A QUASI-EXPERIMENTAL QUANTITATIVE STUDY OF THE EFFECT OF IB ON SCIENCE PERFORMANCE

By

Margaret Irene Healer

A Dissertation Presented in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education in Curriculum and Instruction

UNIVERSITY OF PHOENIX

August 2012

A QUASI-EXPERIMENTAL QUANTITATIVE STUDY OF THE EFFECT OF IB ON SCIENCE PERFORMANCE

by

Margaret Irene Healer

August 2012

Approved:

Wendy Schulze, Ed.D., Chair

Debra Beebe, Ph.D., Committee Member

Heidi Pace, Ph.D., Committee Member

Accepted and Signed:	Wendy Schulze	8/10/12 Date
Accepted and Signed:	Debra Beebe	8/10/12 Date
Accepted and Signed:	Heidi Pace	8/10/12 Date

Jeremy Moreland, Ph.D.
Executive Dean, School of Advanced Studies
University of Phoenix

8/10/2012
Date

ABSTRACT

The purpose of this quasi-experimental quantitative research study was to investigate the effect of participation in the International Baccalaureate (IB) program on science performance. The findings of the 2x3 mixed ANOVA and Eta square analysis indicated a significant difference (in science CSAP mean scores between the treatment group: IB students (*n* = 50) and the control group: non-IB students (*n* = 50) at the 5th through 10th grade level. The analysis of data concluded that although scores declined between 5th, 8th, and 10th grades with IB and non-IB students, a statistical difference was indicated at each level between the two groups: IB and non-IB in the area of science performance as measured by the CSAP assessment. Educational leaders can use the findings of this study to maximize student science achievement. Further research is recommended through a mixed study to determine the effectiveness of participation in the IB Program and a study of specificity of pedagogical strategies used with science performance with a larger sample size of IB and non-IB students longitudinally.

DEDICATION

This dissertation is dedicated to my wonderful friend and partner, Diane, whose unconditional love and patience provided the support I needed to accomplish this goal. To my sons, Christopher and Brian, thank you for your encouragement and understanding throughout this long process that sometimes took time away from you. My parents, Dan and Gloria Good, I thank you for inspiring me to believe that I can accomplish anything that I set my mind to. Thank you to all my family and friends, without your love, assistance, and support this goal could not have been accomplished.

ACKNOWLEDGMENTS

I want to express my sincere thanks and acknowledgment to those who helped me accomplish the dream of completing this dissertation. Special appreciation is extended to my chair, Dr. Wendy Schulze, and committee members, Dr. Debra Beebe, and Dr. Heidi Pace. Thank you for your mentorship and patience in assisting with the completion of this dissertation. Your support encouraged me to persevere to the end of this academic journey. Special appreciation is also extended to my statistician, Sherry Marshall, for your guidance and mentorship throughout my doctoral journey. I also sincerely want to thank my colleagues and friends who supported me throughout my doctoral journey, especially Denine Jimmerson, Donna Mulder, Todd Morse, and Clark Maxon.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

An increasingly, interconnected workforce and digital, global society requires a learner to become well versed in such skills as global literacy, Information Communication Technology (ICT) literacy, social and cross-cultural skills, problem-solving, innovation, and creativity (Partnership for 21st Century Skills, 2011). To accomplish this, educators must meet the changing learning needs of all students through the integration of a 21st century curriculum framework. One premise behind the No Child Left Behind Act (NCLB) is to prepare students to be successful lifelong learners with the skills and knowledge to navigate the future.

Curriculum leaders have to redefine what it means to be a knowledgeable, successful person in the first quarter of the 21st century (Partnership for 21st Century Skills, 2011). According to Kuhlthau, Maniotes, and Caspari (2007), students and staff need to adapt to a world, which evolves daily. Davey (2011) shared schools have a need to prepare students for a challenging future through a constructivist, inquiry approach to learning. The International Baccalaureate Organization (IBO; International Baccalaureate Organization, 2009) proposes to integrate the academic and skill-based learning essential for such a rapid shift in a global society. One requirement for school leaders and educators is to ensure that students can compete in a 21st century international economy as measured by student achievement on state, national, and international assessments (Tarc, 2010). IB provides the kind of framework that supports student achievement, progressive educational reform, and meets the needs of the 21st century learner (Tarc, 2010).

High-stakes concept and skill assessment provides a limited analysis of student knowledge. Boykin and Noguera (2011) shared that comprehensive curriculum models that systemically meet the needs of the whole child have proven successful in increasing student achievement. A need for effective educational practices, particularly in reading, math, and science, are critical as schools dedicate efforts to meeting the NCLB requirements (American Association for the Advancement of Science [AAAS], 2005).

The National Science Teachers Association (NSTA) posed that less content and more of a focus on “process, inquiry, deep understanding, and communicating ideas” supports student success in science (Kuhlthau et al., 2007, p. 102). Partnering with students through pedagogical practices, such as constructivist, inquiry-based learning engages students in taking more ownership in acquiring knowledge and understanding (Prensky, 2011). Inquiry-based instruction supports the construction of knowledge and understanding with science concepts and has shown to be an effective model to engage students (Hermann & Miranda, 2010; Minner, Levy, & Century, 2010; University of Innsbruck, 2011). Scientific concept acquisition progresses over-time through questioning and discovery. However, research has indicated that teachers believe the integration of inquiry as a strategy may preclude students from performing well on state assessments (Hermann & Miranda, 2010).

Chapter 1 summarizes the quasi-experimental research study, including a background of the problem, purpose, and method designed to analyze the longitudinal effect between participation in the IB Primary and Middle Years Program (MYP) and student science achievement. In addition, the significance of the problem to the field of education, the nature of the study, a list of definitions and the research design was evident. The research question guided the method, assumptions, limitations, and delimitations of the study.

Background of the Problem

The current NCLB Act of 2002 requires schools to report student progress based on data-driven instruction and therefore, demonstrate greater accountability (Pritchett, 2005). Successful results on student assessment require a new paradigm including a focus on research based curriculum that supports differentiated instruction and more student partnership in learning. The IB Pre-K-12 continuum supports the acquisition of knowledge accumulated over-time within a particular framework with a common vision. By means of a constructivist approach, students explore the world, seeing life through more than one lens. Students gain an understanding of themselves in relation to time, place, and others through the process of constructing meaning. According to Morse (2007), IB participation supports standardized test performance in math and reading when compared to a more traditional approach to learning.

Andain and Murphy (2008) posited that a rapidly changing world requires a new learner and a new teacher. Educators are “challenged to engage, motivate, and prepare students for their new world” (p. 11). Inquiry-based learning through an Inductive approach may provide students with the necessary challenge needed in a changing world

(Prince & Felder, 2007). A traditional science method from the past used a teacher-guided lecture approach with text books. One solution for addressing one's role in a global society and to achieve excellence is to change one's paradigm about how humans learn (Zhao, 2009). Wiggins and McTighe (2005) shared that curriculum focused on a core set of big ideas within each subject area and across disciplines supports construction of meaning. Walker (2011) discussed the premise that teaching concepts through inquiry and critical thinking supports the development of lifelong learners within a stimulating learning environment.

Silva (2009) noted that models of instruction and assessments that teach and measure both skills and concepts bridge the gap between science knowledge and application. The IB Diploma program provides evidence that science core content within a standards-based curriculum can demonstrate successful student achievement on international evaluations (Silva, 2009). Colorado public IB schools integrate state science standards centering on conceptual ideas within a program of inquiry. Formative assessments for learning transpires throughout the teaching of units of inquiry, whereas summative assessments occur at the end of each unit and through the Colorado Student Assessment Program (CSAP) science tests given during the spring of each school year to all 5th, 8th, and 10th grade students. The National Governors Association (2008) stated that a world class education includes a global measurement of state-level education performance in science, mathematics, and literacy by examining student acquisition of skills and concepts to compete nationally and internationally in a 21st century economy.

