2006

Investigating the impact of teachers' implementation practices on academic achievement in science during a long-term professional development program on the Science Writing Heuristic

Murat Gunel
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Science and Mathematics Education Commons

Recommended Citation
Gunel, Murat, "Investigating the impact of teachers' implementation practices on academic achievement in science during a long-term professional development program on the Science Writing Heuristic" (2006). Retrospective Theses and Dissertations. 1257.
https://lib.dr.iastate.edu/rtd/1257

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Investigating the impact of teachers’ implementation practices on academic achievement in science during a long-term professional development program on the Science Writing Heuristic

by

Murat Gunel

A dissertation submitted to the graduate faculty in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY

Major: Education

Program of Study Committee:
Brian M. Hand, Co-Major Professor
Joanne Olson, Co-Major Professor
Michael Clough
Mack Shelley
Frederick O. Lorenz

Iowa State University

Ames, Iowa

2006
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
Graduate College

Iowa State University

This is to certify that the doctoral dissertation of

Murat Gunel

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-Major Professor

Signature was redacted for privacy.

Co-Major Professor

Signature was redacted for privacy.

For the Major Program
DEDICATION

To my wife Ozlem Gunel, for her endless support, motivation and encouragement.

To my son Ege and my unborn daughter for their love and inspiration.
# TABLE OF CONTENTS

| LIST OF FIGURES | vii |
| LIST OF TABLES | viii |
| ABSTRACT | x |
| CHAPTER ONE | 1 |

**Introduction**  
*General Overview*  
*Background*  
*Inquiry and Impact of Teacher Practices on Students' Academic Achievements*  
*Writing-to-Learn in Science*  
*Purpose*  
*Research Questions*  
*Outline of the Dissertation*  
*Chapter 1: Introduction*  
*Chapter 2: Literature Review*  
*Chapter 3: Journal Article*  
*Purpose and research setting.*  
*Participating teachers.*  
*Research framework.*  
*Qualitative research design.*  
*Ranking mechanism.*  
*Quantitative research design.*  
*Variables.*  
*Chapter 4: General Conclusion, Implications, and Limitations*  

| CHAPTER TWO | 18 |

**Literature Review**  
*Cognitive Process of Writing*  
*Writing-to-Learn in Science & Science Literacy*  
*Constructivism and Learning Science*  
*Social Learning Theory*  

*
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danielle</td>
<td>83</td>
</tr>
<tr>
<td>Bill</td>
<td>83</td>
</tr>
<tr>
<td>Matt</td>
<td>84</td>
</tr>
<tr>
<td>Nick</td>
<td>84</td>
</tr>
<tr>
<td>Results</td>
<td>85</td>
</tr>
<tr>
<td>Results for John's Case</td>
<td>85</td>
</tr>
<tr>
<td>Results for Gary's Case</td>
<td>86</td>
</tr>
<tr>
<td>Results for Danielle's Case</td>
<td>88</td>
</tr>
<tr>
<td>Results for Matt's Case</td>
<td>89</td>
</tr>
<tr>
<td>Results for Bill's Case</td>
<td>90</td>
</tr>
<tr>
<td>Results for Nick's Case</td>
<td>92</td>
</tr>
<tr>
<td>Effect Size Results for all Teachers</td>
<td>94</td>
</tr>
<tr>
<td>Discussion</td>
<td>96</td>
</tr>
</tbody>
</table>

CHAPTER FOUR

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Conclusions, Implications, and Limitations</td>
<td>102</td>
</tr>
<tr>
<td>General Overview</td>
<td>102</td>
</tr>
<tr>
<td>Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>Literature Review</td>
<td>102</td>
</tr>
<tr>
<td>The Study</td>
<td>104</td>
</tr>
<tr>
<td>Implications</td>
<td>108</td>
</tr>
<tr>
<td>Implications for In-service Program</td>
<td>108</td>
</tr>
<tr>
<td>Implication for Science teachers</td>
<td>109</td>
</tr>
<tr>
<td>Limitations</td>
<td>110</td>
</tr>
</tbody>
</table>

REFERENCES CITED

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>112</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. The science writing heuristic: A template for students ........................................... 7
Figure 2. The science writing heuristic: A template for teachers ........................................... 8
Figure 3. The science writing heuristic, Part II: A template for student thinking .................... 30
Figure 4. The science writing heuristic, Part I: A template for teacher-designed activities to promote laboratory understanding ................................................................. 31
Figure 6. Model for Writing-to-Learn ......................................................................................... 32
Figure 7. The relationship between professional development and student achievement (Supovitz & Turner, 2000, p. 965) ................................................................. 51
Figure 8. The relationship between professional development and student achievement (Supovitz & Turner, 2000, p. 965) ................................................................. 64
Figure 9. The science writing heuristic, Part II: A template for student thinking .................... 69
Figure 10. The science writing heuristic, Part I: A template for teacher-designed activities to promote laboratory understanding ................................................................. 69
LIST OF TABLES

Table 1. Four Faces of Constructivism (Yore, 2001) ................................................................. 35
Table 2. Outline of Levels of Implementation Characteristics ......................................................... 79
Table 3. Participating teachers and student information ................................................................. 82
Table 4. Descriptive statistics for ITED Science Scores for John .................................................. 86
Table 5. Descriptive statistics for ITED Science Scores for Gary .................................................. 87
Table 6. Descriptive statistics for ITED Science Scores for Danielle ............................................. 89
Table 7. Descriptive statistics for ITED Science Scores for Matt ................................................. 90
Table 8. Descriptive statistics for ITED Science Scores for Bill .................................................. 91
Table 9. Descriptive statistics for ITED Science Scores for Nick .................................................. 92
Table 10. Pairwise Comparisons for Nick’s Students ..................................................................... 94
Table 11. Summary of Changes of Teacher Implementation and Effects of Such Changes on Student Scores .............................................................................................................. 95
ACKNOWLEDGMENTS

I would like to extend my appreciation and sincere thanks to my major professor, Dr. Brian M. Hand, for his support, encouragement, advice, modeling, and guidance throughout my graduate education and dissertation work. Special thanks to all the other members in my committee, Dr. Joanne Olson, Dr. Mack Shelley, Dr. Michael Clough, and Dr. Frederick O. Lorenz for their constructive feedback on the dissertation.

I also wish to thank my colleagues and lifelong friends Dr. Irene Grimberg, Dr. Sozan Omar, and Recai Akkus for their dedicated work and support. My great thanks and appreciation are for all the science teachers and their students who participated in the study for without their participation this study would not existed. Finally, I would like to thank the Turkish Ministry of Education for sponsoring my graduate education.
ABSTRACT

With the reform movement in teaching science, the National Science Education Standards has pushed for more emphasis on science inquiry and for students to be involved in activities that promote critical thinking, reasoning, and argumentation skills. According to the National Science Education Standards, science educators have been calling for the insertion of inquiry-based approaches in science classrooms as a change for science instruction. To achieve such a reform vision, researchers in science education, higher education, and adult learning and literacy argued for the need for well-articulated, organized, implemented, and evaluated in-service professional development programs. Further, it has been argued that programs also need to be longitudinal to assist science teachers in refining both their beliefs and practices. Finally, an emphasis has been placed on programs that are supported and evaluated by a combination of teacher reflection, classroom observation, and ongoing student assessments.

This study is part of a bigger project known as the Science Writing Heuristic (SWH) Partnership Professional Development Project, conducted at Iowa State University and the University of Iowa in association with the Iowa Department of Education to help improve science teaching. Overall, the goal of the project is to help in-service science teachers understand and apply a student-oriented instructional method, using the SWH approach. The purpose of this research study was to examine the link between teachers’ implementation of a student-oriented teaching method through the SWH approach with embedded non-traditional writing practices and students’ performances on standardized tests over a 3-year period. This study investigated the impact of 6 teachers’ (3 high school teachers and 3 middle school teachers) implementation of the SWH approach on student standardized test scores over the
3-year period. The study was an extension of a previous study. In an earlier study, Omar (2004) studied the change in pedagogical practices of eighteen teachers who implemented SWH approach. In her work she focused on the changes in teachers’ beliefs and attitudes toward learning and teaching and the changes in their pedagogical practices when implementing the SWH approach over a two-year period.

This study was an extension of the original study by Omar; however, a subset of six teachers only was used in this reported study. The major difference between the two studies is that this study extended the observation period for an additional year and measured how students’ performances on standardized tests were affected by various implementation levels of the SWH approach by individual teachers over the 3-year long period where possible students’ scores for each teacher from the year prior to first year of the project were obtained and used as a baseline to analyze progress of each teacher.

There are several emerging themes in the results of this study. The first thesis is that, in general, teachers’ change and the rate of change differ by teachers, and an individual-level change is not a linear process for teachers. Results of the study indicated different cases where teachers improved their pedagogical skills related to the SWH approach and were rated as moving to higher implementation levels over the 3-year period. Another emerging theme from this work is that the SWH in-service program did have an impact on participating teachers’ pedagogical practices. The majority of the participating teachers improved their pedagogical practices of implementing science inquiry through the SWH approach over a 3-year professional development program. The structure and content of the program appear effective in achieving teacher change. Further, when teachers’ rankings were correlated against students’ standardized test scores, the results indicated that as teachers’ SWH
implementation levels increased, their students’ test achievements also increased. The study indicates the importance of assisting science teachers for implementing the SWH science inquiry approach with effective pedagogical strategies as a means to promote a significant impact on students’ academic achievement.
CHAPTER ONE

Introduction

General Overview

This chapter presents a brief overview of the study included in the dissertation, provide a context for the study, and identify the major research questions. The dissertation contains a chapter of extended literature review, one research journal article, and a general conclusion chapter. The paper investigates teachers' change process and the impact of such changes on students' academic achievement case by case. The research paper is a quasi-experimental, mixed-method study providing some empirical evidence of the impact of teachers and writing on learning. The final chapter of the conclusion, which presents the study's limitations and implications, discusses results from the research paper included in this set and attempts to theorize further about the impact of teachers on students' learning.
Background

Inquiry and Impact of Teacher Practices on Students’ Academic Achievements

With the reform movement in teaching science, the National Science Education Standards (NRC, 1996) have pushed for much more emphasis on science inquiry and for students to be involved in activities that promote critical thinking, reasoning and argumentation skills. According to the National Science Education Standards, science educators have been calling for the insertion of inquiry-based approaches in science classrooms as a change for science instruction. Thus, science educators were called to use a student-oriented setting along with an inquiry approach in which students inquire about scientific phenomena by working collaboratively with their peers as they engage in science investigations. Luft (2001) argues that well-articulated, organized, implemented, and evaluated in-service professional development programs are crucial to archive this reform-based vision. Further, such professional development programs not only need to target teachers’ efficacy belief toward science teaching (Eshach, 2003), professional knowledge (Tillema, 1995), and pedagogical content knowledge (van Driel, Verloop, & de Vos, 1998), but the programs also need to be longitudinal (Luft, 2001) by assisting science teachers in refining both their beliefs and practices. In addition, the programs need to be supported and evaluated by a combination of teacher reflection, classroom observation, and ongoing student assessments (Fishman, Marx, Best, & Tal, 2003).

In educational research, the impact of professional development programs, and, consequently, teachers’ practices and epistemological beliefs toward teaching and learning, are investigated by using qualitative methodology (e.g., observations, interviews); on the
other hand, the other important outcome, students’ academic achievement, provides strong criteria to measure the accomplishments of in-service programs and the effectiveness of teachers (Chinn & Malhotra, 2002; Songer, Lee, & McDonald, 2003). In particular, students’ performances on standardized tests becomes crucial to evaluating educational settings and the impact of professional development programs by considering teachers’ classroom practices (Wright, Horn, & Sanders, 1997)

However, literature in educational research is limited in terms of such research findings that focus teachers’ change in beliefs and practices over time through professional development programs and the impact of such changes on students’ standardized test scores. The need is for empirical studies in which in-service programs help teachers implement effective teaching approaches, not target only additive skills. Such in-service programs should promote fundamental changes of beliefs, knowledge, and habits of practice to accomplish national standards for teaching science. Teacher’s classroom practices should be evaluated by multiple approaches, including the relationships among students’ academic achievement and teacher change.

*Writing-to-Learn in Science*

Students’ ability to speak, read, and write about science and their ability to unify concepts of science, the nature of science, and the relationship among science, technology, society, and the environment, is the target for both contemporary and interactive constructivist science teaching and learning (Yore, Bisanz, & Hand, 2003). The view that individuals create meaning and knowledge by interacting with their environment, and by reflecting on and making sense of these interactions, is the accepted interactive constructivist
position that forms the basis for the writing-to-learn science movement (Hand & Prain, 2002). Studies of writing-to-learn strategies generally involve using different writing tasks within investigative science to prompt construction of knowledge in active-learning environments such as inquiry, whereby students construct personal meanings within the classroom community (Rivard & Straw, 2000).

Teachers who provide opportunities for students to articulate, defend, and explain their own ideas within the social context of classroom change the classroom environment from being teacher-oriented and text-dominated to being more student-oriented, in which language is used by students (Prain & Hand, 1996). The importance of the language, especially written language for science learning, has been emphasized and discussed by many scholars (Halliday & Martin, 1993; Hand, Hohenshell, & Prain, 2002; Keys, 1999b, 2000; Keys, Hand, Prain, & Collins, 1999; Prain & Hand, 1996; Sutton, 1992; Yore et al., 2003). Various cognitive writing models such as Bereiter and Scardamalia (1987), Hayes and Flower (1980) and Torrance and Galbraith (1999) are the basis for research studies in writing-to-learn in science. While, researchers in cognitive psychology have analyzed the process of writing from different perspectives, the consensus among them is that writing can serve as a learning tool (Alamargot & Chanquoy, 2001; Nelson, 2001).

Scholars have adopted different positions regarding the value of using writing-to-learn strategies in helping students understand science. The main conflict between the positions lies in the purpose of using writing in science. Halliday and Martin (1993) have argued that the implementation of traditional writing in science is necessary because students need to use proper technical scientific language and types of genre to learn science. For Halliday and Martin, one has to learn micro and macro structures of the genres of science
writing to be able to understand science. However, Prain and Hand (1996) proposed that students should be encouraged to write their understanding of science concept in a variety of ways using their own language. By adopting student-oriented views of learning, Keys et al. (1999) have emphasized that students need experiences with a variety of writing genres to communicate ideas. Consequently, students would construct their own science conception, through interacting with other students, materials, and the teacher in the classroom context under the teacher's guidance.

Comparing and contrasting others' writing, rewording others' ideas through their own words, and speculating about possible explanations, provides students opportunities to sort out what they understand (Prain & Hand, 1996). Meaningful science learning needs to have similarities with the methods used by scientists in practice (Yore et al., 2003). Keys (1999) argued that connections do exists between the conventions of English used by scientists over several hundred years and students' everyday language, which has more personal meaningfulness for students. She quotes Lemke (1994) in pointing out this crucial connection:

I think the most important issue here is to understand why science registers show the grammatical and other linguistic peculiarities that they do [and] what specific functions they serve. I suspect that it is when learners see a need to perform these functions, when the functions make sense to them (e.g., classification), that they will be able to adopt the linguistic means of doing so that has evolved historically in modern European culture.

The issue raised by Lemke, using writing as a tool for meaningful science learning, brings about some critical questions, such as: How can one learn to read, write, and speak the
language of science? What are the protocols and expectations for reporting one's research? How can we communicate our ideas effectively to experts, scientists in other fields, and laypeople? How can we promote effective science learning by using investigative science methods and by using writing-to-learn or by using any possible combination of investigative science method and writing-for-learning science?

Keys et al. (1999) argued that in addition to formal science writing genres, informal writing genres, such as journals writing, question reflection, cartoons, and narratives, writing laboratory reports also need to be considered as effective science learning tools. The authors proposed the Science Writing Heuristic (SWH) as a tool to guide teachers and students. The SWH approach has two templates for each audience. In the teacher template, the teacher uses a series of writing, reading, and small- and large-group discussion activities to support students in meaningful thinking. Thus, the teacher template illustrates the necessary pedagogy to support student learning. In the student template, students are encouraged to investigate their own question(s) about an activity and use scientific methods during investigations; however, students are encouraged to use their own language to share their findings. Figure 1 and 2 shows student and teacher templates of the SWH approach.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1-</td>
<td>Beginning ideas – What are my questions?</td>
</tr>
<tr>
<td>2-</td>
<td>Tests – What did I do?</td>
</tr>
<tr>
<td>3-</td>
<td>Observations – What did I see?</td>
</tr>
<tr>
<td>4-</td>
<td>Claims – What can I claim?</td>
</tr>
<tr>
<td>5-</td>
<td>Evidences – How do I know? Why I am making these claims?</td>
</tr>
<tr>
<td>6-</td>
<td>Reading – How do my ideas compare with others?</td>
</tr>
<tr>
<td>7-</td>
<td>Reflection – How have my ideas changes?</td>
</tr>
</tbody>
</table>

Figure 1. The science writing heuristic: A template for students.
1. Exploration of pre-instruction understanding through individual or group concept mapping.
2. Pre-laboratory activities, including informal writing, making observations, brainstorming, and posing questions.
3. Participation in laboratory activity.
4. Negotiation phase I- writing personal meanings for laboratory activity (For example, writing journals).
5. Negotiation phase II- sharing and comparing data interpretation in small group (For example making a group chart).
6. Negotiation phase III- comparing science ideas to textbooks or other printed recourses (For example, writing group notes in response to focus questions).
7. Negotiation phase IV- individual reflection on writing (For example, creating a presentation such as a poster report for larger audience).
8. Exploration of post instruction understanding through concept mapping.

Figure 2. The science writing heuristic: A template for teachers.

In practice, a teacher who implements the SWH approach starts the unit by using a demonstration activity, a short reading, or some sort of visual representation followed by a set of question or cues to open a large group/small group discussion depending upon the purpose. Student ideas are generated and noted in some visual form such as a concept map or list which can then be referred to throughout the unit. The teacher encourages students to generate questions about the topic or phenomenon under investigation. Students' questions regardless of the relevance level are noted on the board. During the large group discussion students are asked to think about the significance of these questions to the topic and whether or not they are testable questions. Such discussions are held within small and/or large groups depending upon the progress of the students and the purpose.

A similar structure was followed for identifying testing procedures for the research questions. Students are encouraged to make connections between research questions that they have generated and the materials available to test those questions. During the testing phase
working with small groups, teacher makes effort to identify, challenge and connect students' ideas, research questions and activity. After testing and data collection students are asked to make claims and evidence based upon the perceived pattern of data and observation. These claims and evidence are then publicly articulated and defended in front of the whole class with significant amount of discussion time being provided. At the end of this stage, students are provided opportunities to modify their research questions, test procedures, claims and evidence if they believe it to be necessary. Students are then required use textbooks or other sources to compare their findings. Toward end of the SWH cycle, students are asked to reflect how their ideas about the unit have been changed throughout this activity. While students conduct their research question, test procedure, claim, evidence, expert reaction and reflection within the outline of the SWH student template, at the end of the unit they are asked to write a non-traditional piece to communicate what they learned about the unit. Such non-traditional writing tasks can include, constructing a brochure, article to local newspaper, letter to parents, grandparents, peers or younger grades, a poster, poem, or song.

**Purpose**

This dissertation has two purposes. First, I briefly examined the pattern of teacher change with the SWH approach implementation over the 3-year in-service program. The second purpose is to investigate the impact of the teacher change and SWH approach implementation level on panel date of students' academic achievement. Through the investigation, I generated cases for each participating teacher about their changes in classroom practices prompted by a longitudinal professional development program they participated in. In addition, academic achievement of those students who took place in
participating teachers’ classes were evaluated to investigate the relationship between change of teacher practices and its impact on students’ achievement.

*Research Questions*

Two research questions were pursued with this study. Investigation of each research questions were presented in a journal article format. Those research questions are as follows:

- How have participating teachers’ pedagogical practices changed over the 3 years of implementing the SWH approach?
- How have standardized test scores of those students in the teacher’s classes been impacted by such changes?

*Outline of the Dissertation*

*Chapter 1: Introduction*

The research proposal was used as an introduction chapter.

