

The Effect of an Earth–Science Learning Program on Students' Scientific Thinking Skills

Nir Orion

Science Teaching Department Weizmann Institute of Science, Rehovot, Israel
76100

Yael Kali

Technion, Haifa, Israel

ABSTRACT

This study explored junior high school students' understanding of essential concepts of scientific thinking "observation", "hypothesis" and "conclusion" and the effect of the learning of the program "The Rock Cycle" on the development of such understanding. The study sample consisted of 582 students of the 7th and 8th grade, who learned in 21 classes, with 14 teachers from 8 schools in Israel. The data collection was based on a quantitative research tool that was specifically developed for this study and qualitative tools such as observations and interviews.

The findings indicated that the students have considerable difficulties in understanding the basic concepts underlying the scientific inquiry, and that the "The Rock Cycle" has a potential to develop such understanding. An unexpected gender difference was found. Girls outperformed boys in scientific thinking, both in the pre and the post tests. The unique character of geoscience methodology, together with structured-inquiry and metacognitive activities, served as an appropriate framework for students to develop basic scientific thinking. The co-interpretation of quantitative and qualitative analysis indicated that the type of teacher (openness to innovative methods, enthusiasm and scientific background) was a crucial factor in students' ability to exploit the potential of "The Rock Cycle".

INTRODUCTION

The "Science for all" reform identified the inquiry as a principal method in science education (e.g., AAAS, 1993; National Research Council, 1996). Scientific inquiry in different disciplines, although much alike in many fundamental aspects of the scientific method, is very different in others. The authors of Science for All Americans, in Project 2061 (AAAS, 1990) claimed that: "scientists differ greatly from one another in what phenomena they investigate and in how they go about their work; in the reliance they place on historical data or on experimental findings and on qualitative or quantitative methods; in their recourse to fundamental principles; and in how much they draw on the findings of other sciences." (p. 4).

Inquiry in the geosciences has a unique characteristic, which derives from their involvement with the "experiments" that have already been conducted by nature. Consequently, many geological inquiries are of a retrospective type - trying to unravel what happened in the past, using "fingerprints" left on the earth. Frodeman, (1995) describes geology as an interpretive, and historical science, which "embodies distinctive methodology within the sciences". He further argues that "the geologist picks up on the clues of past events and processes in a way analogous to how the physician interprets the signs of illness or the detective builds a circumstantial case against a defendant" (p. 963). Edelson

et al. (1999) described several aspects of the geosciences as being "observational sciences". We believe that the focus of geology in making inferences from observations, rather than a focus in experimentation (Ault, 1998), is a suitable milieu for internalizing the meaning of some of the most basic constructs of scientific thinking, i.e., observations, conclusions and hypotheses. An additional advantage in geology as a context for learning science, is the possibility to provide students with concrete materials of the earth (e.g., rocks, minerals, soils), from which they can conclude upon exciting phenomena like volcanic eruptions, earthquakes or formation of mountain ranges.

In the current study, the effect of a new learning program, on students' scientific thinking skills was studied. "The Rock Cycle" is a 30-hour learning program developed for the new Israeli curriculum, "Science and Technology", for junior high school students. The program, which focuses on geological processes that transform the materials found within the crust of the earth, takes advantage of the retrospective type of inquiry of the geosciences, in order to enhance students' general scientific thinking skills. The main learning method in the program is guided-inquiry, designed as structured activities.

The goal of the study was to examine the effect of different implementations of "The Rock Cycle", on students' understanding of the essence of scientific inquiry. Specific objectives were the following:

- A) To compare students' ability to discriminate between observations, conclusions and hypotheses, before learning the program and afterwards, in different implementations, on a quantitative basis.
- B) To suggest possible causal relationships between students' achievements (assessed quantitatively), and qualitative characterization of the different implementations.