In 2005, two different studies evaluated the correlation of the IB Program on math and reading student achievement (Kiplinger, 2005a, 2005b). The first study investigated the longitudinal effects of the IB on reading achievement and growth from 2000-2004 in grades 5 through 8 and grades 8 through 10. The second study explored longitudinal math achievement of 5th through 10th grade students from 2001-2004, as a result of participation in IB. Such studies indicated that student participation in the IB Program has positive effects on academic achievement in reading and mathematics. A modified replication of the two previously mentioned comparable studies correlated the effect of IB to state science assessment data results. This quasi-experimental study adds an element of validity and reliability to previous data. A literature review indicates a lack of research conducted on the relationship between IB Program student participation and achievement in science based on state standardized assessments.

Statement of the Problem

Business leaders throughout the world demand a different kind of employee in the 21st century. The problem this quasi-experimental study addressed is that students in North America are not demonstrating sustained proficiency in science as measured by state assessments. Educators must find ways to address low science performance to meet NCLB requirements by 2014 through innovative means. Leading careers in today's global society require students to have strong math and science skills; however, students in North America are not demonstrating proficiency and application of 21st century skills, specifically those related to science (Hancock, 2009). Fortune 500 company employees share a collective set of attributes that contribute to their success (Richards, 2011). Creativity, critical thinking, teamwork, and collaborative inquiry enable individuals to formulate innovative ideas to problems not even known to exist at this time. The IB Program, including the Learner Profile, compliments this list of Fortune 500 21st century skills with a focus on preparing successful students for the 21st century.

According to Getchell (2010), the United States no longer has a competitive edge in the workforce over other nations in the areas of science and technology. Reports indicate that science competency in the United States is at an unacceptable level of performance compared to other countries (National Commission on Mathematics and Science Teaching for the 21st Century, 2007; Schroeder, Scott, Tolson, Huang, & Lee, 2007). If students are to successfully participate in an international economic environment, the United States must take into account the current pedagogical framework of schools needed to progress toward a different paradigm.

The NCLB Act of 2001 and Colorado Department of Education (CDE; 2010) require all students to show growth in student achievement, as measured by CSAP. NCLB follows such a stance with an overlying goal that all students will be proficient or advanced by 10th grade in all reading, writing, math, and science CSAPs by 2014. In a high stakes accountability environment, schools are reticent to implement a new program without empirical research that supports student achievement (Kaniuka, 2009).

Although research has shown that the goals of IB align with industry requirements for the 21st century employees, no studies about the effect of IB participation on state science performance scores are evident. Educational organizations must find a science model that will address the integration of research-based strategies and concepts within the written, taught, and assessed curriculum framework (Gewertz, 2008). By examining the inferred correlation of participation in the IB Program with state science assessment performance, this study provided school administration with information about the integration of the IB Program, support of student achievement, and long-term growth in science within a school culture and specific framework.

Purpose of the Study

The purpose of this quasi-experimental design approach was to examine whether students fully participating in the IB Primary Year Program (PYP) and MYP demonstrated higher scores in science performance as measured by the Colorado state science assessment compared to students not participating in IB PYP and MYP based upon attendance in a school that met the study criteria through a longitudinal growth perspective. Specifically, an investigation took place into the science achievement composite scores as evidenced by Colorado state science assessment scores of 50 randomly chosen students who participated in IB and 50 randomly chosen non-IB students in 5th through 10th grades in one Colorado school district (see Appendix A and B). Through the data analysis, a comparison was made with the differences in the science CSAP scores of the experimental group that received instruction using the pedagogical framework of the IB Program, and the control group which received instruction using traditional teaching methods.

One possible instructional framework to support increased student achievement on school, district, and state, national, and international assessments in the area of science is the IB Program. The purpose of this quasi-experimental study examined the degree of association between participation in IB and test scores of 5th through 10th grade student achievement in science as measured by the CSAP science data from 2006-2010 using a coefficient of determination. Currently, little prior academic research delineates such influence on student growth of participation in an IB Program that may be associated with science academic performance or longitudinal growth (Gilliam, 1997).

Significance of the Problem

The significance of this study to the field of education addressed the NCLB mandate that all children will meet or exceed state approved standards by 2014. One goal of the NCLB Act of 2002 is for educators to provide students with the proficiency to acquire knowledge and skills to become successful lifelong learners. The state of Colorado student assessment data indicates more than 52% of students are performing below proficient on the science assessment at the 5th, 8th, and 10th grade levels, indicating there is currently a problem with achievement in science (CDE, 2010). Not in doc, either cite or indicate remove). Educators must find ways to address low science performance in innovative ways to support science achievement and future success.

Significance of the Study

A literature review indicated a shortage of research exists on the relationship between the IB Program and student achievement in science at the primary and middle years levels. A connection to a model, such as the IB, which incorporates research-based pedagogical practices and epistemology within the educational field to current policy initiatives promoting student achievement, supports an original contribution to educational leadership. This raises the question of whether the implementation of such facets of pedagogy and epistemology within any school district continuum could achieve the same results with longitudinal student achievement growth in the area of science.

IB research has focused primarily on the IB Diploma Program from 1991-2007 (Hanover Research Council, 2010). Findings indicated that students who participate in the IB Program at the IB Diploma level score significantly higher than their peers on standardized tests and are accepted into more postsecondary institutions (IBO, 2010). One such study explored the achievement of IB students in the discrete scientific disciplines of physics, chemistry, and biology with a comparison to students participating in programs that are deemed more traditional by design (Poelzer & Feldhusen, 2006). Findings concurred that IB students earned higher achievement levels than non-IB students in science performance at the high school level (Poelzer & Feldhusen, 2006); however, to date, no study has explored the effect of participation in IB at the PYP and MYP levels on state science assessment performance.

Significance of the Study to Leadership

The focus of this quasi-experimental study on science curriculum models and pedagogical practices contributed to educational leadership by adding to the body of literature regarding teacher effectiveness and research-based best science procedures (Creswell, 2008). Relationships between the uses of research-based science strategies evident within the IB Program aligned with student science achievement outcomes may provide a framework for administrators to consider in addressing a gap in achievement.

Other studies have explored the use of inquiry as a method to increase student science achievement with positive results (Headrick, 2009; Longo, 2010). Both Headrick (2009) and Longo (2010) focused on inquiry and a constructivist approach to science learning as a pedagogical tenet of the studies. One principal strategy of the IB PYP and MYP Programs is inquiry. The results of this quasi-experimental study may assist school administrators in making a decision about implementation of the IB Program as a viable option to support improved student science achievement. According to Kaniuka (2009), administrators and educators are better equipped to make school reform decisions within their district or schools with the foundation of research-based best practices, models, and curricular frameworks.

Nature of the Study

According to Salkind (2007), a correlation between two variables reflects the variability and common elements. Quantitative research is appropriate to explain a hypothesis about the probable relationship between two variables and using statistics to analyze data (Creswell, 2009). A quasi-experimental design is appropriate when a group has already been pre-determined by certain criteria (Salkind, 2009). For the purposes of this research, pre-determined criteria of participating in IB or not, longitudinal district attendance from 2006-2011, and achievement of valid Science CSAP scores during those years of attendance were taken under consideration. Data collection was retrospective from existing district and state archived data; therefore, control for differences between intervening variables was limited to the data analysis used to examine the effect and relationships between variables and group differences (Muller, Judd, & Yzerbyt, 2005).

