DOCUMENT RESUME

ED 303 366 SE 050 356

AUTHOR	Carey, Susan; And Others
TITLE	"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.
INSTITUTION	Educational Technology Center, Cambridge, MA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	ETC-TR-88-19
PUB DATE	Jun 88
CONTRACT	OERI-400-83-0041
NOTE	37p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Cognitive Processes; Cognitive Structures; *Experiential Learning; *Inquiry; *Junior High Schools; Learning Processes; *Learning Strategies; Middle Schools; Process Education; Science Activities; *Secondary School Science
IDENTIFIERS	*Constructivism

ABSTRACT

Developing a constructivist view of scientific inquiry and knowledge is considered to be important to the training of future scientists, as well as to the understanding of scientific information by all citizens. The research reported targets the junior high school grades. Curricular materials that introduce seventh graders to the constructivist view of science have been developed and implemented. Students' initial epistemological stance concerning scientific knowledge is that knowledge is a passively acquired, faithful copy of the world, and that the inquiry process is limited solely to observing nature, rather than constructing explanations of phenomena in nature. The assumption guiding this curricular intervention is that if students are to gain a better understanding of the nature of scientific inquiry and knowledge, they must be actively involved in constructing and evaluating explanations for natural phenomena, and they must be engaged in metaconceptual reflection on that process. (CW)

TR88-19

"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

June 1988

Educational Technology Center
Harvard Graduate School of Education
337 Gutman Library Appian Way Cambridge MA 02138
(617) 495-9373

"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

Draft Article

June 1988

Prepared by:

Susan Carey
Massachusetts Institute of Technology
and
Educational Technology Center

Risa Evans
Maya Honda
Eileen Jay
Chris Unger
Educational Technology Center

Nature of Science Project

Group Members

Diane Bemis
Susan Carey
Risa Evans
Mary Greaves
Maya Honda
Eileen Jay
Wayne O'Neil
Sandra Smith
Chris Unger
Al Weinstein

Preparation of this report was supported in part by the Office of Educational Research and Improvement (Contract # OERI 400-83-0041). Opinions expressed herein are not necessarily shared by OERI and do not represent Office policy. (To be submitted to Cognition and Instruction.)

Acknowledgments

We want to thank the following people for their participation in the Nature of Science Group and for their invaluable contributions to this research: Diane Bemis, Carol Chomsky, Mary Greaves, Wayne O'Neil, Sandra Smith and Al Weinstein. We particularly want to thank Sandra Smith and her students for participating in this study. We also gratefully acknowledge the cooperation of the Watertown Public Schools with this work. Finally, we want to thank Mary Maxwell Katz, Martha Stone Wiske and Judah Schwartz for their comments on an earlier draft of this paper.

INTRODUCTION

The primary aims of science education are to help students understand the nature of the scientific enterprise itself, as well as to help them understand theories and concepts of biology, physics, and chemistry. Most secondary school science curricula fail to impart an understanding of the constructed and fluid nature of scientific knowledge and of the inquiry process that supports this knowledge. Developing a constructivist view of scientific inquiry and knowledge is considered important to the training of future scientists, as well as to the understanding of scientific information by all citizens (National Science Board Commission on Precollege Education in Mathematics, Science and Technology, 1983).

The research reported here targets the junior high school grades. We have developed and implemented curricular materials that introduce seventh graders to the constructivist view of science. As we will demonstrate, students' initial epistemological stance concerning scientific knowledge is that knowledge is a passively acquired, faithful copy of the world, and that the inquiry process is limited solely to observing nature, rather than constructing explanations (i.e., theories) of phenomena in nature. The assumption guiding our curricular intervention is that if students are to gain a better understanding of the nature of scientific inquiry and knowledge, they must be actively involved in constructing and evaluating explanations for natural phenomena, and they must be engaged in metaconceptual reflection on that process. Thus, we have designed an instructional unit that engages students in scientific inquiry -- in the tentative groping toward a deeper understanding of the world, in the cumulative and intellectual process of theory construction.

THE STANDARD CURRICULAR APPROACH TO TEACHING ABOUT SCIENTIFIC INQUIRY

Much of current educational practice grows out of curriculum reform efforts that have emphasized the teaching of the "process skills" involved in the construction of scientific knowledge -- such diverse skills as observation, classification, measurement, conducting controlled experiments, and constructing data tables and graphs of experimental results. These skills are typically covered in the junior high school science curriculum, beginning with the introduction of "the scientific method" in the seventh grade. The standard curricular unit on the scientific method, for example, contains many exercises to teach students about the design of controlled experiments, such as identifying independent and dependent variables in experiments, and identifying poorly designed experiments in which variables have been confounded. Although students then go on to design and conduct controlled experiments, often the possible hypotheses and variables (and thus, experimental outcomes) for a given problem are prescribed by the curriculum.

Certainly, process skills are important elements of a careful scientific methodology. Junior high school students do not spontaneously measure and control variables or systematically record data when they attempt experimental work. Yet, the standard curriculum fails to address the motivation or justification for using these skills in constructing scientific knowledge. Students are not challenged to utilize these process skills in exploring, developing and evaluating their own ideas about natural phenomena. Rather, instruction in the skills and methods of science is conceived of outside the context of genuine inquiry. Thus, there is no context for addressing the nature and purpose of scientific inquiry, or the nature of scientific knowledge.

MOTIVATION FOR THE STANDARD CURRICULAR APPROACH

The standard curricular approach to scientific methodology, with its emphasis on the teaching of process skills, is geared towards remedying the dramatic deficiencies preadolescent children have in designing controlled experiments, and in drawing conclusions from experimental evidence. These deficiencies are amply documented in Inhelder's and Piaget's (1958) classic work exploring scientific reasoning in young children and adolescents. Inhelder and Piaget argue that before the ages of 13 to 15 years, children are not able to entertain or evaluate hypotheses because the logic of confirmation is not available to them. Their studies confounded domain-specific knowledge of particular science concepts with scientific reasoning more generally. For example, suppose the child's task is to determine experimentally the effects of weight, size, and density on whether an object will float. If the child does not completely differentiate these concepts, then the child will certainly be unable to separate them from one another in forming and testing hypotheses (see Carey, 1985). However, the differences between children and adults go beyond differences between domain-specific understanding of the concepts being reasoned about. Recent studies of the development of scientific reasoning reveal striking developmental differences even when knowledge of the domain is equated.

Kuhn, Amsel, and O'Loughlin (1988) asked children in the third, sixth, and ninth grades, and adults to evaluate and generate evidence about the effects of features of sports balls (e.g., color, size, texture) on the quality of a player's serve. Subjects first articulated their own views (e.g., large balls would be better than small ones; the color of the ball would not make a difference). They were then asked to state whether the results of various experiments supported their view (called "theory" by Kuhn et al.), disconfirmed it, or provided insufficient information. Subjects of all ages, even adults, found the task difficult, but there were marked developmental differences. To give one example, preadolescent subjects were unable to generate possible data that would disconfirm their theory. Kuhn et al. argue that their subjects' faulty reasoning revealed a lack of differentiation, at a metaconceptual level, of the notions of *theory* and *evidence*. Before adolescence, children have no concept of evidence as independent of the theory bearing on it; pieces of evidence are considered only as instances illustrating the theory.

In earlier work, Kuhn and Phelps (1982) studied 10- and 11-year-olds attempting to identify the ingredients critical to producing a chemical reaction (i.e., a color change) when mixed together. The children's "experimentation" was unsystematic, and the conclusions drawn were often invalid. Kuhn and Phelps comment that subjects commonly behaved as if their goal was not to find the cause of the color change, but rather to produce the color change. Just as children do not distinguish theory from evidence, they do not seem to distinguish between understanding a phenomenon and producing the phenomenon.