THE DESIGN OF INQUIRY ACTIVITIES IN THE PROGRAM

The design of "The Rock Cycle" activities is derived from the constructivistic and social constructivistic epistemologies. All the activities in the program are performed in an inquiry method, the main resources of which are concrete items - natural materials of the earth, brought to the lab, or studied in the field. The inquiry is guided by means of a booklet, which includes mainly questions, and only a minimal amount of textual information. In this manner, groups of three or four students work in collaboration to "discover" the geological processes by themselves, following the "bread-crumbs" the designers have left for them on the way. The inquiry is guided in a similar manner to that described by Karplus (1979) in his "learning cycle" strategy. Each chapter in the booklet starts with student observations, which create a certain cognitive conflict. To resolve this conflict, students initially express their own

1. Summarize the Activity in the Following Manner:
List of Observations
 A) When we poured water on the marl it became malleable.
 B) When we poured drops of diluted acid on the marl it fizzled

List of Additional Information
 C) Clay minerals are the only minerals that are malleable.
 D) Calcite is the only mineral that fizzes in diluted acid.

List of Conclusions
 E) The marl is comprised of the following minerals: clay and calcite.

2. Represent (verbally or graphically) the scientific thinking route you have passed in this activity.
Verbal representation of the scientific thinking route
Observations A and B, together with additional knowledge C and D, brought us to conclusion E.

Graphic representation of the scientific thinking route

A + B + C + D → E

Figure 1. Examples of the third type of MRS activities (including example answers).

hypotheses and then follow a route of inquiry that was designed for them in the booklet. The role of the teacher is therefore, to mediate between the students and scientific knowledge, by helping students to use the inquiry method to investigate the earth and its processes (Kali, Orion and Eylon, 2003).

To foster students' awareness of this scientific thinking route, summary activities of Metacognitive Scientific Reconstruction (MSR) are performed at the end of each inquiry activity. In these activities linguistic terms are used as organizing schemes in a metacognitive process. Students examine their investigation with "scientific inquiry spectacles", by characterizing the different stages of the inquiry, using terms like "observation", "hypothesis" and "conclusion".

The nine MSR activities that are included in the booklet are presented in three alternative variations, which the students can choose from. Each variation is a different combination of solutions adjusted to different students' needs and difficulties. The first variation starts with a short verbal summary of the inquiry route, with missing words for the student to complete. It continues with using the linguistic schemes (terms like "observation", "hypothesis" and "conclusion"), for organizing the inquiry process. This variation might be adequate for students who feel comfortable with verbal and structured environments, and who have difficulties in following the scientific route without assistance. Shymansky, and Yore (1980) indicated that field-dependent students are more likely to favor such structured environments. The second variation includes verbal expressions, for which the student is asked to decide upon the scientific concept expressed in them. The expressions are arranged in a sequence representing the scientific thinking route. This variation is very similar

| School | Socio-economic Background | Class no.* | N sum |
|--------|---------------------------|------------|-------|
| a*g | Average | 9*c | 30 |
| | | 3 | 15 |
| | | 15 | 25 |
| | | 8 | 19 |
| | | 19 | 23 |
| b*g | Above average | 14*c | 24 |
| | | 7 | 0 |
| c | Below average | 21*c | 6 |
| d | Average | 16 | 10 |
| e | Average | 20 | 26 |
| f | Below average | 13 | 16 |
| | | 6 | 22 |
| g | Average | 17 | 20 |
| | | 2 | 16 |
| h | Average | 4 | 16 |
| | | 10 | 24 |
| i | Average | 5 | 20 |
| | | 1 | 18 |
| | | 18 | 14 |
| | | 11 | 18 |
| Sum | | | 411 |

Table 1. Details of the sample (reduced data). * Class number was determined by the magnitude of pre-post differences in Observations sub-test scores. *g Schools that received our guidance throughout implementation. *c Case study schools.

to the first variation. However, since our preliminary observations showed that some students felt insulted by having to fill in missing words, this variation enabled them to focus on the scientific terms, without literally presenting their understanding of the scientific thinking route. The third variation (Figure 1) includes instructions for organizing lists of all the scientific stages of the inquiry, using the organizing schemes (the scientific terms), and to freely present the scientific thinking route with verbal or graphic means. According to Shymansky and Yore (1980), this variation, which is less structured than the first two, is adequate for field-independent students with either a graphic or a verbal cognitive style.

METHOD

Sample and Data Sources - The sample consisted of 582 students of the 7th and 8th grade, who learned in 21 classes, with 14 teachers from 8 Junior high schools in Israel, representing urban, suburban and rural societies. The analysis was conducted on a smaller sample, including only 411 of those students (221 males and 190 females) from whom we obtained full data (pre-tests and post-test) (Table 1). This procedure of reducing the sample did not introduce any bias, since it was a random process that resulted from students' absence in either of the tests.