A comparison of groups includes statistical analysis and interpretation of data in consideration of the research study hypothesis. An inferred correlation was evident between the two variables upon completion of this study. A stratified sampling of science state assessment achievement scores of students from one Colorado school district who participated in the IB Program from 2006-2011 and those who did not participate in the IB Program from the same district formed the basis for statistical analysis. Such private data were transferred from the Colorado Department of Education (2010) to district data archives.

CDE's stance is that the purpose of assessment should be aligned with the use of the data. CSAP's provide a yearly snapshot of student performance in math, reading, and writing in grades 3 through 10 and in Science in grades 5, 8, and 10. Each assessment provides a view of student understanding and performance on the state standards relative to the discipline assessed (CDE, 2010) each individual year and over time. Currently, state assessment data are not disaggregated by participation in the International Baccalaureate Program at the state level; however, such data organization is by gender, ethnicity, Free/Reduced Meals eligibility status, IEP status, Language Proficiency status, migrant/immigrant status, and Gifted and Talented Status (CDE, 2010). Student achievement measurements at the federal, state, and local educational levels typically involve an analysis of the percentage of students scoring proficient or advanced on assessments (CDE, 2010). A qualitative design using surveys, including trends, attitudes, opinions, behaviors or characteristics are not determined appropriate for this study (Creswell, 2009).

Overview of the Research Method

The design of this quasi-experimental study was a retrospective quasi-experimental research technique to examine if a significant statistical difference existed between the analysis of CSAP science scores of 5th, 8th, and 10th grade students who have participated in IB and the performance of 5th, 8th, and 10th grade students not participating in IB. Measurement of student achievement on an analysis of the mean percentage of 5th, 8th, and 10th grade students of the specific study population achieving scores in science on the CSAP student assessment within one school district in Colorado served as a basis for the study.

The study included a 2x3 mixed ANOVA analysis of relationship between the IB Program and student achievement over a period of 6 years. A quasi-experimental design supported this study because archived data were analyzed from an educational setting that cannot be manipulated at a later time (Osbourne, 2008). The school district of the source of the data removed all identifiers and produced double-blinded data results for analysis.

Overview of the Design Appropriateness

A qualitative approach was not appropriate for this study, as the focus was not inductive and a statistical correlation between the mean in student achievement in science scores and participation in the IB PYP and MYP Programs formed the core of this study. The action of connecting to a central phenomenon did not take place, nor did interviewing or ethnographic research as a part of the study (Creswell, 2008). Statistical analysis supported an investigation into the correlation between two variables; unlike a descriptive analysis of text or trends found in data. This study did not address an exploration or search for understanding of the process of teaching or learning science. Data collection involved a pre-established instrument (CDE, 2010) to measure distinct variables, as opposed to data analysis based on text segments and development of a description of the central phenomena under study (Creswell, 2009).

A pre-experimental design was not appropriate for this study as the treatment had already been administered and the study included a control group with a comparison to a treatment group. This study reflected on data from prior assessments from a retrospective stance. In addition, a true-experimental design was not appropriate for this study as stratification of both the control and the treatment group was necessary to align with the particular criteria of the proposed study. A retrospective quasi-experimental research design was more appropriate for the study because the research data stemmed from the past science state assessment scores of two integral student groups from two different stratifications (Creswell, 2009; Salkind, 2009).

Research Questions

This study was the first to analyze the difference in CSAP science scores between 5th through 10th grades from the perspective of the effect of participation in IB PYP and MYP as compared to those students not participating in the program. The purpose of the quasi-experimental design approach was to evaluate whether students fully participating in the IB PYP and MYP demonstrate higher scores in science performance as measured by the Colorado state science assessment compared to students not participating in IB PYP and MYP based upon attendance in a school that meets the study criteria.

If IB demonstrates a statistical positive significance on student achievement in the area of science, more schools may choose to implement an IB Program as a viable curricular model. However, if IB had little or a null effect on student achievement, school administration may choose not to engage in the process of implementation of such a program. According to Rothstein, Wilder, and Jacobsen (2007), the process of learning builds upon prior learning gleaned from previous grade levels. The research question focused on the statistical significance of the inferred correlation between participation in the IB PYP and MYP Program and science achievement of 5th, 8th, and 10th grade students. The following question guided the quasi-experimental study: Does participation in the IB Program from 5th through 10th grades make a statistically significant difference in state science test performance of students participating in IB compared to those students not participating in IB?

Hypotheses

The previous question guided the design for the hypothesis of this study. Comparing those students participating in IB PYP and MYP to those students who had participated in a more traditional program within a medium-sized school district in Colorado provided a means for correlation of whether the IB PYP and MYP written, taught, and assessed curriculum is making a difference in longitudinal student achievement in the area of science between the 5th through 10th grade levels. The hypotheses stemming from the research question was as follows:

$H1_a$: Students from 5th through 10th grade who participated in IB PYP and IB MYP will indicate a statistical difference in science achievement in CSAP compared to those students who did not participate in IB PYP and IB MYP.

$H1_0$: Students in 5th through 10th grade who participated in IB PYP and IB MYP will indicate no difference in science achievement in CSAP compared to those students who did not participate in IB PYP and IB MYP.

Theoretical Framework

Theories of learning connect to a curriculum framework and instructional strategies that may prove to be important to instructional leaders (Oliva, 2008). The constructivist model supports the construction of meaning that needs to occur to facilitate understanding within the IB Program (IBO, 2010). Such a framework requires teachers to craft a climate that promotes students and teachers to think and explore. Instruction begins with an assessment of students' understanding of a concept (Dewey, 1938). The primary objective of the constructivist approach supports the belief that learners construct meaning of their world by linking new learning to prior knowledge or previous understanding of concepts.

Constructivist theories emphasize a transdisciplinary and interdisciplinary approach to curriculum development that result in learners actively participating in the cognitive learning process through varied disciplines (Schunk, 2004). Within a constructivist model, such as the IB Program, students have choices and work on different tasks. The teacher provides a learning experience that challenges students' thinking and encourages them to rearrange their beliefs (Schunk, 2004). A sense of ownership is evident throughout the process of metacognition. The constructivist and cognitive theories endorse that learning is most effective when a student has an opportunity to connect to that which is in memory (Schunk, 2004). IB teachers help students develop their meta-cognitive skills and develop an understanding about what skills and strategies would be best to learn a particular concept (Schunk, 2004). Perceived relationships between patterns of information support understanding of concepts and ideas. IB teachers organize information around big ideas of concepts designed to engage students (IBO, 2009). As maintained by Marzano (2007), students need to construct meaning of their newfound knowledge to apply such information in another context or to use critical thinking to solve problems. According to Powell and Kalina (2009), the constructivist model is key to educational reform and promotes student-centered learning, collaboration, and learning habits of mind that support acquisition of skills to become a lifelong learner.

Students must face challenges within a world where knowledge develops continuously (IBO, 2009). Within the science domain, students are connecting concepts to a changing world. The IB Program maintains such practices and beliefs support student achievement and growth. The constructivist approach to learning connects meaning to social and physical experiences within the environment (Schunk, 2004). The interrelated theories of Piaget and Vygotsky blend a rationalist and an empiricist approach to learning (Gredler, 2005). Edward L. Thorndike believed that it was important to provide opportunities for learners to make connections through trial and error and inquiry (Schunk, 2004). When an event connects with a satisfying response that supports understanding, the connection strength increases (Schunk, 2004).

Definition of Terms

Analysis of variance (ANOVA) is a test that measures the difference between two or more means (Salkind, 2007).

