1. INTRODUCTION

The National Science Education Standards [NSES] and the National Association for the Education of Young Children [NAEYC] affirm that accessible and high-quality science education for 3- to 10-year-old children is the central underpinning for future science learning (NAEYC, 2001; NRC, 1996). These documents highly emphasize that teaching and assessment should be redesigned to promote student understanding through scientific inquiry. They also point to fact that effective science teaching requires understanding of what students know and need to learn and then challenging and supporting them to learn it well. In the same vein, assessment should support children’s science learning and provide useful information to various stakeholders, including teachers, students, and community.

Teaching and assessment may be practiced in many different ways, depending on the perspective one takes. Every teacher develops a particular way of progressing about the demanding task of teaching and assessment. One introduces a topic by raising questions to capture students’ existing knowledge or misconception(s) about a specific topic whereas another teacher only gives a lecture and then a quiz to evaluate students’ understanding. One puts more emphasis on knowing and learning basic skills in science whereas other implements blended teaching and assessment method not only focusing on fact of knowledge but also on understanding students’ higher-order learning skills in science whereas other implements blended teaching and assessment method not only focusing on fact of knowledge but also on understanding students’ higher-order learning skills in science such as problem solving and critical thinking. Understanding is a complicated term “stand-alone” and self-understanding is, therefore, a complex process that is not easily measured and is directly related to the nature of human mind. The question of How do people think and learn? has been viewed as one of the most influential questions of science since the 20th century (Pellegrino, Chudowsky, & Glaser, 2001).

Researchers from a variety of disciplines—such as anthropology, linguistics, cognitive science, developmental psychology, and biology—have been studying the nature of human mind and developing theories to expand our understanding of such matters: how mind works, how a child develops conceptual understanding, and how people think, know and learn, in particular. Concerning science teaching and assessment, all advances in the theories of knowing and learning, now, assist us to realize diverse perspectives and understandings of human thinking with differing implications for what should be taught and how should be assessed (Greeno, Pearson, & Schoenfeld, 1996; Pellegrino et al., 2001). According to Greeno et al. (1996), the four perspectives on the nature of the human mind have considerable influence in the history of research and theory: Behaviorist, Differential, Cognitive, and Situative Perspectives. The first two perspectives reflect the most widely used traditional teaching and assessment ways. Teachers having these perspectives generally employ teacher- or textbook-based tests to measure student science comprehension. They also put more emphasis on remembering or memorizing scientific facts instead of assessing higher-order skills in science. Alternatively, cognitive and situative learning perspectives reveal the importance of the development of knowledge and sociocultural dimension of learning (Pellegrino et al., 2001). As described by Webb and Mason (2003), these teaching and learning perspectives are Being recognized as alternative views of science teaching practices. Science teaching and assessment of learning should be based on higher-order thinking skills so that students can cope with real-life related problems.
Influence of Science Teaching and Assessment Modalities on Students’ ... (Winking, 1997). These skills in science are seen as the key for learning in that students make certain learning possible and make acts of carrying out certain tasks in science. In this respect, a framework that is still regarded as being functional was devised by Bloom (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956). Concerning his taxonomy of knowledge, summarized in Figure 1, moving from knowledge acquisition (usually considered as a lower order thinking skill) up to analysis, synthesis, and evaluation (often considered as higher-order thinking skills) phases, knowledge progresses from simpler to more complex forms. During this process, we want to foster students’ abilities not only to understand, but also to engage in inquiry and to apply scientific thinking to problems in life. Probably an unquestionable issue, herein, is the fact that science teaching should be based on experimentation (hands-on, minds-on), higher-order process skills, and new discoveries.

![Figure 1. Bloom’s taxonomy of knowledge and associated scientific process skills](image)

The learning perspectives in terms of traditional and alternative forms of teaching and assessments are not seen as independent viewpoints in the NSES and the NAEP because both assist educators to find out what students know and can achieve. However, an increasing number of research have been recommending the use of cognitive and situative views of knowing, herein referring to alternative teaching techniques, in an attempt to expand the knowledge on students’ science learning (NRC, 1999; NAEP, 2001; Pellegrino, et al., 2001; Webb & Mason, 2003). This study therefore is an attempt to examine the influence of alternative and innovative science instruction and assessment techniques on elementary students’ science achievement utilizing the data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K). Specifically, the study addressed the following research questions:

1. What is the association between science teaching factors and students’ science achievement at elementary level?
   a. What is the association between the student level factors including gender and grade level and elementary students’ science achievement?
   b. What is the association between the time allocated for science related activities and elementary students’ science achievement?
   c. What is the association between the degree of emphasis on teaching higher order skills and elementary students’ science achievement?
   d. What is the association between the frequency of using science related instructional activities and elementary students’ science achievement?
   e. What is the association between the frequency of using test-based assessments and elementary students’ science achievement?

