INT. J. SCI. EDUC., 1989, VOL. 11, SPECIAL ISSUE, 514-529

`An experiment is when you try it and see if it works': a study of grade 7 students' understanding of the construction of scientific knowledge

Susan Carey, Massachusetts Institute of Technology and Educational Technology Centre, Harvard University, and Risa Evans, Maya Honda, Eileen Jay, and Christopher Unger, Educational Technology Centre, Harvard University, Boston, USA

The research reported here focuses on grade 7 (12-year-old) students' epistomological views prior to and after exposure to a teaching unit especially developed to introduce the constructivist view of science. A clinical interview was used to assess students' understanding about the nature of scientific knowledge and inquiry. Students' initial epistemological stance is that scientific knowledge is a passively acquired, faithful copy of the world, and that scientific inquiry is limited solely to observing rather than constructing explanations about nature. We found that it is possible to move students beyond this initial view.

Introduction

One important goal of science education is to help students to understand the nature of the scientific enterprise itself. To fulfil this goal, we must agree on the epistemological view of science we want to impart, and, in order to successfully engage students, we must assess their epistemological views, diagnosing their misconceptions and alternative conceptual frameworks in this domain.

Much of current educational practice about scientific inquiry 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 grade 7. The standard curricular unit on scientific method 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, possible hypotheses and variables (and thus, experimental outcomes) for a given problem are often prescribed by the curriculum.

Certainly, process skills are important elements of careful scientific methodology. Junior high school students do not spontaneously measure and control variables or systematically record data when they first 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

relativism-there is no true knowledge and everybody is free to believe whatever they want. Finally, some people reach a mature epistemology that recognizes the impossibility of absolute truth, and recognizes the relativity of belief to interpretative frameworks, but also recognizes that there are canons of rational justification of belief. Given our concern with junior high school students, the earliest, naive realist stage is what is relevant.

The second source of evidence concerning young adolescents' epistemology of science derives from studies in which they are asked to design experiments and/or draw conclusions from experimental evidence. Dramatic deficiencies in scientific reasoning are amply documented in the classic work of Inhelder and Piaget (1958). 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, but it is equally likely that the problem is understanding the point of experimentation. Recent work by Kuhn and her colleagues supports this latter interpretation. Kuhn and Phelps (1982) studied 10- and 11-year-olds attempting to identify the substances critical to producing a chemical reaction (i.e., a colour 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 colour change, but rather to produce the colour change. Just as children do not distinguish theory from evidence, they do not seem to distinguish between understanding a phenomenon and producing the phenomenon.

In a similar vein, Kuhn *et al.* (1988) asked children of ages 8, 11 and 14 years, and adults to evaluate and generate evidence about the effects of features of tennis balls (e.g., colour, 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 colour of the ball would not make a difference). They were then asked to state whether a given set of data supported their view (called `theory' by Kuhn *et al.* 1988) refuted it or provided no evidence regarding it. Subjects of all ages, even adults, found the task difficult. To give one example, the two youngest groups were unable to generate possible data that would refute their theory. Kuhn *et al.* (1988) argue that their subjects' faulty reasoning revealed a lack of differentiation, at a metaconceptual level, of the notions of *theory* and *evidence.* They argue that 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.

This may be so, but these experiments provide only indirect evidence on this point, for they may also reflect nothing more than subjects' lack of knowledge of statistical inference rules. The `theories' offered by the subjects of Kuhn *et al.* (1988) are actually hypotheses about causal relations. The process skills explored by these studies and by Inhelder and Piaget (1958) 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 the goals for student understanding of scientific knowledge outlined above.

To provide more direct evidence concerning young adolescents' epistemology of science, we devised a clinical interview to probe specifically their views on the nature of scientific knowledge and inquiry. This interview also served as a pre-test and post-test to evaluate the effectiveness of an instructional unit devised to move grade 7 students beyond their initial epistemology. Before presenting the interview results, we turn to a description of our nature of science unit.

Our curricular approach: the nature of science unit

We assume that 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. We also assume that such knowledge is important in its own right, and that it can be gained only by actively constructing such knowledge and reflecting on this process. These assumptions motivate a curricular approach that emphasizes theory building and reflection on the theory building process. Thus, we have developed an instructional unit to replace the typical junior high school unit on the scientific method. The heart of our three-week-long nature of science unit is a two-week series of lessons in which students formulate and test their theories about the nature of yeast. This follows a week of introductory lessons that orient students.