Causal-Comparative design is a study that involves subjects being assigned to a group based beyond the control of the researcher (Salkind, 2009).

Colorado Student Assessment Program (CSAP) represents the state's accountability measure for 3rd through 10th grade students, and school and district performance related to the state content standards in reading, writing, mathematics, and in select grades in science and is employed to measure adequate yearly progress (AYP) in student attainment toward the No Child Left Behind (NCLB) goal of proficiency for all students by 2014 (Colorado Revised Statute 22-7-102, cited by the CDE, 2010).

Constructivism focuses on the belief that "learners are active constructors of their own knowledge" (Gredler, 2005, p. 84).

Correlation is a statistical measure of the variance of two or more variables.

Correlation coefficients are a measure of variance that can range from -1.00 to +1.00 (National Center for Educational Statistics, [NCES], 2010).

Eta Squared (η^2) is a variance indicator that measures the effect size of a population sample (Salkind, 2009).

Inquiry-based learning is a pedagogical strategy that promotes discovery learning through a constructivist and socio-constructivist model (Dewey, 1938; Eick & Reid, 2002, Kuhlthau et al., 2007).

International Baccalaureate Middle Years Program is a curriculum framework for students 11-16 that supports a connection between the real world and the study of discrete disciplines (IBO, 2009).

International Baccalaureate Primary Years Program is a student-centered holistic framework designed for students ages 3-12 to explore the world through inquiry (IBO, 2009).

International Baccalaureate Program is an international framework designed "to develop inquiring, knowledgeable and caring young people who create a better more perfect world through intercultural understanding" (IBO, 2009, p. 1).

Levene's Test for Equality of Variance is an inferential statistic performed to assess the quality of variance in different samples.

Longitudinal: A survey design that collects data from the same population sample defined by specific indicators over time (Creswell, 2008; NCES, 2010).

NAEP: National Assessment of Educational Progress, a national assessment given to 4th, 8th, and 12th grade students in the area literacy, mathematics, science, and social studies (NCES, 2010).

No Child Left Behind Act of 2001, 2002 provides accountability for results in the area of student achievement and places an emphasis on scientific research in the classroom (U.S. Department of Education, 2006).

PISA: Program for International Student Assessment- An international assessment given to 15-year-old students every 3 years in reading, math, and science (Hassard, 2011)

21st century skill: A set of skills defining success in the 21st century. Such skills include: Problem solving, critical thinking, creativity, and cooperation.

Zone of proximal development theory defined by Vygotsky in 1978 (as cited in Schunk, 2004) as "the distance between the actual developmental level as determined independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 295).

Assumptions

An assumption that the IB philosophy and methods influence student achievement formed the foundation for this study. Effective student achievement stems from valuable IB professional development and a result of teacher utilization of research-based science teaching strategies. Science academic growth measured by state standards-based assessment scores is valid and accurate. In addition, the IB assumes that "science throughout the IB continuum contributes to the development of students as global citizens who will think critically and creatively when solving problems and making decisions affecting themselves, others and the environment" (IBO, 2011, p. 3).

An empirical assumption supported the view that the sample size was adequate to make generalizations to a larger population of students participating in the IB and not participating in the IB. A belief that the test proctors administering the CSAP science assessment tool had the expertise to administer the test and adhere to protocol was upheld. An assumption supported the belief that the district scores within this study are accurate regarding 5th, 8th, and 10th grade students who participated in IB and the cohort who did not. Student science assessment success was not related to the gender, ethnicity, skill, and experience of the teacher. Teacher participation in professional learning community meetings and ongoing staff development workshops provided a viable approach to constructing meaning and effective pedagogical practices to adequately use IB instructional strategies and methodologies to support science conceptual understanding in the classroom.

Scope and Limitations

This quasi-experimental study researched the impact of IB on science academic achievement of 5th, 8th, and 10th grade students who attended an IB school in the same district from 5th through 10th grade, compared to those who had not participated. To begin, the study had the limitation of a pure random selection of IB students not being possible as students had to meet the requirement of participating in an authorized IB Program in the same district from 5th through 10th grade. The teaching experience and gender of the teaching staff was not considered in this study. Student science performance was limited to the results of only the CSAP test at 5th, 8th, and 10th grade with this study between the years 2006 and 2011, which cannot provide every essential piece of information necessary to evaluate school science curricula and programming.

A possibility exists that students in non-IB schools may have been influenced by teachers using similar strategies such as, inquiry-based instruction, effective science instructional strategies, and groupings of students who include a variety of socio-economic levels and academic levels. Another possibility may be that non-IB schools used a pull-out program as a structure for teaching science. Teachers at the 8th and 10th grade level may have not used the IB methods with fidelity throughout the program of study. One advantage to the structure of this quasi-experimental design was the students within the time of the study brought the same demographic background and experiences to the study situation. According to Salkind (2009), such a factor minimizes the intra-individual variability within the study. A predetermination of the group make-up through the use of retrospective data supported no drop-out of study participants.

Delimitations

The research focused on the long-term inferred correlational effect on science performance comparing those students who participated in the IB Program and those students who did not participate from 2006-2010. A conceptual framework of inquiry, constructivism, transdisciplinary, and integrated curriculum approaches were the basis of this study and formed a boundary to other studies outside of the scope. The results were only as good as the science CSAP test; however, the CSAP test has been administered to thousands of students throughout Colorado over the past 5 years. Alignment evaluation of the adopted curricula to the Colorado Model Content Standards at the standard and benchmark level derived from generalizations from Colorado state assessment data.

According to CDE (2010), generalizations from the Colorado state data report were used to evaluate the alignment of the adopted curricula to the Colorado Model Content Standards at the standard and benchmark level.

A delimitation to 5th, 8th, and 10th grade IB and non-IB students who took the CSAP science assessment in a medium-sized suburban school district in Colorado between 2006-and 2011 was evident. The study sample included only those students who participated in an authorized IB Program continuously during the years of the 5th, 8th, and 10th grade Science Assessment.

Summary

The U.S. Department of Education federal mandates have provided an impetus for increased research and best practice in educational practices and curriculum. Such research focuses on addressing student achievement in the area of science, the current gap between students of different ethnicities and socio-economic levels, and the acquisition of 21st century skills. Through a quasi-experimental design, an examination took place between 5th through 10th grade science scores, as measured by the CSAP and participation in the IB Program. Education is essential in preparing individuals for challenges in the future. The IB Program is one such model that may address education for the future challenges within the world.

Chapter 2 presents an exchange of ideas about significant literature related to science student achievement and the IB Program. Effective instructional strategies necessary to address the needs of a 21st century learner, specifically in the area of science were evident in this narrative, along with elements of the IB Program. Such an examination of literature delineates the history of the IB Program and student science achievement from a federal and state educational requirement perspective. In closing, Chapter 2 initiates a comparison between structural, pedagogical, ideological consistency, and the IB Program to research-based best science practices.

CHAPTER 2: REVIEW OF THE LITERATURE

Currently, literature review findings do not indicate evidence of research correlating participation in the IB Program to science performance. Chapter 2 supports a historic overview of the IB Program and science instructional philosophy. An analysis based on a literature review connected to the research question including literature comparing and contrasting different points of view regarding research in the field. The literature review justified and depicted a need for a study on how specific programs may support science achievement, along with the limited amount of information about the connection between the integration of the IB Program as a means to support increased science achievement.