2. METHODOLOGY

2.1. Data source and sample

The data of the study is derived from the Early Childhood Longitudinal Study-Kindergarten Class of 1998-99 (ECLS-K), conducted by National Center for Educational Statistics (NCES). ECLS-K is a “multisource, multimethod study that focuses on children’s early school experiences beginning with kindergarten” (US Department of Education, 2000b, p. 1-1). The ECLS-K is a longitudinal study that gathers data from a nationally representative sample of nearly 22,000 children from kindergarten through the eighth grade. The children in ECLS-K came from both public and private schools and diverse socioeconomic and racial/ethnic backgrounds. Parents, teachers, and school administrators are also involved in the study. ECLS-K is not a standardized or mandatory state or national test to be performed. The idea for the study is to produce a comprehensive and reliable data source that can provide researchers, teachers, parents, and policy makers to better portray and to understand children's progress from kindergarten through middle school years. The data set of the current study was released in 2004 for researchers’ use.

The working sample was generated from child data if a child had complete records on study variables in 2nd and 3rd grade levels. The last working sample included 9,463 elementary students and 3,556 teachers. Since ECLS-K is not a simple random sample, which means not all schools, teachers, and children had an equal probability of selection; an appropriate student weight was initially normalized and then adjusted for the analysis. Using SPSS 13.0 (Statistical Program for the Social Sciences), the standard errors were corrected with average root design effect (DEFT) to calculate standard errors, assuming the data were collected with a simple random sample (SRS). In the SPSS, the standard errors were corrected using DEFT. The standard error of an estimate under the actual sample design was approximated with the following formula;

$$SE_{DESIGN} = \sqrt{DEFT \times Variance_{SRS}} = DEFT \times SE_{SRS}$$

After using the appropriate study weight for children data, the sample was nationally representative of 1,205,271 children.

2.2. Study variables

The science assessment domain in ECLS-K study includes items that measure knowledge and skills in the natural sciences. During assessment, it is reported that equal emphasis was placed on earth and space science, physical science, and life science. The variables in the current study included both student and teacher level characteristics. Variables and their operational definitions are listed in Table 1. Third graders’ science achievement is the outcome of this study (dependent variables). The variables related to students’ and teachers’ characteristics were used as predictors (independent variables) in the model.
Influence of Science Teaching and Assessment Modalities on Students’ ...
distance <1) was observed. Regarding assumptions, all predictor variables and the outcome were measured at the interval level. All predictors in the model had some variation in value (see Table 2). There was no evidence of multicollinearity considering linearity statistics (e.g., VIF and the tolerance values). Specifically, the largest VIF was estimated as 1.530 that is less than 10 and the average VIF was 1.17 which is not substantially greater than 1, therefore, there was no cause for concern (Bowerman & O’Connell, 1990; Myers, 1990). Moreover all the tolerance values were greater than 2, which indicates no collinearity problem in the data (Menard, 1995).

### 3.2. Model summary

Regression analysis revealed that the model significantly predicted students’ science achievement, \( F(7, 9455) = 171.84, \ p < .001 \), \( R^2 \) for the model was .113, and adjusted \( R^2 \) was .112, indicating that the proposed model including seven explanatory variables related to student characteristics and teachers’ science teaching modalities explained 11.3 \% of the variability on students’ science achievement.

The list of interpretations for each of the predictors in the model

The following equation reflects estimates, \( \beta \) obtained through the regression model given in Table 3.