Thus, beginning with the work of Inhelder and Piaget, studies show that preadolescent children lack certain fundamental metaconceptual notions, such as *theory*, *evidence*, and *experiment*, or, at the very least, they interpret these notions differently from adults. It is important to see that metaconceptual understanding of such notions is distinct from, but related to, process skills. Children may be unsystematic in evaluating and generating evidence for two reasons: they do not have the target process skills, or they fail to understand the task of examining the fit between data and theory because of their epistemological stance towards the nature of knowledge. The above studies document children's conceptions of theories and evidence indirectly; that is, children's metaconceptual understanding of scientific inquiry is inferred from their attempts to engage in it. In the present study, we adopt a complementary approach. We employ a clinical interview to probe seventh graders' understanding of the nature and purpose of scientific inquiry.

Even though Kuhn et al. discuss children's differentiation of theories from evidence, their tasks, like those of other researchers probing the development of formal operational thought (e.g., Inhelder and Piaget, 1958), do not engage subjects in true theory building. The "theories" offered by Kuhn et al.'s subjects are actually hypotheses about causal relations. The process skills explored by these studies and by Inhelder and Piaget, concern causal inference from covariation data. While such skills are an important component of scientific inquiry skills, their mastery constitutes only a small part of a constructivist appreciation of the nature of scientific knowledge. It would be quite possible for a student to master the logic of experimental argument, or to differentiate hypothesis from evidence, without having any understanding of the intellectual construction of theory-like explanations of natural phenomena. If on the other hand, students had a sense in which theories are tentative intellectual constructions that provide explanatory frameworks for understanding nature, it is very possible they would be in a better position to see the differences between theories and evidence.

OUR APPROACH

The research reported here rests on several working assumptions:

(1) Process skills will be more easily and better learned if they are embedded in a wider context of metaconceptual points about the nature of scientific knowledge.

(2) Such metaconceptual knowledge is important in its own right.

(3) Metaconceptual understanding of the nature of scientific inquiry and scientific knowledge can be gained only by actively constructing such knowledge and reflecting on this process.

(4) As in any case of science education, curricular materials must be aimed at the students' beginning conceptions.

These assumptions motivate a curricular approach which emphasizes theory building and reflection on the theory building process. We have developed an instructional unit to replace the typical junior high school unit on the scientific method. While our Nature of Science Unit has been used in seventh grade science classes, it would be appropriate for the sixth through the ninth grades. For a detailed discussion of the development and initial field testing of the Nature of Science Unit and the clinical interview, see Carey et al. (Note 1).

THE STUDY

The two goals of our study are (1) to probe junior high school students' initial understanding of the nature and purpose of scientific inquiry, and (2) to explore whether it is feasible to move students beyond their initial conceptions with a relatively short classroom-based intervention. Our report has three parts: first, to portray the students' initial understanding, we characterize their responses to our clinical pre-interview; second, we describe our curricular intervention; and third, we discuss the results of the clinical post-interview in order to assess the intervention's effectiveness.

DESIGN OF THE STUDY

The study was conducted in a K-8 public school in a middle income, ethnically-mixed suburb of Boston, Massachusetts. Seventy-six students in five mixed-ability seventh grade science classes participated in the study. All classes were taught our three-week-long Nature of Science Unit by their regular teacher. Each of the lessons was observed by one or two research assistants.

Twenty-seven of the students were randomly selected to be interviewed both prior to and after participating in the Unit. The individual clinical interviews were administered by research assistants. All interviews were tape-recorded for later coding.

RATIONALE FOR DEVELOPING THE CLINICAL INTERVIEW

There are a number of instruments which assess some aspects of students' metaconceptual understanding of science and/or the scientific method. Of these, two address students' understanding of the nature of scientific inquiry and knowledge. The Test of Understanding Science for junior high school students (TOUS; Klopfer and Carrier, 1970) is a multiple-choice, written test which assesses students' understanding of science as a human endeavor and as a social institution. Its rather ambitious scope may be a consequence of the fact that it was originally designed for high school students (see Cooley and Klopfer, 1961). About the endeavor of science, TOUS probes whether students understand that the purpose of scientific inquiry is to develop explanations about phenomena in nature. TOUS characterizes scientific knowledge as creative, cumulative, and testable; that is, such knowledge is the product of the scientist's mind, is built on previous work, and is always subject to revision. In addition, TOUS probes knowledge of selected science vocabulary (e.g., hypothesis, theory, law). The Nature of Scientific Knowledge Scale (NSKS; Rubba and Andersen, 1978) is a scaled-response, written measure designed for secondary school students.

Students judge 48 statements about scientific knowledge along a five-point Likert scale (i.e., from "strongly agree" to "strongly disagree"). The NSKS represents scientific knowledge as amoral, creative, developmental, parsimonious, testable, and unified.

Both TOUS and NSKS offer an excellent analysis of the nature of scientific inquiry and knowledge, and thus are very clear about possible student endpoints. However, such tests have a clear limitation: multiple-choice and scaled response assessment measures necessarily place constraints on what can be revealed of students' own initial conceptions. Further, it is not possible to know what students understand about the terminology used on the test. Thus, we turned towards developing an interview which would allow students to give their own answers, and which would allow us to probe the meanings of critical terms and ideas.[1]

THE CLINICAL INTERVIEW

Our half-hour clinical interview probes students' understanding of the following: (1) the nature and purpose of science; (2) the main elements of scientific work, including ideas, experiments, and results/data; and (3) the relation among these elements. In addition, follow-up questions probe what students mean when they use key words or phrases, such as "discover," "try out [an idea or invention]," "proof," "explanation." Appendix 1 is a copy of the clinical interview

THE CODING PROCEDURE

Responses to the questions on the interview were coded into categories that reflected three general levels of understanding, which are described below. When students answered

---

[1] We also developed a 24-item, multiple choice, written pre/posttest which included some items from TOUS and other standardized instruments. Our written test attempted to evaluate students' understanding of the nature of scientific inquiry and knowledge, and of experimental design. Detailed results from our written test are not presented here because we found that as a research tool the test was a less sensitive measure than our clinical interview.

However, we note that the analysis of the written test scores from the 59 students who completed both the pretest and the posttest showed a small but significant pre to post gain in average scores on a 1-tailed paired comparison t-test ($p < .0005$, $df = 58$). The pretest mean of 68.5 percent correct increased to a posttest mean of 74.4 percent correct. (See Carey et al., Note 1.)

The written pre/posttest scores of the 27 students whose clinical interview data are described in this study are not significantly different from the scores of students who were not interviewed. Thus, we consider the students who were interviewed to be representative of the whole group of students.

"I don't know" to a question, the response was not scored. This contrasts with responses which were scored Level 0; Level 0 responses reveal actual misconceptions or misunderstandings (e.g.,"an experiment is when you try something new"; or, "a hypothesis is a kind of snake"). Students received a score for each question; questions with more than one part were counted as a single item (e.g., see Appendix 1: Ideas 4a, 4b and 4c).

The interviews were coded on the basis of listening to the interview tapes. Each interview was coded by at least two research assistants, who were unaware of whether it was a pre-interview or post-interview. Interscorer reliability was modest (74 percent agreement); disagreements virtually always involved only one level, such as one scorer judging a response to be Level 1, while another judged it to be Level 2. Scoring disagreements were resolved by discussion.

On both the pre-interview and post-interview, an overall score for each student was calculated as the mean of all interview items. In addition, questions were grouped into six sections: (1) *Nature/Purpose of Science and Scientific Ideas*; (2) *Nature of a Hypothesis*; (3) *Nature/Purpose of an Experiment*; (4) *Guiding Ideas and Questions*; (5) *Results and Evaluation*, and (6) *Relationships* (between particular elements of scientific work, e.g., ideas and experiments). For each section, every student received a mean section score. In addition, the highest score a student attained in each section was noted.