Title Searches, Articles, Research Documents, and Journals

Title searches, articles, research documents, and journals used in the study focused on a review of literature from the IBO and science best practice and related strategies. University of Phoenix online library search engines (EBSCOhost, ProQuest, and Gale Power Search databases) provided access to peer-reviewed journal articles, documents, and texts. The IB Research Organization provided specific inquiries in alignment with the IB. Key words , such as: *International Baccalaureate Continuum (IB), IB Primary Years Program, IB Middle Years Program, science instruction, constructivism, inquiry, and 21st century skills* provided a foundation for the literature review search.

Literature Review

The literature review for the study encompassed two areas: (a) International Baccalaureate framework and (b) science methodologies. The scope of the study explained a historical relationship among the variables and yielded insight into how the IB Program influences student science achievement.

International Baccalaureate Program

The IBO began in 1968 as a way of teaching for peace in Geneva, Switzerland as a nonprofit organization to provide international secondary schools with a comprehensive curriculum respected worldwide at the university level (Van Oord, 2007). The creation for the IB Program originated from a group of teachers from the International School of Geneva. A key focus of the IB Program is to develop international-mindedness (Davey, 2011). There are three program divisions of the IBO: IB PYP, IB MYP, and the IB Diploma Program. The IBO believes that developing into a successful citizen of the planet requires exploring knowledge and world issues from different perspectives and disciplines (IBO, 2009).

IBO Regional offices began in the 1970s and 1980s, the first being in New York City, IB North America. In fall 2011, the North and South American Regional offices merged, became IB Americas', and moved to Bethesda, Maryland. The Latin American region set up office in 1982 in Buenos Aires. The Regional office for Africa, Europe, Asia, and Middle East opened in 1986 and later settled in Geneva and Singapore in 1994.

Early beginnings of the program focused more on pre-university courses and external examination for students. The majority of IB world schools in the United States are state-funded public entities with varied socio-economic levels, compared to similar IB Programs in other parts of the world primarily supported through private funds for an elite group of students (Bunnell, 2009; Hill, 2008). As of 2012, the IB Programs collectively have over 3,000 authorized schools with more than 1,000 of those schools in

North American state or provincial public schools (IBO, 2011). The IB framework supports integration of a school's scope and sequence within the written, taught, and assessed curriculum. According to Hill (2008), from 2004-2006, the IBO branch of the America's targeted a number of low-income schools to form a partnership linking middle and high School frameworks through a Department of Education Advance Placement Incentive (API) grant.

A second API grant extended from 2007-2010 designed to increase the IB Diploma program in a number of Title I schools. Initial results demonstrated that such an endeavor increased the number of low income students graduating with an American and IB Diploma (Hill, 2008). The American Competitive Initiative (ACI) grant from 2007 and President Bush's 2006 State of the Union Address mentioned the IB as a viable program to address school academic achievement reform for diverse learners, increased accountability, and a promotion of proven educational methods for all students.

The IB Program incorporates a backward design model to support curriculum implementation (IBO, 2009). The Backward Design model is a curriculum planning sequence that begins with the educational objective or goal of a lesson or unit of study (Wiggins & McTighe, 2005). Each stage of the model builds upon achievement of the objective. Stage One identifies the enduring understanding from the concept. Stage Two incorporates formative and summative assessment tools to determine acceptable evidence, followed by Stage Three, in which staff determines research-based learning experiences to support understanding. Assessment of prior knowledge sets the stage for new learning and experiences within the lesson framework. Opportunities to explore new ideas and concepts give students a platform to build on understanding and construct new meaning.

Classroom teachers who work in IB schools throughout the world have more autonomy from the stance that they are focusing on in-depth exploration of concepts and skills focused on a transdisciplinary or interdisciplinary set of themes while creating authentic opportunities for students to make connections to their learning (IBO, 2009). The IB Program provides the opportunity for learners to engage in the construction of meaning through a concept-based approach (IBO, 2009). Student-centered learning and inquiry methods emphasize continual interactions between the professor and student.

Teachers who focus on construction of meaning with students within the science realm experience a sense of fascination and awe of student's ability to articulate understanding of science concepts (Jerome, 2010; Longo, 2010). Traditional science instruction incorporates a teacher-directed textbook approach of instruction through lecture and worksheets, with an emphasis on scientific facts. Inquiry-based science instruction is taught from a student-centered approach, including manipulation of materials, peer dialogue, lab work, data analysis and problem solving (Hung, 2009). The use of inquiry provides ownership and responsibility to learn science from both a teacher and student perspective. The IB science continuum supports seamless alignment and articulation between the three IB Programs from a common philosophical perspective (IBO, 2011).

The IB Program incorporates peer dialogue and cooperative learning as ways to solidify a student's understanding. Construction of meaning is achieved through social interdependence in a particular community. Teachers provide a framework for the subject being taught, and students develop questions that will further shape their understanding. In alliance with Gredler (2005), IB believes that “knowledge is transactional; learning socially constructed and is distributed among the co-participants” (Gredler, 2005, p. 85). For example, students work in small groups or pairs developing ways to solve science problems. Understanding is articulated to classmates once confidence in goal accomplishment transpires. According to Hayes-Jacobs (2010), relevant inquiry-based science instruction supports student engagement. Along with their classmates, students try to make sense of dialogue from comments and questions to shift toward understanding of concepts. Students apply such understanding to a real life problem.

The IB Primary Years Program. The IB PYP program framework supports the primary level learner between the ages of 3 and 11. The use of structured inquiry promotes the interaction with the environment physically, socially, and intellectually so that the student can construct and refine meaning and understanding. According to the IBO (2009), a community of learners works collaboratively to develop sustainable dynamic, caring learners who endeavor to make the world a socially, emotionally, and physically healthier place to live. Five essential elements provide the program structure and curriculum framework. Such elements focus on action, knowledge, skills, concepts, and attitudes (IBO, 2009). Each element of the IB PYP Program supports the goal of students becoming a successful global citizen through a transdisciplinary model of inquiry.

The IB PYP (IBO, 2009) science knowledge element supports the view of science as a global study of the natural world, and the linked associations. The PYP concept element drives the inquiry-based concept-driven curriculum framework through a cluster of ideas: "form, function, causation, change, connection, perspective, responsibility, and reflection" (IBO, 2009, p. 98-99). The transdisciplinary skills element supports the higher level learning process of construction of meaning (IBO, 2009). The PYP attitude element supports a frame of mind and approach to learning and life. The PYP attitudes include: "appreciation, commitment, confidence, cooperation, creativity, curiosity, empathy, enthusiasm, independence, integrity, respect, and tolerance" (IBO, 2009, p. 24). The action element of the PYP prompts students to respond to learning through a "choose, act, and reflect cycle" (IBO, 2009, p. 26).

The term transdisciplinary stems from the idea that an investigated concept spans across disciplines with the goal of developing a new vision and understanding of knowledge that connects to the real-world (IBO, 2009). Students explore IB PYP concepts through a set of transdisciplinary themes and actively constructing meaning through inquiry. A set of transdisciplinary skills: "social skills, communication skills, thinking skills, research skills, and self-management skills" (IBO, 2009, p. 21) support acquisition of knowledge and construction of meaning. The IB PYP elements are important to this study as each element forms the foundation for academic success (see Figure 1 and Appendix C).