Introductory lessons

In the initial lessons, 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. Students first observe and speculate about whether or not a small, unfamiliar object purportedly from Mars is alive. They compile a list of attributes of 'aliveness', and discuss ways to test the object for these attributes. The teacher helps the students to recognize that their ideas about living organisms come from their own experience, reside in their minds, and can be made explicit for inspection and evaluation.

Next, students view video material showing animals that disguise themselves in various ways. Using their own ideas about the basic needs of animals and the possible functions of different disguises, students organize the different kinds of animals disguises into categories. The teacher points out that categories and classification systems, like other scientific ideas, are constructed to help us make sense of the world.

Finally, students watch a video item of Linus Pauling working out the shape of an object in a closed box by systematically isolating and testing one feature of the object's shape at a time. Given a similar black box problem, students engage in formulating, testing, and revising their own hypotheses about the shape of their assigned object. Since the students are never allowed to actually see the object, they cannot determine which of their hypotheses is `right'. Instead, 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 that they may not be able to observe directly, and which may never be definitively proven.

Yeast lessons

The two-week series of yeast lessons involve 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. This 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?' Discussions lead the class to realize that they have not yet determined which ingredients are necessary for the mixture to bubble. In carrying out their own experiments to determine the essential ingredients, students' first efforts are unsystematic, reflecting their lack of process skills. They do not measure ingredients, nor do they even keep a record of which ingredients they used. In addition, their experiments display their limited understanding of the nature and purpose of experiments. Their view of the task is limited to trying to produce the bubbling phenomenon, which they attempt rather haphazardly. To them, experimentation consists of simply trying things out. Their view of the goal is to reproduce the bubbling phenomena rather than identifying what ingredients are necessary.

When the teacher challenges the class to draw conclusions from their experiments, none can be supported. The stage is set for standard lessons about the scientific method. The class then constructs a series of controlled experiments, which 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. Thus, the unit moves beyond simply considering how to collect reliable data and towards 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. Using what they know about water, sugar, yeast, and gases, students consider two mechanisms to explain why the yeast mixture produces a gas: (1) the bubbles are caused by some kind of chemical reaction between the yeast, water and sugar, 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; some help from the teacher is often required for the second hypothesis to even emerge from the discussion.

It is here that the class begins to learn that systematic experimentation has a purpose; it is in the service of constructing a deeper explanation of the phenomenon. Students are challenged to design controlled experiments that will help to 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 refute 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 carbon dioxide as a product of metabolism, the students hypothesize that if the yeast is alive, perhaps it too gives off carbon dioxide. A bromthymol blue experiment demonstrates that, indeed, the gas given off by the yeast mixture is carbon dioxide. In discussing their conclusions, students appreciate that this outcome is consistent with the hypothesis that yeast is alive, but is certainly does not prove this since carbon dioxide is the product of chemical reactions as well. Another experiment explores the effects of extreme heat and cold on yeast. Students expect that boiling or baking a living organism would kill it; they are less sure of the effect of extreme temperature on a chemical reaction. The results show that when yeast is baked before being mixed with sugar and water, the mixture does not produce a gas. This is consistent with their hypothesis that yeast is alive.

Other experiments, including *gedanken* experiments, are performed, and gradually students come to accept the mechanism they did not originally favour. 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 final lesson concludes the unit with a general discussion about the interplay of thought and experimentation in science, with special emphasis on the motivations for experimentation as an aid to theory building. Students are reminded that some of their basic notions about living things were challenged in the course of their investigations.

The study

The two goals of our study are to probe grade 7 students' initial understanding of the nature and purpose of scientific inquiry and to explore whether it is feasible to move students beyond their initial conception with a relatively short classroom-based intervention using our nature of science unit.

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 grade 7 science classes participated in the study. All classes were taught the 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.

The clinical interview

There are a number of written instruments that 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), a multiple-choice test; and the Nature of Scientific Knowledge Scale (NSxs; Rubba and Andersen 1978), a scaled-response measure designed for secondary school students. While both TOUS and NSKS offer a constructivist analysis of the nature of scientific inquiry and knowledge, and thus are very clear about possible student end-points, 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 during the test. Thus, we turned towards developing an interview that would allow students to give their own answers, and which would allow us to probe the meanings of critical terms and ideas.

Our half-hour clinical interview protocol 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'. The clinical interview protocol is reproduced in the appendix.