Because the point of the interview was to elicit students' conceptions and misconceptions about the nature of scientific inquiry, rather than to test their vocabulary or make an inventory of the things they had no ideas about, questions to which the students answered "I don't know" and thus received no score were not included in the calculations of any of the means. While this may have slightly raised the overall means on the pre-interview, it did not have an appreciable effect on the scores of the individual sections, since nearly all such responses appeared in the pre-interview *Nature of a Hypothesis* section.

THE CODING SCHEME - GENERAL LEVELS[2]

The students' ideas about the nature of science ranged from a notion that doing science means discovering facts and making inventions, to an understanding that doing science means constructing explanations for natural phenomena. Student responses were coded along two

[2] The coding scheme was developed from the clinical interviews of students who did not have the full battery of written tests and clinical interviews (n=12). Data from these students were not included in the final analysis reported here.

main dimensions. The first is the degree to which ideas, experiments, and results are defined and differentiated from one another. The second is the degree to which the relationships among these elements are articulated and understood.

In General Level 1, the students make no clear distinction between ideas and activities, especially experiments. A scientist "tries it to see if it works." The nature of "it" remains unspecified or ambiguous; "it" could be an idea, a thing, an invention, or an experiment. The motivation for an activity is the achievement of the activity itself, rather than the construction of ideas. Scientists discover facts, invent things, or find out what happens when they do "X," and the goal of science is equivalent to the activity itself.

In General Level 2, students make a clear distinction between ideas and experiments. The motivation for activity is verification or exploration of an idea, and more specifically, the purpose of doing an experiment is to test an idea to see if it is right. There is an understanding that an idea or phenomenon may be operationalized and explored. The goal of science is to find out how things work.

In General Level 3, as in Level 2, students make a clear distinction between ideas and experiments, and understand the motivation for activity as verification or exploration. Added to this is an appreciation of the relation between the results of an experiment (especially unexpected ones) and the idea being tested. An idea is evaluated in terms of the results of a test, and may be changed or developed in accordance with these results. Thus, Level 3 understanding recognizes the cyclic, cumulative nature of science, and identifies the goal of science as the construction of ever-deeper explanations of the natural world.

These general levels provide the framework for the coding levels within each section of the clinical interview.

PRE-INSTRUCTION RESULTS: STUDENTS' INITIAL UNDERSTANDING

In the following discussion, we present the pre-instruction results by section, characterizing the questions and coding levels for each of the six sections of the clinical interview, as well as describing typical student answers. We describe Level 0 misconceptions for those sections where such answers were frequently given. The overall mean scores and mean high scores for each section are shown in Table 1. The number of students scoring equal to or above specific levels overall and in each section is shown in Table 2. At present, please consider only the pre-interview scores given in these tables.

The *Nature/Purpose of Science and Scientific Ideas* section included questions about the goals of science and the kinds of ideas scientists have (Appendix 1: Introductory Questions 1, 2, 3; Ideas 2, 3).

Level 1 answers in this section focus on the activities themselves. The point of science is to find out or discover facts and answers, or to invent cures or contraptions, and scientists have ideas about how to carry out these activities. In typical student answers, the goal of science is "to discover new things," "to find new cures for diseases," "to discover all the facts there are about the world." Scientists have ideas about how to discover or invent things, about "finding out about the stars and nature," about "teaching us and doing experiments."

Level 2 answers focus on the development of a mechanistic understanding of the world. The point of science is to find out or discover how things work, where things come from, how things might behave in the future. Scientists have ideas, questions and predictions about how things work. Typical student answers often include Level 1-type statements, but in addition refer to more specific goals such as "to find out how something works," or "to find out how animals get oxygen." Scientists have ideas or questions about how things work or "how we got here," and predictions about "what will happen in a certain experiment."

Level 3 answers focus on the development of an explanatory understanding of the world. The point of science is to construct explanations for why things happen; scientists have ideas and questions about why things happen. In typical student answers, the goal of science is to find out "why something happens," "why the leaves change color," "why the dinosaurs became extinct," or "why the yeast bubbled in the tube."

The average pre-interview score on this section was 1.09 (Table 1). Scores ranged from .33 to 1.67. The median score was 1.0; the modal scores were 1.0 (n=8) and 1.33 (n=8). On the pre-interview then, students saw the purpose of science as discovering facts, making inventions, and developing cures.

The *Nature of a Hypothesis* section included a single question about what a hypothesis is (Appendix 1: Hypothesis 1).

In Level 1 answers to this question, a hypothesis is an idea or guess. Typical answers were vague; a hypothesis is "an idea about something," or an "educated guess."

In Level 2 answers, a hypothesis is also an idea or guess, but it is explicitly related to an experiment or phenomenon. The hypothesis is something that can be tested, or it is a prediction about an experiment or phenomenon. In typical Level 2 answers, a hypothesis is "an idea about something, but you can test it," or "an if-then statement about what you think might happen," or "an idea of what will happen in an experiment."

In Level 3 answers, the hypothesis is not only related to an experiment, but aids in interpreting the results of the experiment, or is evaluated in terms of the results.

Only two students knew the word "hypothesis" on the pre-interview; each gave a Level 1 answer. The remaining students received no score.

The *Nature/Purpose of an Experiment* section included questions about what experiments are and why scientists do them (Appendix 1: Experiment 1, 2a, 2b).

In Level 0 answers, an experiment is described as a disembodied process. It is an activity that is not guided by an idea, a question, or an implicit assumption. Typical Level 0 answers include statements like, "an experiment is when you mix two chemicals together and see what happens," or "when you try something new."

In Level 1 answers, there is no clear distinction between experiments and ideas. The motivation for doing an experiment, implied or explicitly, is to find something out about the thing being experimented with. In typical Level 1 answers, a scientist "tries something" to see if it "works" or "reacts," or to find something out. In an experiment, a scientist "takes samples to find out more about it." Scientists do experiments "to check it out . . . find out about the thing they're experimenting on . . . . [Checking it out is] taking a good look at it . . . When you experiment something, you do something to it, like if you're gonna experiment of putting steel into water, you're putting steel into water. That's the experiment -- finding out."

In Level 2 answers, the distinction between the idea and activity is clear. The experiment is a test of a scientist's idea, or an operationalized exploration of an phenomenon. In typical Level 2 answers, scientists do experiments "to test to see if their idea is right" or "to test their hypothesis." An experiment is "trying to figure out your hypothesis . . . like yeast, you wanna figure out if it's alive. You cook it, then you put it in a [flask] and then you put the rubber stopper on and then if it gasses out it's probably alive and if it doesn't do anything, it isn't." (Note that although the student's interpretation of the possible results is backwards, it is clear that he or she understands that the experiment is motivated by a question.)

In Level 3 answers, there is a distinction between the idea and the activity, and the experiment plays the same role as it does in Level 2. In addition, however, the relationship between the results of an experiment (especially unexpected ones) and the idea being tested is clearly articulated. Results aid in the evaluation or development of an idea, and ideas may change as a result of the work a scientist does. In a Level 3 answer, scientists do experiments "to test what they were thinking of . . . if it was the same thing they guessed it tells them they were right, and if it was different, it tells them they were wrong, and they need to change their idea."

On the pre-interview, the mean score was 1.0 (Table 1), with a range from 0 to 3. The median was 1.0, as was the mode (n=20). On the pre-interview, then, students saw experiments as activities that support the goal of science -- finding things out and discovering facts.

The *Guiding Ideas and Questions* section included questions about how scientists do their work, where they get their hypotheses, and how they decide which experiments to do (Appendix 1: Introductory Questions 4; Hypothesis 2; Experiment 3).