The IB Middle Years Program. The IB MYP program framework supports the skills and concepts that make up the foundation of the IB Diploma Program through an integrated approach to learning (see Appendix C). According to the IBO (2009), the MYP framework facilitates consistency between the MYP and national and international curriculum and assessment overviews. One method of curriculum and assessment

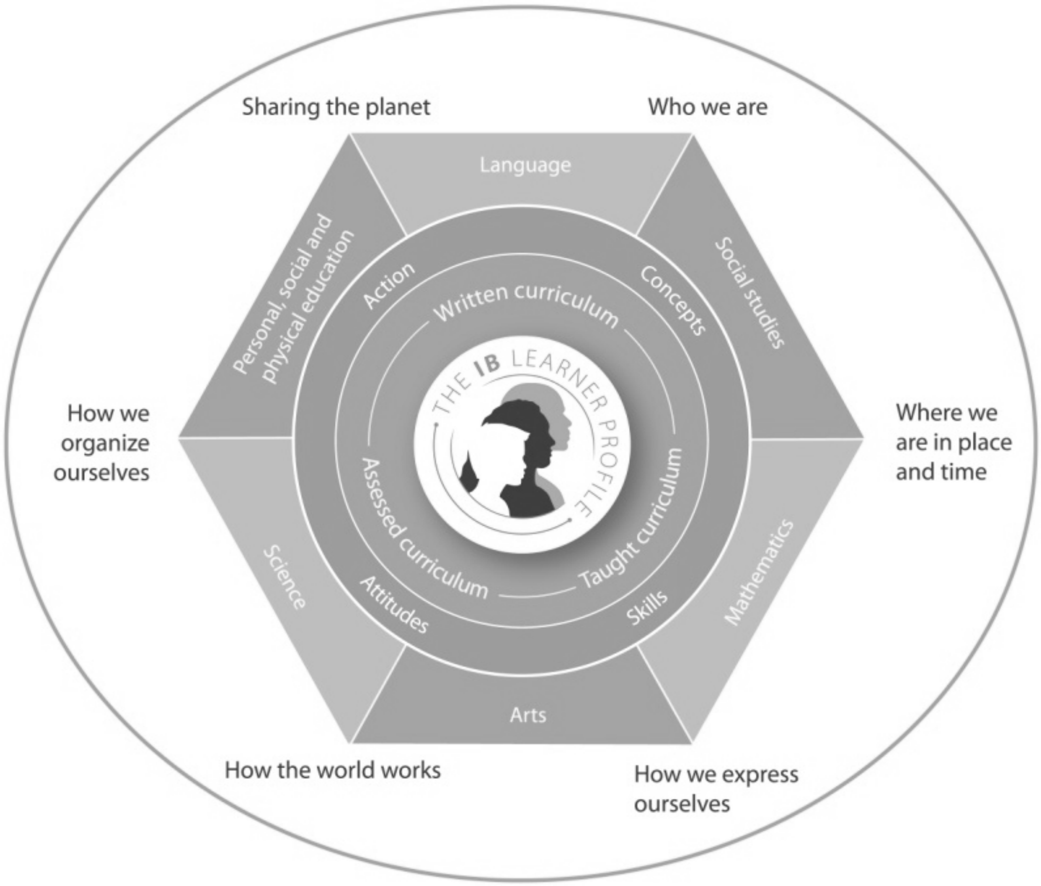

Figure 1. IB PYP Curriculum Cycle (IBO, 2009).

informs another, supporting an acquisition of skills and knowledge relevant within a global environment. Students in the IB Middle Year's Program transition to interdisciplinary learning through an integrated approach to construction of meaning. The term integration stems from the idea that an investigation of a concept takes place through two or more disciplines moving toward students constructing new meaning. The interdisciplinary approach is embedded in the disciplines, unlike a transdisciplinary approach to learning (IBO, 2009). As students transition into the IB Diploma Years, an understanding of the theory of knowledge develops through a focus on the individual disciplines.

The MYP program curriculum cycle supports the middle level learner from 11-16 years. An integration of eight subject groups takes place around five areas of interaction. The five areas support different approaches to learning, including: health and social education, community and service, human ingenuity, approaches to learning, and environments (IBO, 2009; see Figure 2 and Appendix C).

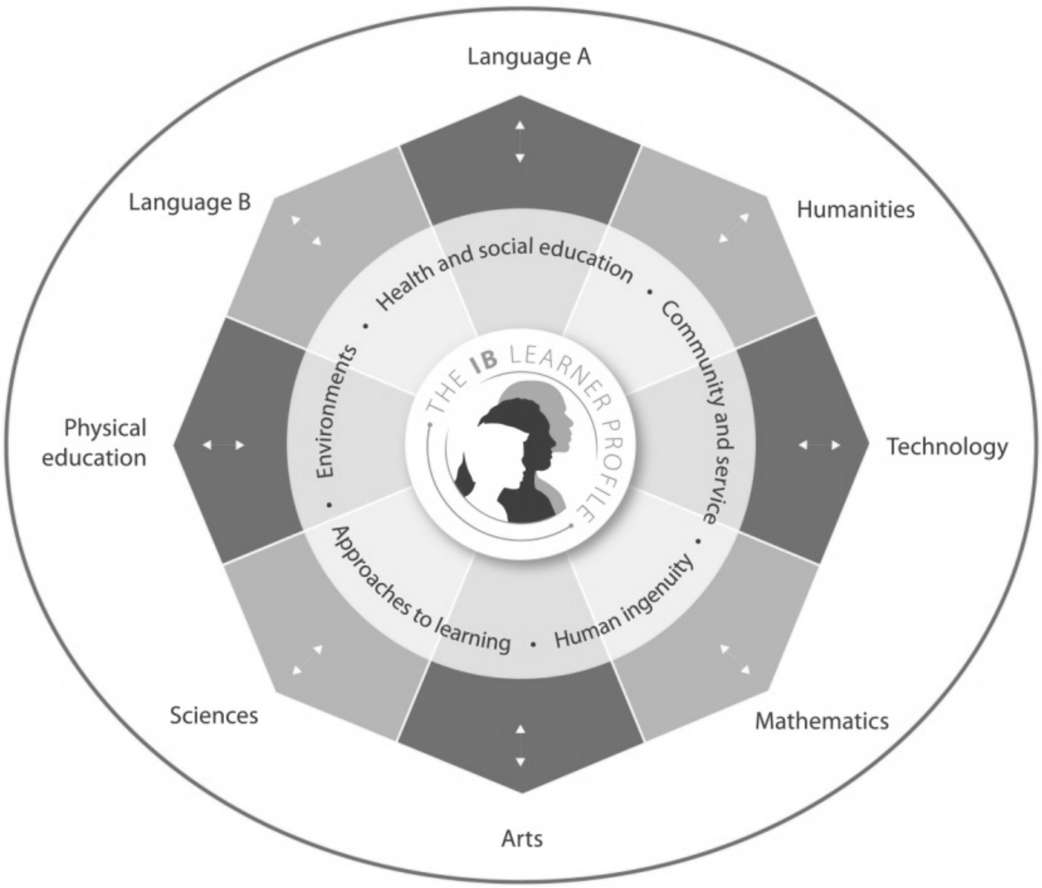

Figure 2. IB MYP Curriculum Cycle (BIO, 2011).

IB and Science. Determining the most effective means to teach science and achieve proficiency in high stakes state assessments can be challenging (Brookhart, 2010). Science instruction and high-stakes assessments are typically not aligned with best practice in science. The IB (IBO, 2011) takes a unique approach to the study of science. A global scientific perspective supports understanding of a diverse world stemming from a variety of beliefs and values. Students explore their world and construct meaning through inquiry and investigation of an array of dimensions that provide a framework to connect theory and practice.

Reflection on scientific knowledge supports the development of global citizenship and critical thinking evolving into positive action. The IBO maintains the assumption that IB is a complement to the American national system (Tarc, 2010). Beginning with the PYP program, students three through eleven explore science concepts from a transdisciplinary approach to gaining understanding. A curriculum that focuses on concepts supports a connection for the learner to construct meaning (IBO, 2009). As students transition to the MYP program, student understanding builds upon science concepts through an integrated model. Upon transition to the Diploma program, students shift into a study of science from a discrete discipline approach with the exception of the environmental studies and society. Students' knowledge-base of science shifts from transdisciplinary, to interdisciplinary, to disciplinary throughout the continuum.