Student Science Achievement = \( \beta_0 + \beta_1 \) Gender + \( \beta_2 \) Grade Level + \( \beta_3 \) Science Area + \( \beta_4 \) How Often Science + \( \beta_5 \) Higher-order Skills + \( \beta_6 \) Scientific Activities + \( \beta_7 \) Test-based Assessment

\[ \hat{Y} = 17.53 - 2.346 X_1 + 9.4 X_2 + .724 X_3 + .732 X_4 + 1.43 X_5 + .921 X_6 - 2.319 X_7 \]

According to Table 3, in terms of individual relationships between science achievement and each predictor, all predictors significantly predicted the outcome. Initially, the predicted average science score for male students were 2.46 units higher than that of female students, controlling for all other predictors in the model (\( \beta_1 = -2.346 \); \( t = -12.645, \ p < .001 \)). The estimated science achievement for second graders were lower than that of third grade students by 9 points, controlling for all other predictors in the model (\( \beta_2 = 9.4; \ t = 27.033, \ p < .001 \)). Incorporating science or nature areas into science activities was positively related to student science achievement (\( \beta_3 = .724 \); \( t = -3.606, \ p < .001 \)). In other words, students benefitted more in science when they experienced science in nature areas. There is an expected .7 increase in science achievement when students frequently conducted science related projects (\( \beta_4 = .732; \ t = 9.493, \ p < .001 \)). When teachers put greater emphasis on higher-order scientific process skills in science (i.e., problem solving, data analysis, reasoning, interpreting), their students’ scores on science were higher than that of other students whose teachers did not give importance on using these skills in lessons (\( \beta_5 = 1.430, \ t = 5.918, \ p < .001 \)).

### Table 3. Results of Science Achievement Regression Model

<table>
<thead>
<tr>
<th>Unstandardized coefficients</th>
<th>( t )</th>
<th>95% Confidence Interval for ( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>17.53</td>
<td>1.007</td>
</tr>
<tr>
<td>Gender (( \beta_1 ))</td>
<td>-2.346</td>
<td>.186 -12.645*</td>
</tr>
<tr>
<td>Grade Level (( \beta_2 ))</td>
<td>.724</td>
<td>.201 3.606*</td>
</tr>
<tr>
<td>Science or Nature Area (( \beta_3 ))</td>
<td>.732</td>
<td>.124 5.918*</td>
</tr>
<tr>
<td>How often science (( \beta_4 ))</td>
<td>.724</td>
<td>.289 4.943*</td>
</tr>
<tr>
<td>Higher Order Scientific Process Skills (( \beta_5 ))</td>
<td>.921</td>
<td>.270 3.415*</td>
</tr>
<tr>
<td>Science Enriched Instructional Activities (( \beta_6 ))</td>
<td>.921</td>
<td>.270 3.415*</td>
</tr>
<tr>
<td>Test-based Assessment (( \beta_7 ))</td>
<td>.921</td>
<td>.270 3.415*</td>
</tr>
</tbody>
</table>

When teachers frequently use enriched science activities, their students achieved more in science (\( \beta_6 = .921, \ t = -3.415, \ p < .001 \)). Finally, the results revealed that alternative assessment modes were superior to traditional techniques. More frequently using standardized tests, commercially produced test, and teacher-made tests/quizzes (\( \beta_7 = -2.319, \ t = -12.622, \ p < .001 \)) resulted a decrease in science achievement.

### 4. DISCUSSION AND IMPLICATIONS

Research spanning several decades underline that effective teaching needs to supervise and perform a wide range of activities and skills which have been recognized as potential determinants of students’ academic success. With respect to the teaching of science, scientific thinking skills, and assessment of learning, the results of this study replicate other studies in which innovative science instruction anchored in hands-on activities and performance based assessment was clearly superior to traditional lecture based science teaching in facilitating children’s acquisition of scientific concepts (Chang & Barufaldi, 1999; Dimitrov, 1999; Dori, 2003; Genc, 2005; Martin, Mullis, Beaton, Gonzalez, Smith & Kelly, 1997; Stohl-Hunt, 1998; Wenglinsky, 2000).

Taken together, the hypothesized model described here is intended to develop a sense of what the effects of different science teaching and assessment modalities on elementary science achievement are. More specifically, it is believed that inspecting different examples of science teaching in a large scale data set (ECLS-K) provided insight into practical and innovative aspects of the teaching science such as improving students’ observational skills and reasoning abilities, orchestrating meaningful discussions, and continually providing constructive feedback to learners.