*Chapter 2: Literature Review*

In the literature review chapter relevant literature that guided this particular research study was examined as outlined below:

I- Cognitive Process of Writing
   a. Writing Models
      - Hayes and Flower (1980)
      - Bereiter and Scardamalia (1987)
      - Galbraith (1999)

II- Writing-to-Learn in Science and Science Literacy
   a. The connection between the process of writing and using writing as a learning tool
b. Writing as a learning tool in science education

c. Science literacy and science education

d. The role of language in science literacy

e. The role of writing in science literacy

III- Constructivism

a. Constructivism as a learning theory

b. Implications of constructivist learning theory in the learning environment

c. Constructivist learning theory and science education

IV- Social Learning Theory

a. Tenants of the social learning theory

b. Role and applications of the social learning theory in education

V- Inquiry in Science Education

a. Inquiry in science and education

b. Meaning of inquiry in science education

c. The teacher and students in inquiry-based classrooms

d. Needs and calls for inquiry

e. Current stage of science education and science inquiry in the U.S.

VI- Professional Development for Science Teachers

a. Relationship between professional development for science teachers and their use of inquiry-based teaching

b. Overview of approaches on professional development programs for teacher change

c. Impact of in-service programs on teachers' practices

d. Assessment of teacher inquiry implementation

VII- Impact of the Teacher on Students’ Academic Achievements

a. Overview of empirical finding on the effect of teachers on students’ academic achievement.

b. The impact of the teacher change over time on students’ performance of test scores.
Chapter 3: Journal Article

The chapter attempts to examine the link between change in teachers’ implementation of the SWH approach and students’ performance on a standardized test, the Iowa Test of Educational Development (ITED), on a yearly basis over 3 years. The data were drawn from a larger-scale professional development project. As indicated above, within the same research project other reported works (Gunel, Akkus, Hohenshell, & Hand, 2004; Gunel, Omar, & Akkus, 2003; Maroney, Finson, Beaver, & Jensen, 2003; Omar, 2004; Omar & Gunel, 2004) investigated the levels of teachers’ pedagogical implementation of strategies recommended by the SWH approach and of an interactive constructivist teaching/learning approach. Thus, this particular study examined how the particular teacher’s classroom practices are changed during the 3-year project and how students’ test scores are influenced by such changes.

Purpose and research setting.

The data used for this study were part of the SWH partnership project. This particular research focused on a group of 6 science teachers and examined their change within the professional development program. Further, this study investigated the effect of each teacher’s change in pedagogical practices on students’ academic achievement. A total of 6 science teachers who taught general science, biology, chemistry, and physical sciences subjects to 1988 students across grade 7-10 were selected from one school district that participated in this 3-year study. Data for student participants were collected from ITED. Although detailed teacher observations, interviews, debriefing, and questioners were collected to identify the level of implementations as part of the project, findings from those
qualitative data analyses were reported in previous works, as mentioned above. Thus, in this report, criteria for ranking teachers' implementation levels have been outlined briefly.

**Participating teachers.**

Six science teachers were selected out of 18 participants in the large SWH project due to availability of their students' longitudinal test scores over the 3-year period of this study. Four teachers out of 6 participating teachers implemented their SWH unit(s) during the 2003-2005 academic school years, and 2 of the teachers implemented SWH units during the 2003 and 2004 academic school years. The teachers involved teach several science disciplines, including general science, chemistry, physical science, and biology, to a wide range of grade levels in middle school and secondary school. Each teacher holds a different educational degree with different experience in traditional teaching. The population included 1 female and 5 male teachers. All teachers attended the 3-day professional development workshop that initiated the study.

**Research framework.**

In this study, while a quantitative research approach was used predominantly, qualitative data analysis results and some interpretive case study findings from previous works were used in the quantitative data analyses. By using a qualitative approach, individual implementation profiles were constructed for each teacher over the 3-year period of the study. For each teacher, his or her students' performance on the ITED for each year across the 3-year period of the study was compared. The ITED test scores from the year prior to the study constituted a baseline. Finally, individual teacher cases were evaluated by comparing students' standard test scores, to explore trends across the 6 teachers.
Qualitative research design.

Since the main purposes of this paper were to construct a holistic understanding of a particular group of teachers' implementation of the SWH approach, and to measure the impact of those implementations on students' standardized test scores within an in-service professional development program, an interpretative case study was part of the research design. The same methodology was adopted by Omar (2004) to illustrate teachers' implementation levels of the SWH approach. Yet, her main focus was teachers' implementation of the SWH approach. While this study used her findings of the SWH approach implementation levels for each teacher to extend this study for a third year, the same instruments were used to identify the level of implementation for each teacher who participated.

In this study, observations and video tapes were used to identify teachers' implementation level. 3 researchers and one independent observer used a criteria matrix to analyze each teachers' implementation level of the SWH approach. The author, two other researchers and the independent observer were participated in Omar's study as an observer. Prior and during Omar's study necessary trainings and discussions were held in order to achieve internal consistency for ranking. Throughout this study, weekly meetings among the researchers and monthly meetings which included the external observer took place. The purpose of these meetings were sharing active context of each lesson and teacher's progress were maintained and shared.
Ranking mechanism.

Each teacher was ranked based on his/her level of implementation. Adopted from Omar (2004), the ranking mechanism consisted of three composite rankings. The first ranking was a 10-point scale specifically targeting the first criterion—promoting dialogical interaction—with 1 indicating low interaction and 10 indicating high interaction. The promotion of dialogical interaction and student argument is considered a major demand of the SWH approach and reflects a shift in instructional practice.

The second and third scales target the second and third criteria, respectively—focus of learning and connection. The second scale has a 5-point scale, with 1 point reflecting the first level, 3 points reflecting the second level, and 5 points reflecting the third level, while 2 and 4 points indicate transition levels. The 5-point focus of learning scale reflects the degree to which the teacher is in focus of learning in his or her classroom, with the lowest level implying knowledge transfer. Further, the 10-point connection scale reflects teachers’ success in preparing for the unit(s) from a constructivist teaching perspective (Simon, 1995).

Quantitative research design.

The quantitative approach in this study was used to answer the research question of how the academic achievements in terms of ITED of those students taught by participating teachers change over four years, with standardized test scores for the 2001-2002 academic year used as a baseline. To answer this question it is necessary to compare test scores for students those teachers taught in different years for each teacher. Therefore, files for individual teachers were created with their students’ scores over the 4-year period to be able to compare scores year by year.
To ensure the accuracy of the data collected, frequency distributions and descriptive statistics were obtained using the SPSS Frequencies procedure (Mertler & Vannatta, 2002). The SPSS Explore procedure was employed to examine whether outliers possibly could affect the results of the study. To control for other variables that might impact students’ science achievement, using covariates in the statistical analysis is crucial. Therefore, analysis of covariance (ANCOVA) was chosen as a statistical method where students’ ITED science scores are the dependent variable, the year those students participated in a particular teacher’s class is the independent variable, and those students’ ITED social sciences, mathematics, and language scores are the covariates.

Variables.

2001-2002 ITED Total Score: Total score for each student obtained prior to the study were used to account for students’ knowledge prior to the study.

2002-2003, 2003-2004, and 2004-2005 ITED Science Total Scores: Total science scores were obtained throughout the project to investigate the effects of teachers’ implementations. One variable was created for each year.: 02-03 ITED Science, 03-04 ITED Science, and 04-05 ITED Science.

Year the ITED Taken/Teacher’s Implementations of Pedagogy for 2002-2003, 2003-2004, and 2004-2005 Years: Each year teachers were ranked according to criteria given above to investigate the effects of pedagogy on pupils’ standardized science test scores.

Chapter 4: General Conclusion, Implications, and Limitations

The last chapter of the dissertation presents a general conclusion, implications, and limitations based on the literature review. Emerging themes of the literature on science
inquiry, professional development, and implementation quality on students’ learning outcomes were outlined. Second, findings of the study, their implications to practice, and their relation to current literature were discussed. Then, surrounding limitations of the study were delineated.
CHAPTER TWO

Literature Review

This literature review examines the relevant literature to provide a theoretical framework for the research questions. This chapter is organized in segments. In the first part of the review the theoretical underpinnings of the SWH approach are discussed. Writing-to-learn, science literacy, constructivism, and social learning theory are the areas of interest examined in detail. The second part of the review focused on reforms in science teaching with regard to science inquiry approaches. Further, in this section, the needs and calls for changing the way of teaching from traditional didactic approaches to inquiry-based student-oriented approaches is discussed. Later, structures of effective in-service programs and characteristics of teacher change are the topics examined. Finally, attention is drawn to relevant literature on teacher change through in-service programs and the effect of such changes on students' achievements.
Cognitive Process of Writing

Writing is a sophisticated task that requires the application of several cognitive processes. The complexity of writing is associated with the nature of the task, the writer's goal, and the syntactic and grammatical rules required by the tasks (Alamargot & Chanquoy, 2001). Added to these demands is the need for students to relate with their audience and to formulate and present relevant ideas. All of these factors combine through the written word to produce a satisfactory text. The writer generally must make a number of drafts, corrections, deletions, and additions to the text after considering writing cues such as goals, state of writing, and topic to produce the final product. Such translation of thoughts into a written format involves a complex cognitive mechanism, and, as a consequence, diverse models to explain the process have been proposed in both the psychology and the linguistic disciplines (Bereiter & Scardamalia, 1987; Galbraith, 1999; Hayes & Flower, 1980, 1987; Klein, 1999).

The different models of written composition, called blueprints by Alamargot and Chanquoy (2001), allow researchers to focus on different dimensions of writing while retaining consideration for the complexity of whole cognitive action. The common feature of these blueprints is that they tend to propose an analytic definition, or map, of the writing process in which the processes and sub-processes, and definitions and arrangements of them, are articulated by providing rules, constraints, and limitations. Even though numerous experimental research studies guided by different models have been conducted, there has not been a clear formulation of the complex cognitive mechanism of writing (Alamargot & Chanquoy, 2001; Klein, 1999).
Hayes and Flower (1980) attempted to explain the writing process by adopting a cognitive approach in which writing is perceived as a set of mental processes that are not only rigidly graded but also are embedded in nature (Hand, 2004). Furthermore, in their model, writing is perceived as a confirmed and multifaceted activity since it requires writers to process and implement an adequate amount of knowledge through several mental activities (Alamargot & Chanquoy, 2001). In his extensive review of writing models Hand (2004) identified four foundation stones necessary to understanding Hayes and Flower's model:

1. The process of writing is best understood as a set of distinctive thinking processes that writers orchestrate or organize during the act of composing.

2. These processes have a hierarchical, highly embedded organization in which any given process can be embedded within any other.

3. The act of composing itself is a goal-directed thinking process, guided by the writer's own growing network of goals.

4. Writers create their own goals in two key ways: by generating both high-level goals and supporting sub-goals which embody the writer’s developing sense of purpose, and then, at times, by changing major goals or even establishing entirely new ones based on what has been learned in the act of writing (p. 366).

The model has three main components: task environment, the writer’s long-term memory, and the general writing process. Hayes and Flower have defined each of these as: (a) task environment; where all outside factors, such as topic, audience, motivation, and text produced so far, can influence writing and must be considered; (b) the writer’s long-term memory, in which the writer’s conceptual, pragmatic, linguistic, and lexical knowledge are stored and can be used to guide and complete the task; and (c) the general writing process,
which refers to the process of translating writer’s knowledge of the topic into linguistic form, which consists of four essential parts—planning, translating, reviewing, and monitoring.

The planning process includes generating ideas from the writer’s long-term memory, organizing the ideas according to the given task, and creating goals to evaluate the written text. Translation comprises turning retrieved ideas from long-term memory into a text format, influenced by goals already put in place. Reviewing scrutinizes the quality of the text by reading and editing. The important feature of these processes is that they are not linear cognitive operations that occur in an ordered fashion; instead, they are recursive and concurrent. That is, at any given moment planning, translating, and revisiting can occur in a different order during the writing process (Tynjala, Mason, & Lonka, 2001).

The model Hayes and Flower proposed attempted to incorporate the information, knowledge, and cognitive processing required for writing. Writing was not considered as a product-based, linear activity; instead, writing was referred to as a cognitive process based on a means to monitor the control of planning, writing, and editing at any moment during writing (Galbraith, 1999). In other words, writing was perceived as a goal-driven activity wherein, to accomplish writing task, a writer needs to balance components of the task environment with his or her content knowledge in the environment to be presented, and, thus, the overall result of interactions among those environments operates as a learning tool (Hand, 2004). Finally, the model proposed is reviewed as a problem-solving metaphor, whereby writing is controlled by the writer’s general problem-solving skills, rhetorical knowledge, and content knowledge (Galbraith, 1999; Hayes & Flower, 1980).

While Hayes and Flower’s model was not adequate enough to explain the particular procedures, functions, and relationship among sub-procedures or the differences between
expert and novice writers, their model was the first attempt at creating a framework of the writing process (Alamargot & Chanquoy, 2001). Extending Hayes and Flower’s model, Bereiter and Scardamalia (1987) proposed a more defined version of the problem-solving process in writing. What is more, they proposed two types of writing: “knowledge telling—psychology of the nature” and “knowledge transforming—psychology of problematic,” recognizing the second type of writing as a learning mechanism (p. 5).

Writing as a knowledge-telling process involves the use of naturally gained (through social interaction) language capability and skills, whereas knowledge transformation through writing encompasses studied ability and skills. By studied skills Bereiter and Scardamalia (1987) meant that, unlike a more naturally developed ability such as casual reading and talking, studied skills involve conscious, strategic control over parts of process such as in critical reading and oratory. The knowledge telling process is employed by mostly children and novice writers and the process entails putting down everything that writer knows about the topic, what Hayes and Flower (1980) called “Get it down as you think” (p. 20). Bereiter and Scardamalia argued that when an audience is not present in the conversation as is the case in writing, children and novice writers encounter the problem of what to say and how to say it during the text production process. While they are constrained by the absence of immediate feedback, they need to use a limited number of cues such as content retrieval from long-term memory, topic, discourse schema, and the text already produced. In other words, “the knowledge-telling model is a model of how discourse productions can go on, using only these sources of cues” and the quality of text produced in this fashion is affected largely by the writers’ content knowledge and level of activation in their long-term memory (p. 7).
There are three essential components in the knowledge-telling model. The first component involves mental representation of the assignment, whereby the writer creates the mental representation of the assignment that allows him or her to define the text topic and function. Also, mental representation of the assignment guides the whole rhetorical writing activity. The second component refers to the two types of knowledge stored in the long-term memory, content knowledge and discourse knowledge, both of which need to be articulated. The content knowledge refers to the topic and is what the writer knows related to the topic, and the discourse knowledge is the nature of the task, such as the linguistic, lexical, grammatical, narrative, or argumentative structures that are necessary for producing a written text (Bereiter & Scardamalia, 1987).

The third component in the knowledge-telling model is the process of knowledge telling, which has a close relationship with the other two components (mental representation of the assignment and content and rhetorical knowledge). Knowledge telling presents writing as a display resulting from pre-existing knowledge about the assignment and knowledge of the topic stored in long-term memory, and it is not conceptualized as a process that can modify either the assignment or content knowledge during the act of writing. In summary, knowledge telling is essentially a “think-say” method, in that ideas are retrieved directly from memory in response to a topic and then translated into text. The condition of the succession of the text is attributed to ideas stored in the memory, similar to the Hayes and Flower model. On the other hand, the knowledge-transforming model conceptualizes writing as an act that stimulates thinking; thus, mental representation of the assignment and knowledge of the topic can be modified through the act of writing.
The knowledge-transforming model suggests that writing at the expert level can be a complex problem-solving activity, although it first requires the writer to move through a knowledge-telling phase comparable to that required for knowledge-transforming activities in Hayes and Flower’s model. The additional components required for knowledge to be modified involve problem-analyzing and goal-setting activities, which allow the writer to understand the task. The knowledge-transforming model also involves an interaction between the content problem space and rhetorical problem space that allows the writer to explore the content setting and rhetorical setting, that is, respectively, what to tell and to whom and how to tell it. Also, the model proposed by Bereiter and Scardamalia (1987) suggests that there is a possible interaction among content knowledge, the content problem space, discourse knowledge, and the rhetorical problem space that can transform and modify the writer’s thought. More specifically, there is a dynamic relationship between where content is stored, thought about, and worked out, and the rhetorical space where goals for the text are worked out, which provides the stimulus for reflection in writing and problem solving (Keys, 1999a).

In the knowledge-transforming model, it is assumed that the retrieval and translation of ideas is mediated by active problem solving. However, writing not only is taking information from memory or translating this information. Writing also involves working out new content when existing content does not satisfy goals, and, when realized, writing allows further development of the writer’s understandings of the topic (Galbraith, 1999). The importance of writers understanding content may be realized in two ways. First, when already-existing knowledge is not enough to complete the task, the writer has to develop new understandings. Second, the new content, born as a result of dissatisfaction with existing
content, is now a new construction of knowledge rather than simply a reflection of old knowledge (Galbraith, 1999).

In brief, while this model recognizes that not all writing can lead to learning, as Langer and Applebee’s (1987) found in their empirical study, the model proposes an important structure to illuminate the writing process and pathways to recognize writing as a learning tool. Yet, the way Bereiter and Scardamalia captured the thinking and learning mechanism during the process of writing was criticized by other researchers as being simplistic and unable to capture the unpredictability of thinking (Hand, 2004).

Rather than viewing writing using a problem-solving metaphor as in the knowledge transforming model, cognitive psychologists Galbraith and Torrance (1999) have suggested that writing should be viewed as a text-production model called a romantic position. They have proposed a knowledge constitution model to analyze the process of text production (Alamargot & Andriessen, 2002). The fundamental conflict between the knowledge-transforming or problem-solving model and the knowledge constitution model is that the former accepts writing as text production. In other words, the problem-solving model assumes knowledge is stored in a uniform way so no difference exists between the searching and retrieval processes, which are used during problem solving and text production. In contrast, the knowledge constitution model asserts that the knowledge encoded in sentences is represented within a distributed network of conceptual relationships.

Moreover, in the knowledge constitution model ideas are synthesized by constraint satisfactions within this network, rather than simply being retrieved as is (Galbraith, 1999). The overall synthesis of an idea to satisfaction is affected by two factors: first, the constraint satisfaction within the disposition, which is responsible for the formation of the message,
and, second, the constraint satisfaction within the linguistic network, which is responsible for the expression of the message in words.

The writer's conceptual knowledge is embodied in the connections between the units within their disposition (content knowledge base) and cannot be accessed directly. Instead, to make their understanding explicit, writers have to articulate their dispositional response to the topic, but this cannot happen in a single utterance. To capture the understanding as a whole, the writer must continue to synthesize the dispositional response as it unfolds. This means that the writer must constitute thought, obliquely and unpredictably, over a series of cycles. At the end of each cycle, a product, the written text, which may consist of only a phrase, or a few words or sentences, allows writer to feed back through the system again where different activations take place to produce further writing (Hand, 2004). Through each cycle less activation occurs within the writer's content knowledge base. The resulting dialect is a stable and satisfied product that requires no more activation. This final product is different than what was originally stored in the writer's content knowledge base since throughout the production process some meanings are omitted and some are added according to constraint satisfactions within the disposition linguistic network.

Writing-to-Learn in Science & Science Literacy

Being able to speak, read, and write about science and to unify concepts of science, the nature of science, and the relationship among science, technology, society, and environment are targets for both contemporary and interactive constructivist science teaching and learning (NRC, 1996; Project 2061 [American Association for the Advancement of Science], 1993; Yore et al., 2003). The view that individuals create meaning and knowledge
by interacting in their environment and by reflecting on and making sense of these interactions is the accepted interactive constructivist position that forms the basis for the writing-to-learn science movement (Hand & Prain, 2002). Studies of writing-to-learn generally involve using different writing tasks within investigative science to prompt construction of knowledge in active learning environments, whereby students construct personal meanings within the classroom community (Rivard & Straw, 2000).