The coding procedure

Questions on the interview protocol were divided into six sections, and students' responses for each section were coded into categories that reflected three general levels of understanding, which are described below. When students answered `I don't know' to a question, the response was not scored, and did not enter into the student's overall score. Interviews were coded on the basis of listening to the interview tapes. Each interview was coded by at least two coders, who were unaware of whether it was a pre-instruction or post-instruction interview. Interscorer reliability was modest (74% agreement); disagreements virtually always involved only one level difference, and were resolved by discussion.

The coding scheme general levels of understanding

The students' ideas about the nature of science range 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 according to the degree to which ideas, experiments, and data/results are defined and differentiated from one another, and according to the degree to which the relationships among these elements are articulated and understood.

Three general levels of response were identified. In 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. The goal of science is to discover facts and answers about the world and to invent things.

In level 2, students make a clear distinction between ideas and experiments. The motivation for experimentation is to test an idea to see if it is right. There is an understanding that the results of an experiment may lead to the abandonment or revision of an idea; however, there is yet no appreciation that the revised idea must now encompass all the data-the new and the old. The goal of science is understanding natural phenomena-how things in the world work.

In level 3, as in level 2, students make a clear distinction between ideas and experiments, and understand that the motivation for experiments is 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. 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.

In addition, for some sections, level 0 responses were recorded. These reflected misconceptions in which students seem not to consider the information-seeking aspects of science at all.

Results: students' initial understanding and post-instruction improvement

For each section, every student received a mean section score. These were averaged, yielding a group mean score for each section. In addition, the highest score a student attained in each section was noted, and these scores were also averaged, yielding a mean high score for the section. The overall mean scores and mean high scores for each section are shown in table 1. For the pre-interview, the overall mean was 1·0. Of the 27 students interviewed, only four students had overall mean scores of over 1·5. 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 scientists' 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* natural phenomena. 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.

The overall mean score increased from 1·0 for the pre-instruction interview to 1·55 for the post-instruction interview ($P<0{\cdot}001$, Wilcoxon signed ranks test, 1-tailed). Every student improved, and improvement averaged half a level. Now 16 students achieved overall scores of 1·5 or better, and 5 scored 2·0 or better, a score nobody achieved at the pre-interview. The meaning of these results, and of the levels, can best be seen as each section of the interview is considered in turn.

Nature/purpose of science and scientific ideas

This section included questions about the goals of science and the kinds of ideas scientists have (appendix: Introductory Questions 1, 2, 3; Ideas 2, 3). Level 1 answers in this section focus on the activities themselves-the inventions of cures or contraptions-scientists have ideas about how to carry out these activities. Students are unable to elaborate the goal of science beyond statements such as, `to discover new things', `to find new cures for diseases', and so on. Level 2 answers focus on the development of a mechanistic understanding of the world. Scientists have ideas, questions and predictions about how things work, and predictions about the

Table 1.

Mean scores and mean high scores by section for the clinical interviews

n-27

	Mean scores: Pre-interview	Post-interview	Difference	Significance	Mean high scores: Pre-interview	Post-interview	Difference	Significance
Nature/purpose of science/ scientific ideas	1·09	1·28	+0·19	*p<0.03*	1·52	1·74	+0·22	NA
Nature of a hypothesis	1·0	1·37	+0·37	NA	1.0	1·37	+0·37	p < 0·002
Nature of an experiment	1·0		+0·52	p<0·002	1·19	1¯52	+0-52	p<0·001
Guiding ideas and questions	0·65	1·45	+0·80	P<0·001	1·41	222	+1·03	p<0·001
Results and evaluation	1·06	1.8	+0·74	p<0·001	0-93	2·19	+1-78	p<0·001
Relationships	0·91	1·69	+0-78	p<0·001	NA	2v3A	+1·37	NA
Overall score	1.0	1·55	+0·55	p<0·001			NA	

outcomes of experiments. More specific goals are mentioned as examples (e.g., `to find out how animals get oxygen'). Level 3 answers focused on the development of an explanatory understanding of the world. The point of science is to construct explanations for why things are the way they are, examples being `why the leaves change colour', `why the dinosaurs became extinct'.