Teacher Training - All the teachers in the study, participated in an in-service teacher-training course (INSET) of 20 hours, prior to the implementation. In this course, teachers carried out the same inquiry learning activities, which they were supposed to teach in the

| Questionnaire Component | Pre/Post | All Mean Score (N=411) | Boys' Mean Score (N=221) | Girls' Mean Score (N=190) | t | P |
|-------------------------|----------|------------------------|--------------------------|---------------------------|------|--------|
| Observations | Pre | 74 (SD=33) | 67 (SD=33) | 78 (SD=30) | 3.26 | 0.001 |
| | Post | 87 (SD = 25) | 81 (SD=28) | 91 (SD=19) | 3.57 | 0.0004 |
| Hypotheses | Pre | 80 (SD=24) | 77 (SD=25) | 81 (SD=21) | 1.65 | NS* |
| | Post | 84 (SD=23) | 80 (SD=24) | 86 (SD=20) | 2.55 | 0.001 |
| Conclusions | Pre | 47 (SD=26) | 46 (SD=26) | 46 (SD=25) | 0.16 | NS* |
| | Post | 49 (SD=25) | 44 (SD=26) | 52 (SD=23) | 2.60 | 0.009 |
| Overall Scores | Pre | 69 (SD=18) | 66 (SD=19) | 72 (SD=17) | 2.72 | 0.008 |
| | Post | 76 (SD=18) | 72 (SD=18) | 79 (SD=16) | 3.91 | 0.0001 |

Table 2. Pre and post scores of the entire sample and t-test results comparing scores of boys and girls in the different sub-tests. *None Significant.

Assume that holes in apples are created only by worms

What does each of the following expressions represent? (Circle the correct answer.)

- There is a hole in the apple.
A) Observation B) Conclusion C) Hypothesis
- A worm created the hole in the apple.
A) Observation B) Conclusion C) Hypothesis
- The worm made the hole in the apple after it was rotten.
A) Observation B) Conclusion C) Hypothesis

Figure 2. Example tasks from scientific thinking-skills set.

following weeks. Focus was put on the MSR activities, in which teachers were encouraged to discuss different possibilities of describing scientific thinking-routes, with the terms "observation", "hypothesis" and "conclusions". It was emphasized that whether an idea expresses an observation, a hypothesis or a conclusion, is dependent in many cases on the context, and that scientists themselves sometimes disagree about such methodological issues. Using this context, the constructivistic and social-constructivistic teaching methods were emphasized.

Research Tools - In order to examine students' understanding of some of the basic concepts underlying scientific inquiry, we developed the HOC (Hypothesis, Observation, Conclusion) test. In this test, the student is presented with information consisting of an illustration and a related verbal statement. Based on this information, the student is required to sort several expressions according to three categories, and decide whether they represent an observation, a hypothesis or a conclusion (Figure 2). The test includes seven hypotheses expressions (the "hypotheses" sub-test), eight observation expressions (the "observations" sub-test) and five conclusion expressions (the "conclusions" sub-test), which are mingled together in each of the illustrations.

Cronbach's reliability coefficient for the HOC test was 0.80. This value was calculated using pretest scores of five classes (N=175) from one of the schools in the full sample. The rationale for using data from only one of the

schools was to rely upon scores of students with a relatively similar background. Using the same sample, values for the subtests were 0.84 for the observations subtest, 0.72 for the hypotheses subtest and 0.55 for the conclusion subtest (notice that this subtest included only 5 items).

Expert judgment including 10 experts (5 senior science education researchers and 5 expert science teachers) indicated that the three item types of the HOC-test (hypotheses, observation and conclusion items) belong to a common category of scientific inquiry constructs.

Procedure and Data Analysis - In order to examine the effect of implementing "The Rock Cycle" on students' scientific thinking skills, a zoom-in type of analysis approach was conducted. Large samples were used for quantitative assessment of different implementations as to their effect on students' understanding of basic scientific concepts, and small samples were used for qualitative characterization of three case-study implementations. Finally, deeper insight was gained by using the qualitative characterization, to interpret the quantitative effect of the program on students' scientific thinking skills.