Teachers who provide opportunities for students to articulate, defend, and explain their own ideas within the social context of the classroom change the classroom environment from a teacher-oriented and text-dominated environment to a more student-oriented environment, in which language is used by students (Hand & Prain, 1996). The importance of the language, especially written language for science learning, has been emphasized and discussed by many scholars (Gunel, 2004; Halliday & Martin, 1993; Hand et al., 2002; Hand & Prain, 1996; Keys, 1999b, 2000; Sutton, 1992; Yore et al., 2003). Various cognitive writing models (Bereiter & Scardamalia, 1987; Galbraith, 1999; Hayes & Flower, 1980), which have been described earlier in this literature review, are the basis for research studies in writing-to-learn in science.

Different positions have been adopted by scholars regarding the value of using writing-to-learn strategies in helping students understand science. The main conflict between the positions lies in the purpose of using writing in science. Halliday and Martin (1993) have argued that the implementation of traditional writing in science is necessary because students need to use proper technical scientific language and types of genre to learn science. For Halliday and Martin, one has to learn the micro and macro structures of the genres of science writing to be able to understand science. However, Prain and Hand (1996) proposed that
students should be encouraged to write their understanding of science concepts in a variety of ways using their own language. By adopting student-oriented views of learning, Keys, Hand, Prain, and Collins (1999) have emphasized that students need experiences with a variety of writing genres to communicate ideas. Consequently, students construct their own science conceptions, through interacting with other students, materials, and the teacher in the classroom context under the teacher's guidance.

Comparing and contrasting others' writing, rewording others' ideas in their own words, and speculating about possible explanations, provides students opportunities to sort out what they understand (Prain & Hand, 1996). Moreover, meaningful science learning has similarities with the methods used by scientists in practice (Yore, Hand, & Prain, 2002). Keys (1999b) argued that there is a important connection between language conventions in science that scientists use and the everyday language of students that is personally meaningful. She used quotes from Lemke (1994) to point out this crucial connection:

I think the most important issue here is to understand why science registers show the grammatical and other linguistic peculiarities that they do [and] what specific functions they serve. I suspect that it is when learners see a need to perform these functions, when the functions make sense to them (e.g., classification), that they will be able to adopt the linguistic means of doing so that has evolved historically in modern European culture.

The issue raised by Lemke, using writing as a tool for meaningful science learning, brings about some critical questions: How can one learn to read, write, and speak the language of science?, What are the protocols and expectations for reporting one's research?, How can we
communicate our ideas effectively to experts, scientists in other fields, and laypeople?, and How can we promote effective science learning by using investigative science methods and by using writing-to-learn or by using any possible combination of investigative science methods and writing-for-learning in science? Using language of science and partial answer to such questions addressed by standards (NRC, 1996);

The language and practices evident in the classroom are an important element of doing inquiries. Students need opportunities to present their abilities and understanding and to use the knowledge and language of science to communicate scientific explanations and ideas. Writing, labeling drawings, completing concept maps, developing spreadsheets, and designing computer graphics should be a part of the science education. These should be presented in a way that allows students to receive constructive feedback on the quality of thought and expression and the accuracy of scientific explanations (p. 144-45).

Keys et al. (1999) argued that in addition to informal writing genres, such as journal writing, question reflection, cartoons, and narratives, writing laboratory reports should be considered an effective science learning tool. The authors proposed the Science Writing Heuristic (SWH) as a tool to guide teachers and students. The SWH approach has a template for each of its two audiences. In the teacher template, the teacher uses a series of writing, reading, and small- and large-group discussion activities to support students in meaningful thinking. Thus, the teacher template illustrates the necessary pedagogy to support student learning. In the student template, students are encouraged to investigate their own question(s)
about the activity and use scientific methods during investigations; however, they are encouraged to use their own language to share their findings. Figures 3 and 4 give templates for students and the teacher.

Several empirical studies have been carried out to investigate the influence of the SWH approach on the learning process in both qualitative and quantitative aspects (or dimensions) across different grade levels. Studies by Hand, Hohenshell, and Prain (2002), Hand, Prain, and Hohenshell (2001), and Keys et al. (1999) showed that the implementation of the SWH approach had an impact on students' use of metacognition and reflection to understand knowledge, students' abilities to generate meaning from data in relation to specific knowledge, students' abilities to extend science ideas, and students' understanding of the nature of science. In addition to explicit evidence from the use of the SWH approach for meaningful learning in science in terms of the reshaping of the traditional laboratory report to more productive activities that require more cognitive and meta-cognitive activities through the use of writing activities, Prain and Hand (1996) and Hand and Prain (2002) assert the need for broadened implementation of writing-to-learn strategies.

1- Beginning ideas – What are my questions?
2- Tests – What did I do?
3- Observations – What did I see?
4- Claims – What can I claim?
5- Evidences – How do I know? Why I am making these claims?
6- Reading – How do my ideas compare with others?
7- Reflection – How have my ideas changes?

Figure 3. The science writing heuristic, Part II: A template for student thinking.
1- Exploration of pre-instruction understanding through individual or group concept mapping.
2- Pre-laboratory activities, including informal writing, making observations, brainstorming, and posing questions.
3- Participation in laboratory activity.
4- Negotiation phase I- writing personal meanings for laboratory activity (For example, writing journals).
5- Negotiation phase II- sharing and comparing data interpretation in small group (For example making a group chart).
6- Negotiation phase III- comparing science ideas to textbooks or other printed recourses (For example, writing group notes in response to focus questions).
7- Negotiation phase IV- individual reflection on writing (For example, creating a presentation such as a poster report for larger audience).
8- Exploration of post instruction understanding through concept mapping.

Figure 4. The science writing heuristic, Part I: A template for teacher-designed activities to promote laboratory understanding.

A Model for Writing for Learning in Science proposed by Prain and Hand (1996) (see Figure 5) is designed to guide teachers in planning writing tasks for secondary science topics. The crucial elements of the model include a theoretical base in that there are strong interactions between the demands of different writing tasks, subject-topic-task, and student learning outcomes; and a practical base in that teachers need to develop their understanding of writing-to-learn and which types of writing should be used (Hand & Prain, 2002).

In brief, to improve students’ conceptual understanding of science, educators need to focus on students’ conceptions of “what language is,” rather than on “what science is,” by using diverse types of writing in the classroom environment (Sutton, 1993, p. 1224). Implementing different types of writings for different purposes, different audiences, and different science contexts from beginning of the unit to the end will promote students’
conceptual understanding of science. However, more research studies are needed to explore what type of writing serves for which type of conceptual understanding and to investigate the most effective implementation of the these non-traditional writings in science context (Sutton, 1993).

<table>
<thead>
<tr>
<th>Method of Text Production</th>
<th>Audience</th>
<th>Purpose</th>
<th>Type</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Individuals</td>
<td>- Peers</td>
<td>Start</td>
<td>- Narrative</td>
<td>- Key concepts</td>
</tr>
<tr>
<td>- Pairs</td>
<td>- Younger</td>
<td>- Review</td>
<td>- Travelogues</td>
<td>- Linking themes</td>
</tr>
<tr>
<td>- Groups</td>
<td>- Students</td>
<td>- Hypothesizes</td>
<td>- Reports</td>
<td>- Factual understanding</td>
</tr>
<tr>
<td>- Looping</td>
<td>- Textbook</td>
<td>- Explore</td>
<td>- Instructions</td>
<td>- Apply concepts</td>
</tr>
<tr>
<td>- Computer</td>
<td>- Parents</td>
<td>- Devise plan</td>
<td>- Concept maps</td>
<td></td>
</tr>
<tr>
<td>- Pen</td>
<td>- Teachers</td>
<td>During</td>
<td>- Letters</td>
<td></td>
</tr>
<tr>
<td>- Redrafting</td>
<td>- Visitors</td>
<td>- Clarify</td>
<td>- Brochures</td>
<td></td>
</tr>
<tr>
<td>- Other</td>
<td>- Consumers</td>
<td>- Revise</td>
<td>- Poetry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Government</td>
<td>- Revise</td>
<td>- Posters</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Self</td>
<td>- Consider</td>
<td>- Journals</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Persuade</td>
<td>- Explanations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Interpret</td>
<td>- Diagrams</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Demonstrate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Revise</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Design solution</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Apply</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. Model for Writing-to-Learn

Constructivism and Learning Science

Constructivism, as a philosophical position, asserts that we, as human beings, cannot reach an objective reality directly, because reality is independent of our way of knowing it (Piaget, 1970; Simon, 1995). Further, the constructivist epistemology denies those fundamental wishes: there is a reality and secrets of this reality are to discover, reality exists
independently of us, those secrets are lawful so reality can be predicted and controlled, and those discoveries about reality are true (Segal, 2001). What is more, as Brooks and Brooks (1993) point out, not only the meaning of reality but also the meaning of knowledge and way of knowing are also challenged by constructivist epistemology.

While, in general, constructivism accepts the assertion that the structure of the mind is the source of our understanding of the world, there are different conceptions about the mechanism of learning. For instance, Piaget argued that the source of knowledge is the cognitive disequilibrium followed by either assimilation or accommodation. On the other hand, others [Bruner, (1986); Ernest, (1998); Vygotsky (1986)] disputed that the construction of knowledge is more closely related with prior experiences and language than is cognitive equilibrium. Since the effect of the constructivist epistemology appeared much later on education than on psychology and philosophy since education was strongly connected to behavioral psychology until the late 1990s, such diversity in schools of thought within the constructivist epistemology made different impacts on educational philosophy and practice (Brooks & Brooks, 1993).

How we perceive knowledge and the process of coming to know provides the basis for educational practice. As Schunk (2000) argued, if we believe that learners passively receive information, then the priority in instruction will be on knowledge transmission. If, on the other hand, we believe that learners actively construct knowledge in their attempts to make sense of their world, then teaching likely will emphasize the development of meaning and understanding. Moving from theory to practice always presents challenges, in education or in any other domain. When there are multiple positions within the theory, as happened in constructivist views of learning, the task of moving from theory to practice becomes even
more demanding. Further, since constructivism is a learning theory rather than a teaching theory and it is difficult for practitioners to make this distinction, being able to understand what is meant by constructivism and to distinguish branches of constructivism is very important for educators (Windschitl, 2002b).

Henriques (1997), in reviewing the literature on constructivism, used the labels Information Processing, Interactive Constructivism, Social Constructivism, and Radical Constructivism to represent forms of constructivism. Adopting from Henriques' (1997) work, Yore (2001) provided a synopsis of the constructivist spectrum (see Table 1). In such a spectrum, information processing, as a form of constructivism, assumes students learn from both the teacher and through experiences. In discussing a representational view of the mind, Cobb, Yackel, and Wood (1992) conceptualized the teacher in this particular orientation as being focused on three main concerns:

1. The goal of the instruction is to show students relationships that are located outside of their minds.

2. Instruction should help students to construct their internal relationship.

3. External instruction materials are the main sources for their knowledge.

While an information processing view accepts that learning occurs when there is a conflict between a newly encountered situation and an already existing conceptual framework, the level of success for learning is directly related with how the new situation is represented in relation to previous experiences of the learners (Fisher & Lipson, 1985).

Social constructivism, however, is based on the premise that the main source of knowledge is language and the society in which meaning is communicated and shared through language (Ernest, 1998). The crucial arguments put forward by social constructivism
are that: (a) without language individuals cannot have meaning, (b) language is dependent upon sociological and historical structures and occurrences, (c) individuals in the community can function only using language, which has locally established rules, and therefore language governs our functions in the community (Staver, 1998). Thus, individuals can construct knowledge, which is comprised by the culture and social norms, when they interact with other people through language (Schunk, 2000).

Table 1. Four Faces of Constructivism (Yore, 2001)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Information Processing</th>
<th>Interactive-Constructivist</th>
<th>Social Constructivist</th>
<th>Radical Constructivist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worldview</td>
<td>Mechanistic</td>
<td>Hybrid</td>
<td>Contextualistic</td>
<td>Organistic</td>
</tr>
<tr>
<td>Ontological View</td>
<td>Realist</td>
<td>Naïve Realist</td>
<td>Idealist</td>
<td>Idealist</td>
</tr>
<tr>
<td>Epistemic View</td>
<td>Absolutist (traditional)</td>
<td>Evaluativist (modern)</td>
<td>Evaluativist (postmodern)</td>
<td>Relativist (postmodern)</td>
</tr>
<tr>
<td>Judgment Criteria</td>
<td>Nature as Judge</td>
<td>Nature as Judge</td>
<td>Social Agreement as Judge</td>
<td>Self as Judge</td>
</tr>
<tr>
<td>Psychological Locus of Mental Activity</td>
<td>Private</td>
<td>Public and Private</td>
<td>Public</td>
<td>Private</td>
</tr>
<tr>
<td>Pedagogical Structure</td>
<td>Teacher</td>
<td>Shared: Teacher and Individuals</td>
<td>Group</td>
<td>Individual</td>
</tr>
<tr>
<td>Linguistic Discourse</td>
<td>One-Way: Teacher to Student</td>
<td>Two-Way: Negotiations to Surface Alternatives and to Clarify</td>
<td>Two-Way: Leading to Consensus</td>
<td>One-Way: Individual to Self (inner speech)</td>
</tr>
</tbody>
</table>

By contrast, radical constructivism asserts that knowledge is constructed by the individual in trying to make sense of interactions with his or her environment (Henriques, 1997). Knowledge construction/build up is perceived as an active cognitive process independent from any kind of communication or passively received inputs through the senses.
(Staver, 1998). While cognition and social interaction play central roles in knowledge construction, von Glasersfeld (1993) indicated that cognition serves to organize individuals' experiential world rather than to discover objective reality. On the other hand, social interaction is perceived as part of the environment or in the range of lived experiences that effect reorganization of individual cognition (Simon, 1995).

The crucial assertion of both epistemic views is that learning is an act of making meaning. Further, the environmental interactions, community, culture, and language are perceived to be central for individuals’ knowledge construction between radical and interactive constructivism. Yet, the difference appears when the social perspective asserts that the knowledge is inhabited in the society or in the culture in which individuals live in and share and when the cognitive perspective affirms that knowledge is created by individuals through experiences by living in a historical and cultural context. In short, the way radical and social constructivists perceive learning is similar to “frame of reference” phenomenon in physics. The frame of reference for the social cultural view is public aspects, while the frame of reference for the cognitive view is private aspects of knowledge. Indeed, as Simon (1995) argued, the issue is not finding out the primary view; rather, the issue is what we can learn from combining those two perspectives. That is, recognizing the dual faces of learning, like in the theory of light, can lead to advancements in the field of learning as it happened in the field of physics.

Adopting the premises of both social and cognitive perspectives, the interactive constructivist view asserts two dimensions of knowledge construction: “public knowledge,” whereby students construct knowledge and learn when they are able to interact with the physical world and other people, and “private knowledge,” whereby meaning is formed when
students reflect on and make sense of their interactions with their environment (Henriques, 1997). As Yore (2001) called it, a hybrid worldview of learning argues that knowledge construction and cognition differ from person to person, and the view recognizes that this knowledge construction has a public component. That is, not only cognition, but also social structure, language, and culture has impact on knowledge construction.

Advocates of the interactive constructivist view in the education field such as Cobb et al., (1992), Driver and Oldham, (1986), Driver, (1990), Simon, (1995), and Yore, (2001) argue that there is a need for taking account of the fact that individuals try to make their ideas clear to themselves and to others with whom they interact. The importance of such considerations is two-fold in terms of teaching. The first is that considering how students present their ideas to themselves and to others allows the teacher to have some understanding of personal models that the learners have – to learn something about the learner (Watts & Pope, 1989): “It is very easy to fall into thinking that I told them that, so they will remember it” or “we had an item about that on the last test, so they will know how to do this one.” (Maloney, 1990, p. 389) called these types of statements the cluttered states. He claimed that these states are based on teachers’ belief in the need to tell students what is true and wrong and to force students to erase their earlier concepts about topics so they can learn the correct one. These frameworks are not in line with a constructivist learning environment that combines sociological and psychological aspects of learning. Furthermore, teachers may accept some primitive and/or oversimplified model of learning, which ignores most of the context of the learner and teaching situation and gives little acknowledgment to the nature of scientific knowledge being learned (Gunstone, 1988).
The second consideration involves meta-cognition because, in the process of communicating ideas, the learner has an opportunity to clarify thought so that he or she can recognize how it might be changed – for learners to think about their own thinking (Watts & Pope, 1989). By using social negotiation, members of the classroom community share some aspects of understanding, although they have no direct access to each other’s understanding, and, thus, knowledge is not entirely a communal experience (Simon, 1995). Furthermore, as Strike and Posner (1982) indicated, knowing the cognitive dimension of the learning process is also crucial for teaching. At the individual level, learning is an interaction between what the student is taught and his/her current ideas or concepts; learning is not simply the acquisition of a set of correct responses, but rather an assimilation or accommodation (Posner, Strike, Hewson, & Gertzog, 1982; Strike & Posner, 1982; Thorley & Stofflett, 1996).

Accordingly, the conceptual change model proposed by Posner et al. (1982) and examined by many scholars (Driver & Oldham, 1986; Gil-Perez & Carrascosa, 1990; Hand & Treagust, 1991; Lijnse, 1990; Watts & Pope, 1989) takes into account students’ current ideas as a first step to constructing meaning. According to the conceptual change model, there are two steps for learning. First, the variation phase occurs, in which students use existing concepts to deal with a new problem—a stage called assimilation. During this time, students use their pre-existing knowledge to grasp new phenomena. In the second phase, if students need to replace or reorganize their central concepts, the stage is called accommodation, whereby conceptual change occurs (Posner et al., 1982). When we closely examine the conceptual change ecology, learners’ pre-existing knowledge plays a central role for either assimilation or accommodation.
Learners’ current ideas, which often are referred to as alternative frameworks or misconceptions, are associated with intuitive ideas or preconceptions constructed from experiences prior to school learning (Driver, Asoko, Leach, Mortimer, & Scott, 1994; Gil-Perez & Carrascosa, 1990). Children’s learning in science is analogous to scientists’ advancement of ideas, hypotheses, and principles when faced with new phenomena. Their prior knowledge and initial theories therefore are important as a part of the world around them. Driver and Oldham (1986) and Simon (1995) emphasized that what students know should serve as the basis for the teaching plan. In essence, the idea of science learning as a process of knowledge construction starts, necessarily, from prior knowledge.

The conceptual change model (Posner et al., 1982) presents prior knowledge as that which triggers students’ development of ideas. There are several empirical studies in science education based upon constructivist learning theory and the conceptual change model, including Gil-Perez and Carrascosa (1990) on classical mechanics, Watts and Pope (1989) on light, Lijnse (1990) on energy, Hand and Treagust (1991) on acids and bases, and Maloney (1990) on force. These researchers reported improvement not only in learning environments for students, but also for teachers’ personal development and pleasure. However, as Windschitl (2002a) pointed out, implementing constructivist instruction is quite difficult and challenging for teachers since most of the teachers have difficulty in making personal sense of constructivism and in reorienting the cultures of classrooms in harmony with constructivist view of learning.

Finally, in practice widely discussed instructional approach/design, student-oriented learning environment, is largely influenced by constructivist view of learning, social learning theory, and information processing theory. The student centered learning environment
demonstrated several crucial characteristics that are not embraced by traditional didactic or
teacher-oriented approach. Some of those characteristics are: students were
allowed/encouraged to take responsibility of determining the topics and method of
investigation. The teacher becomes consultant to help student structure their objectives; the
students were promoted to display their prior ideas about the unit under investigation which
serves as base for instructional decision making. In additions students were asked to engage
in multiple modes of representations within the classroom environment among peers to create
classroom discourse and make connection clear with social environment in large to apply and
communicate their learning/ideas (Brooks & Brooks, 1993; Honebein, 1996).

On the other hand, in a broad sense the environment where students were perceived as
passive receivers of the information given by teacher and the social reality by school
recognized as teacher-centered (Anyon, 1981; Halperin, 1994). Further, Catalano and
Catalano (1997) argue that teacher centered accepts the following assumptions;

1- an(y) educational process is considered culturally neutral as well as linear and rationale

2- language serves as a conduit for transmission of information

3- the teacher becomes manager of the classroom with the learning process

heavily dependent upon the pronounced and enforced rules (p. 95)

*Social Learning Theory*

The social learning theory claims that culture is the main determinant of individual
development. Humans are the only species to have created culture, and every human child
develops in the context of a culture (Schunk, 2000). Therefore, a child's learning
development is affected in ways large and small by the culture, including the culture of the family environment (Doolittle, 1997). Aligned with communist views of society, Russian psychologist Vygotsky’s ideas are widely associated with social learning theory.

