	SA	A	U	D	SD
26.	Scientists seek problems where the answer can be easily and quickly discovered.	SA	A	U	D	SD
27.	Observation in science is influenced by personal opinion.	SA	A	U	D	SD
28.	An essential ability of the scientist is the ability to ask the right questions.	SA	A	U	D	SD
29.	If you are ever going to be able to understand a scientific statement, it will make sense to you the first time you hear it.	SA	A	U	D	SD
30.	Seeking expert opinion is not a good strategy for solving scientific problems.	SA	A	U	D	SD

Key

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

	STATEMENTS	HOW YOU ACTUALLY FEEL				
31.	Once accepted, scientific knowledge is no longer subject to change.	SA	A	U	D	SD
32.	The statements of science represent the best approximations of the natural world available at the present time.	SA	A	U	D	SD
33.	Scientific knowledge is a product of human imagination.	SA	A	U	D	SD
34.	Test results that can be repeated by other scientists are required for the acceptance of scientific knowledge.	SA	A	U	D	SD
35.	A statement is true when it corresponds to established scientific knowledge.	SA	A	U	D	SD
36.	Rules used when solving scientific problems were originally formulated by prominent scientists and can be applied without question.	SA	A	U	D	SD
37.	A theory may be modified when new evidence is discovered.	SA	A	U	D	SD
38.	A scientific statement is accepted when it accurately represents natural phenomena.	SA	A	U	D	SD

Key

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

STATEMENTS	HOW YOU ACTUALLY FEEL				
39. A true scientific statement will not contradict established scientific knowledge.	SA	A	U	D	SD
40. When trying to solve a difficult scientific problem, it may be more reasonable to seek expert opinion than to seek a solution without help.	SA	A	U	D	SD
41. A scientist will reject a scientific theory if she or he has doubts about it.	SA	A	U	D	SD
42. Factual information does not change.	SA	A	U	D	SD
43. Reliable observations are the means by which scientific laws, theories and concepts are tested.	SA	A	U	D	SD
44. Advice from scientific experts should be questioned.	SA	A	U	D	SD
45. Scientists consider scientific theories to be changeable.	SA	A	U	D	SD
46. Scientific theories are discovered, not created by scientists.	SA	A	U	D	SD
47. A scientific statement is true when it describes the world the way it really is.	SA	A	U	D	SD

Key

Strongly agree	SA
Agree	A
Undecided	U
Disagree	D
Strongly disagree	SD

	STATEMENTS	HOW YOU ACTUALLY FEEL				
48.	If a person can't understand a scientific statement within a short amount of time, they should keep trying.	SA	A	U	D	SD
49.	Original thinking is necessary for scientific work.	SA	A	U	D	SD
50.	Working hard on a difficult science problem for an extended period of time rarely pays off.	SA	A	U	D	SD
51.	True scientific statements accurately describe relationships in nature.	SA	A	U	D	SD
52.	Today's scientific laws, theories and concepts may have to be changed in the face of new evidence.	SA	A	U	D	SD
53.	Knowing when to seek an expert opinion is important when trying to solve a scientific problem.	SA	A	U	D	SD
54.	The solutions to scientific problems are developed gradually over time.	SA	A	U	D	SD
55.	Factual information is subject to change.	SA	A	U	D	SD
56.	A scientific statement is true if most scientists believe it.	SA	A	U	D	SD

APPENDIX B

Essay Test of

Inductive Reasoning Strategies: Part A

Essay test of

INDUCTIVE

REASONING

STRATEGIES

PART A

Are There Living Creatures on Zed?

This test asks you to write down what you are thinking and what you plan to do as you work on a problem.

You and two scientists must search the planet Zed for living creatures. You have only four days to explore before returning to Earth. Some of the things that happen each day are described.

Directions

Read about the things that happen on each day, and think about what they mean for your search for living creatures.

Then, keeping in mind your search for living creatures:

(1) write what you are thinking about the things that happened on that day and previous days; and

(2) write what you plan to do on your search because those things happened.

STEPHEN P. NORRIS and JAMES RYAN
Institute for Educational Research and Development
Memorial University of Newfoundland

Day 1 (An example)

Your spaceship lands where you think your search might be successful. With life support suits on, you and the scientists start to walk away from the ship. You hear some high-pitched sounds which last for about twenty or thirty seconds. Then, your radio buzzes a warning. You return to the spaceship to check the instruments. You find a tape recorder is on. When the scientists return to the ship, neither remembers having turned on the recorder.

Here is an example of what you might write:

It might be important to find out more about the high pitched sounds. They might be made by some living creatures, but many other things could have made them. I will listen for them again, and see whether there is any pattern or whether they come from a particular place.

I wonder why my radio buzzed a warning? Did it buzz because the recorder was on or was there some problem in the instruments? I will check out the instruments to see whether everything is working properly.

I shall also check for other things. Did some creature enter the ship? I will check to see whether anything is disturbed. Why was the recorder on? I will ask the scientists to try to remember whether they were using it before they went out. I will also check the recorder for malfunctions. It is also possible that some creature came in and turned on the recorder. I will set up something like a camera that will turn on if anything enters on one of the next days. This way, we will be able to see what it is.

CONTINUE IN THE SAME MANNER FOR DAYS 2, 3, AND 4
WHEN YOU HAVE READ THE EXAMPLE

YOU MAY GO BACK TO READ

Day 2

Today, you set out for a nearby hill. When you reach the top you see a valley below. There is a river flowing through the valley with a variety of vegetation and rock formations on either bank. You take some photographs of the landscape and begin to follow the river downstream. The water here is very clear and you can easily see to the bottom. In the water there are several cone-shaped, brownish objects which look like broken pieces of large eggshells. You pick one from the water. It is thin and you are able to crack it easily like an eggshell. You take it back to the ship.

DO NOT GO AHEAD UNTIL FINISHED DAY 2

YOU MAY GO BACK TO READ BUT NOT TO WRITE

Day 3

First thing today the chemist analyzes the object you took from the water yesterday. She finds that it is made up of a combination of elements similar to those found in eggshells on earth.

You then return to the river and notice that the water is muddy today. Just behind you is a cave. The chemist crawls in and returns a few minutes later. She says she heard many noises she did not recognize, and saw little trenches that went for a few meters and then disappeared beneath the walls. They were like those made by a rodent or some small animal.

You have the rest of the day to explore the area before heading back to the ship.

DO NOT GO AHEAD UNTIL FINISHED DAY 3

YOU MAY GO BACK TO READ BUT NOT TO WRITE

Day 4

This morning the chemist does another analysis on loose material collected from the floor of the cave yesterday. The analysis shows that the material could contain animal waste. The biologist examines a bag of material taken from the small trenches. He thought there was animal hair or fur in the bag but discovers that the material is a hairlike leaf from a plant which grows in the area.

Since you must leave for Earth today you decide to explore the area near the ship. Before long you begin to hear high-pitched sounds like the ones you heard the day you arrived. There seems to be a pattern to the sounds, and they are coming from a number of places. You command the computer to record the sounds and play them back slowly. It plays a series of beeps and spaces, repeated over and over. You command the computer to broadcast the sounds into the atmosphere. You then hear another series of sounds which the computer says originated on Zed.

You only have a few hours left to explore before returning to Earth.

Unfortunately, there is no time left to explore, even though there is still a lot to learn. You must return to Earth.

School: ______________________________

Name: ______________________________

Grade: ______________________________

Sex: ______________________________

ID number: ______________________________