Such a structure and framework sustain a transdisciplinary approach to learning science concepts. Boyer (1995) supported the belief that learning best occurs through the process of making connections across disciplines. Science context scaffolds a transdisciplinary approach to constructing meaning. Students examine and observe living and nonliving items over time and record such data to formulate predictions through the use of scientific vocabulary application, tools, and techniques to articulate understanding and conceptual understanding. Students explore concepts and skills through questioning and posing problems within the context of the natural world (Longo, 2011). An investigation and manipulation of variables support the construction of meaning and the testing of predictions through trial and error. Information gleaned from such an investigation generates conclusions upon interpretation and evaluation of data.

Science skills and concepts build upon each other and spiral throughout the written and taught curriculum (IBO, 2009). Hancock (2009) suggested that traditional teaching practices have not kept up with the demands of advanced technology and global competitor requirements in the mathematics and science disciplines, nor do they provide opportunities for students to relate to a global society. According to Marx and Harris (2006), standards-based science inquiry sustains gains on student assessment through student engagement, motivation to learn and application of concepts. Students are making connections to science concepts through prior knowledge and construction of meaning through inquiry. Investigation of science concepts with peers and teachers, researching and communicating data findings, and inquiring about new ideas supports new learning and advancement in the scientific process (Marx & Harris, 2006). Such strategies are similar to those found in the IB PYP and MYP programs and expanded into discrete disciplines in the Diploma Program (see Figure 3 and Appendix C).

No Child Left Behind (NCLB) Act

The NCLB Act bases student success on proficient or advanced scores on a standardized criterion referenced assessment in reading and math in 3rd through 8th grade. According to Zhao (2009) math, reading, and science have become the most important content of education. Conversely, Marx and Harris (2006) posed that

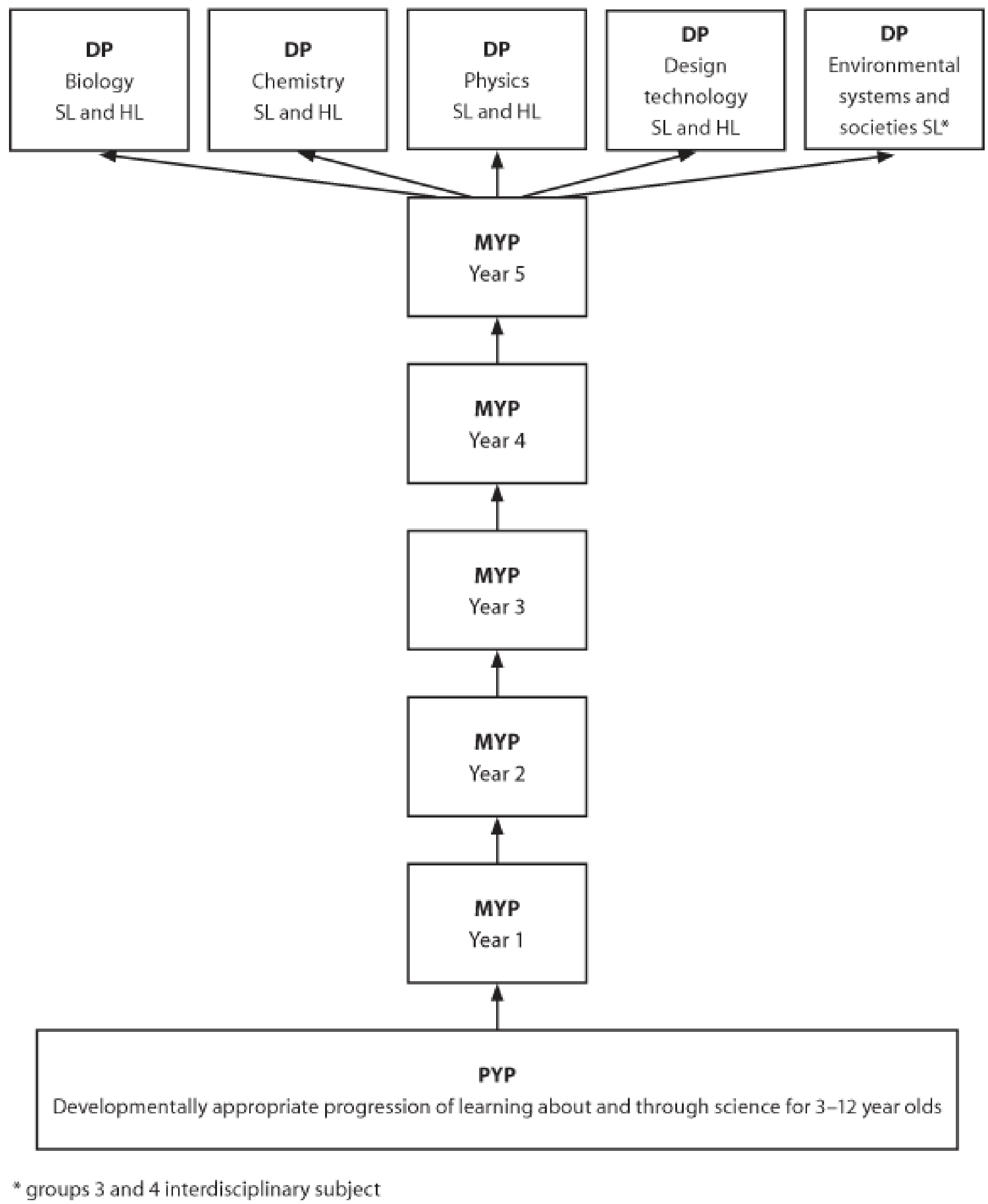

Figure 3. IB Science Curriculum (IBO, 2011).

educators experience conflict by focusing on science inquiry at the expense of reading or mathematics instruction and not doing as well on high stakes assessments. NCLB requires that states measure students' progress in science annually between grades three through five beginning in 2006; however, *NCLB* science is not documented as a portion of AYP reports (Marx & Harris, 2006).

Federal changes have focused on the highly qualified status of teachers. The NCLB era has led educators to link science learning to language arts and math instruction at the elementary level (Lee & Luykx, 2005). According to the U.S. Department of Education (2006), quality science instruction is of critical importance to increased student science achievement. Although teacher quality and preparedness is beyond the scope of the current study, quality influences student achievement.

21st Century Skills

Schools from Pre-K-college can provide and require opportunities for students to pursue their quest for knowledge and thirst to answer personal questions through the attainment of 21st century skills. New ideas and knowledge must include the skills to synthesis and analyze information in a critical manner. According to the Partnership for 21st Century Skills (2011), the acquisition of creativity, innovation, and critical thinking skills creates a better prepared future citizen and scholar. Scott and O'Sullivan (2005) suggested that 21st century skills span across all disciplines and all educational levels and are integrated in such programs as the IB.

Society has a need for restructuring of the learning process, including skills to think critically and effectively use information (Zabel, 2004). Lauer and Yodanis (2004) proposed that schools must prepare students to have the skills to live in a global economy. While there seems to be much debate about what defines success, President Bush (as cited in Zhao, 2009) defined success and excellence as good test scores in math, reading, and science. Bush (as cited in Zhao, 2009) added the means to achieve excellence is through standards and accountability.