The major theme of Vygotsky's theoretical framework is that social interaction plays a fundamental role in the development of cognition. Vygotsky (1978) stated:

Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people and then inside the child. This applies equally to voluntary attention, to logical memory, and to the formation of concepts. All the higher functions originate as actual relationships between individuals. (p. 57)

A second aspect of Vygotsky's theory is the idea that the potential for cognitive development is limited to a certain time span, which he calls the “zone of proximal development” (ZPD) (Schunk, 2004). Furthermore, full development during the ZPD depends upon full social interaction. The range of skills that can be developed with adult guidance or peer collaboration exceeds what can be attained alone.

Vygotsky’s idea of social learning has found wide acceptance in education since the theory has been proposed (Schunk, 2000). Further, one would argue that a branch of constructivism, social constructivism, had been strongly influenced from social leaning theory and Vygotsky’s ideas. While numerous research studies were carried out in relation to social learning theory and education, the focus of this literature review will include a brief overview of educational applications of the theory.
Dixon-Krauss (1996) and Doolittle (1997) outlined the need to consider the following requirements of a learning environment when we look at learning through the lens of social learning theory:

- **Full cognitive development requires social interaction:** Through culture children acquire much of the content of their thinking, their knowledge. Also, culture teaches children both what to think and how to think. That is, the surrounding culture provides a child with the processes or means for their thinking, so curricula should be designed to emphasize interaction between learners and learning tasks.

- **Cognitive development results from a dialectical process whereby a child learns** through problem-solving experiences shared with someone else, usually a parent or teacher but sometimes a sibling or peer.

- **The person interacting (teacher and other students in the learning environment) with the child assumes most of the responsibility for guiding the problem solving, but gradually this responsibility transfers to the child.**

- **Internalization refers to the process of learning—and thereby internalizing—a rich body of knowledge and tools of thought that first exist outside the child. This happens primarily through language.**

- **A difference exists between what child can do on his or her own and what the child can do with help.** So teachers need to have a rich understanding of ZPD to design a learning environment. With appropriate help from a teacher (scaffolding), children can often perform tasks that they are incapable of completing on their own. Scaffolding, where the teacher continually adjusts the level of his or her help in response to the child's level of performance, is an effective form of teaching.
Since much of what a child learns comes from the culture around him or her and much of the child's problem-solving is mediated through an adult's help, it is wrong to focus on a child in isolation.

- Interactions with surrounding culture and social agents, such as parents, teachers, and peers, contribute significantly to a child's intellectual development.
- Assessment should reflect both the Level of Independent Performance and the Level of Assisted Performance in a learning environment.
  - **Level of Assisted Performance**: The maximum one can achieve with help.
  - **Level of Independent Performance**: The best one can do on a given task without help.

*Inquiry in Science Education*

Long after the scientific revolution that shifted the belief that the source of knowledge and the way we come to know is through individual discovery and construction rather than through divine intervention, the impacts on the education community included important paradigm shifts in methods of teaching science (Finley & Pocovi, 2000; Schwab, 1960). Starting in the early twentieth century with Dewey's progressive movement in education, science thought has been criticized as being a process of giving an accumulation of information rather than a method of thinking and habits of mind or, in other words, inquiry (Alberts, 2000a; Bybee, 2000). Further, the condemnation of science teaching is supported by the science community. Bruce Alberts, former president of the National Academy of Sciences, pointed out that the way science and scientists work and think are not reflected upon in science classrooms (Alberts, 2000b). Indeed, Alberts and others argued that
fallaciously portrayed and misperceived concepts of science not only prevent the young generations from pursuing careers in science, but these false concepts also negatively impact their curiosity, creativity, critical thinking, and lifelong learning skills (Alberts, 2000b; Franken, 1994; Penick, 1995).

In search of finding and describing the purpose and methods of teaching science, a great number of works are noted in the literature since the publication of Dewey's (1910, 1916) argument for teaching science as an inquiry. Although education and teaching practices have always been affected by ideas in field of psychology—e.g., the behaviorist school of thought until the middle of the twentieth century, and the constructivist view of learning from the late twentieth century until the present—the current stage of science education, by amalgamating such learning theories, is anchored in science as inquiry by the National Science Education Standards and Benchmarks for science literacy (Finley & Pocovi, 2000; NRC, 1996, 2000; Project 2061, 1993).

In general, as a term, scientific inquiry, which embraces a wide range of mental and intellectual activities, such as problem posing, hypothesis testing, argumentation, explanation, modeling, and Socratic dialogs, is used commonly in science education (Abell, Anderson, & Chezem, 2000; Windschitl, 2003). To provide guidelines to educators, the National Science Education Standards—the current vision of science teaching—not only provides what students should know about science and science literacy by grade 12, but also provides a summary definition of inquiry for educators:

Scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they
develop knowledge and understanding of scientific ideas, as well is an understanding of how scientists study the natural world. (p. 23)

A definition from a different perspective is also offered:

Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations. (p. 23)

As Bybee pointed out, the Standards attribute two meanings to inquiry: content, that is, what should students understand about inquiry (stated with first quotation above); and, abilities, or what should students be able to develop with inquiry (stated with second quotation above) (Bybee, 2000). Changing the objectives of teaching science to science inquiry, where inquiry means content, should be acquired and a sets of skills should be developed during the investigation, has significant impact on scientific thought. Gerald Wheeler (2000), Executive Director of the National Science Teachers Association, remarked on the changing role of the teacher from just providing some level of engagement with material to that of developing knowledgeable thinkers about inquiry. However, in general, Wheeler’s (2000) and Windschitl’s (2002b) calls for inquiry in science are perceived as implementing more hands-on activities among science teachers.

The National Science Education Standards emphasize that “science as inquiry is basic to science education and a controlling principle in the ultimate organization and selection of
students' activities” (p. 105), whilst the Standards also suggest that inquiry is not the recommended single science teaching approach (NRC, 1996). Furthermore, the position statement published by The National Science Teacher Association's aligns with the Standards initiatives about teaching science as “hands-on, minds-on” inquiry-based science instruction where curiosity about the natural world is fostered (NSTA, 2003).

In addition, such publications underline some critical aspects of inquiry science teaching to guide teachers. Those critical aspects include the following:

- Focus and support inquiries while interacting with students.
- Orchestrate discourse among students about scientific ideas.
- Challenge students to accept and share responsibilities for their own learning.
- Recognize and respond to student diversity and encourage all students to participate fully in science learning.
- Encourage and model the skills of scientific inquiry, as well as the curiosity, openness to new ideas and data and skepticism that characterize science. (NRC, 1996, pp. 32-37)

Yet, results from large and small scale research studies and national assessments indicate that, in general, inquiry is yet not a characteristic of science classrooms in United States (NCES, 2002; O'Sullivan & Weiss, 2000; Tobin, Tippins, & Gallard, 1994; Weiss, 1994).

While U.S. Department of Education reports from grade 4, 8, and 12 unveiled the current state of using inquiry, such as more than 50% of those teachers who participated in this study use hands on activities no more than one or two times in a week (NCES, 1999, 2002), there were other reports focusing on the status of science teaching across kindergarten through 12 grade to identify characteristics of teachers, science teaching, and instructional objectives. For example, Fulp's (2002a) report on the current status of elementary science
teaching revealed that nearly two-thirds of the 655 participating elementary science teachers indicated they are not familiar with science standards and the documents. Moreover, only 14% of K-5 teachers perceived themselves to be well prepared to implement science aligned with the Standards. Thus far, the majority of participants in the surveys indicated that they spend very little time on such professional development programs for science teaching where they can get the help necessary to reach a stage of preparedness and confidence in using inquiry in the classroom.

The picture from the middle grades is somewhat similar to the findings above. In another report prepared by Fulp with cooperation of the National Science Foundation, results were the reverse of the elementary findings; that is, two-thirds of the 529 participating middle school science teachers indicated that they are familiar with the Standards in different levels (somewhat to very familiar). More than 80% of participating teachers stated that they are well prepared to listen and ask question to promote students' understanding, create learning environments in which students engage with hands-on science activities and cooperative learning groups, and implement a number of practices aligned with the Standards (Fulp, 2002b). Interestingly, conflicting with the findings above, more than 50% of the participating teachers indicated that they need help in learning how to use inquiry-based instructional strategies through professional development. Further, in practice, those teachers indicated that at least 63% of their classes include instructional approaches such as lecture, solving worksheets, and working on textbook problems. Finally, when students were asked to work on hands-on activities and investigation, there is a high chance (76%) that such work requires following specific instructions for completing activities, which is very similar to the findings about elementary settings (Fulp, 2002a, 2002b).
As for high-school science teaching, Banilower (2002), Smith (2002), and Wood (2002) reported key findings from the 2000 National Survey of Science and Mathematics Education, in which 5,728 science teachers across the United States participated. The results depicted the current state of high school physics, chemistry, and biology education. While around 40% of the participating teachers in the three subject areas indicated that they are not familiar with the Standards at all, oddly at least 75% of the participants in each area felt prepared to implement the Standards-based instruction. Further, about half of the physics, chemistry, and biology teachers indicated they need professional development associated with using inquiry in science teaching. What is more, in practice percentages of class time spent on working with hands-on, manipulative, or laboratory materials were rather low (around 20%) when we compared them with time spend on other instructional activities (lecture/discussion around 40%, reading textbooks, completing worksheets around 13%, non-lab small group work and daily routines around 25%) (Banilower, 2002; Smith, 2002; Wood, 2002).

In summary, teachers from elementary and secondary levels indicated moderate levels of familiarity about the Standards and inquiry in science teaching. On the contrary to some teachers’ strong confidence in their own preparedness to teach Standards-based and inquiry-oriented science classes, their practices indicated that those classes focused on more traditional didactic instructional approaches. Nonetheless, those results appear to be correlated with teachers’ call and need for professional development programs to learn how to implement inquiry. Indeed, understanding the needs of practitioners, and, more importantly, the origin of teachers’ beliefs and pedagogical practices that ground the current
landscape of science education, would help us as researchers to develop in-service training agendas (Fullan, 2001).

In the process of this quest, ideas put forward by Schwab appear relevant and important to take into consideration:

Much of what can be sketched in outline is already obvious. A teacher whose own study has been dogmatic and doctrinaire will be unprepared to teach science as inquiry. A teacher whose own training has demanded, or done little to discourage, acquiescence, dependency, and passivity will, in all likelihood, demand the same of his students. The first of these strictures bears mainly but not exclusively on the teacher’s training in subject matter. He will need to have a substantial part of that training in the form of inquiry into inquiry—enough of it to equip him with the ability to read and understand reports of inquiry. He will need to become familiar with the sorts of questions whose answers illuminate such materials. He will, himself, need to understand the ways in which invention and observation, datum and conception, interpenetrate to form the growing fabric of scientific knowledge. It is in this way, by developing the competence to participate in the movement of scientific inquiry, that the teacher of science can, in truth, be a scientist.

(Schwab, 1960, p. 192)

*Professional Development for Science Teachers*

Staff development or professional development in education ventures, as in all other business enterprises, plays an important role in teachers’ personal improvement and, as a
result, school improvement and the outcome of the quality of students' learning (Kennedy, 1998). In their review of research literature, Supovitz and Turner (2000) pointed out that professional development programs were the main source of implementing educational reforms since other possible ways of reaching and delivering reforms or innovations to teachers were found to be less effective than professional development.

While traditionally professional development or in-service programs were recognized as activities whereby teachers improved their specific teaching skills and knowledge through courses and presentations, recently this perception has changed toward viewing the school as a learning environment for all and an environment for ongoing inquiry for administrators, teachers, and students (Fullan, 2001; Knight, 2002; Loucks-Horsley, 1995; Loucks-Horsley, Stiles, & Hewson, 1996). That is, with long-term, strategically planned in-service programs educators focus on teaching and learning as more of an inquiry: individual and systematic development, job-embedded learning, and student learning outcomes (Loucks-Horsley, 1995). Furthermore, such contemporary visions of in-service programs are recognized to be an effective method of bringing reform movements into school environments.

In the publications of the National Science Standards (1996) and Inquiry and the National Science Standards (2000), the National Research Council emphasized the vital role of professional development programs whereby the teacher can experience inquiry as a learner and develop understanding of teaching science as an inquiry. Also, such reports pointed out the importance of long-term efforts to bring changes with the collective actions of teachers and administrators in an environment in which learning is an ongoing activity for everyone. Generally, with this holistic notion of professional development, it was expected that well-planned and implemented in-service programs, along with school administration
support will make changes in teachers' practices and such positive changes will bring improvement to students' performances (Kennedy, 1998; Supovitz & Turner, 2000). Figure 1 illustrates the theoretical relationship between professional development programs and student academic achievement.

Figure 6. The relationship between professional development and student achievement (Supovitz & Turner, 2000, p. 965)

Such calls for changing the orientation of in-service programs and promoting the implantation of science inquiry through such programs resulted in substantial amounts of federal, state, and other funding, through which researchers from science education, adult education, and sociology studied different dimensions of effective professional development programs (Darling-Hammond, 1996; Knight, 2002). While there is a general consensus on the needs of well-articulated, organized, implemented, and evaluated in-service professional development programs to archive this reform-based vision, the path ways that researchers followed to achieve such a vision differed (Luft, 2001).
In terms of the content of the program, Eshach (2003) studied teachers' efficacy belief toward science teaching, Tillema (1995) investigated the importance of the professional knowledge, and Van Driel and his colleagues (1998) explored the function of the pedagogical content knowledge. Also, regarding the theory-program connection, Knight (2002) analyzed the association between learning theories and ways to set up long-term effective in-service programs. Windschitl (2002a) examined teachers' understanding of and struggles with the constructivist learning theory and their connections to teacher change and inquiry implementations. Finally Fishman and his colleagues and others examined the structure of such in-service programs and concluded that programs need to be supported and evaluated by a combination of teacher reflection, classroom observation, and ongoing student assessments (Fishman et al., 2003; Marx et al., 2004).

A review of literature yields that teachers' epistemological beliefs and related classroom practices are crucial elements that need to be addressed with in-service programs to make significant changes (Kennedy, 1998; Luft, 2001). Researchers such as Luft (2001), Osborne and Wellington (2001), Simon (1995), Simon and Schifter (1987), and Windschitl (2002a) urged that not only robust understanding of learning theories (especially constructivist learning theory) and understanding the underpinnings of their educational practices, but also teachers' willingness and motivation to analyze and alter their understanding of teaching, learning, and classroom practices, are vital for bringing change. Therefore in-service programs should address such epistemological elements. Windschitl (2002a) argued that teachers' "epistemology must become an explicit target of change. Without such change as a priority, efforts directed at teacher development become narrowly
focused on changing the kinds of attributes and skills that may be added to, subtracted from, or modified” (p. 143).

Franke and her colleagues, based on the findings from their longitudinal professional development programs, suggested that teachers will demonstrate generative and sustainable change with programs in which they engaged with ways of students’ thinking and learning (Franke, Carpenter, Fennema, Ansell, & Behrend, 1998; Franke, Carpenter, Levi, & Fennema, 2001). The importance of student thinking and classroom interaction is also stressed by others. For example, Simon, (1995) highlighted that teachers’ hypothetical ideas about how learning occurs influence the way they teach and the pedagogical decisions they make (Millar, Osborne, & Nott, 1998; Osborne & Millar, 1998). Therefore, professional development programs need to pay attention to teachers’ beliefs toward student learning and thinking which interact with their practices (Richardson & Placier, 2001).

With the contemporary approach to professional development, researchers started to recognize the role and the importance of the teachers’ knowledge that they bring into in-service programs, similar the importance of students’ ideas about science and prior knowledge that they bring into learning environment (Loucks-Horsley, 1995). According to Marx, Freeman, Krajcik, and Blumenfeld (1998), to be able to pursue successful teacher change abstract propositional knowledge about learning and teaching is necessary but not sufficient enough. They argue that most of the teachers’ knowledge about teaching comes from their practical work, which is situated, personal, and not connected to abstract propositional knowledge. From their perspective:

... teachers construct their knowledge by integrating new learning with prior knowledge and beliefs, by applying ideas to practice, and by reflecting on the
results. Understanding is facilitated when teachers collaborate on authentic problems in a manner that requires them to justify their practices in light of emerging understandings of theory. (p. 669)

One would argue that what Marx and his colleagues portray for teacher change is not different than what Posner and his colleagues, and Driver and her colleagues argue for learning in general as a conceptual change and learning and teaching framed by constructivism (Driver et al., 1994; Posner et al., 1982).

In summary, combining the findings of professional development and teacher assessment literature suggests that special attention should be paid to inner elements of the Supovitz and Turner's (2000) model—structure of high quality in-service programs, implementation of such programs, and student learning outcome—to analyze systematically the effects of the innovation and the effects on teachers. The high-quality, effective in-service programs are recognized as being sensitive to diverse learning; emphasizing inquiry-based learning, problem posing, problem solving, and application of knowledge; considering the tacit and propositional knowledge of teachers; and providing long-term support and collaboration (Knight, 2002; Loucks-Horsley, 1995; Loucks-Horsley et al., 1996). The level of the teachers’ implementation of innovation is generally recognized in two categories: cognitive and pedagogical. Cognitive change describes the extent of changes in beliefs, attitudes, and habits of mind toward science and science teaching, and being able to recognize such changes as a result of the program, where as pedagogical change describes the extent of changes in tacit knowledge about teaching and in pedagogical practices toward proposed direction by the program (Flecknoe, 2002; Knight, 2002).
Impact of the Teacher & Innovations on Students’ Science Achievement

With the current reform movement in science education, there is strong emphasis on the idea that learning science is an inquiry process whereby students actively engage with natural phenomena by asking questions, constructing possible explanations, developing ways of testing them, and communicating their ideas with the members of their community. While the prominence of students’ active engagement, called hands-on and minds-on, is a widely accepted view in science learning, the way that science currently is taught has been criticized to be teacher-oriented and ineffective (Anderson, 1997; Goodlad, 1983; National Commission on Excellence in Education, 1983; NRC, 1996). Reform documents, such as the National Science Standards (NRC, 1996, p. 20), argues that “learning science is something students do, not something that is done to them.”

To be able to achieve such reform, science learning with understanding, teachers need to move away from the traditional didactic approach of teaching science in which students are passive receivers of information and vocabulary of science topics are presented by teacher and textbooks (Driver et al., 1994; Franke et al., 2001; NRC, 1996; Prawat, 1989). Instead, the reform movement encourages teachers to promote and implement student-oriented inquiry in which students’ current ideas of science are connected with scientific knowledge found in different sources through inquiry process, small- and large-group negotiations, and problem solving (Driver et al., 1994; NRC, 1996).

Moving from theory to practice in understanding the demands of the innovation is difficult for teachers (Bitan-Friedlander, Dreyfus, & Milgrom, 2004; Windschitl, 2002a). As pointed out by Fullan (2001), implementation of innovation in an educational setting can be affected by numerous aspects of the teacher, such as attitudes, knowledge, and skills, and of
the environment, such as support from professionals, administration, and society. Further, the adaptation of such fundamental changes in beliefs, attitudes, and skills is not only a non-linear process for each teacher, but the process also differs significantly from teacher to teacher (Bitan-Friedlander et al., 2004). Adding those dimensions together, understanding or measuring the effect of the innovation, or, in essence, the effect of the teacher and teacher's implementation level of the innovation becomes difficult and controversial.

In their historical analysis of teacher evaluation, Shinkfield and Stufflebeam (1995) pointed out that findings from previous research, where teachers' behaviors were correlated with students' outcomes, are somehow perplexing and summarizing the casual relationship between them is difficult. While finding a form of assessment that can evaluate the totality of a student's learning or the impact of an innovation or teacher on students' learning is not practical, using standardized and alternative forms of assessments and triangulating them can be an indicator of learning and impact of innovations and of the teacher. In practice, the impact of innovations on students' outcomes, usually academic achievement, has been predominantly accepted and used for funding and policy making (Darling-Hammond, 1996, 1999; IES, 2003; Shinkfield & Stufflebeam, 1995). However, as Sanders and Horn (1995) claimed, instead of arguing the use of one standardized measurement over alternative ones, or vice versa, it is more important to choose and use the most proper indicator variables and measurements available for our purposes.