In Level 0 answers, there is no sense that the scientist is seeking information or has any other guiding purpose for the activities that he or she pursues, and no sense of the relationship between what the scientist does or thinks and anything other than the scientist's own whims and desires. A scientist does his/her work by "reading," "doing experiments," "doing research." A scientist does a certain experiment because he or she "feels like it."

In Level 1 answers, the focus is on activities such as thinking, observing, or exploring, and the goal of these activities is to gather information. In typical Level 1 answers, scientists do their work by "looking it up in books," "putting things under microscopes to see how they behave," and asking other scientists. They get a hypothesis "from their mind."

In Level 2 answers, exploration is guided by a particular idea, question, object, or phenomenon. In a typical student answer, a scientist "walks through a forest and finds something new and tries to find out more about it." Another scientist "goes and looks for things. For example, if he goes to the moon, he finds a rock, he brings it to earth and experiments. [I mean] he puts -- he has to know if it's alive or dead, if it needs energy."

In Level 3 answers, either the guided exploration of Level 2 is elaborated to include operationalization or reflection on prior knowledge and experience, or there is an understanding of evaluation and development of ideas (as discussed in Section 3). In a typical Level 3 answer, a scientist "probably thinks up an idea, and then he builds an experiment out of the idea, and if he's right or wrong he keeps building up more questions to see, to find out even more stuff than he knows." Scientist's choice of experiment is based on "an idea . . . and the stuff he already knows."

On the pre-interview, the mean score was .65, with a range from 0 to 1.5. The median and modal scores (n=14) were .5. Thus, on the pre-interview, students revealed the misconception that a scientist's choice of hypotheses and experiments is mostly capricious.

The *Results and Evaluation* included questions about when and why scientists change their ideas, and what scientists do when they get unexpected results when they are testing their ideas (Appendix 1: Ideas 6; Results 1).

In Level 1 answers, the scientist is trying to get a result; if the experiment doesn't come out "right," it is because something is not working and should be checked or changed. The "something" is not clearly specified as an idea. In typical Level 1 answers, the scientist "checks it to see what he did wrong and tries to fix the problem," or changes the idea/experiment a little by "adding stuff or taking stuff away."

In Level 2 answers, the scientist is testing an idea; if the results of the experiment are unexpected, then, as in Level 1, something requires attention. In Level 2, however, the idea and experiment are clearly distinguished. In a typical Level 2 answer, "the person might think something went wrong in the way they did the experiment, so they go back to fix it," or, "they would change their idea."

In Level 3 answers, there is an understanding that an idea is modified because of a conflict between the idea and experimental results or other evidence, and the modified idea takes these data into account. In a post-interview Level 3 answer, "he'd probably test it one more time and see if he came up with that again, and if he did, he'd probably have to change his hypothesis a little to fit in with the new data." The crucial part of this answer is the notion that the modified hypothesis must "fit" the experimental data.

The mean for the pre-interview was 1.06 (Table 1), with a range from 0 to 3. The median score was 1.0, as was the mode (n=8). Nine students scored less than 1, thirteen scored between 1 and 2, and five scored 2 or better (Table 2). While there was one student who demonstrated the understanding that a conflict between results and ideas may lead to revisions in the ideas, the majority gave Level 1 responses in which even hypotheses and experiments are not clearly differentiated.

The *Relationships* section included questions about the relation between a scientist's ideas and the rest of the work (such as observing and testing or experimenting) that a scientist does (Appendix 1: Ideas 4a, 4b, 4c, 5; Hypothesis 3).

The levels in this section are similar to those in the experiment section. We coded the two sections separately because they address the same issue (the relationship between ideas and experiments) from opposite points of view.

In Level 0 answers, there is no relationship between a scientist's ideas and the rest of the work he or she does. Students at this level typically answer that there is no relationship, or else give a very incomplete rendering of it; scientists "report their ideas," or "write them down." It was sometimes difficult for us to gauge whether a student really had no understanding of the relationship, or whether he or she did not fully understand the questions.

In Level 1 answers, there is no clear distinction between ideas and experiment. A scientist tries an idea to see if it works, "does it," or uses it as a guide or a blueprint. In typical Level 1 answers, scientists "make their ideas work," "see if they're accurate, if they can really do them," or "fulfill them by experimenting on them."

In Level 2 answers, there is a clear distinction between the idea and the experiment. The idea is tested to see if it's right, or the idea is used to predict the outcome of an experiment. In a typical Level 2 answer, "they test them [ideas] . . . in experiments, and see if they're right."

A basic Level 3 response goes beyond the Level 2 notion that ideas are tested in experiments to include the understanding that they are evaluated or developed in accordance with the results of these tests. A full Level 3 response would include, in addition to this, an understanding that the modification of an idea may entail reorganizing and re-interpreting the data on which the idea was originally based. None of the students gave such a full response, and only a few students were given Level 3 credit at all for their basic responses.

The mean score on the pre-interview was .91 (Table 1); the range from 0 to 3. The median and modal (n=16) scores were 1.0. Only three students scored 2 or better (Table 2). Thus, the overwhelming majority of the students saw ideas, at best, as blueprints for action, or interchangeable with the things they are about.

Discussion: Seventh Graders' Initial Understanding

The overall mean on the pre-interview was 1.0. Only four students had overall mean scores over 1.5. Of the 27 students, only 11 had scores of 2 or more scattered through the six sections of the interview. According to our instrument, then, the seventh graders in the study may be characterized as having a Level 1 understanding of the nature of science and scientific inquiry prior to instruction.

Perhaps the critical feature of Level 1 is the absence of an appreciation that ideas are distinct, constructed and manipulable entities. There is no understanding that a scientist's ideas motivate the scientist's other, perhaps more tangible work, such as gathering data and doing experiments, or that the ideas, in turn, are affected by this work. Instead, ideas are confused with experiments, or with whatever else they are about (an invention, cure, and so on), and there is no acknowledgement of the theoretical motivations behind scientists' experiments and other activities.

More generally, in Level 1 understanding, nature is there for the knowing. Accordingly, scientists "discover" facts and answers that exist, almost as objects, "out there." In typical Level 1 fashion, there is no understanding that "facts" and "answers" are actually constructed ideas *about* objects. Other goals of science include inventing new things and finding cures for diseases. Here, too, ideas are equated with things, or else with simple plans of action (e.g., "they have an idea for a rocketship, so they do it"). Scientists work towards their goals by observing things and looking for discoveries, or by trying things out to see if they work. Scientists' ideas themselves, however, are never the object of scrutiny.

DESCRIPTION OF THE UNIT[3]

Overall Goal

One goal of our Unit is to teach students about the *mechanics* of the scientific method -- that is, that experimentation usually follows a path of defining the problem, gathering information, making a hypothesis, identifying dependent, independent, and control variables, and then performing the experiment. Beyond this, we want students to begin to appreciate the *motivation* for these activities in the service of building scientific ideas. We hope to help students see that ideas are mutable mental constructs which offer the best explanation of a natural phenomenon given the information available at a particular time.

[3] Lesson plans are available on request from the Nature of Science Project, Educational Technology Center, Harvard Graduate School of Education, 337 Gutman Library, Appian Way, Cambridge MA 02138.

Ideas are constructed through observation, information gathering, and frequent reflection upon our own knowledge and experience. In this context, experimentation is a way of systematically evaluating ideas. Whether or not the results of an experiment support or disconfirm an idea, they contribute to the development of that idea, and may well lead to new questions and new ideas about the phenomenon under investigation. In this way, it is always possible to build ever deeper explanations for a phenomenon, as more questions are raised in the process of inquiring into the original problem. Indeed, through this process our initial conception of a particular phenomenon may be significantly changed.