The average pre-interview score for this section was 1·09 (see table 1). Scores ranged from 0·33 to 1·67. The median score was 1·0, the modal scores were 1·0 ($n = 8$) and 1·22 (n=8). At the pre-interview then, students saw the purpose of science as discovering facts, making inventions and developing cures. The unit had a small, but statistically significant, impact on the students' responses to the questions in this section (see table 1). As at the pre-interview, the average score hovered slightly above 1 (1·28), and the median and modal score (n=ll) were again 1·0. The range, however, extended to include higher scores-from 0·33 to 3·0. Whereas no students scored 2·0 or higher at the pre-interview, four scored 2·0 or higher at the post-interview.

Nature of a hypothesis

This section included a single question about what a hypothesis is (appendix: Hypothesis 1). In level 1 answers to this question, a hypothesis is an idea or a 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 clearly related to an experiment or phenomenon, and it is explicitly something that can be tested, e.g., `an if-then statement about what you think might happen'. In level 3 answers, the hypothesis is not only related to an experiment, but aids in interpreting the results of the experiment, and is evaluated and developed in terms of the results.

Only two students knew the word `hypothesis' at the pre-interview; each gave a level 1 answer. The remaining students received no score. All students answered at the post-interview, where the mean score was 1·37 (table 1), with a range from 0 to 3. The median score for 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).

The nature/purpose of an experiment

This section included questions about what experiments are and why scientists do them (appendix: 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 included statements like, `when you try something new'. In level 1 answers, there is no clear distinction between experiments and ideas. The motivation for doing an experiment, either implied or explicitly stated, 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 out about the thing they're experimenting on'. In level 2 answers, the distinction between the idea and activity is clear. The experiment is a test of a scientist's ideas, or an operationalized exploration of a phenomenon. In typical level 2 answers, scientists do experiments `to test to see if their idea is right'.

Experimentation is also seen as playing this role in level 3 answers. In addition, the relationship between results 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.

At 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).* At the pre-interview, then, students saw experiments as activities that support the goal of science-finding things out and discovering facts. There was a considerable increase in the mean score to 1·52 (see table 1). Mean scores ranged from 0 to 3·0, and the median and model scores *(n=10)* 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.

Guiding ideas and questions

This section included questions about how scientists do their work, where they get their hypotheses, and how they decide which experiments to do (appendix: Introductory Question 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, but there is no specific question or phenomenon that guides these activities. In typical level 1 answers, scientists do their work by `putting things under microscopes to see how they behave'. In level 2 answers, exploration is guided by a particular idea, question, object of phenomenon, e.g., a scientist `walks through a forest and finds something new and tries to find out more about it'. In level 3 answers, either the guided exploration of level 2 is elaborated to include reflection on prior knowledge and experience, or there is an understanding of evaluation and development of ideas, e.g., 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'.

At the pre-interview, the mean score was 0·65, with a range from 0 to 1·5. The median and modal scores *(n=14)* were 0·5. Thus, at the pre-interview, students revealed the misconception that a scientist's choice of hypotheses and experiments is mostly capricious. The improvement at the post-interview was dramatic (see table 1). The mean rose to 1·45, with a range from 0·33 to 2·5. The median score was now 1·33, and the modal score was 1 *(n = 8).* While none of the students scored 2 or better at the pre-interview, seven did so at the post-interview. Following the unit, these students understood that the activities of science are guided by particular ideas and questions.

Results and evaluation

This section 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: 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 I answers, the scientist `checks it to see what he did wrong and tries to fix the problem', or changes the 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 typical level 2 answers, `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'. However, level 2 responses provide no account of what constrains these changes. 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, e.g., `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 hypothess 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; 13 scored between I and 2. Thus, the majority of students gave level 1 responses in which even hypotheses and experiments are not clearly differentiated. The improvement in the mean to 1·*8* at the post-interview was also dramatic (table 1). Scores for the post-interview ranged from 0·5 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, 19 did so on the post-interview, demonstrating a clear understanding of the distinction between idea and experiment, and in some cases, of the relations between idea and results.

Relationships

This 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; 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 the opposite point 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 either state that there is no relationship, or else they give a very incomplete rendering of it; scientists `report their ideas', or `write them down'. In level 1 answers, there is still no clear distinction between ideas and experiments. A scientist tries an idea to see if it works, `does it', or uses it a guide or a blueprint. Scientists, `make their ideas work', `see if they are accurate, if they can really do them', or `fulfil 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. A level 3 response goes beyond the 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.