In search of the teacher's and/or innovation's effectiveness on students' learning of subject matter, researchers used different approaches. Seeker and Lissitz (1999) used a large data set from the nationwide High School Effectiveness Study (HSES) to investigate the impact of teachers' understanding level (called implementation) of the inquiry reform and the
Standards on students' academic achievement by using Hierarchical Linear Modeling (HLM). Further, their investigation also sought to find possible relationships among teachers' implementation of the Standards, school achievement levels, students' demographic: gender, ethnic, and socioeconomic backgrounds. While the measurement they used to identify implementation level, a teacher questionnaire, has been found to be problematic for identifying teacher practices without actual observation data, their findings indicated that there was an increase in students' science achievement across different achievement, demographic, and socioeconomic levels, and the increase was associated with the instruction in which laboratory inquiry was emphasized. Nonetheless, their overall findings indicated that insufficient evidence existed to suggest student-centered practices are more beneficial than teacher-centered practices in terms of learning outcomes. Further, they dispute that the consequences of using Standards-based instruction are not self evident, and that the benefits of Standards-based instruction are largely anecdotal. Finally, they acknowledge the fact that there is a need for rigorous assessment of practices and identification of evidentiary criteria for implementation.

Following a different path, Sanders and Horn developed a process of measuring the impact of the school system, classroom characteristics (homogenous or heterogeneous achievement distributions, size, etc.), students' achievement levels, and, most importantly, the teacher on students' standardized test scores for large data sets (Sanders & Horn, 1994, 1995, 1998). They called this process The Tennessee Value Added Assessment Analysis System (TVAAS), which uses mixed-model methodology and multivariate statistical approaches to examine students' longitudinal standardized test data. The main objective of TVAAS is to find elements of the learning environment that have made an impact on
students’ gains over time. Keeping this rationale in mind, Sanders and Horn’s approach neither uses any questionnaire or interviews as in Seeker and Lissitz’s work, nor observational data to identify the impact of teachers’ practices on students’ gain over time.

The results from the research studies conducted in different states by using a value-added approach yielded that the teacher is the major factor influencing the students academic achievement gain over time, followed by students’ achievement levels (Carey, 2004; Doran & Izumi, 2004; Sanders & Horn, 1998; Sanders & Rivers, 1996; Wright et al., 1997). Further, the way Sanders and Horn identified an effective teacher and teaching was by analyzing the pattern of students’ gain scores over the time for each teacher. That is, the high gain scores in effective teachers’ classrooms were consistent over the years. Another interesting finding emerging from those studies is that not only low achieving students had more gain with effective teachers than they did with other teachers, but also their gains was greater than high achieving students’.

The significance of the value-added model for educational research is two-fold. First, without making the identification of “the best teacher” or “best way of teaching” (p. 10), this approach provides methods to evaluate the impact of different practices and implemented innovation on students’ leaning outcome (Sanders & Horn, 1998, p. 10). Second, results from primary and secondary educational settings suggest that effective teaching practices have significant impact on students’ achievement in general, and superior impact on low-achieving students’. Nonetheless, to be able to use such tools and approaches in educational settings, as Sanders and Horn urged, there is a need for standardized and normed longitudinal data that are expensive and labor intensive for schools and researchers. As a result, alternative assessments such as teacher generated pre-post tests could be used to analyze the impact of
innovation or other variables on students' learning if those assessments validity and reliability are ensured (Sanders & Rivers, 1996; Thum, 2003).

On the contrary to Secker and Lissitz's argument about the lack of evidence of the impact of standard based science teaching, Lee and her colleagues demonstrated significant findings of enhancement in science learning through using/implementing inquiry and Standards-based science instruction in large, culturally, socially, and economically diverse inner city settings (Lee, Deaktor, Hart, Cuevas, & Enders, 2005). Their focus was investigating the impact of Standards-based science and literacy teaching, students' demographic characteristics, and students' achievements on students' pre-post measures of content and inquiry skills. Similar to Sanders and Horn's method, Lee and her colleagues used sophisticated multivariate statistical approach and growth curve modeling to compare gains of different subgroups within the study sample and to compare findings of this study to nationwide assessment findings (e.g., NAEP).

Results not only suggested a significant cumulative impact of implementing Standards-based interventions on elementary students' science and literacy scores over the 3-year project, but the results also suggest that intervention had a larger effect after a year of the project. Analysis of teacher observations, teacher questionnaires, and teacher interviews supported such qualitative findings. Teachers who participated in this longitudinal study reported more struggles and difficulties in implementing the intervention throughout the first year of the study in comparison to second year (Lee et al., 2005). Further, during the intervention teachers' practices were improved by changes in their beliefs, attitudes, knowledge, and content understanding, and such changes and improvements showed a larger impact on students' achievement (Lee, Hart, Cuevas, & Enders, 2004).
In general, researchers argue that there is large gap in all levels of educational research literature about how in-service programs and implementation of such innovations can affect student outcomes (Flecknoe, 2002; Kennedy, 1998). Further, as Lee and others who work in the area of teacher professional development and student learning suggested, those needed research studies should focus on the relationship between teacher change/level of change and student science achievement to evaluate interventions and draw conclusions about accountability of such innovations (Chinn & Malhotra, 2002; Eshach, 2003; Fishman et al., 2003; Songer et al., 2003; van Driel, Beijaard, & Verloop, 2001; van Driel et al., 1998).
CHAPTER THREE

Journal Article

Abstract

This study is a part of a bigger project known as the Science Writing Heuristic (SWH) Partnership Professional Development Project, conducted at Iowa State University and the University of Iowa in association with the Iowa Department of Education to help improve science teaching. Overall, the goal of the project is to help practicing science teachers understand and apply a student-oriented instructional approach, using the SWH. The purpose of this research study was to examine the link between teachers’ implementation of a student-oriented teaching approach through the SWH approach with embedded non-traditional writing practices and students’ performances on standardized tests over a 3-year period. This study investigated the impact of 6 teachers’ (3 high school teachers and 3 middle school teachers) implementation of the SWH approach on student standardized test scores over the 3-year period. A mixed method approach was adopted as a research method.

A major premise underpinning this study is that in the rate of change differs by teachers, and that change is not a linear process for teachers. Results of the study indicated a differential across teachers in terms of improvement in pedagogical skills related to the SWH approach. Further, results showed that the SWH approach in-service program did have an impact on participating teachers’ pedagogical practices. The majority of the participating teachers improved their pedagogical practices of implementing science inquiry through the SWH approach over the 3-year period of the professional development program. Further, when teachers’ rankings were correlated against students’ standardized test scores, the results
indicated that as their implementation levels increased their students' test achievements also increased.

**Background**

*Science Inquiry and Inquiry in Science Teaching*

The term “scientific inquiry,” which embraces a wide range of mental and intellectual activities, such as problem posing, hypothesis testing, argumentation, explanation, modeling, and Socratic dialogs, is used commonly in science education (Abell et al., 2000; Windschitl, 2003). To provide guidelines to educators, the National Science Education Standards provide a summary definition of inquiry for educators:

Scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world. (p. 23)

A definition from a different perspective is also offered:

Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations (p. 23).
As Bybee (2000) pointed out, the Standards attribute two meanings to inquiry: content, or what students should understand about inquiry (as stated with the first quotation above), and abilities, or what students should be able to develop with inquiry (stated with the second quotation above. Changing the objectives of teaching science to science inquiry, where inquiry means content that should acquired and sets of skills that should be developed during the investigation, has significant impact on scientific thought. For example, the Executive Director of the National Science Teachers Association remarked on the role of teachers changing from just providing some level of engagement with material to developing knowledgeable thinkers about inquiry (Wheeler, 2000). However, as Wheeler (2000) and Windschitl (2002b) urged, calls for inquiry in science generally are perceived as implementing more hands-on activities among science teachers.

The position statement published by The National Science Teacher's Association aligns with the Standards initiatives about teaching science as “hands-on, minds-on” inquiry-based science instruction where curiosity about the natural world is fostered (NSTA, 2003). In addition, such publications underline some critical aspects of inquiry science teaching to guide teachers. Those critical aspects include the following:

- Focus and support inquiries while interacting with students.
- Orchestrate discourse among students about scientific ideas.
- Challenge students to accept and share responsibilities for their own learning.
- Recognize and respond to student diversity and encourage all students to participate fully in science learning.
- Encourage and model the skills of scientific inquiry, as well as the curiosity, openness to new ideas and data, and skepticism that characterize science (NRC, 1996, pp. 32-37).
Yet, results from large-scale and small-scale research studies and national assessments indicate that, in general, inquiry is not yet a characteristic of science classrooms in the United States (NCES, 2002; O'Sullivan & Weiss, 2000; Tobin et al., 1994; Weiss, 1994).

Reports from large-scale research studies such as Banilower (2002), Fulp (2002a, 2002b), NCES (2002), Smith (2002), and Wood (2002) pointed out that teachers from elementary and secondary levels indicated moderate levels of familiarity about the Standards and inquiry in science teaching. Contrary to some teachers' strong confidence in their own preparedness to teach Standards-based and inquiry-oriented science classes, their practices indicated that those classes focused on more traditional didactic instructional approaches. Nonetheless, those results appear to be correlated with teachers' call and need for professional development programs to learn how to implement inquiry. Indeed, understanding the needs of practitioners, and, more importantly, the origin of teachers' beliefs and pedagogical practices that ground the current landscape of science education, would help us as researchers to develop in-service training agendas (Fullan, 2001).

**Professional Development for Science Teachers**

While traditionally professional development or in-service programs were recognized as activities whereby teachers improved their specific teaching skills and knowledge through courses and presentations, recently this perception has changed toward perceiving the school environment as a learning environment for all and an environment for on-going inquiry for all parts of the system: administrators, teachers, and students (Fullan, 2001; Knight, 2002; Loucks-Horsley, 1995; Loucks-Horsley et al., 1996). That is, with long-term, strategically-planned in-service programs, educators focus on teaching and learning as more of an inquiry:
individual and systematic development, job-embedded learning, and student learning outcomes (Loucks-Horsley, 1995). Furthermore, such contemporary visions of in-service programs are recognized to be an effective method of bringing reform movements into school environments.

In the publications of the National Science Standards (1996) and Inquiry and the National Science Standards (2000), the National Research Council emphasized the vital role of professional development programs whereby the teacher can experience inquiry as a learner and can develop understanding of teaching science as inquiry. Also, such reports pointed out the importance of long-term efforts to bring changes with the collective actions of teachers and administrators in an environment in which learning is an ongoing activity for everyone.

Figure 7. The relationship between professional development and student achievement (Supovitz & Turner, 2000, p. 965)

The findings of the professional development and teacher assessment literature suggest that special attention should be paid to inner elements of the Supovitz and Turner (2000) model—structure of high-quality in-service programs, implementation of such
programs, and student learning outcomes—to analyze systematically the effects of the innovation and the effects on teachers. High-quality, effective in-service programs are recognized as being sensitive to diverse learning; emphasizing inquiry-based learning, problem posing, problem solving, and application of knowledge; considering the tacit and propositional knowledge of teachers; and providing long-term support and collaboration (Knight, 2002; Loucks-Horsley, 1995; Loucks-Horsley et al., 1996). Furthermore, not only the quality of the program but also the implementation level of the innovation or the teacher change due to programs becomes important in the proposed Supovitz and Turner model (2000). The level of the teachers' implementation of innovation is generally recognized in two categories: cognitive and pedagogical. Cognitive change describes the extent of changes in beliefs, attitudes, and habits of mind toward science and science teaching, and being able to recognize such changes as a result of the program, where pedagogical change describes the extent of changes in tacit knowledge about teaching and in pedagogical practices toward proposed direction by the program (Flecknoe, 2002; Knight, 2002).

Impact of the Teacher & Innovations on Students' Science Achievement

In their historical analysis of teacher evaluation, Shinkfield and Stufflebeam (1995) pointed out that findings from previous research where teachers’ behaviors were correlated with students’ outcomes are perplexing and summarizing the casual relationship between them is difficult. While finding a form of assessment that can evaluate the totality of a student's learning or the impact of an innovation or teacher on students’ learning is not practical, using standardized and alternative forms of assessments and triangulating them can be an indicator of learning and the impact of innovations and of the teacher. In practice, the
impact of innovations on students’ outcomes, academic achievement in most the cases, has been largely accepted and used for funding and policymaking (Darling-Hammond, 1996, 1999; IES, 2003; Shinkfield & Stufflebeam, 1995). However, as Sanders and Horn (1995) claimed, instead of arguing for the use of one standardized measurement over alternative ones, or vice versa, it is more important to choose and use the most proper indicator variables and measurements available for our purposes.

In search of the teacher’s and/or innovation’s effectiveness on students’ learning of subject matter, researchers have used different approaches. Seeker and Lissitz (1999) used a large data set from a nationwide research study, the High School Effectiveness Study (HSES), to investigate the impact of teachers’ understanding level (called implementation) of the inquiry reform and the Standards on students’ academic achievement by using Hierarchical Linear Modeling (HLM). Further, their investigation also sought to find possible relationships among teachers’ implementation of the Standards, school achievement levels, and students’ demographic (gender, ethnic, and socioeconomic) backgrounds. While the measurement they used to identify implementation level—a teacher questionnaire—has been found to be problematic for identifying teacher practices without actual observation data, their findings indicated that there was an increase in students’ science achievement across different achievement, demographic, and socioeconomic levels, and the increase was associated with instruction in which laboratory inquiry was emphasized.

Following a different path, Sanders and Horn developed a process of measuring the impact of the school system, classroom characteristics (homogenous or heterogeneous achievement distributions, size, etc.) on students’ achievement levels, and, most importantly, the impact of the teacher on students’ standardized test scores for large data sets (Sanders &
Horn, 1994, 1995, 1998). They called this process The Tennessee Value Added Assessment Analysis System (TVAAS), which uses mixed-model methodology and multivariate statistical approaches to examine students’ longitudinal standardized test data.

Contrary to Seeker and Lissitz’s argument about the lack of evidence of the impact of standard-based science teaching, Lee and her colleagues demonstrated significant enhancement in science learning through using/implementing inquiry and Standards-based science instruction in large culturally, socially, and economically diverse inner city settings (Lee et al., 2005). Their focus was on investigating the impact of Standards-based science and literacy teaching, students’ demographic characteristics, and students’ achievements on students’ pre-post measures of content and inquiry skills. Further, the results suggested that intervention had a larger effect in the second year of the project. Analysis of teacher observations, teacher questionnaires, and teacher interviews supported such qualitative findings. Teachers who participated in this longitudinal study reported more struggles and difficulties in implementing the intervention throughout the first year of the study in comparison to the second year (Lee et al., 2005). Moreover, during the intervention teachers’ practices were improved by changes in their beliefs, attitudes, knowledge, and content understanding, and such changes and improvements showed a larger impact on students’ achievement (Lee et al., 2004).

In general, researchers argue that there is large gap in all levels of the educational research literature about how in-service programs and implementation of such innovations can affect student outcomes (Flecknoe, 2002; Kennedy, 1998). Further, as Lee and others who work in the area of teacher professional development and student learning suggested, those needed research studies should focus on the relationship between teacher change/level
of change and student science achievement to evaluate interventions and draw conclusions about the accountability of such innovations (Chinn & Malhotra, 2002; Eshach, 2003; Fishman et al., 2003; Songer et al., 2003; van Driel et al., 1998, 2001).

The Science Writing Heuristic

Hand and Keys (1999), in an attempt to develop a rich teaching/learning approach by combining ideas from constructivism, writing-to-learn, science literacy, science inquiry, and learning cycle, constructed the science writing heuristic (SWH), which they consider to be an inquiry-based approach that links writing, reading, and science laboratory activities. The framework for designing the SWH approach includes the shift to constructivist theory, understanding the nature of science, and promoting scientific literacy. The SWH approach consists of two templates—one for the teacher and the other for the student.

In the teacher template, the teacher uses a series of writing, reading, and small- and large-group discussion activities to support students in meaningful thinking. Further, the teacher template presents important phases of suggested activities to enhance learning by promoting negotiation of meaning among students and/or among students and teachers in both small- and large-group activities (Keys et al., 1999). In essence, the teacher template provides strong pedagogical focus for implementing and conducting scientific investigation and science inquiry as a means to learn the scientific methods and procedures aligned with the National Standards.

In the student template, students are encouraged to investigate their own question(s) about the activity and use scientific methods during investigations; however, they are encouraged to use their own language to share their findings. The student template is to
scaffold student understanding of scientific concepts while writing the laboratory report by relating claims to evidence (Hand & Keys, 1999). The heuristic is constructed upon an epistemological view that allows students to think about their claims and how they might interpret the data to provide supporting evidence. The SWH approach is based on the assumption that science writing genres in school should reflect some of the characteristics of scientist's writing and be shaped as a pedagogical tool that encourages students to distinguish scientific meaning from reasoning.

1- Beginning ideas – What are my questions?
2- Tests – What did I do?
3- Observations – What did I see?
4- Claims – What can I claim?
5- Evidences – How do I know? Why I am making these claims?
6- Reading – How do my ideas compare with others?
7- Reflection – How have my ideas changed?

Figure 8. The science writing heuristic, Part II: A template for student thinking.

1- Exploration of pre-instruction understanding through individual or group concept mapping.
2- Pre-laboratory activities, including informal writing, making observations, brainstorming, and posing questions.
3- Participation in laboratory activity.
4- Negotiation phase I—writing personal meanings for laboratory activity (for example, writing journals).
5- Negotiation phase II—sharing and comparing data interpretation in small group (for example, making a group chart).
6- Negotiation phase III—comparing science ideas to textbooks or other printed resources (For example, writing group notes in response to focus questions).
7- Negotiation phase IV—individual reflection on writing (for example, creating a presentation such as a poster report for larger audience).
8- Exploration of post instruction understanding through concept mapping.

Figure 9. The science writing heuristic, Part I: A template for teacher-designed activities to promote laboratory understanding.
Several empirical studies have been carried out to investigate the influence of the SWH approach. Studies by Hand, Hohenshell, and Prain (2002); Hand, Prain, and Hohenshell (2001); and Keys et al. (1999) showed that implementation of the SWH approach had an impact on students' use of meta-cognition and reflection to understand knowledge, students' abilities to generate meaning from data in relation to specific knowledge, students' abilities to extend science ideas, and students' understanding of the nature of science. Further, other studies by Gunel, Akkus, Hohenshell, and Hand (2004) and Omar and Gunel (2004) pointed out two important points in relation to implementation of the SWH approach. First, using the SWH approach as opposed to traditional instruction had a significant impact on students' learning of science. Second, when teachers' implementation level of the SWH approach was higher, students' scores were significantly higher on teacher generated tests. Results drawn from a recent study (Poock, Burke, Greenbowe, & Hand, 2004) that used the SWH approach in a chemistry course for the tertiary level provided evidence of closing the gap in gender achievement, measured from students' responses to the American Chemical Society California Diagnostic Exam at the end of the semester; while the difference in gender improvement scores was significant in the beginning of the semester, the difference was not significant at the end of the semester, with females scoring larger improvement score.

Integrating the previous findings of studies with the SWH approach and the current literature gap in professional development and the impact of science inquiry implementation, there is a need for studies where not only the process of teacher change through the SWH in-service programs is investigated in detail but the impact of such practice changes on students' learning outcomes also is explored. Further, these studies should be in a longitudinal format so that change and both its short-term and long-term effects can be investigated. Guided by
current literature and findings from previous studies on the SWH approach this, study will investigate the following research questions.

Research Questions

Two research questions were pursued with this study.

- How have participating teachers’ pedagogical practices changed over the 3 years of implementing the SWH approach?
- How have standardized test scores of those students’ in the teacher’s classes been impacted by such changes?