Gender differences in science have been widely investigated in the science education research, comparing on a range of variables such as achievement, attitude, motivation, interest, and performance behaviors. In general, males are found to be more interested and successful in science as opposed to females (Osborne & Collins, 2003; Steinkamp & Maehr, 1983). The student level predictors of the present study also revealed consistent findings with previous research. More specifically, the results
yielded that both boys and third graders had significantly higher mean science achievement than girls and second graders. By the same token, Martin, et al. (1997) reported that girls performed significantly less than boys at both the third and fourth grades internationally and in about half of the TIMSS (Trends in International Mathematics and Science Study) countries. Dimitrov (1999)’s study focused on gender differences as well but he further examined the differential effects of gender, ability, response formats, and learning outcomes on student science achievement. The results showed that boys did better than girls on the open-ended format in physical sciences at the high ability level, but no gender differences were observed in nature of science, earth and space sciences, and life sciences in particular. Furthermore, he reported no gender differences in science achievement for the low and medium ability students. Therefore, it is suggested that the current data may be reexamined in an attempt to determine if ability or performance levels in science is differentiated by gender.

The most important result of this study is the high degree of consistency between students’ science achievement and the frequency of science experience, specifically hands-on activities. Previous studies also addressed this fact that hands-on activities, if regularly incorporated to classroom instruction, enhance cognitive learning (Echevarria, 2003; Freedman, 1997, Gerstner & Bogner, 2009; Stohr-Hunt, 1998). According to Stohr-Hunt (1998)’s study on the relationship between the amount of time students spent experiencing hands-on science and science achievement measured through the National Education Longitudinal Study of 1988, the significant differences existed across the hands-on frequency variable with respect to science achievement. Specifically, students who engaged in hands-on activities every day or once a week scored significantly higher on a standardized test of science achievement than students who engaged in hands-on activities once a month, less than once a month, or never (p.101). On the one hand, Gerstner and Bogner (2009) cautioned that “spontaneously implemented hands-on instructions did not evoke positive effect on achievement scores”, although hands-on approaches motivate students’ science learning (p.2). For this reason, it is suggested that teachers should evolve and incorporate activity-based learning settings in regular plans beforehand, to build up students’ fascination for science.

The findings of the study also supported the common notion in research that conveying higher-order thinking skills lead to improved student science achievement (Chang & Barufaldi, 1999; Wenglinsky, 2000; Zohar & Dori, 2003). The results of Chang and Barufaldi (1999)’s study proved that the problem-solving-based instructional model did significantly improve students’ science achievement comparing to the traditional-lecture approach. They also underlined that application of science concepts rather than rote learning or memorization should be emphasized in the science classroom to help students develop higher-level thinking skills (Chang & Barufaldi, 1999). In other words, teachers should pay more attention to engaging students in discussing, questioning, researching, problem-solving, and communicating their scientific ideas to enhance their learning in science.

On the subject of assessment practices, the data yielded similar results with previous studies as well. Specifically, when teachers tend to implement alternative assessment techniques rather than testing, students benefited more in science (Bol & Strage, 1996; Dori, 2003; Genc, 2005). Meanwhile, previous research investigated the potential impact of high-stakes testing on teaching science in elementary to high schools (Genc, 2005; Hilliard, 2002; Pringle & Martin, 2005; Shepard, 2002). It is underscored that teachers constantly feel the pressure of high-stakes testing, even though they are willing to implement innovative assessment methods. For instance, the majority of science teachers in Florida claimed that the state-mandated accountability tests hindered their actual teaching (Genc, 2005). Furthermore, it is indicated that they are trying to align their classroom instruction with the science standards when aggressively searching for test preparatory materials (Pringle & Martin, 2005). Pragmatically, we must move beyond any type of tests including standardized tests to teacher-produced ones to more authentic measures of student understanding in science such as individual/group projects, self and peer assessments, oral reports, written assignments, and portfolios, etc.

Overall, it is believed that the results of this study, utilizing such a nationally representative data of “ECLS-K”, would help policy makers plan more effective science programs so that young children’s science learning and development would be enhanced. Still, such results could provide evidence-based guidance to teachers for implementing new instructional approaches to early science instruction. It is highly recommended that researchers continue to examine the data and the new releases of it and to generate reliable information for the future.

REFERENCES


Influence of Science Teaching and Assessment Modalities on Students’ ...