In order to understand such points, students must be engaged in constructive theory building, and they must be helped to reflect on the process. To this end, the heart of our three-week-long Unit is a two-week series of lessons on the nature of yeast. This follows a week of introductory lessons which orient students. We have also developed materials on linguistic theory which were not used in the present study (see Carey et al., Note 1).

Introductory Lesson

In the introductory lesson, students begin to reflect upon their own inquiry process -- to think about how they come to understand something and to think about where their ideas come from. In order to ground this reflective discussion, students observe and speculate about a "martian object" (i.e., a small, unfamiliar object) which the teacher displays before the class. The students come up with questions about the nature of the object, eventually focusing on the question of whether or not it is alive. Students then brainstorm a list of attributes of "aliveness," and discuss ways they could test to see which (if any) of these attributes the object possesses. Through this discussion, the teacher helps the students recognize that their ideas about living organisms come from their own experience, reside in their own minds, and that these ideas can be made explicit for inspection and evaluation.

Embedded in the lesson are the points that:

(1) The basis for scientific inquiry is mental work; and,

(2) Experiments are tests of ideas.

The Videodisc Lessons: Seeing the Unseen

In 1985-86, Harvard University's Educational Technology Center Videodisc Project produced Seeing the Unseen, a four-segment interactive videodisc designed to teach middle school students about the nature of science. Storey and Mellin (Note 2) explored the videodisc's use by individual students, pairs of students, and whole classes, and concluded that its effectiveness would be enhanced if the points it contained about the process of inquiry were explicated by the teacher and discussed by the class as a whole. Two segments of the videodisc engage students in pieces of theory building. We incorporated their use in the next two introductory sections of our Unit. Subsequently, these lessons have been redesigned so that use of the videodisc is optional.

Animal Mimicry Lessons

In the animal mimicry lessons, which are designed to support the videodisc segment called "What Disguises Do Animals Use?" students use their initial ideas about the basic needs of animals to construct their own categories of different kinds of animal disguises. On the first day, after viewing a video clip showing animals which disguise themselves in various ways, students are asked to generate ideas about the needs of animals and to begin applying what they know toward understanding the functions of different kinds of disguise. With these ideas in mind, they then view six other animal disguise clips with the task of devising categories. On the second day, they discuss the categories they devise and check them against additional information. At the end of these exercises, the teacher points out that the process they have gone through is similar to the way a scientist might go about categorizing the same data.

Points embedded in these lessons include:

(1) Ideas, including categories or systems of classification, help us to make sense of the world; and,

(2) Categories are mental constructs that are developed by observing and thinking.

The Black Box Problem

In the lessons for the Black Box Problem, students try to figure out the regular, geometric shape of a three-dimensional object in a closed box. First, the class brainstorms about methods which may help them to figure out what the shape of the object is, and asks the teacher to perform those functions on the box. Students then view the interactive videodisc segment called "How do Scientists Study Things They Can't See?" which shows Linus Pauling verbalizing his thought process while he works on the black box problem. The teacher stops the videodisc at several points to discuss how Pauling systematically approaches the problem by considering one feature of the object's shape at a time, such as whether the object is cylindrical. In this case, for example, he hypothesizes that if the object is cylindrical, it will roll smoothly in the box. After testing the hypothesis, he reflects upon the meaning of the result and any previous results for that hypothesis, and then, proceeds to consider and test other aspects of the object.

On the second day, groups of two or three students are given sealed boxes with an object inside and are asked to work on the problem themselves. This involves students in the ongoing process of formulating, testing, and revising their hypotheses about the shape of the object. At the end of the lesson, each group must present their final hypothesis to the rest of the class, and defend it on the basis of their tests and results. Since the students are not allowed to actually see the shape of the object in the box, they cannot determine which group's hypothesis is "right." Rather, they must decide which hypothesis offers the best account of the evidence. The teacher draws the analogy that scientists use systematic experimentation in order to develop and test ideas about phenomena which may never definitively be proven true.

Embedded in these lessons are the following points:

(1) Scientific inquiry can proceed even when the object under investigation cannot be seen or touched;

(2) A useful way of solving a problem is to break it down into smaller, more investigable parts;

(3) We make a hypothesis before doing an experiment;

(4) An experiment is a test of a hypothesis, the results of which will either support or disconfirm the hypothesis; and,

(5) We may never know the "right" answer.

Yeast Lessons

The two-week series of yeast lessons engage students in constructing an ever-deepening theoretical understanding of a natural phenomenon -- in this case, the phenomenon of bread dough rising. The students make and test hypotheses, perform experiments, reflect upon what they are doing, and reflect on *why* they are doing what they are doing.

The exploration begins by observing and discussing the difference between a piece of bread and a piece of unrisen bread dough. Eventually, the question "What makes bread rise?" is raised. Brainstorming usually leads to a list of the ingredients in bread, so the teacher brings the phenomenon "into the laboratory" by making a mixture of yeast, flour, sugar, salt, and warm water in a flask with a corked top. The students observe this mixture bubbling up in the flask (in fact, the cork soon flies off), and correctly infer that a gas is produced by the mixture. They see that this provides a tentative answer to the original question of what makes bread rise. One reason the answer is satisfactory is that they can even understand properties of bread that they did not set out to explain -- for example, the texture of bread reflects gas bubbles.

Although a satisfactory answer to the original question has been obtained, it leads the students to ask yet another question: "Why do yeast, flour, sugar, salt, and warm water produce a gas?" Discussion leads the class to realize that they do not even know which of these ingredients are necessary. Students are charged to carry out their own experiments to determine which ingredients are required for the mixture to bubble.

Their first efforts reflect their Level 1 understanding of the nature and purpose of experiments (see above). Their experiments are unsystematic. They do not measure how much of each ingredient they are using, nor do they even keep a record of which ingredients they use. Consistent with their Level 1 understanding, their view of the task is limited to producing the bubbling phenomenon. In fact, some simply try to duplicate exactly the conditions of the teacher's demonstration. Most proceed more haphazardly. Experimentation is, after all, trying things out. They do not accept the task of trying to find out what ingredients are necessary; indeed, they do not even seem to understand this goal.

When the teacher challenges the class to draw conclusion from their experiments, none can be supported. The stage is set for standard lessons about the scientific method. The class then constructs a controlled series of experiments, which, when the data are pooled, reveal that yeast, sugar and water are necessary for the mixture to bubble. The question then becomes which other variables may have an effect (e.g., amount of ingredients, temperature of the water), and which of those are most worth exploring. With this question, the Unit moves to considerations that go beyond the standard curricular lessons on how to collect reliable data. The task now is to reflect on how we know what data are worth collecting.

The teacher points out to the students that the aim of their experimentation is to try to understand *why* these ingredients produce a gas. Students brainstorm about what they know about water, sugar, yeast, and gases. This leads them to consider two mechanisms for why the yeast mixture produces a gas: (1) the bubbles are caused by some kind of chemical reaction between the yeast, water and sugar, on analogy to what happens when baking soda is put in vinegar, and (2) yeast is alive; the yeast eats the sugar and the gas is a product of metabolism. Students almost universally prefer the first hypothesis; most often some help from the teacher is required for the second to even emerge from the discussion.

The class finds itself at a juncture where systematic experimentation may help in constructing a deeper explanation for how and why bread rises. Students are challenged (and helped, in a classroom brainstorming exercise) to design controlled experiments that will help decide between the two possible mechanisms. To do so, they must draw on what they know of living things and of chemical reactions, and they are shown that the results of their experiments will challenge their understanding of both types of entities.