Method

Research Context and Participants

The purpose of this research study is to examine the link between teachers’ implementation of a student-oriented teaching approach through the SWH approach with embedded non-traditional writing practices and students’ performances on standardized tests over a 3-year period. This study investigated the impact of 6 teachers’ (3 high school teachers and 3 middle school teachers) implementation of the SWH approach on student standardized test scores over the 3-year period. The study was an extension of a previous study. In an earlier study, Omar (2004) studied the change in pedagogical practices of 6 teachers who implemented the SWH approach. Her work focused on the changes in teachers’ beliefs and attitudes toward learning and teaching and the changes in their pedagogical practices when implementing the SWH approach over a 2-year period. This study was an extension of the original study by Omar and adopted a case study approach for individual teachers. The major
difference between the two studies is that this study extended the observation period for an additional year and measured how students’ performances on standardized tests were affected by various implementation levels of the teacher over the 3-year-long change process. Students’ scores for each teacher from the year prior to the first year of the project were obtained and used as a baseline to analyze the progress of each teacher.

Setting

Participating School District

The school district that participated in this study is located in a moderate sized federally designated poverty town in rural Iowa. The school district is classified as a 3-A district with 106 teachers in middle and high school buildings. Total student enrollment was 2,276. The distribution of students’ ethnicities was: 96% white, 1.6% African American, 0.37% Asian, 1.85% Hispanic, and 0.18% Native American. Six science teachers were divided equally between middle school and high school. The population included 1 female and 5 male teachers.

Interventions

Workshops.

Several workshops were conducted: 3 3-day events during the summer, and 4 half-day events during the school year, held over the 3-year period of the project. The first workshop laid down the basis of constructivism as a learning theory and teaching science as an inquiry. Further, this workshop mainly emphasized students’ epistemological beliefs toward teaching and learning, and modeled a scientific activity using the SWH approach as a
first phase of their professional growth. Later workshops were conducted at the local university campus (2 3-day events) and at the school site (4 half-day events) based on teachers' need and accessibility. Throughout these meetings, teachers carried out science inquiry activities from biology, chemistry, and physics areas using the SWH approach. After completing the investigations, discussions were held about the students' experiences with the SWH approach learning environment based on the teachers' observations and discussions of the effective and desired role and of the concept of the student-oriented teaching approach.

Unit preparation was another portion of the workshops. Teachers were asked to prepare concept maps for units they were planning to teach each year. The purpose of concept mapping was to recognize and place more emphasis on the "big" ideas for each unit rather than on student goals. These concept maps were used for aligning the topic with district standards and benchmarks and for selecting appropriate activities. Further, the construction of maps assisted the teachers in creating their assessment tools, where an emphasis was paced on developing extended response conceptually-oriented test questions. Researchers who specialize in the different subject areas were present during unit preparation to assist and challenge teachers. Finally, every year researchers provided detailed reports where the previous year's analyses of students' performances on pretest and posttest, and analyses of teacher's implementation levels were outlined.

On-Site support.

Recognizing that teachers need adequate time for understanding, implementing, evaluating, and reflecting upon the new teaching approach, several one-on-one sessions designed to support each teacher in implementing a unit also were held. Researchers visited
each teacher at least once a week in their classrooms and observed his or her implementation of the SWH approach by taking rich field notes. Upon a request from the instructor, the researcher(s) modeled teaching or co-taught with the teacher. At the end of each session, the researcher(s) had short debriefing sessions with teachers where constructive feedback was provided after teachers’ identification/self evaluation of strengths, weaknesses, and difficulties with implementation. Such debriefings targeted several areas of interest such as increasing the level of confidence and trust between the teacher and observer, promoting the teacher’s awareness of certain behaviors observed, highlighting the pedagogical areas that needed improvement, and suggesting some strategies that the teachers might use to improve their implementation of the required student-oriented approaches in the future.

*Data Collection*

*Observations*

Observations were the primary data source in this study, by which teachers’ third-year implementation levels were recognized. Observations included rich descriptive data of each participant’s teaching practices, debriefing summary notes, and a video recording of at least one lesson in each semester per teacher. Merriam (1998) emphasized that an effective observation needs to capture as many details as possible to draw meaningful conclusions about what is observed. Yet, McMillan and Schumacher (1997) argued that because “… it is impossible to observe everything that occurs, the researcher must decide on the variables or units of analysis that are most important and then define the behavior so that it can be recorded objectively” (p. 270). Hence, comprehensive observations were guided mainly by observational units: the implementation criteria provided above, while within the project
framework semi-structured teacher interviews and teacher questioners were collected from participants. The analyses of such data were conducted and reported by Omar (2004).

In this study, observations and videotapes were used to identify teachers’ implementation level. This included three researchers, including two Ph.D. students and a professor, and one independent observer—a retired science teacher who successfully implemented the approach several years. This group used a criteria matrix to analyze each teacher’s implementation level of the SWH approach. The author, two other researchers and the independent observer participated in Omar’s study as an observer. Prior to and during Omar’s study, necessary trainings and discussions were held in order to achieve internal consistency for ranking. Prior to conducting any observations in Omar’s study, all observers attended a two-day workshop in regards to implementing of the SWH approach. The focus of the workshop was twofold: training the observers how to observe an SWH implementation and conduct a debriefing session afterward to help the participating teachers overcome weaknesses and defining strengths of the performance. Throughout this study, weekly meetings among the researchers and monthly meetings including the external observer took place and the active context of the each lesson and teacher’s progress were maintained and shared during the meetings.

Students’ Test Scores

For each teacher, Iowa Test of Educational Development (ITED) test scores from the year prior to SWH professional development project, baseline (2001-2002), the first year (2002-2003), the second year (2003-2004), and the last year (2003-2004) of the project were collected. Such data files consisted of students’ composite achievement in different subject
areas: science, mathematics, social studies, and language arts. However, baseline year scores for two teachers, John and Gary, were not available.

Method of Analysis

In this study a mixed method approach was adopted to add more robustness to the results generated from the study and to ensure generalization of its findings. Using a qualitative approach, individual implementation profiles were constructed for each teacher over the 3-year period that the professional development program took place. In essence, there were 6 cases, one for each teacher, that are made up of a continuing profile of his or her progress with his/her students' ITED scores in each year explored. Further, such scores were weighing against test scores prior to the in-service program, in the 2001-2002 academic year, to investigate the impact of the SWH approach professional development program.

Assessment Instruments

Criteria Matrix for Assessing Teacher Implementation

Each teacher was ranked based on his/her level of implementation. Using the same scale as Omar (2004), the ranking mechanism consisted of 3 composite ranking scales. The first was a 10-point ranking specifically targeting the first criterion—promoting dialogical interaction—with 1 indicating low interaction and 10 indicating high interaction. The promotion of dialogical interaction and student argument are considered the major demand of the SWH approach and reflect a shift in instructional practice.

The second and third scales targeted the second and third criteria, respectively—focus of learning and connection. Each was a 5-point ranking, with 1 reflecting the first level, 3 the
second level, and 5 the third level, while 2 and 4 indicate transitional levels. The 5-point focus of learning ranking reflects the degree the teacher is in focus of learning in his or her classroom, where a low level implies knowledge transfer. Further, the 10-point connection ranking reflects teachers' success in preparing for the unit(s) from a constructivist teaching perspective (Simon, 1995).

Most of the sessions were observed by individual researcher. After each observation observers provided detailed observational notes about the session watched, and mean score of three criteria mentioned above. Some of the sessions were recorded and those recordings were analyzed by independent observer. During the monthly meetings independent observer was informed about the context of the lesson for which to analyze. The percentage of agreement or the inter-rater reliability among the observers and independent rater was 90%. In the cases of ranking disagreement, researcher(s) and independent observer held a meeting where they watched video tape of the teacher in question and further discussed their observations and articulated rationales behind their ranking. Such discussion sessions resulted with 100% agreements on ranking. Since the same ranking procedures were also followed by Omar (2004) the consistency was established between rankings from the first two years and the last year of the project.

**Iowa Test of Educational Development (ITED) Standardized Test Scores**

The ITED test is a district-wide, annually-administrated standardized test that was used in the participating school district. The test was administrated around the same time every year (in February) across the district. The ITED covers a wide variety of subject areas, such as vocabulary, reading comprehension, spelling, capitalization, math concepts and
estimation, math problem solving and data interpretation, social studies, reference materials, and science. While science as a subject is assessed for grades 4 to 12, the middle and high school levels of the test’s science assessment emphasize the methods and processes used in scientific work. Further, the ITED targets knowledge and skills of selected content areas such as life science, earth and space sciences, and physical sciences. Students are required to use the concepts and principles of science to explain, infer, hypothesize, measure, and classify.

Data Analysis

Analysis of Observational Data

The aim of observational data was to assess the quality of teacher-oriented pedagogy reflected by their implementation of the SWH approach as an inquiry-based approach. Teachers’ first 2 years of implementation levels were drawn from Omar’s study (2004), and the same norms were used for third-year implementation level evaluation. As pointed out above, through on-site and video recording observations, based on 3 criteria—dialogical interaction, focus of learning, and connection—developed to distinguish the quality level of teacher implementation of the SWH approach, a composite ranking score was constructed for each teacher.

In total, 4 observers worked on observation data and ranked teachers: 3 researchers and 1 independent rater, a retired science teacher who had previously implemented the SWH approach. Each observer ranked the teacher implementation on a 10-point ranking, with 1 indicating the lowest and 10 the highest level of implementation for the first and third criteria. Further, for the second criteria, evaluators used a 5-point ranking, with 1 indicating the lowest and 5 indicating the highest level of implementation. Adding the scores of the
second and third criteria and averaging this score with the first criterion score, a 10-point ranking implementation level score was created for each teacher. Additionally, the same scales were used to calculate the percentage of interrater reliability with the independent rater.

Table 2. Outline of Levels of Implementation Characteristics

<table>
<thead>
<tr>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dialogical Interactions</strong></td>
<td>• Communication is mostly from teacher to student, but rarely from student to student</td>
<td>• Communication is usually from teacher to student, but occasionally from student to student</td>
</tr>
<tr>
<td></td>
<td>• Teacher uses IRE pattern (initiates, responds, evaluates) of questioning</td>
<td>• Teacher asks open-ended questions</td>
</tr>
<tr>
<td></td>
<td>• Known answer</td>
<td>• Teacher response to students’ answers are non-evaluative but also non-probing</td>
</tr>
<tr>
<td><strong>Management (Focus of learning)</strong></td>
<td>• Teacher plans only whole-class instruction</td>
<td>• Teacher plans whole-class instruction, but occasionally uses small groups</td>
</tr>
<tr>
<td></td>
<td>• Teacher has difficulty with unexpected results</td>
<td>• Teacher begins to accept unexpected results</td>
</tr>
<tr>
<td></td>
<td>• Teacher-centered</td>
<td>• Teacher-centered, but occasionally student centered</td>
</tr>
<tr>
<td></td>
<td>• Teacher-controlled</td>
<td>• Displays little confidence in SWH process</td>
</tr>
<tr>
<td></td>
<td>• Displays little confidence in SWH process</td>
<td>• No student sharing of knowledge</td>
</tr>
</tbody>
</table>
Table 3. Outline of Levels of Implementation Characteristics

<table>
<thead>
<tr>
<th>Connections</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Teacher does not recognize opportunities to make connections outside science</td>
<td>• Teacher recognizes opportunities to make connections outside science, but doesn’t follow through</td>
<td>• Teacher recognizes opportunities to make connections outside science and capitalizes on them</td>
<td>• Teacher gathers opportunities to make connections outside science, but doesn’t follow through</td>
</tr>
<tr>
<td>• Connection to big idea is absent or difficult to see</td>
<td>• Connection to big idea is mechanical</td>
<td>• Connection can be seen from beginning to end and articulated by students</td>
<td>• Connection can be seen from beginning to end and articulated by students</td>
</tr>
<tr>
<td>• Science activities do not promote big ideas</td>
<td>• Science activities promote big ideas in a vague way</td>
<td>• Science activities promote big ideas clearly and extend students’ learning</td>
<td>• Science activities promote big ideas clearly and extend students’ learning</td>
</tr>
<tr>
<td>• Language activities are add-ons</td>
<td>• Language activities flow naturally throughout the SWH approach</td>
<td>• Language activities, both planned and unplanned, evolve and enrich learning as a result of SWH activities</td>
<td>• Language activities, both planned and unplanned, evolve and enrich learning as a result of SWH activities</td>
</tr>
<tr>
<td>• Teacher does not build or activate student prior knowledge</td>
<td>• Teacher builds or activates student prior knowledge but does not use information to make instructional decisions</td>
<td>• Teacher builds or activates student prior knowledge with evidence of using to make instructional decisions</td>
<td>• Teacher builds or activates student prior knowledge with evidence of using to make instructional decisions</td>
</tr>
<tr>
<td>• Assessment does not align with intended or taught curriculum</td>
<td>• Assessment somewhat aligns with intended or taught curriculum</td>
<td>• Assessment aligns clearly and strongly with intended and taught curriculum</td>
<td>• Assessment aligns clearly and strongly with intended and taught curriculum</td>
</tr>
</tbody>
</table>

In the range of implementation scores, using this method, teachers were clustered in 3 categories, low (1-4 points), medium (5-7 points), and high (8-10 points) levels of implementation. While the distinctiveness of each implementation level was reported in detail by Omar (2004) and others, due to the scope of this study a brief outline of cluster characteristics is provided below in Table 1 (Akkus, 2005; Gunel, Hand, & Norton-Meier, 2005; Omar & Gunel, 2004).
**Statistical Analysis**

To ensure the accuracy of the data that were collected, both frequency distributions and descriptive statistics were obtained using the SPSS Frequencies procedure (Mertler & Vannatta, 2002). The SPSS Explore procedure was employed to examine whether outliers possibly could affect the results of the study. To control for other variables that might impact students' science achievement, covariates are crucial. Therefore, analysis of covariance (ANCOVA) was chosen as a statistical method, whereby students' ITED science scores were the dependent variable, teachers' level of implementation for each year was the independent variable, and students' performance on other subject areas, such as ITED math and ITED social science, were the covariates.

In this study, we reported effect sizes to recognize the magnitude of intervention on students' learning. There are three advantages of reporting effect sizes. First, reporting effect size makes meta-analyses possible for a given report. Second, effect size reporting allows a researcher to determine more appropriate study expectations in future studies. Third, reporting and interpreting effect sizes facilitate assessment and comparison of a study's results across existing related studies (Wilkinson & Affairs, 1999).

**Assumptions**

To ensure the validity of statistical results, three general statistical assumptions were examined in this study with analysis of variance (Mertler & Vannatta, 2002 pp.341-42):

- Normality: The assumption that each variable and linear combination of variables is normally distributed.
• Linearity: The assumption that there is a straight line relationship between two variables.

• Homogeneity of variance: The assumption that the variance in scores for one continuous variable is roughly the same for all other continuous variables.

*Participating Teachers*

While all teachers participated during the first 2 years of the project, 2 of the teachers were assigned to different positions during the last year of the project. For those 2 teachers, data were not available. A summary distribution of students, grade level by teachers (Table 2), and a brief introduction to each participant are provided below.

Table 4. Participating teachers and student information.

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Grade</th>
<th>Area</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>10-11</td>
<td>Biology</td>
<td>NA</td>
</tr>
<tr>
<td>Gary</td>
<td>10-11</td>
<td>Biology/Chemistry</td>
<td>NA</td>
</tr>
<tr>
<td>Danielle</td>
<td>9-10</td>
<td>Physical science</td>
<td>97</td>
</tr>
<tr>
<td>Matt</td>
<td>7-8-9</td>
<td>General Science</td>
<td>105</td>
</tr>
<tr>
<td>Bill</td>
<td>8</td>
<td>General Science</td>
<td>111</td>
</tr>
<tr>
<td>Nick</td>
<td>7</td>
<td>General Science</td>
<td>124</td>
</tr>
</tbody>
</table>

*John*

John is the most experienced science teacher, and has taught high school biology for 28 years. Before joining his current school 18 years ago as a biology teacher, he taught 5 different science subjects for 10 years in a small school district in Iowa. In addition to his bachelor's degree in biology, John has a master's degree in education. John participates in school extracurricular activities, coaching football during the fall and basketball during the
winter. John taught biology, including the topics of ecology, cells, cell energy, and genetics, to 10th and 11th grade high school students.

Gary

Gary is an experienced biology teacher with 16 years experience teaching grade 10 biology and grade 11 chemistry in his current school. He has a bachelor’s degree in zoology and a master’s degree in education. Gary worked as a manager of a swine farm operation for 13 years before he began his teaching career. He was holding the responsibility of head of the science department at the school. While Gary taught chemistry at the grade 11 level, he used the SWH approach in his biology classrooms with the topics of ecology, cells, and genetics units with mainly 10th grade students distributed over 4 class periods.

Danielle

Danielle is an experienced teacher who has been teaching physical science to 9th graders for 11 years. She holds a bachelors’ degree in extended physical science. Before she became a teacher, she worked in the food service industry for 13 years. She taught friction, spring balance and the force of gravity, speed and acceleration, and physical properties units to 6 classes of 9th and 10th graders. She decided to work collaboratively with Danielle, who is one of the new science teachers in her school. Although both teachers agreed on the big idea and the conceptual questions, they ended up implementing different SWH activities.

Bill

Bill is an experienced teacher who has been involved in classroom teaching for 27 years. He had taught several grade levels from the upper-elementary to high school level and
several disciplines, including physics, chemistry, biology and field biology, physical science, earth science, life science, general math, 8th grade math, and health and physical education. In addition to his bachelor's degree in physical education with a minor in science, he has a masters' degree in education administration of secondary education. Bill has been teaching science and math in his current school for 24 years. He implemented consecutive units of astronomy and machines, and work and energy with the SWH approach to 6 class periods of 8th graders.

**Matt**

Matt is an experienced science teacher, with 21 years experience in teaching middle school science and high school biology. He has a bachelor's degree in biology science and a master's degree in education. He has been teaching 7th and 8th grades in his current school for 14 years. Matt participates in many school extracurricular activities, such as coaching football, wrestling, and track. Matt started to participate in the SWH approach project by implementing a unit on force and motion in 7th grade science.

**Nick**

Nick is a young teacher with only 7 years experience teaching. He has a bachelor's degree in elementary education, and began his career by teaching science to 6th and 7th grades in his current school. In addition to his bachelor's degree, Nick recently received a master's degree in educational leadership, and later became a school principal. Nick participates in many school activities, coaching boy's basketball and baseball. Although Nick felt uncomfortable with implementing the innovation at the early stage, he decided to contribute to the SWH approach study by implementing two units: solutions and
classification as part of 7th grade general science. Nick later used the SWH approach as a main tool for his instruction.

Results

Results of analyses were reported case by case for each teacher. Prior to conducting any ANCOVA, possible violations of assumptions were investigated. A simple graphical method and normal probability plots of model residuals, along with the Kolmogorov-Smirnov test, were used to examine the normality assumption for every teacher's data files by year. The linearity assumption is addressed by plotting standardized residual values against the predicted values. Examination of the Normal Q-Q Plots obtained through the SPSS Explore procedure enabled the researcher to monitor the patterns of lines in which resemblance of linearity was judged. The analysis results of linearity and normality assumptions and transformation procedures were not reported, unless one or both are violated. Finally, the homogeneity assumption was examined by using Levene's test for equal variances within each ANCOVA analysis. Levene's test results were provided with each set of analysis of covariance results.

Results for John's Case

At the first year of the project John was ranked as a medium implementer, with a composite score of 7. At the second and third years of the project his level of the SWH approach implementation was high, with composite scores of 8 and 9, respectively.

While standardized test scores for all years of the project were obtained for John, the baseline year's scores were not available due to technical difficulties. Testing company failed to provide results for this particular year. Analyses focused on scores from the 3 years that
John was involved with the project. A one-way ANCOVA model was estimated with ITED Science Scores as the dependent variable, year the ITED exam was taken (Years) as an independent variable, and ITED mathematics (ITED Math) total, language total (ITED Lang), and social studies (ITED Sstudy) scores as covariates. Results indicated that all 3 covariates, ITED Math ($F(1, 254) = 23.299, p < .001, \eta^2 = .084$), ITED Lang ($F(1, 254) = 17.999, p < .001, \eta^2 = .066$), and ITED Sstudy ($F(1, 254) = 96.617, p < .001, \eta^2 = .276$), did significantly influence ITED Science Scores. Yet, the main effect for the variable Years was not significant ($F(2, 254) = 2.093, p = .125, \eta^2 = .016$) (see Table 3 for adjusted means and standard errors). Finally, an analysis of pairwise comparisons showed no significant mean differences among scores obtained in different years. Mean Square Error was 400.481, and adjusted $R^2$ was .763 for this model. Finally, Levene's test of equality of error variance showed non-significant results ($F(2, 257) = 1.661, p = .192$), which confirms the assumption that the error variance of the dependent variable is equal across groups.