The pre test was administered a week before starting the program, and the post test, a week after it ended. The period between the pre-tests and the post-tests varied from four to six months, depending on the intensity of implementation. In order to obtain broad information about students' acquisition of scientific thinking skills following the intervention, data were initially analyzed on a general level (the whole sample). Since the implementations differed from each other in several criteria (population, teaching styles, length of implementation etc.) a single-class analysis was additionally applied. The single class cases were randomly selected. In each implementation, t-tests were used to examine the effect of the program on students' scientific thinking skills. Wilcoxon's signed rank test was used in cases of small samples.

RESULTS

General Analysis - Table 2 presents the mean score values of the whole sample, derived from the scientific thinking-skills test. The data show that before learning the program, the students' ability to recognize observations and hypotheses was mediocre and, their ability to recognize conclusions was very low. On a general level of analysis, certain improvements can be seen, especially in the scores of the "observation" sub-test. However, due to a very large heterogeneity

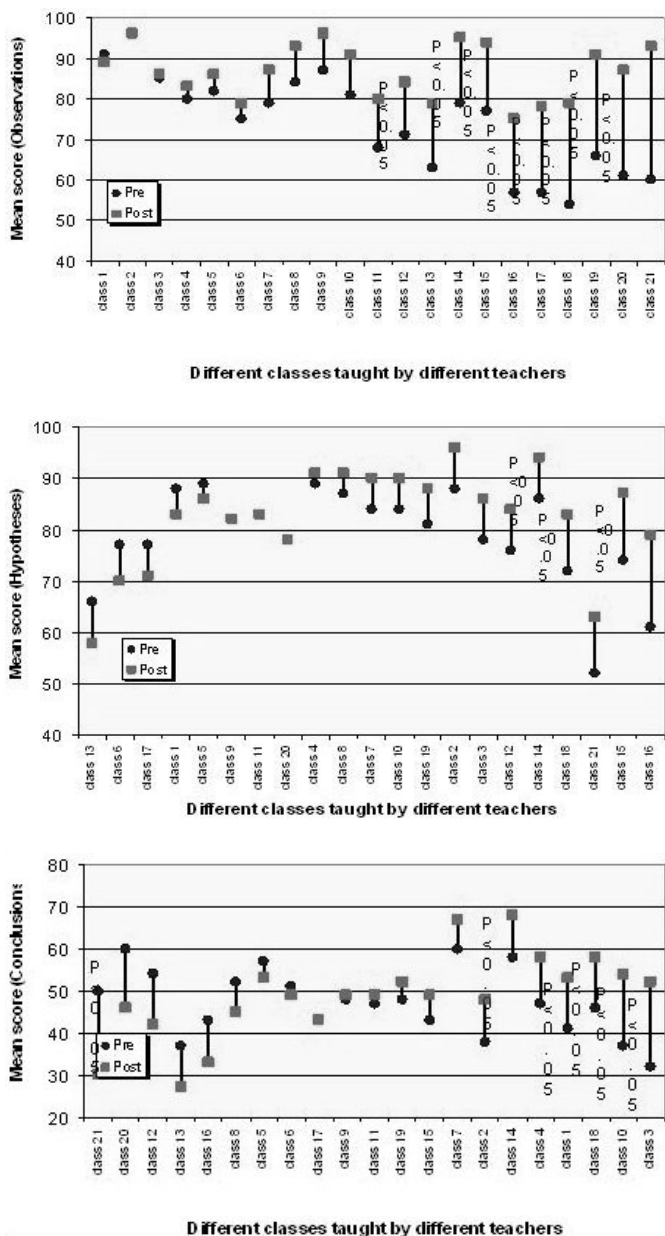


Figure 3. Scores of pretests and posttests of the three sub-tests (single class analysis).

between the classes, the large-scale improvements, are masked at this level of analysis.

Single-class Analysis - A single class analysis shows a great deal of improvement in the scores of all three sub-tests in many of the classes (Figures 3a, 3b and 3c). In these figures, the classes are arranged by the size of the pre-post differences, with the higher differences towards the right side. The most outstanding feature is the large variability between the classes in the differences between pretest and posttests. Such variability exists even between classes taught by different teachers within the same school. This is illustrated in table 1 by the diversity of class-numbers within each school. Since class-numbers were determined by the degree of improvement in the "observations" sub-test, such diversities indicate large variability in the improvement of classes within the same school.