Several tests that might support or disconfirm one or the other mechanism are designed by the students and performed by the teacher in front of the class. For example, after first considering the fact that people produce CO_2 as a product of metabolism, they turn to the problem of identifying what kind of gas is produced by the yeast mixture. The students hypothesize that if the yeast is alive, perhaps it too gives off CO_2. A bromthymol blue experiment demonstrates that, indeed, the gas given off by the yeast mixture is CO_2. In discussing the conclusion that may be drawn, students appreciate that this outcome is consistent with the hypothesis that yeast is alive, but that it certainly does not prove it since CO_2 is the product of chemical reactions as well. Another experiment explores the effects of extreme heat and cold on yeast (i.e., boiling or baking would be expected to kill a living organism; it is unknown what effect that would have on a chemical reaction, although students' intuitions are that it would either intensify it or leave it unchanged). Students discover that if the yeast is baked before being mixed with the sugar and water, the mixture does not produce a gas. They also see this result as consistent with the hypothesis that yeast is alive.

Other experiments, including gedanken experiments, are performed, and gradually students come to accept the mechanism they did not originally favor. In the course of this exercise, their very notion of a living organism is challenged; it must be expanded to include what looks like an inert brown powder, which can survive being frozen, remaining dormant until conditions support activity and growth.

The metaconceptual points about the nature of science embodied in these lessons are explicitly drawn by the teacher. These include:

(1) One of the goals of science is to arrive at explanatory understandings of natural phenomena;

(2) Scientists construct explanatory understandings through the interplay of their ideas, the experiments they choose to test those ideas, and new ideas that arise from the results of the experiments;

(3) The answer to one question often leads to other deeper questions; and,

(4) In exploring a phenomenon, scientists may be forced to challenge some of their basic notions about the way things are.

Wrap-up Lesson

The wrap-up lesson concludes the Unit with a general discussion about the interplay of thought and experimentation in science. Students review not only the specific *mechanics* of scientific inquiry -- defining a problem, making a hypothesis (or several competing hypotheses), testing the hypothesis, evaluating the results, and perhaps, revising the hypothesis given the accumulated evidence -- but also the *motivation* for engaging in this process, making specific references to previous lessons. Moreover, the teacher reminds the students of the ever-deepening and constructed nature of their ideas, and the role of their present knowledge in coming to understand, inquire into, and build explanations about the natural world.

The general points covered in this lesson include:

(1) The point of exploring problems is not just to solve the problem, but rather to develop a deeper understanding of things in nature;

(2) Scientific inquiry involves an interplay of thinking and experimenting;

(3) Questions change and deepen during the course of scientific inquiry; and

(4) Results of investigations often challenge our basic notions about the way things are.

OBSERVATIONS AND TEACHER /STUDENT REACTIONS TO THE NATURE OF SCIENCE UNIT

In general, both the teacher and the students enjoyed the lessons and felt they were educationally approriate. Table 3 shows student reactions to this effect; these data were gathered from all the students who participated in the study (n= 76). Below, we report the observations and reactions of the teacher, students, and observers, for all of the lessons.

Introductory Lesson

Throughout the class, the students were very curious and involved in speculating about what the martian object might be: a rock, an "orbitee," a burnt brownie, and so on. They reflected on the fact that their ideas about what it means to be "alive," like scientists' ideas, came from their own experience and resided in their own minds. When the teacher asked how they could find out whether the object possessed any of the attributes of life that they had listed, some students responded, "Do an experiment!" Although they seemed to understand the logic of testing ideas, they came up with no examples of well-designed experiments, and were interested to find out how they could test their ideas.

Videodisc Lessons

Animal Mimicry Lessons

Whereas the main point of the introductory lesson was made quite clear to the students, the main points of the animal mimicry lessons, unfortunately, were not as clearly articulated. The students were engaged by the videodisc, but did not grasp the metaconceptual points being made. Some of the students did not explicitly understand the concept of "category," and a lack of class time precluded thorough student discussion. When asked where the categories came from some students answered, "from books and computers," even after the teacher had pointed out that the categories came from the students own minds and the observations they made. In addition, students did not see beyond the categorizing task to the usefulness of the categories in building an explanation for animal disguise.

The Black Box Problem

As was predicted, when students were asked for suggestions as to how to go about finding the shape of the object inside the enclosed box, they suggested randomly shaking it or moving it in one direction or another. After this exercise, students watched the video-clips of Pauling methodically testing for aspects of the object's shape. They enjoyed guessing which hypothesized shapes were supported and not supported by the results of Pauling's tests.

On the second day, however, the students initially had difficulty understanding why they had to proceed with the task one step at a time. In answer to student questions, the teacher usually replied, "If you just shook the box up, how would you know what was inside of it? But if you test one aspect of it at a time, you can find something out about it which will help you know more about your problem." All students seemed to accept this explanation.

Yeast Lessons

The majority of the students enjoyed the yeast lessons the most. Throughout the lessons, students showed a great deal of exuberance for doing experiments. The content of the lessons provoked student curiosity and provided the teacher a number of opportunities to draw on students' pre-instructional knowledge about bread and the function of yeast. An important feature of the lesson plans was the time allocated for student questions and detailed discussion. Students were highly engaged, animated, and often freely offered their ideas.

The lessons provided experiences that could lead students beyond their initial Level 1 understanding of the nature of science. The first experiences were necessarily "hands-on," giving the students an opportunity to practice their naive notions of experimentation, and then to reflect on the inadequacies of their methods. We found, however, that subsequent experiments performed by the teacher as demonstrations were as effective in communicating points about the process of inquiry, because the logic of the experiments was made clear ahead of time through careful discussion of the motivation for the experiments and the meanings of the possible results. In the class discussion that followed the bromthymol blue experiment, for example, the students were able to articulate the logic behind the experiment and to interpret the results of the experiment without having manipulated any of the materials themselves.

Observations of the yeast wrap-up lesson suggested that the lessons successfully communicated the points we wanted to make about the nature of science. In this lesson, students were reminded of how the questions they had been asking of nature changed through the process of constructing, testing and reflecting on their ideas. The students were challenged to think of a number of plausible explanations for the fact that bread dough stops rising at some point. Typically, three explanations were contrasted, one consistent with the chemical reaction mechanism (i.e., the yeast or the sugar is depleted, so there is no more reaction), and two consistent with the living thing mechanism (i.e., the yeast eats all the sugar; and/or, the heat of the oven kills the yeast). Students were able to think of possible experiments to decide among these explanations. They were also able to think of other properties of bread dough that require explanation (e.g., why frozen dough does not rise; and, why we put dough in a warm place to rise) and to offer explanations.

POST-INSTRUCTION RESULTS: DID STUDENTS MOVE BEYOND THEIR INITIAL CONCEPTIONS?

The overall mean score increased from 1.0 to 1.55 on the post-interview. This increase was statistically significant ($p<.001$, Wilcoxon signed ranks test, 1-tailed). Every child improved, and improvement averaged half a level. Now 16 students achieved overall scores of 1.5 or better (Table 2, $X^2=7.84$, $p<.01$), and five scored 2 or better on the post-interview (Fisher Exact Test, $p<.03$), a score nobody achieved on the pre-interview. There were 41 scores of 2 or better on the various sections of the interview, whereas there were only 11 such scores prior to instruction.

In the *Nature/Purpose of Science and Scientific Ideas* section there was a slight, significant increase in the mean scores on the post-interview (Table 1). As in the pre-interview, the average score hovered slightly above 1 (1.28), and the median and modal scores (n=11) were again 1.0. This means that many of the students retained their belief that the goal of science is to discover facts and answers, and to invent things. The range, however, extended to include higher scores -- from .33 to 3.0. Whereas no students scored 2 or higher on the pre-interview, four scored 2 or higher on the post-interview (Table 2, Fisher Exact Test, $p<.06$). Of these four, two students saw the goal of science as the development of a mechanistic understanding of the world (Level 2), and two saw it as the development of an explanatory understanding of the world (Level 3).