Table 5. Descriptive statistics for ITED Science Scores for John

<table>
<thead>
<tr>
<th>Year ITED taken</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002-2003</td>
<td>87</td>
<td>281.330</td>
<td>2.165</td>
</tr>
<tr>
<td>2003-2004</td>
<td>81</td>
<td>282.466</td>
<td>2.316</td>
</tr>
<tr>
<td>2004-2005</td>
<td>92</td>
<td>287.147</td>
<td>2.110</td>
</tr>
</tbody>
</table>

Results for Gary's Case

At the first year of the project Gary was ranked as a medium implementer, with composite score of 6. During the second year of the project he decided not to collaborate with researchers. Although he continued to use the SWH approach in his classrooms and researchers observed him, Gary was not provided with the researchers' feedback. During this
year it was observed that his overall pedagogical practices moved down toward a low level, with composite score of 4. At the third year of the project, Gary again was eager to work and communicate with researchers. In this year, his level of the SWH approach implementation was high, with composite score of 8.

While standardized test scores for all years of the project were obtained for Gary, the baseline year's scores were not available due to the same reason explained above, as in the case of John. Analyses focused on scores from the 3 years that Gary was involved with the project. A one-way ANCOVA model was estimated with ITED Science Scores as the dependent variable, year that the ITED exam was taken (Year) as an independent variable, and ITED mathematics (ITED Math) total, language total (ITED Lang), and social studies (ITED Sstudy) scores as covariates. Results indicated that 2 of the 3 covariates, ITED Math \((F(1, 258) = 22.762, p < .001, \eta^2 = .081)\), and ITED Sstudy \((F(1, 258) = 128.446, p < .001, \eta^2 = .332)\), did significantly influence ITED Science. Yet, neither the third covariate, Lang \((F(1, 258) = .881, p = .349, \eta^2 = .003)\), nor the main effect for year \((F(2, 258) = 1.474, p = .231, \eta^2 = .011)\) was significant (see Table 4 for adjusted means and standard errors). Finally, through the analysis of pairwise comparisons, no significant mean differences were detected among scores obtained in different years. Mean Square Error was 535.908, and adjusted \(R^2\) was .677 for this model. Finally, Levene’s test of equality of error variance showed non-significant results \((F(2, 261) = 1.121, p = .327)\), which confirms the assumption that the error variance of the dependent variable is equal across groups.
Table 6. Descriptive statistics for ITED Science Scores for Gary

<table>
<thead>
<tr>
<th>Year ITED taken</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002-2003</td>
<td>98</td>
<td>284.422</td>
<td>2.357</td>
</tr>
<tr>
<td>2003-2004</td>
<td>73</td>
<td>280.959</td>
<td>2.723</td>
</tr>
<tr>
<td>2004-2005</td>
<td>93</td>
<td>287.275</td>
<td>2.438</td>
</tr>
</tbody>
</table>

Results for Danielle's Case

At the first year of the project Danielle was ranked as a low implementer, with composite score of 2. During the second year of the project she improved her implementation and moved up to medium level, with composite score of 6. However, the last year of the project Danielle decided not to use the SWH approach and not to collaborate with researchers. Researchers were allowed to observe her classroom, although feedback was not provided to her. During this year it was observed that Danielle's overall pedagogical practices regressed back to a low level, with composite score of 3.

Both standardized test scores for all years of the project and scores for the baseline year were available for Danielle. A one-way ANCOVA model was estimated with ITED Science Scores as a dependent variable, year that the ITED exam was taken (Year) as an independent variable, and ITED mathematics (ITED Math) total, language total (ITED Lang), and social studies (ITED Sstudy) scores as covariates. Results indicated that all covariates, ITED Math ($F(1, 375) = 39.580, p < .001, \eta^2 = .095$), ITED Lang ($F(1, 375) = 33.750, p < .001, \eta^2 = .083$), and ITED Sstudy ($F(1, 375) = 87.928, p < .001, \eta^2 = .190$), did significantly influence ITED Science. Further, the main effect for year ($F(3, 375) = 3.254, p = .022, \eta^2 = .025$) was also significant (see Table 5 for adjusted means and standard errors).
The significant main effect of year was due to students at the second year of the project (2003-2004), $M = 281.903$, $SD = 2.099$ outperforming all other years: baseline year (2001-2002), $M = 274.720$, $SD = 2.165$, ($t$ (199) = 2.382, $p < .05$), the first year (2002-2003) $M = 273.844$, $SD = 2.076$, ($t$ (207) = 2.990, $p < .05$), and the third year (2004-2005) $M = 274.128$, $SD = 2.487$, ($t$ (178) = 2.389, $p < .05$). Mean Square Error was 452.342, and adjusted $R^2$ was .662 for this model. Finally, Levene's test of equality of error variance showed non-significant results ($F$ (3, 378) = .821, $p = .483$), which confirms the assumption that the error variance of the dependent variable is equal across groups.

Table 7. Descriptive statistics for ITED Science Scores for Danielle

<table>
<thead>
<tr>
<th>Year ITED taken</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-2002</td>
<td>97</td>
<td>274.720</td>
<td>2.165</td>
</tr>
<tr>
<td>2002-2003</td>
<td>105</td>
<td>273.844</td>
<td>2.076</td>
</tr>
<tr>
<td>2003-2004</td>
<td>104</td>
<td>281.903</td>
<td>2.099</td>
</tr>
<tr>
<td>2004-2005</td>
<td>76</td>
<td>274.128</td>
<td>2.487</td>
</tr>
</tbody>
</table>

Results for Matt's Case

At the first year of the project Matt was ranked as a medium implementer, with composite score of 5. During the second year of the project his implementation level moved down, to low, with score of 3. Matt was reluctant to improve his teaching in this specific year. While he continued to use the SWH approach and allowed researchers to observe him and provide feedback, he preferred not to try and change his practices. At the last year of the project he moved from teaching 7th grade to 9th grade at the high school. Even though his last year of teaching was observed and students' test data for him teaching in this grade level were available, in this analysis the data from the third year were not included due to nonequivalence of grade levels.
Standardized test scores for the first 2 years of the project and for the baseline year were available for Matt. A one-way ANCOVA model was estimated with ITED Science Scores as the dependent variable, that year that the ITED exam was taken (Year) as an independent variable, and ITED mathematics (ITED Math) total, language total (ITED Lang), and social studies (ITED Sstudy) scores as covariates. Results indicated that all covariates, ITED Math \(F(1, 320) = 37.211, p < .001, \eta^2 = .104\), ITED Lang \(F(1, 320) = 3.964, p = .047, \eta^2 = .012\), and ITED Sstudy \(F(1, 320) = 120.044, p < .001, \eta^2 = .273\), did significantly influence ITED Science. Further, the main effect for year \(F(2, 320) = 11.302, p < .001, \eta^2 = .066\) was also significant (see Table 6 for adjusted means and standard errors).

The significant main effect of year was due to students at the first year of the project (2002-2003), \(M = 278.269, SD = 1.786\) outperforming all other years: baseline year (2001-2002), \(M = 265.911, SD = 1.882, (t (218) = 4.763, p < .05)\), and the second year (2003-2004) \(M = 272.872, SD = 1.869, (t (219) = 2.088, p < .05)\). Further, the performance difference between the second year and the baseline year was also significant \(t (209) = 2.624, p < .05\). Mean Square Error was 365.155, and adjusted R\(^2\) was .640 for this model. Finally, Levene’s test of equality of error variance showed non-significant results \((F (3, 323) = 1.575, p = .209)\), which confirms the assumption that the error variance of the dependent variable is equal across groups.

Table 8. Descriptive statistics for ITED Science Scores for Matt

<table>
<thead>
<tr>
<th>Year ITED taken</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-2002</td>
<td>105</td>
<td>265.911</td>
<td>1.882</td>
</tr>
<tr>
<td>2002-2003</td>
<td>115</td>
<td>278.269</td>
<td>1.786</td>
</tr>
<tr>
<td>2003-2004</td>
<td>106</td>
<td>272.872</td>
<td>1.869</td>
</tr>
</tbody>
</table>
Results for Bill's Case

At the first year of the project Bill was ranked as a low implementer, with composite score of 4. During the second year of the project there was a slight improvement on his implementation level that moved him up to medium level with score of 5. Due to his position change to a principal at another school building, data for the final year of the project were not available for Bill.

Standardized test scores for the first two years of the project and scores for the baseline year were available for Bill, as in the case of Matt. A one-way ANCOVA model was estimated with ITED Science Scores as a dependent variable, the year that the ITED exam was taken (Year) as an independent variable, and ITED mathematics (ITED Math) total, language total (ITED Lang), and social studies (ITED Sstudy) scores as covariates. Results indicated that 2 covariates, ITED Math \( F(1, 324) = 54.272, p < .001, \eta^2 = .143 \), ITED Sstudy \( F(1, 324) = 108.185, p < .001, \eta^2 = .250 \), did significantly influence ITED Science, whereas ITED Lang \( F(1, 324) = .252, p = .616, \eta^2 = .001 \) was not significant. Further, the main effect for year \( F(2, 324) = 9.194, p < .001, \eta^2 = .054 \) was also significant (see Table 7 for adjusted means and standard errors).

The significant main effect of year was due to students at the first year of the project \( (M = 268.903, SD = 2.034) \) and at the second year of the project both \( (M = 266.524, SD = 1.988) \) outperforming the baseline year \( (M = 257.597, SD = 1.988) \), \( t(217) = 3.984, p < .05 \), and \( t(217) = 3.124, p < .05 \), respectively. Mean Square Error was 413.583, and adjusted \( R^2 \) was .661 for this model. Finally, Levene's test of equality of error variance showed non-
significant results ($F(3, 327) = 1.580, p = .208$), which confirms the assumption that the error variance of the dependent variable is equal across groups.

Table 9. Descriptive statistics for ITED Science Scores for Bill

<table>
<thead>
<tr>
<th>Year ITED taken</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-2002</td>
<td>111</td>
<td>257.597</td>
<td>1.943</td>
</tr>
<tr>
<td>2002-2003</td>
<td>108</td>
<td>268.903</td>
<td>2.034</td>
</tr>
<tr>
<td>2003-2004</td>
<td>111</td>
<td>266.524</td>
<td>1.988</td>
</tr>
</tbody>
</table>

Results for Nick's Case

At the first year of the project Nick was ranked as a medium implementer with composite score of 5. During the second year of the project there was an improvement on his implementation level that moved him up to a high level with score of 8. Nick stayed in high implementation level during the last year of the project as well, with average score of 8. Yet, Nick was appointed as a halftime principal in the third year of the project. Due to time commitments to his new position, more than half of his classes were taught by a substitute teacher who was did not implement the SWH approach.

Standardized test scores for all years of the project and scores for the baseline year were available for Nick. A one-way ANCOVA model was estimated with ITED Science Scores as the dependent variable, the year that the ITED exam was taken (Year) as an independent variable, and ITED mathematics (ITED Math) total, language total (ITED Lang), and social studies (ITED Sstudy) scores as covariates. Results indicated that all covariates, ITED Math ($F(1, 396) = 33.979, p < .001, \eta^2 = .079$), ITED Lang ($F(1, 396) = 13.227, p < .001, \eta^2 = .032$), and ITED Sstudy ($F(1, 396) = 109.112, p < .001, \eta^2 = .216$) did
significantly influence ITED Science. Further, the main effect for year \((F (3, 396) = 19.903, p < .001, \eta^2 = .131)\) was also significant.

Table 10. Descriptive statistics for ITED Science Scores for Nick

<table>
<thead>
<tr>
<th>Year ITED taken</th>
<th>N</th>
<th>Adj. Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001-2002</td>
<td>124</td>
<td>241.050</td>
<td>1.943</td>
</tr>
<tr>
<td>2002-2003</td>
<td>95</td>
<td>253.586</td>
<td>2.193</td>
</tr>
<tr>
<td>2003-2004</td>
<td>98</td>
<td>262.717</td>
<td>2.149</td>
</tr>
<tr>
<td>2004-2005</td>
<td>86</td>
<td>247.021</td>
<td>2.290</td>
</tr>
</tbody>
</table>

The significant main effect of year was due to several significant performance differences among students who took the test in different years (mean scores and results of pairwise comparison are provided in Tables 9 and 10). First of all, students who participated in Nick’s classes in the first year of the study outscored those at the baseline year and the third year of the project. Students from the second year scored significantly higher than all other years, that is, baseline, first year, and last year. Finally, students at the last year of the project scored significantly higher than did students in the baseline year. Mean Square Error was 446.105, and adjusted \(R^2\) was .668 for this model. Finally, Levene’s test of equality of error variance showed non-significant results \((F (3, 399) = 2.697, p = .046)\), which confirms the assumption that the error variance of the dependent variable is equal across groups.
Table 11. Pairwise Comparisons for Nick’s Students

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean</th>
<th>Difference (I-J)</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>Baseline year</td>
<td>12.535*</td>
<td>217</td>
<td>4.273</td>
</tr>
<tr>
<td>Year 3</td>
<td>Baseline year</td>
<td>6.564*</td>
<td>179</td>
<td>2.065</td>
</tr>
<tr>
<td>Year 2</td>
<td>Baseline year</td>
<td>21.667*</td>
<td>220</td>
<td>7.479</td>
</tr>
<tr>
<td>Year 1</td>
<td>Baseline year</td>
<td>9.132*</td>
<td>191</td>
<td>2.980</td>
</tr>
<tr>
<td>Year 3</td>
<td>Baseline year</td>
<td>15.696*</td>
<td>182</td>
<td>2.998</td>
</tr>
<tr>
<td>Year 3</td>
<td>Baseline year</td>
<td>5.971*</td>
<td>208</td>
<td>1.998</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the .05 level.

Effect Size Results for all Teachers

In this analysis effect size was reported using Cohen’s $d$ index, which is widely used in social science, because it enables us to measure “the difference between two means expressed in standard deviation units” (Sheskin, 2000, p. 835). The criteria for identifying the magnitude of an effect size were: (a) a small effect size is between .2 and .5 standard deviation units; (b) a medium effect size is between .5 and .8 standard deviation units; and (c) a large effect size is .8 or more standard deviation units (Rosenthal & Rosnow, 1984; Sheskin, 2004). Effect sizes smaller than .2 standard deviation units are named trivial (Kulik, 2002). Table 10 provides summary results of effect sizes for all teachers.
<table>
<thead>
<tr>
<th>Teacher</th>
<th>Compared Years</th>
<th>Change of implementation level</th>
<th>Cohen’s $d$</th>
<th>Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>1\textsuperscript{st} to 2\textsuperscript{nd}</td>
<td>Medium to high</td>
<td>.1</td>
<td>Trivial</td>
</tr>
<tr>
<td></td>
<td>2\textsuperscript{nd} to 3\textsuperscript{rd}</td>
<td>High to high</td>
<td>.2</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 3\textsuperscript{rd}</td>
<td>Medium to high</td>
<td>.3</td>
<td>Small</td>
</tr>
<tr>
<td>Gary</td>
<td>1\textsuperscript{st} to 2\textsuperscript{nd}</td>
<td>Medium to low</td>
<td>-.1</td>
<td>Trivial</td>
</tr>
<tr>
<td></td>
<td>2\textsuperscript{nd} to 3\textsuperscript{rd}</td>
<td>Low to high</td>
<td>.3</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 3\textsuperscript{rd}</td>
<td>Medium to high</td>
<td>.1</td>
<td>Trivial</td>
</tr>
<tr>
<td>Daniela</td>
<td>Baseline to 1\textsuperscript{st}</td>
<td>Traditional to low</td>
<td>0</td>
<td>Trivial</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 2\textsuperscript{nd}</td>
<td>Low to medium</td>
<td>.4</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>2\textsuperscript{nd} to 3\textsuperscript{rd}</td>
<td>Medium to low</td>
<td>-.4</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>Baseline to 2\textsuperscript{nd}</td>
<td>Traditional to medium</td>
<td>.3</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 3\textsuperscript{rd}</td>
<td>Low to low</td>
<td>0</td>
<td>Trivial</td>
</tr>
<tr>
<td></td>
<td>Baseline to 3\textsuperscript{rd}</td>
<td>Traditional to low</td>
<td>0</td>
<td>Trivial</td>
</tr>
<tr>
<td>Matt</td>
<td>Baseline to 1\textsuperscript{st}</td>
<td>Traditional to medium</td>
<td>.6</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 2\textsuperscript{nd}</td>
<td>Medium to low</td>
<td>-.3</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>Baseline to 2\textsuperscript{nd}</td>
<td>Traditional to low</td>
<td>.4</td>
<td>Small</td>
</tr>
<tr>
<td>Bill</td>
<td>Baseline to 1\textsuperscript{st}</td>
<td>Traditional to low</td>
<td>.5</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 2\textsuperscript{nd}</td>
<td>Low to medium</td>
<td>.1</td>
<td>Trivial</td>
</tr>
<tr>
<td></td>
<td>Baseline to 2\textsuperscript{nd}</td>
<td>Traditional to medium</td>
<td>.4</td>
<td>Small</td>
</tr>
<tr>
<td>Nick</td>
<td>Baseline to 1\textsuperscript{st}</td>
<td>Traditional to medium</td>
<td>.6</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 2\textsuperscript{nd}</td>
<td>Medium to high</td>
<td>.4</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>2\textsuperscript{nd} to 3\textsuperscript{rd}</td>
<td>High to high</td>
<td>-.7</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Baseline to 2\textsuperscript{nd}</td>
<td>Traditional to high</td>
<td>1</td>
<td>Large</td>
</tr>
<tr>
<td></td>
<td>1\textsuperscript{st} to 3\textsuperscript{rd}</td>
<td>Medium to high</td>
<td>-.3</td>
<td>Small</td>
</tr>
<tr>
<td></td>
<td>Baseline to 3\textsuperscript{rd}</td>
<td>Traditional to high</td>
<td>.3</td>
<td>Small</td>
</tr>
</tbody>
</table>

Baseline indicates the baseline year, 2001-2002, 1\textsuperscript{st} indicates first year of the project, 2002-2003, 2\textsuperscript{nd} indicates the second year of the project, 2003-2004, 3\textsuperscript{rd} indicates the final year of the project, 2004-2005. Baseline year implantation recognized as a traditional.
Discussion

Before moving to the discussion of the results, the researcher would like to point out 3 major limitations that surround this study and the generalization of the results obtained. The first limitation is the measurement tool. While there is recognition and widespread use of standardized tests for measuring learning outcomes, standardized tests cannot capture the complexity of the learning process and the intricacy of the learning environment. Standardized tests are limited in length and in complexity to measure skills, understanding, and factual knowledge. Yet, many researchers argue that standardized tests are stronger measurement tools than are other alternative assessment tools, such as teacher-generated tests. Further, the ITED, in particular, targets not only content areas such as life science, earth and space sciences, and physical sciences, but it also accents the methods and process skills used in science, such as the nature of science and science inquiry. While ITED is limited in terms of measuring number of the science inquiry skills, it does focus on measuring skills such as critical thinking and problem solving.

The second important limitation is the lack of uniformity of students’ ITED data. While the baseline year data were available for most teachers, such data were not available for the two grade 10 teachers, John and Gary. Likewise, ITED data from the last year of the project were not available for Bill and Matt. Analysis with those areas of missing data prevented a comparison of all teachers’ progress before and during the SWH project. Further, due to incomplete data, the ability to draw a general conclusion about relationships between teachers’ implantation levels and students’ academic progress was limited.

The final limitation is subjectivity bias. The prolonged interaction between the researcher and the participating teachers may have influenced the researcher’s view and
driven conclusions. However, the researcher attempted to limit this effect by using multiple data collection methods, triangulation of the data, and scoring from independent observers. Further, one would argue that subjectivity, in some degree, is part of all research. The crucial part of research is methods followed to cope with possible subjectivity. As a final point, the issue of the validity of coding instruments needs to be stated as a limitation. While the instrument is constructed over the years with theoretical and pragmatic underpinnings, as an instrument it is created and influenced by researchers' views and beliefs. Further, the matrix has been constructed specifically for the SWH approach implementation so the validity and reliability of the instrument outside of the SWH settings remains unanswered.