Beside some very impressive improvements on the right side of each of these figures, negative differences can also be seen in some of the classes (left side of figures 3b and 3c).

Gender Differences - An unexpected gender difference, in which girls outperformed boys, was found. The general scores of girls, both in the pre-test and the post-test, were higher than those of boys (Table 1). A closer examination shows that the difference in the pre-test scores are due to higher scores of girls in the "observations" sub-test. However, after learning "The Rock Cycle", girls' scores were higher in all of the sub-tests (observations, hypotheses and conclusions). These findings are in contrast to many studies, in different scientific domains, which indicate the opposite trend. An extensive meta-analysis of 30 studies about gender differences in science achievement, at different age levels, did not reveal any advantage of girls over boys in different scientific domains (Becker, 1989). On the contrary, boys were found to be superior at biology, general science, and physics studies. According to that meta-analysis, no significant differences are found in geology and earth sciences studies. Similar findings are reported on science achievement in other fields of study (e.g., Erickson, and Erickson, 1984; Lock, 1992).

The Role of Teachers - One of the striking findings of this study is their variability across the classes. In order to control the influence of the implementation of the unit by the different classes an in-depth profile was conducted for each teacher through interviews and classes' observations (at least in 6 hours teaching of each teacher). The observations documented both the teacher's way of implementing the unit and the activity of the students.

The following is a description of three representing case studies which describe three types of teachers can be characterized: a) the senior teacher with a firmly fixed teaching approach, b) the beginner, open-minded teacher and c) The enthusiastic teacher lacking sufficient scientific knowledge.

Case Study 1: A Senior Teacher "Z" with Firmly Fixed Teaching Approach

Although Z felt as if she got used to the teaching methods that we had advocated (i.e., constructivistic, and social-constructivistic approaches), her behavior indicated that the transmissionist teacher-centered approach was a very basic component of her teaching style. This was expressed in several manners, which could be detected as a result of our observations and interviews with Z:

- A) Her tendency to extend whole class introductions and summaries at the expense of independent group-work of students.
- B) Her anxiety of "losing control" over the students.
- C) Her belief that there is only one "correct" knowledge, which is the one that is found in her booklet.
- D) Her belief that she can transmit this "correct" knowledge to the students by making explanations to the whole-class, or correcting their written homework.

During the interview Z express her duty as a teacher by using the metaphor "Mother". Z's behavior towards her students as evident from scolding, preaching, being proud of them or insulted on their behalf, suggests that

the type of mother she referred to, was one that loves and cares for her children, but uses her power to impose her views on them and direct their steps where she considers appropriate. Her transmissionist epistemology is congruent with specific functions of a mother, such as being the source of right-and-wrong type of knowledge for the child, about the world.

The students' reaction to Z's teaching style was acceptance of her authority, at least, as far as knowledge was concerned. This attitude was expressed in their tendency not to argue with Z's corrections of their homework answers. Such an attitude, which caused them to seek answers that would satisfy Z, seems to have reduced their motivation to understand the phenomena studied in the classroom.

The quantitative analysis shows that students in Z's class significantly improved their scores in the "observations" sub-test, but did not do so in the "hypotheses" and the "conclusions" sub-tests. We suggest that Z's teaching style was one of the main reasons for the potential of the curriculum materials not to be fully exploited in this implementation.

CaseStudy 2: The New and Open-minded Teacher "L" - The characterization of case study 2 shows that L did not have fixed ideas about teaching approaches, and therefore, did not have to adjust herself to the constructivistic approach more than to any other approach. This was evident in the following findings:

- A) Her rapid appropriation of the new teaching approaches, which were introduced during the INSET and the guidance meetings.
- B) The large number of opportunities she provided, for collaborative inquiry and metacognitive reconstruction activities to the students.
- C) Her encouragement of debates in a congenial atmosphere, in which several answers can be considered as "correct" provided that appropriate reasoning stands behind them.
- D) The fact that she gave students credit, for being able to continue their laboratory experiments on their own, with the aid of the booklet, on occasions when she was absent of school.