In the *Nature of a Hypothesis* section, the only Level 3 answer given was on the post-interview. It is quoted here along with follow-up questions and answers.

> **Student:** [A hypothesis is] when you're trying to figure it out . . . like the guy [Linus Pauling] who was tipping the box on the screen . . . he thought it was a round thing, a cylinder. He was tipping the box and rolling it and he said it was a cylinder.
>
> **Interviewer:** Does he have just one hypothesis?
>
> **Student:** No, at the beginning he said it could be a cylinder with even sides or uneven sides, but when he tipped it, it made the same noise, the sides, so he knew it was a cylinder with even ends . . . his conclusion was that it was a cylinder with equal sides.
>
> **Interviewer:** Why did he have the hypothesis?
>
> **Student:** So you would know if what you would expect, but if you get the wrong answer, than you expected, you'd do it again, and if you got out the same answer, you'd know it was that, not the thing you thought it was gonna be.

All students answered on the post-interview, where the mean score was 1.37 (Table 1), with a range from 0 to 3. The median score on the post-interview was 1; the modal score was 2 (n=12). By the post-interview, then, almost all of the students understood that a hypothesis is an idea about something, and nearly half the students were able to relate hypotheses to experiments or tests (Level 2).

In the *Nature/Purpose of an Experiment* section the considerable increase in the mean score to 1.52 on the post-interview was significant (Table 1). Mean scores ranged from 0 to 3, and the median and modal (n=10) scores were now 2.0. By the post-interview, then, over half of the students saw experiments as tests of ideas, and some could articulate how unexpected results lead to revisions of ideas. This stands in marked contrast to the pre-interview, where only 3 students attained a score of 2 or higher (Table 2, X^2=8.90, $p<.005$).

In the *Guiding Ideas and Questions* section the improvement on the post-interview was dramatic and significant (Table 1). The mean rose to 1.45, with a range from .33 to 2.5. The median score was now 1.33, and the modal score 1 (n=8). Fifteen students scored 1.5 or better, as opposed to only 1 on the pre-interview (Table 2, X^2=17.42, $p<.001$). While none of the students scored 2 or better on the pre-interview, seven did so on the post-interview (Table 2, Yates' X^2 =5.91, $p<.02$). Following the Unit, these students understood that the activities of science are guided by particular ideas and questions.

A similarly clear picture emerges when one considers the mean high scores, which improved significantly from near 1 on the pre-interview to over 2 on the post-interview (Table 1).

In the *Results and Evaluation* section, the improvement in the mean to 1.8 on the post-interview was also dramatic (Table 1). Scores for the post-interview ranged from .5 to 3. The median was 1.5, as was the mode (n=8). While ten students scored 1.5 or better on the pre-interview, nineteen did so on the post-interview (Table 2, X^2=6.04, $p<.02$). In contrast with the two students who scored 2.5 or better on the pre-interview, eight students scored 2.5 or better on the post-interview (Table 2, Fisher Exact Test, $p<.04$), demonstrating a clear understanding of the distinction between idea and experiment, and in some cases, of the relationship between idea and results.

In the *Relationships* section, the improvement in the mean score to 1.69 on the post-interview was dramatic and significant (Table 1). The median score was now 1.5, and the modal score 2.5 (n=8). Fourteen students scored 2 or better on the post-interview (Table 2, X^2=22.74, $p<.001$). Again, mean high scores increased from around 1 to over 2, which is also significant (Table 1). These students made clear distinctions between ideas and experiments, and understood that experiments are tests of ideas (Level 2); in some cases, students also understood that an idea is evaluated according to the results of a test (Level 3).

Discussion: Suggested Influence of the Nature of Science Unit

The greatest score increases occurred in the sections on *Guiding Ideas and Questions, Results and Evaluation*, and *Relationships*. Viewed in terms of the Nature of Science Unit, these results make sense. While the Unit did incorporate lessons that focused specifically on hypotheses and experiments, its emphasis was on the relation between these and other elements (e.g., results/data), and the highest scoring sections of the post-interview al. made reference to these relationships.

The main points of the unit's introductory lesson were that the basis for scientific inquiry is mental work, and that experiments are tests of ideas. These points were put into practice in the black box lessons and the yeast lessons, where the students conducted experiments or tests only *after* they had explicitly formulated a hypothesis about a certain problem or phenomenon. The *Guiding Ideas and Questions* section of the interview, in which the mean score increased from .65 to 1.45, emphasized this relationship between ideas and experiments.

In addition, in both the black box problem and the yeast lessons, students practiced evaluating their ideas by looking at the experimental data that they had generated. The yeast lessons made the additional point that new ideas and questions may arise from the results of an experiment. The *Results and Evaluation* section of the interview, in which the mean score increased from slightly above Level 1 almost to Level 2, emphasized the relationship between an idea and the results of an experiment.

The *Relationships* section of the interview emphasized the role of ideas in motivating and guiding experiments, and the role of data and results in evaluating ideas. Students were essentially asked to integrate the points that had been made throughout the Unit. The mean score increased from .91 to 1.69.

In conclusion, all of the students interviewed moved beyond their initial Level 1 understanding of scientific knowledge. Many showed solid appreciation of a Level 2 differentiation between ideas and experiments, seeing the point of experimentation as testing ideas. Some showed a tentative grasp of aspects of a Level 3 understanding of the cumulative revision of ideas in the light of unexpected results.

GENERAL DISCUSSION OF LEVELS

Consistency of Scores Across Sections

In order to address the question of how consistent students were in their scores across the six sections, we noted the highest score a subject received on each section, and analyzed how tightly clustered these scores were. We found the clustering to be moderate; see Table 4.

In the pre-interview, over half of the students received the same score for all six sections (in every case, this was Level 1), with at most one section receiving the adjacent score (Level 0 or Level 2). Even on the post-interview, half of the subjects located their high scores on one main level, or on two adjacent levels. On the post-interview, these almost always included Level 2 or Level 3 responses. Consistency was greater on the pre-interview than on the post-interview because of the preponderance of Level 0 and Level 1 responses on the pre-interview, and because the curricular intervention affected the students' understanding of the material in some sections more than in others.

Are the Levels "Stages?"

Are the levels stages, as in Kohlberg's stages of moral understanding? No claims are here being made that the interview places students into stages. Before this possibility could even be systematically explored, more careful articulation of the differences between levels (especially between Level 2 and Level 3) would be needed. The scoring procedures would require refinement, so that interscorer reliability could be improved over our current 74 percent level. So, while the present effort indicates that it might be possible to identify stages in the understanding of the nature of science and scientific inquiry , we recognize it is only a first step towards doing so.

CONCLUSION

The results from our clinical interview support the suggestion in the literature that preadolescent children have a different epistemological stance towards scientific knowledge than do scientifically literate adults. Initially, most of the seventh grade students in our study thought that scientists seek to discover facts about nature by making observations and trying things out. This Level 1 understanding of the nature of science might be called a "copy theory " of knowledge: knowledge is a faithful copy of the world that is imparted to the knower when the knower encounters the world. By this view then, the only way scientists can be wrong about some aspect of nature is through ignorance that is, by not having looked at that aspect of nature.

This Level 1 epistemology provides a context for interpreting the literature on children's dramatic failures both at designing experiments to discover causal mechanisms and at interpreting experimental data (Inhelder and Piaget, 1958; Kuhn and Phelps, 1982; Kuhn et al., 1988). As these authors suggest, one source of these failures is children's lack of metaconceptual understanding of the distinction between theory and evidence, and, between the goal of understanding a phenomenon and the goal of producing a phenomenon. In a Level 1 view, knowledge directly reflects reality, so the problem of examining the fit between the two does not arise.