The mean score at the pre-interview was 0·91 (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. Thus, the overwhelming majority of the students saw ideas, at best, as blueprints for action, or interchangeable with the things they are about. The post-interview improvement in the mean score to 1·69 was again dramatic (table 1). The median score was now 1·5 and the modal score 2·5 (n = 8). Fourteen students scored 2 or better at the post-interview.

Discussion

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 all made reference to these relationships.

The significance of the gains at the post-interview may be questioned, since the interview required only that the student repeat points explicitly made by the teacher several times during the unit. While students moved beyond level 1 understanding, they did not approach level 3 understanding, although the lessons included level 3 points on each of the topics probed at the interview. We conclude, therefore, that the gains in understanding are genuine. It is an open question how long lasting these gains would be, if not consolidated in the rest of the curriculum. It is also an open question whether more sustained curricular intervention could induce level 3 understanding in grade 7 students.

Several issues remain for further study. The levels must be better articulated, especially the distinction between levels 2 and 3, and must be related to the students' general epistemological ideas. The relation between metaconceptual understanding of the role of experimentation in theory building and the process skills involved in experimentation should be directly studied. Finally, it is important to explore whether a student's understanding of the nature of science has any impact on his or her learning of science content, especially in those cases where conceptual change is required.

Conclusions

The results from our clinical interview support the suggestion in the literature that pre-adolescent children have a different epistemological stance towards scientific knowledge than do scientifically literate adults. Initially, most of the grade 7 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 from the evidence that supports those ideas. Although grade 7 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. 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 (1958), and by Kuhn *et al.* (1988).

While our three-week nature of science 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 are asked to reflect on the problem under investigation and to examine 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.

Acknowledgements

Preparation of this paper was supported in part by the Office of Educational Research and Improvement (Contract No. OEM 400-83-0041). Opinions expressed herein are not necessarily shared by OEM and do not represent OERI policy. We want to thank the following people for their participation in the Educational Technology Centre's Nature of Science Project 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, and the Watertown Public Schools for their co-operation with this work. Finally, we want to thank Roberta Brawer, Mary Maxwell Katz, Wayne O'Neil, Judah Schwartz, Carol Smith, and Martha Stone Wiske for their helpful comments on earlier versions of this paper.

References

BROUGHTON, J. M. 1978, Development of concepts of self, mind, reality, and knowledge. In W. Damon (ed.), *Social Cognition: New Directions for Child Development*, Vol. 1 (Jossey-Bass, San Francisco, CA).

BROUGHTON, J. M. 1983, The cognitive-developmental theory of adolescent self and identity. In B. Lee and G. G. Noam (eds), *Developmental Approaches to the Self* (Plenum Press, New York, NY).

CHANDLER, M. and BOYES, M. 1982, Social-cognitive development. In B. B. Wolman (ed.), *Handbook of Developmental Psychology* (Prentice-Hall, Englewood Cliffs, NJ).

HODSON, D. 1988, Toward a philosophically more valid science curriculum. *Science Education*, Vol. 72, No. 1, 19-40

INHELDER, B. and PIAGET, J. 1958, *The Growth of Logical Thinking from Childhood to Adolescence* (Basic Books, New York).

KITCHENER, K. S. and KING, P. M. 1981, Reflective judgement: Concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology,* Vol. 2, pp. 89-116.

KLOPFER, L. E. and CARRIER, E. O. 1970, TouS: *Test on Understanding Science (Form Jw)* (Learning Research and Development Centre, University of Pittsburgh, Pittsburgh, PA).

KUHN, D. and PHELPS, E. 1982, The development of problem-solving strategies. In H. Reece (ed.), *Advances in Child Development and Behaviour, Vol.* 17 (Academic Press, New York).

KUHN, D., AMSEL, E. and O'LOUGHLIN, M. 1988, *The Development of Scientific Thinking Skills* (Academic Press, Orlando, FL).

NADEAU, R. and DESAUTELS, J. 1984, *Epistemology and the Teaching of Science* (Science Council of Canada, Ottawa).

PERRY, W. G., JR. 1970, *Forms of Intellectual and Ethical Development in the College Years* (Academic Press, New York, NY).

RUBBA, P. and ANDERSEN, H. 1978, Development of an instrument to assess secondary students' understanding of the nature of scientific knowledge. *Science Education,* Vol. 62, No. 4, pp. 449-458.

Correspondence

Susan Carey, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

Appendix: The clinical interview protocol

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 kind 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. HYPOTHESES

(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]