The students in L's class seemed to have appropriated the learning norms induced by her. An example is the fact that they did not express any surprise or dissatisfaction, even when they were asked to continue working on their own while L was absent.

The quantitative analysis shows that students in L's class significantly improved their scores in all three sub-tests. This indicates that they benefited from the learning situation of this implementation. There is no doubt that the sequential effect from the type of teacher to students' cognitive growth consists of many factors that could have contributed to this improvement. Students' high socioeconomic background, for example, could have created a situation in which students would have been able to exploit the potential of "The Rock Cycle" without the contribution of L's teaching. However, a deeper examination of the quantitative results suggests that this is not the case. Such an examination indicates that while L was teaching in her class, another teacher was doing so, in a different class at the same school. Students of the latter class showed no significant improvement in any of the sub-tests, although the background of the students was similar, indicating

that L's teaching constituted an important factor in their cognitive growth. It is reasonable to assume that the atmosphere and learning opportunities provided by L, were the source of the improvement in her class.

Case study 3: The Enthusiastic Teacher "H" Lacking Sufficient Scientific Background - The quantitative results of this implementation show a very interesting phenomenon. Viewing a cross-section of H's class indicates the largest difference between scores in the "observation" sub-test, and the third largest difference in the "hypothesis" sub-test. However, this class also showed the largest negative difference in the "conclusion" sub-test of all the classes examined.

The qualitative data obtained from H's implementation, led us to interpret the results in the following manner. H, who was very enthusiastic about teaching the program, and was eager to take the challenge of enhancing scientific thinking amongst her students, invested a lot of time and effort in the implementation. We assume that this investment instigated the considerable changes in the students' ability to understand scientific concepts, which were expressed, in a very large improvement in the "observations" and "hypotheses" sub-tests. However, H's own background in science was a little shaky, as noted during the teacher-training course and she used to make mistakes with the students as well. It is probable that imparting mistaken information to the students might have been the cause for large negative differences in the "conclusion" sub-test.

DISCUSSION AND CONCLUSIONS

The outcome of the general analysis as well as that of the single-class analysis, with regard to the pre-tests in all three sub-tests, indicates that the 7th and 8th grade students included in this study, have considerable difficulties in understanding concepts underlying the scientific method. The large pre-post differences, while not consistent across all classes, indicate the potential of "The Rock Cycle" in developing the three basic elements of scientific thinking.

This outcome can contribute to the debate concerning the degree of structure in activity design, appropriate for inquiry learning and teaching at different stages of studying science. Several studies suggest using a higher degree of structure in inquiry activities by constructing a preliminary scaffolding stage for open inquiry (e.g., de Jong and van Joolingen, 1998; Elshout and Veenman, 1992; Germann, 1989; Krajcik et al. 1998). Adding structure to learning programs was also used as a scaffolding method for encouraging students to develop integrated understanding of complex domains (Linn, 1995; Bagno and Eylon, 1997). Indeed, a close examination of students' ability to perform a meaningful exploitation of open environments, brought to an addition of structure and curricular guidance. For example, Edelson et al. (1999) describe an iterative process of software and curriculum design, in which four generations of software were developed, each of which was based on the evaluation of the preceding generations. The first version was a completely open model, based on an entirely student-driven form of inquiry, without any pre-designed curricula to guide it. As the software evolved, a more structured curriculum was incorporated into the program. The authors argue that "by guiding students through short, well-structured

investigations, staging activities can provide students with conceptual models for investigation techniques that the students can draw on in subsequent open-ended investigations" (p. 442).

It is reasonable to assume that students who do not understand the concepts "observation", "hypothesis" and "conclusion", as was found in the current study, are unprepared for engagement in open inquiry activities such as making their own hypotheses, designing experiments, and collecting, analyzing and reporting their own data. Therefore, we suggest that one of the first scaffolds, on the way to open inquiry, should be structured inquiry activities, as those incorporated in "The Rock Cycle". Such activities, in which students are actively engaged in pre-designed scientific thinking routes of inquiry, and are encouraged to reconstruct these routes in a metacognitive manner, are an example of learning activities that can assist students in gaining basic understanding of scientific thinking.