In discussing the findings the researcher would like to point out that this study addresses a gap in the current literature in relation to the effectiveness of in-service programs and their impact on student learning. While some reported studies have correlated teachers' responses to surveys and questionnaire items with their students' academic achievements, other studies have associated classroom observation findings with students' achievement. Nonetheless, the literature is deficient in studies that show a relationship among in-service program, teacher implementation level of the program recommendations, and impact on students' academic achievement. Even though the scale of this study was not very large, the methods followed and data used do provide some evidence of the importance of examining the relationship described above.

In addressing the first research question, the concept and pattern of teacher change, findings of this study pointed out important considerations. In general, the teacher change process is dependent upon the individual teacher and is not viewed as a linear process. Results of the study indicated that implementation varied across the group of teachers. In
some cases teacher implementation of the SWH approach improved throughout the period of the study, resulting in these teachers being rated as high level. There also were cases in which teachers were ranked at a medium level in early years but moved down to a low implementation level. In some cases, teachers struggled to make progress in improving their implementation and they stayed at the same implementation levels over the years.

Such findings align with the literature in the area of teacher change and professional development. For example, Luft (2001) and Omar (2004) have suggested that the nature of teacher change is complex, not linear, and different for every individual. Other researchers indicate that quality, time, structure of the program, and teachers' inclination have the largest impact on change in general (Kennedy, 1998; Smith, Hofer, Gillespie, Rowe, & Solomon, 2003).

In relation to the first research question, the author argues that the SWH in-service program did have an impact on participating teachers' pedagogical practices. The majority of the participating teachers improved their pedagogical practices of implementing science inquiry through the SWH over the 3-year professional development program. The model of workshop, observation, feedback, and reflection appeared to help the teachers engage with some form of understanding of the required pedagogical practices. The author would suggest that in those situations where there was retrogression from the original implementation practices, the three teachers involved were hindered or affected by personal, institutional, and other factors. While this study did not collect data on particular issues facing these three teachers, the earlier study by Omar (2004) did provide evidence of personal and institutional factors.
In addressing the second research question, findings of this study suggest that as teachers' implementation levels of the SWH approach increased, their students' achievement on the standardized test also increased. Three of the four cases for which baseline scores were obtained showed significant improvement in student scores in the first year of implementation. This trend continued into the second year, with three of the four teachers improving in implementation and having improved student scores. However, in the case of Matt, who retrogressed in his level of implementation, his student scores went down significantly. This same result was evident in year 3, when Daniela retrogressed back to a low level of implementation. The reduction in student scores also was evident in Nick's case, even though he was rated as high implementation. During this year he became acting principal; thus, while his teaching was rated high, over half the time his science classes were taught by a substitute teacher who was not trained in the SWH approach. This person, while trying to maintain Nick's approach, was much more didactic in her approach. However, the effect of shifting from traditional approaches to high-level implementation of the SWH approach can be seen in Nick's case, where the difference in baseline to the 2nd year of implementation was a 1.0 effect size gain.

Due to difficulties in obtaining baseline data, comparisons to the baseline are not possible for John and Gary. However, the same trend of movement to high levels of implementation resulting in small effect on student ITED scores achievement was evident for both of them. Similarly, movement to a lower level results in a small negative effect on students' scores, as in the case of Gary. Again, because of problems obtaining the baseline year scores, the researcher was unable to compare the effect size difference for these teachers' traditional strategies and their implementation of the SWH approach. The
researcher would speculate that given the effect size increases for the other 4 teachers, significant deviation on students' standardized test scores would be likely when John's and Gary's implementation levels of the SWH approach and their traditional way of teachings were compared.

In cases such as Daniela's and Nick's (with substitute teacher) third year, and Matt's and Gary's second year, where teachers' implementation level retrogressed but they continued to use a student template as in the form of cookbook activity without necessary pedagogical implementations, students' test scores went down significantly. Such findings point out the importance of the entirety of the SWH student and teacher templates. Using the student component of the SWH without adopting the teacher component did not impact students' performance. Further, as Hand, Hohenshell, and Prain (2004) argued, embedded writing-to-learn activities requires good pedagogy; that is, using the student template of the SWH approach as an embedded writing activity still necessitates appropriate pedagogy. When using the SWH approach, which incorporates both the pedagogical component (teacher template) and the structured writing approach (student template) under the overarching frame of science inquiry strategies, it can be seen that both components of the approach need to be implemented adequately and synchronized to have an impact on students' learning outcome.

Given the limitations described above, this study does begin to provide a picture of the impact of inquiry-based strategies on students' test scores. While much has been discussed in relation to implementing inquiry strategies as stated by the Standards (NTC, 1996), little evidence has been put forward in relation to quality of implementation and impact on student scores. This study does highlight that good implementation of inquiry-
based strategies such as the SWH approach that scaffold critical thinking and scientific argumentation does have a positive impact on students’ standardized test performance. It emphasizes the need for continuing implementation or practice of these strategies. Movement away from these inquiry strategies resulted in students’ test scores reverting into the level associated with the traditional didactic approach. There is a gap in the literature about what happens when teachers go back to a traditional way of teaching.

While longitudinal studies are difficult to do and rare, this study highlights the need for more such studies. Further, a longitudinal study would enable us to track more teachers across time and to determine if the pattern of results with these 6 teachers is typical. Results from these studies would enable us to reshape and monitor in-service activities that could help maintain and advance inquiry practices within science classrooms. Future research studies in which baseline standardized test scores for all teachers and standardized test scorers obtained through the SWH in-service program are compared and correlated against level of the SWH implementation are needed across the elementary and secondary levels to make richer generalizations possible. Further, such studies need to take place in schools that are culturally and ethnically diverse. Moreover, studies need to develop and implement questionnaire and survey instruments that can bridge observed teacher implementation level, student standardized test scores, and responses to such instruments.
CHAPTER FOUR

General Conclusions, Implications, and Limitations

General Overview

The last chapter of the dissertation, general conclusions, implications, and limitations, first addresses conclusions from the literature review. Emerging themes of the literature on science inquiry, professional development, and implementation quality on students’ learning outcomes are outlined. Second, findings of the study, their implications for practice, and their relation for current literature are discussed. Last, surrounding limitations of the study are delineated.

Conclusions

Literature Review

The literature review for science inquiry and its current stage of classroom use revealed some interesting results. First, implementation of science inquiry is the stated goal in many publications about the reform movement. The vision that has been promoted recently is that to achieve meaningful science learning, students need to engage with hands-on and minds-on science activities in which their prior knowledge is perceived to be a basis for learning. Within inquiry science classrooms the role of the teacher is shifted from an expert transferring information to an inquirer right alongside students who guides, focuses, challenges, and encourages students throughout the process of inquiry. While agencies such as NRC, NSF, and DOE and researchers put a strong emphasis on implementing science
inquiry at all levels of science education, large-scale research studies have indicated that the level of science inquiry implementation in U.S. public schools is not yet at the desired stage.

Second, in general, research studies that investigated the current stage of science education indicated that teachers from elementary, middle, and high school levels have a limited understanding of science reforms and science inquiry. However, teachers' classroom practices more closely represent a traditional didactic approach. Furthermore, regardless of degrees of understanding and preparedness of using inquiry in the classrooms, the majority of the in-service teachers requested professional development programs in which they can extend their understanding of the approach, student learning, and engage with inquiry activities.

The findings in professional development literature are aligned with science teachers' requests for in-service programs. The studies in the area of teacher professional development indicated that to be able to reach a desired stage of teacher change in integrating science inquiry, teachers must be offered high-quality professional development programs in which teachers have a chance to apply their newly learned skills into real classrooms, reflect on such practices, receive continued support from researchers and administrators, and communicate with experts and colleagues about their concerns and progress. Furthermore, scholars in the professional development area indicated that there was a lack of research studies examining the relationship among the structures of in-service programs, the impacts of such programs on teachers' pedagogical practices, and students' academic achievement.

In many cases, the studies investigated how teachers' beliefs, attitudes, and understanding of learning and teaching is changed with in-service programs. Further, some of the studies examined how those participating teachers' classroom practices changed after
their participation in the programs. However, few of these studies examined the impact of the program and the implementation quality on students’ achievement. While the current literature lacks studies that synthesize the three crucial components stated above, current legislative movements and funding sources are requiring projects to investigate those components, especially the effectiveness of the programs on students’ learning outcomes, to some extent.

The effectiveness of the programs or the way they are implemented was investigated by using different approaches. While some of the studies correlated teachers’ responses to surveys and questionnaire items with those teachers’ students’ academic achievement, other studies associated classroom observation findings with students’ achievement. While surveys and questionnaires can provide evidence of implementation of innovation, classroom observation and protocols are known to be more accurate for identifying to what extent innovation is implemented. Nonetheless, the literature is deficient in studies that show a relationship among in-service program, teacher implementation level of the program recommendations, and impact on students’ academic achievement.

**The Study**

The aim of this study was to investigate students’ academic achievement associated with teacher’s level of SWH implementation as an inquiry-based approach. Since characteristics that help to delineate levels of SWH implementation were identified previously by Omar (2004), this study employed her guidelines to identify teachers’ implementation levels. Further, data collected for this study built on Omar’s findings; the same participant teachers from the first two years of the 3-year long professional
development program were tracked in this study. A mixed-research method approach was used, with a case study method exploring the level of SWH implementation and a quantitative method correlating students' standardized test scores with teachers' implementation levels.

Several themes emerge from this study. The first theme is related to the pattern of teacher change as raised with the first research question. In general, teachers' change and the rate of change differ by teachers, and at the individual teacher level change is not a linear process. Results of the study indicated different cases where teachers improved their level of SWH implementation and moved to higher levels over the years. Yet, there also were cases in which teachers were ranked at a medium level in early years and later moved down to low implementation. In some cases teachers struggled to improve their implementation, and they stayed at the same implementation level over time. Such findings are adequately aligned with the literature in the area of teacher change and professional development. Scholars such as Luft (2001) and Omar (2004) suggested that the nature of teacher change is complex; that is, teacher change is not a linear process and is different for every individual. Some other researchers indicate that quality, time, structure of the program, and teachers' inclination have the largest impact on change in general (Kennedy, 1998; Smith et al., 2003).

Another emerging theme to come out of this work is that the SWH approach in-service program did have an impact on participant teachers' pedagogical practices. The majority of the participating teachers improved their pedagogical practices of implementing science inquiry through the SWH approach over the 3-year period of professional development programs. In this study the structure and content of the program appeared effective in achieving the desired teacher change. In literature, researchers argue that content
of the program is the most significant element to generating changes in teacher practices (Kennedy, 1998, 2004; Loucks-Horsley, 1995; Loucks-Horsley et al., 1996; Windschitl, 2003). Within professional development programs, content means the topics such as knowledge about subject matter, knowledge about how to teach specific subject matter, knowledge about how kids learn, or knowledge and beliefs toward teaching and learning, not necessarily content of the specific subject matter.

A further emerging premise is related to student academic achievement and implementation quality, as posed by the second research question. The correlation between teachers’ rankings and students’ standardized test scores suggest that as their implementation levels increased their students test achievements also increased. In general, when teachers moved from a low level of the SWH approach implementation to a medium or a medium level to a high, a significant performance increase occurred in students’ standardized test scores. Previously, results from middle school and high school levels indicated that a strong correlation existed between students’ performance on teacher generated tests and teachers’ implementation level of the (Gunel et al., 2004; Omar & Gunel, 2004). Also, results resembling those stated above, where students’ standardized test achievement show a significant relationship with their teachers’ implementation level of the SWH approach, were obtained at the elementary school level.

Analogous to Omar’s (2004) observational findings about the low level of SWH implementation, analysis of the students’ achievement suggested that low implementation in practice does not deviate significantly from traditional teaching. That is, the findings from this study suggest that students’ performances in low SWH approach implementation classrooms are not significantly different than those students’ performances who participated
in same teachers' baseline year classrooms without using the SWH approach. Interestingly, in the cases of Gary and Danielle, when they decided not to be part of program and preferred to go back their regular way of teaching their students' scores went down to baseline year levels. The other significance of Gary's and Danielle's cases is that those teachers continued to use the SWH approach as a template to fill out without incorporating the majority of the pedagogical practices that are suggested by the SWH approach.

In the case of Nick, his implementation level increased over the years and his students' achievements were higher every other year than the previous year until the third year of the project. While he was ranked as a high implementer in this third year implementation of the SWH approach, his students' scores went down to the level of his first year of implementation level. The possible explanation of this situation is that, in his third year of participation in the program he was appointed as a half-time principal and subsequently a number of his classrooms were instructed by a substitute teacher who was not familiar with the SWH approach. As in the case of Gary and Danielle, Nick's substitute teacher used the SWH approach as a template to fill out during the lab activities. Such findings indicate the importance of the SWH teacher template and associated pedagogical practices that should be interwoven with the SWH student template.

Finally, the pattern of participating teachers' improvement over time demonstrates that teachers need time to become proficient in implementing science inquiry with the SWH approach. The adult learning and literacy research and the professional development literature support this finding. Further aligned with that literature, findings of this study indicate that teachers' willingness and readiness to implement and innovations or new
approaches play crucial roles in the overall change process and, consequently, on students’ academic achievement.

**Implications**

The results of this study provided a number of implications that could be categorized in two branches within the field of science education: for in-service science teachers program and for science teachers.

**Implications for In-service Program**

The SWH approach can serve as a tool and scaffolding for teachers to understand science inquiry and the underpinnings of implementing it. The SWH approach in-service program embraces several inquiry activities, where teachers experience science inquiry as a learner using the SWH approach. Then, teachers reflect back to their experiences as a learner to tease out effective teaching strategies and to question epistemological beliefs about learning and teaching. Further, the study suggests the importance of long-term involvement with the SWH approach tool within the in-service program. The long-term commitment of providing teachers support, strategies, and feedback would help build understanding of the foundation theory of the SWH approach tool and enable the translation into practice. Yet, teacher willingness, readiness, and efficacy can hinder significant improvement in teacher practices.

Given that reaching a significant number of teachers to help them use the SWH approach is not feasible, those teachers who improved their practices and had strong understanding of the underpinnings of the approach can model for other teachers in their buildings and districts. Further, one would argue that the approach can be adopted in pre-
service programs to create a larger impact on teacher practices. Yet, limits of the teacher improvement and the improvement of the student performance through teacher implementation quality are areas that require further research studies.

Future research studies are needed across the elementary and secondary levels in which baseline standardized test scores for all teachers and standardized test scorers obtained through the SWH approach in-service program are compared and correlated against level of the SWH approach implementation to be able to make richer generalizations. Further, such studies need to take place in schools that are culturally and ethnically diverse. Moreover, studies need to develop and implement questionnaire and survey instruments that can bridge across observed teacher implementation level, student standardized test scores, and responses to such instruments.

*Implication for Science teachers*

The SWH approach can be an effective tool to use in learning environments to promote science inquiry and Standard-based reforms. Combining previous research findings and findings from this study, one would argue that students' understanding of science is improved when it is measured with standardized tests and teacher generated tests. Further, research studies that investigate the correlation of implementation quality and measured science understanding and skills are needed. That is, further studies are needed to examine how implementation quality may impact students' content understanding, science reasoning, and problem solving.

Since the number of students and grade levels were limited to middle school and high school, future studies on elementary settings would be beneficial to investigate the impact of
the approach on younger students. Moreover, studies conducted in different cultural, ethnical, and social-economical settings can illuminate the value of using the SWH approach for teachers and students. Finally, if the assumption is that better implementation of the SWH approach yields improved science inquiry skills as mentioned above, one would argue that such skills may impact the next year's standardized test scores. Consequently, there is a need for further studies with sophisticated modeling and methods of analysis to explore the direct and indirect effects of the SWH approach implementation levels over the two or more years.

Limitations

There are several statistical limitations surrounding this study. First, one of the assumptions of the statistical analysis, random sampling, was violated. Sampling for the study was restricted to a voluntary school district and voluntary teachers. While the school district, school, and town that school is located in can be considered typical for Iowa, their characteristics may differ from other districts, schools, teachers, and communities in general. So, generalization of the study findings is constrained by the fact that the sample of the study is not a representation of the population in general, but a representation of the population in Iowa. Second, the sampling of the students was not random. In this study, researchers were not able to randomly assign students to teachers and their classes. Student assignment to class was completed by the school district and assumed to be random. In fact, in educational settings random assignment is rarely possible due to several organizational and cultural restrictions.

The other limitation to the statistical procedure is the measurement tool. While there is recognition and wide range of use of standardized tests for measuring learning outcomes,
standardized tests cannot capture the complexity of the learning process and the intricacy of
the learning environment. Standardized tests are limited in length and in complexity to
measure skills, understanding, and factual knowledge. Yet, many researchers argue that
standardized tests are stronger measurement tools than other alternative assessment tools,
such as teacher generated tests. Further, the ITED, in particular, targets not only content areas
such as life science, earth and space sciences, and physical sciences, but it also accentuates
the methods and process skills used in science, such as the nature of science and science
inquiry. While ITED is limited in terms of measuring number of the science inquiry skills, it
does focus on measuring skills such as critical thinking and problem solving.

Another important limitation is the lack of uniformity of students’ ITED data. While
the baseline year data were available for most teachers, such data were not available for the
grade 10 teachers, John and Gary. Likewise, ITED data from the last year of the project was
not available for Bill and Matt. Analysis with those areas of missing data prevented a
comparison of all teachers’ progress before and during the SWH project. Further, due to the
incompleteness, the ability to draw a general conclusion about relationships between
teachers’ implantation levels and students’ academic progress was limited.

The final limitation is subjectivity bias. The prolonged interaction between the
researcher and the participating teachers may have influenced the researcher’s view and
driven conclusions. However, the researcher attempted to limit this effect by using multiple
data collection methods, triangulation of the data, and scoring from independent observers.
Further, one would argue that subjectivity, in some degree, is part of all the research. The
crucial part of research is methods followed to cope with possible subjectivity.
REFERENCES CITED


Akkus, R. (2005, 6-8 October). The characteristics of the implementation of Science Writing Heuristic in fifth grade classrooms. Paper presented at the North Central Region of the Association for Science Teacher Education (NC-ASTE), Cedar Falls, Iowa, USA.


(Eds.), Inquiring into inquiry: learning and teaching in science (pp. xx, 496 p.).


Learning to Improve Professional Development in Systemic Reform. Teaching and
Teacher Education, 19(6), 643-658.


Teachers' Self-Sustaining, Generative Change in the Context of Professional

Change: A Follow-Up Study of Professional Development in Mathematics. American


College Press.

(Prepared for the National Science Foundation).

Fulp, S. L. (2002b). The 2000 National Survey of Science and Mathematics Education:
(Prepared for the National Science Foundation).

Torrance (Eds.), Knowing what to write: conceptual processes in text production.
Studies in writing; v. 4 (pp. 190 p.). Amsterdam: Amsterdam University Press.

Education, 74(5), 531-540.


Hand, B., Prain, V., & Hohenshell, L. (2001). *Students' perceptions of learning when using planned writing-to-learn science strategies within a year 10-biotechnology class.* Paper presented at the National Association for Research in Science Teaching, St. Louis, MO.


Lemke, J. L. (1994). E-mail message to the Language in Science Learning list.


Student Achievement in Science. *Journal of Research in Science Teaching, 36*(10),
1110-1126.

York: Springer.

ed.). Boca Raton: Chapman & Hall/CRC.


*Journal for Research in Mathematics Education, 26*, 114-145.

Educational Leaders in Mathematics Project* (Document (RIE)). Pennsylvania. S.: National
Science Foundation. TEI-8552391.

Change: A Study of Professional Development in Adult Education*. Cambridge, MA:
National Center for the Study of Adult Learning and Literacy.

Status of High School Chemistry Teaching*. Chapel Hill, NC.: Horizon Research, Inc.

Understanding of Inquiry Science beyond One Idealized Standard. *Science
Education, 87*(4), 490-516.


of Science Education, 4*, 231-240.

Teaching, 37*(9), 963-980.