By engaging students in reflecting upon the relationship between ideas and the activities of science, our Unit aims to help them begin to differentiate ideas and the evidence that supports those ideas. Although seventh grade students initially fail to make this distinction, our post-interview results indicate that it is indeed possible to move them beyond their initial understanding. After our Unit, many students clearly understood that inquiry is guided by particular ideas and questions, and that experiments are tests of ideas. These Level 2 notions indicate their improved differentiation of ideas and experiments. Some students gave Level 3 responses to some questions in the post-interview, demonstrating an understanding that ideas may be changed based on experimental results. It is an open question as to what effects such advances in metaconceptual understanding might have on the kinds of process skills probed by Inhelder and Piaget, and Kuhn and her colleagues.

While our three-week Unit is designed to replace the standard unit on the scientific method, we believe that our approach has implications for the structure of the entire science curriculum. In order to reinforce the gains in understanding that students are able to make in a unit such as ours, and to push their understanding further, we believe it is necessary that the rest of the science curriculum reflect a constructivist epistemology. It is vital that the entire curriculum provide opportunities for students to reflect on the process of constructing scientific knowledge as they learn about the theories and concepts of science. In our Unit, students reflected on the problem under investigation and examined the motivation for each step of the process of inquiry. Students should be engaged in this kind of thinking throughout the curriculum. Rather than presenting theories and concepts as static objects, the curriculum should impart an understanding of their development: the questions that provoke them, the data that support them, and the alternatives that challenge them.

REFERENCES

Carey, S. (1985). Are children fundamentally different kinds of thinkers and learners than adults? In S. Chipman, J. Segal, and R. Glaser (Eds.), *Thinking and learning skills*. Vol. 2. Hillsdale, NJ: Erlbaum.

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

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

Klopfer, L. E., & Carrier, E. O. (1970). *TOUS: Test on Understanding Science (Form Jw)*. Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. Orlando, FL: Academic Press.

Kuhn, D., & Phelps, E. (1982). The development of problem-solving strategies. In H. Reese (Ed.), *Advances in child development and behavior*. Vl. 17. New York: Academic Press.

National Science Board Commission on Precollege Education in Mathematics, Science and Technology. (1983). *Educating Americans for the 21st century: A plan of action for improving mathematics, science and technology education for all American elementary and secondary students so that their achievement is the best in the world by 1995*. Washington, D. C.: National Science Foundation.

Rubba, P., & Andersen, H. (1978). Development of an instrument to assess secondary students' understanding of the nature of scientific knowledge. *Science Education, 62*(4), 449-458.

REFERENCE NOTES

1. Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1988). What junior high school students do, can, and should know about the nature of science and scientific inquiry. Cambridge, MA: Educational Technology Center, Harvard Graduate School of Education, March, 1988.

2. Storey, K., & Mellin, C. *A prototype science interactive videodisc: Research on in-school use* (Tech. Rep. 87-4). Cambridge, MA: Educational Technology Center, Harvard Graduate School of Education, April, 1987.

	Mean				Mean High			
	Pre	Post	Diff.	Sig.	Pre	Post	Diff.	Sig.
Nature/Purpose of Science/ Scientific Ideas	1.09	1.28	+.19	p<.03	1.52	1.74	+.22	NS
Nature of a Hypothesis	1.0	1.37	+.37	NA	1.0	1.37	+.37	NA
Nature of an Experiment	1.0	1.52	+.52	p<.002	1.0	1.52	+.52	p<.002
Guiding Ideas & Questions	.65	1.45	+.80	p<.001	1.19	2.22	+1.03	p<.001
Results & Evaluation	1.06	1.8	+.74	p<.001	1.41	2.19	+.1.78	p<.001
Relationships	.91	1.69	+.78	p<.001	.93	2.3	+1.37	p<.001
Overall Score	1.0	1.55	+.55	p<.001	NA	NA	NA	NA

Table 1. Mean Scores and Mean High Scores by Section for the Clinical Interviews (n= 27)

Nature/Purpose of Science

nPre	Level	nPost
22	≥1	22
4	≥1.5	10
0	≥2	4
0	≥2.5	2

Nature of a Hypothesis

nPre	Level	nPost
0	=0	4
2	≥1	23
0	≥2	13
0	=3	1

Nature of an Experiment

nPre	Level	nPost
4	=0	4
23	≥1	23
3	≥2	14
1	=3	4

Guiding Ideas and Questions

nPre	Level	nPost
11	≥1	25
1	≥1.5	15
0	≥2	7
0	≥2.5	2

Results and Evaluation

nPre	Level	nPost
18	≥1	25
10	≥1.5	19
5	≥2	11
2	≥2.5	8

Relationships

nPre	Level	nPost
20	≥1	27
4	≥1.5	21
3	≥2	14
1	≥2.5	8

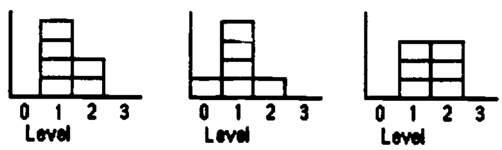
Overall

nPre	Level	nPost
13	≥1	23
4	≥1.5	16
0	≥2	5
0	≥2.5	0

Table 2. Number of Students Whose Mean Score is Equal to or Above a Coding Level, by Section and Overall (n=27)

	Very Good	Good	Fair	Poor	Unsatis-factory
Animal Mimicry	10%	74%	13%	3%	----
Black Box Problem	6%	27%	53%	15%	----
Yeast	73%	22%	3%	----	----
Entire Unit	23%	58%	19%	----	----

Table 3. Student Reactions Regarding Enjoyment and Educational Appropriateness of the Lessons (n= 76)

Clustering of sections by level score	n pre	n post
Single dominant level 	14	2
Tightly clustered distribution 	3	11
Other	10	14

□ = 1 interview section

Table 4. Clustering Patterns of High Scores on the Six Sections of the Clinical Interview (n = 27)

Appendix 1

WORDS TO UNPACK DURING THE INTERVIEW (What do you mean by __?):

Answer	Helps	Theory
Conclusion	Learn	Truth
Discover	Procedure	Try Again
Equipment	Proof	Try Out
Explanation	Test	Understand

INTRODUCTORY QUESTIONS:

1) What do you think science is all about?
2) What do you think the goal of science is?
3) Which statement do you think is a better description of the goals of science?
   - I) The goal of science is to discover new things in the world and the universe.
   - II) The goal of science is to build a better understanding of the world around us.

   Why? Can you give me some examples (of new things, or the kinds of things we try to understand)?
4) How do you think a scientist does this work?

I. IDEAS

1) Where do scientists get their ideas?
2) What kinds of ideas do scientists have?
3) What are scientists' ideas about?

4a) Do scientists do anything with their ideas? What do they do with them?
If TEST then:
   - 4b) How do scientists test their ideas?
   - 4c) What happens to the ideas once they've been tested?

5) Is there a relationship between a scientist's ideas and the rest of the work a scientist does? What is the relationship?
6) Do scientists change their ideas? Why (when) or why not?

II. HYPOTHESIS

1) What is a hypothesis?
2) Where does a scientist get a hypothesis?
3) Is there a relationship between a scientist's hypotheses and the rest of the work a scientist does? What is the relationship?

III. EXPERIMENT

1) What is an experiment? [UNPACK THE ANSWER]

2a) Why do scientists do experiments?
If TO TEST IDEAS then:
   - 2b) How does the test tell the scientist something about the idea?

3) How does a scientist decide what experiment to do?

IV. RESULTS

1) What happens when a scientist is testing his/her ideas, and gets a different result from the one he/she expected? [UNPACK THE ANSWER]