However, as noted by Perkins (1992), in addition to methods, there is considerable importance to contents in curriculum design. We suggest that the large improvement in students' scientific thinking skills, found in many of the classes, might have been a result of students' engagement with the contents of "The Rock Cycle". The unique inquiry methods of geoscience, which enabled students to focus on observations of concrete material of the earth, and to draw conclusions from "experiments" that were conducted by nature in the past, might have served students as a preliminary scaffold towards understanding of the essence of scientific inquiry.

Our findings also indicate that despite the great importance of appropriate curriculum materials they are not sufficient in themselves for inducing cognitive development amongst students. The large variability of the pre-post differences that was found between the classes, indicates that additional factors were involved in students' ability to exploit the potential of "The Rock Cycle" in developing scientific thinking skills. The finding that such differences exist even within schools, between classes taught by different teachers, induced us to focus on factors regarding the effect of the teacher on students' cognitive growth, using the conceptual framework of the aforementioned sequential effect. In spite of the limitations of this conceptual framework, it seems that in each of the implementations discussed, causal effects, between factors regarding the type of teacher, and students' cognitive growth, can be distinguished. Such factors are: a) openness towards innovative teaching methods, b) scientific background and c) enthusiasm, and willingness to invest time and effort in teaching. The type of teacher, therefore, seems to be an important factor in the ability of students to exploit the potential of curriculum materials.

Bearing this conclusion in mind, an important issue is raised, concerning the degree to which teacher-training has an effect on their management of classroom activities. The current study suggests that teachers' management of classroom activities is affected more by factors forging their personality, than by the amount of professional guidance they receive. For example, the teacher presented in case study 1, who received the most extensive guidance, but whose teaching approach was the most firmly fixed, exploited the potential of the curriculum materials, concerning students' scientific thinking, in the least effective manner. In spite of her declarations, which indicated an

epistemological change to a constructivistic approach, her management of classroom activities revealed a transmissionist-based pedagogy. Her firmly fixed teaching approach was eventually stronger than the effect of many hours of guidance.

Teacher training for introducing innovative programs often focus on different teaching methods. The current study indicates that such training is appropriate for the type of teacher who is open-minded, willing to invest much effort in teaching, and has an appropriate scientific background. However, heterogeneity of teachers in INSET courses should receive more attention, especially when junior high school teachers, who have diverse backgrounds, are concerned. Similar pedagogical concerns as those applied with heterogeneity of students, should be applied in teacher training. Distinguishing between different types of teachers, and tailoring appropriate programs for each group, might lead to more effective INSET, which is likely to affect cognitive growth amongst students.

Finally, it is suggested that the findings of this study might be related to the "Nature of science", a subject which is considered to be a primary component of scientific literacy (AAAS, 1993; Bybee, 1997; NRC, 1996). Bell and Lederman (2003) link instruction of the Nature of Science to the ultimate goal of science literacy-improving citizen's abilities to make reasoned decisions in a world increasingly impacted by the processes and products of science. In this context, there is little doubt that the ability of citizens to differentiate amongst observations, conclusions and hypotheses is a primary component of their ability to judge what appear in the media as "scientific facts" and to ask the right questions about such facts' origin, interpretation and acceptance.

LIMITATIONS AND IMPLICATIONS

The main limitation of this study (as in many other studies in the area of science education) is the ability to control variables. Although, we tried in this study to isolate the influence of the teacher from other possible factors, our conclusions and the new research tool that we developed are still not beyond questions. However, we hope that this study and the new research tool will allow earth science education researchers and curriculum developers to test our tool and approach with larger samples and in different ages, various settings and cultures. We hope that such studies will help to convince science teachers to teach the earth science also as a tool for development essential scientific thinking skills.

SUMMARY

- Junior high school students had considerable difficulties in basic scientific thinking skills, and were found to be unprepared for engagement in open inquiry activities.
- Structured inquiry activities, such as those incorporated into "The Rock Cycle" are suggested as a primary scaffold that should be provided to students in order to enhance skills required for independent open inquiry.
- The unique character of geoscience methodology provided an appropriate framework for the development of scientific thinking skills amongst junior high school students.

- An unexpected gender difference, in which girls outperformed boys in scientific thinking, was found before learning the program, and afterwards.
- The type of teacher determined by factors such as, degree of openness to innovative teaching methods, enthusiasm and scientific background, was a crucial factor in students' ability to exploit the potential of "The Rock Cycle".
- In this study, teachers' management of classroom activities was affected more by teacher characteristics, than by the amount of guidance they received.
- It was concluded that heterogeneity of teachers in INSET courses should receive more attention, and that similar pedagogical concerns as those applied with heterogeneity of students, should be applied in teacher training.

Further research is needed in order to answer the following questions:

1. What type of teacher training could enable different types of teachers to learn and internalize innovative teaching approaches?
2. What are all the factors involved in the success of implementation of new learning programs, and what are the proportions between these factors?
3. What is the source of the discrepancy between the results of the current study, which indicate higher achievements of girls in scientific thinking skills, and former research concerning gender differences in scientific achievements?

REFERENCES

- American Association for the Advancement of Science, 1990, *Science for all Americans: Project 2061*, New York, Oxford University Press.
- American Association for the Advancement of Science, 1993, *Benchmarks for science literacy*, New York, Oxford University Press.
- Ault, C.R., 1998, Criteria of excellence for geological inquiry: The necessity of ambiguity, *Journal of Research in Science Teaching*, v. 35, p. 189-212.
- Bagno, E., and Eylon, B., 1997, From problem-solving to a knowledge structure: An example from the domain of electromagnetism, *The American Journal of Physics*, v. 65, p. 726-736.
- Bybee, R. W., 1997, *Achieving scientific literacy: From purpose to practices*, Portsmouth, NH, Heineman.
- Becker, B. J., 1989, Gender and Science Achievement: A re-analysis of studies from two meta-analyses, *Journal of Research in Science Teaching*, v. 26, p. 141-169.
- Bell, L. B., and Lederman, G. N., 2003, *Understanding of the nature of science and decision making on science and technology based issues*, *Science Education*, v. 87, p. 352-377.
- De Jong, T., and van Joolingen, W.R., 1998, Scientific discovery learning with computer simulations of conceptual domains, *Review of Educational Research*, v. 68, p. 179-201.
- Edelson, D.C., Gordin, D.N., and Pea R.D., 1999, Addressing the challenges of inquiry-based learning through technology and curriculum design, *The Journal of the Learning Sciences*, v. 8, p. 391-449.
- Elshout, J.J., and Veenmen, M.V., 1992, Relations between intellectual ability and working method as predictors of learning, *Journal of Educational Research*, v. 85, p. 134-143
- Erickson, G.L., and Erickson, L.J., 1984, Females and science achievement: Evidence, explanations, and implication, *Science Education*, v. 68, p. 63-89.
- Frodeman, R., 1995, *Geological reasoning: Geology as an interpretive and historical science*, *GSA Bulletin*, 107, 960-968.
- Germann, P.J., 1989, Directed-inquiry approach to learning science process skills: Treatment effects and aptitude-treatment interaction, *Journal of Research in Science Teaching*, v. 26, p. 237-250.
- Kali, Y., Orion, N., and Eylon, B., 2003, The development of systems thinking in the context of the rock-cycle: Analysis of the learning processes of one student working with peers on a knowledge integration activity, *Journal of Research in Science Teaching*, v. 40, p. 545-565.
- Karplus, R., 1979, Teaching for the development of reasoning. In A.E. Lawson, Ed. *The 1980 AETS yearbook: The psychology of teaching for thinking and creativity*, Columbus, OH: ERIC clearinghouse for Science, Mathematics, and Environmental Education.
- Krajcik, J. Blumenfeld, P.C., Marx, R.W., Kristin M.B., and Fredricks, J., 1998, Inquiry in project-based science classrooms: Initial attempts by middle school students, v. 7, p. 313-350.
- Lock, R., 1992, Gender and Practical Skill Performance in Science, *Journal of Research in Science Teaching*, v. 29, p. 227-241
- National Research Council (NRC), 1996, *National science education standarts*, Washington, DC, National academic Press.
- Perkins, D., 1992, *Smart schools*, New York, The Free Press.
- Shymansky, J.A., and Yore, L.D., 1980, A study of teaching strategies, student cognitive development, and cognitive style as they relate to student achievement in science, *Journal of Research in Science Teaching*, v. 5, p. 369-382.