Fisher and Frey (2008) posed a continual release of responsibility as a framework to support acquisition of 21st century skills. Guiding students through a framework of learning engagements composed of a balance of collaborative and independent learning, and guided instruction supports student transition as an independent learner. Such a model includes many of the same components as the IB pedagogical framework and current science instruction research. Individual accountability for students includes expanding one's knowledge-base and increasing content mastery. One must acquire skills and knowledge that prove successful in a global economy.

Bybee (2008) claimed full participatory citizenship requires scientific literacy. The American Association for Advancement of Science (AAAS, 2006) organization initiated Project 2061 as a metaphor and vision to reform science, math, and technology education including a reminder that the 21st century will bring profound scientific and technological changes. Students who acquire scientific literacy are more adept at understanding and articulating science information and therefore, more inclined to make informed decisions (Bybee, 2008).

Inquiry Method

Kuhlthau et al. (2007) shared that inquiry requires a student not only to find an answer to a question, but "investigate, explore, search, quest, research, pursuit, and study" (p. 2). Such a process provides greater engagement and ownership of the learner through real life application and involvement in an active process of constructing meaning. The substance of curriculum focuses upon content knowledge, skills, and academic habits that one may need to become a responsible, productive citizen. Essential questions are the heart of such curriculum. A framing of the questions centers on significant content shaped to engage and focus students' inquiry; teacher questions provide the scaffold or backdrop to student inquiry. Teaching units of inquiry begins by posing problems or questions of emerging relevance, which is a guiding principle of constructivist pedagogy (Brooks & Brooks, 2001).

Such questions provide a structure and framework for the unit of study. Not all students arrive in a classroom interested in learning the written or taught curriculum, although a constructivist model can lead one to construct understanding of the importance of such curriculum. Lord and Orkwiszewski (2006) suggested that inquiry supports retention of concepts, critical thinking, and student engagement. As cited in Lindfors' (1999) study on children's inquiry, Piaget's work supports the level of developmental questioning that occurs within the science classroom. According to the National Science Foundation (NSF; 2005), inquiry finds its roots in searching for meaning and new understanding through questioning and exploration of the world. Teachers must make decisions about how to provide such a framework of learning opportunities (Davidson & Carber, 2009).

Constructivism

The constructivist theory prompts much dialogue and debate as a tool for learning. Constructivist theories emphasize a transdisciplinary and interdisciplinary approach to curriculum development that results in learners actively participating in the cognitive learning process through varied disciplines (Schunk, 2004). Within a constructivist model, such as the IB Program, students have choices and work on different tasks at an appropriate cognitive and developmental level. The teacher provides a learning experience that challenges students' thinking and encourages them to rearrange their beliefs (Schunk, 2004). Higher order thinking skills are a tool to transfer information through opportunities to construct meaning. Students can apply new defined knowledge and skills within an innovative context. A sense of ownership is evident throughout the process of metacognition. The constructivist and cognitive theories endorse that learning is most effective when a student has an opportunity to connect to that which is in memory (Schunk, 2004).

According to Brooks and Brooks (2001), the constructivist approach supports critical thinking and making connections to prior knowledge. Much of Brooks and Brooks research focused on Piaget's view of constructivism as a means to construct meaning of one's world. Research supports the use of an integration of critical thinking through the use of higher order thinking skills to increase student achievement (Brookhart, 2010). Collaborative group efforts provide a forum for students to share current understanding of science concepts and construction of new meaning through a metacognitive and hands-on approach. Lamanauskas (2010) maintained that the process of constructivism begs teachers to know the development of knowledge construction at an intimate level. Learning at the most basic levels involves both content and process. An integrated approach to construction of meaning supports such a goal. Motivation and a sense of ownership evolve, as students make sense of the world.

Trial and Error Method

Thorndike was a strong believer in connectionism and learning that occurs by trial and error (Schunk, 2004). Human beings make connections to what they believe to be true and construct meaning through a set of trials that result in a satisfied response by way of a realization of an effect. According to Thorndike (as cited in Schunk, 2004), such an effect may provide a positive or negative consequence. Through the process of inquiry and construction of meaning, students explore the world through a trial and error method and connectionism (Gredler, 2005). Such practice is evident within an IB pedagogical framework.

Social Cognitive Theory

Bandura's social cognitive theory professed that individuals imitate others through modeling and environmental factors on reinforced behavior (as cited in Gredler, 2005). According to Bandura (as cited in Gredler, 2005), the development of learning requires a connection of cognitive processing and decision-making skills of the learner through verbal and visual codes of behavior. Non-verbal mental models and cues support metacognitive processes of understanding. Vgotsky built on such an idea; however, a belief was held that a learner needed to connect to prior knowledge and experience language collaboratively with others at different developmental levels (Gredler, 2005). Student conceptual understanding is built upon articulation of understanding with others and reinforcement from peers and adults.

Interrelationships Between Theories

Much interrelationship occurs between theories within an IB classroom setting. Thorndike's research in the sequence of curricula is an important component of exploration of a concept (Gredler, 2005). Skills and knowledge follow a scaffold through different disciplines to facilitate understanding, often without students even realizing it. The trial and error method endorsed by Thorndike supports a constructivist framework in a constructivist IB classroom. The social cognitive theory and the constructivist theory discuss the importance of social processes in the production of knowledge. Student self-efficacy expands more efficiently through the process of observation and modeling (Schunk, 2004). Teachers' design learning experiences that focus on big ideas and teaching of a particular schema aligned with science conceptual understanding. The teacher is not the center of instruction and includes diverse instructional formats, but instead plays an important role as a facilitator to prompt dialogue and clarify misunderstanding of concepts.

Learning tasks involve real-life problems (Schunk, 2004). The majority of such practices support science conceptual understandings as students construct meaning through inquiry of specific science concepts and skills. Scientific research suggests that student-centered learning along with the learning environment directly affects neuron interaction. Basing learning on experiences provides a means to organize information and construct meaning. Such a theoretical approach not only provides students an opportunity to solidify their comprehension of concepts, but also allows them to apply understanding to real-life experiences. Students experiencing positive results are more willing to take risks with their learning and quickly realize that true learning is often a trial and error process used to construct meaning (Schunk, 2004). The IB Program framework at all three levels, IB PYP, IB MYP, and IB Diploma, begins with a review of students' understanding and preconceptions of explored concepts and relevant vocabulary, followed by a "posed problem of emerging relevance" (Schunk, 2004, p. 245). According to Tarc (2010), the PYP and MYP are well-designed to support increased student achievement reform in schools.

Assessment Findings

As one considers competing in a global world, science and math skills are critical to succeed (Partnership for 21st Century Skills, 2011). The National Center for Educational Statistics (NCES; 2008) indicated that fourth, 8th, and 10th grade students in the United States between the years 1992-2005 suggest no significant growth in science achievement). Although minimal growth was shown at the 4th grade level; the 8th grade level science average score did not change, with the 10th grade science score demonstrating a decline (NCES, 2010). NAEP assesses science concepts through a combination of application of skills and knowledge. An analysis of the 2005 NAEP results indicated that Colorado fourth graders in 2005 scored consistently with the national average in science; however, such scores are basic at best. The Program for International Student Assessment (PISA) assessment indicated in 2006 that American 15-year-olds scored lower than the average of 16 of the other 29 OECD countries and equal to eight other countries in science (Hassard, 2011).

According to Tan and Bibby (2010), a recent study indicated IB students scored higher than their non-IB peers on the International Schools' Assessment (ISA) in Math Literacy, Reading, Narrative Writing, and Expository Writing, with math and expository writing demonstrating the strongest difference. The assessment was cross-correlated with the standards from PISA (Hassard, 2011).

Another recent study from Wilson (2007) compared data from seventh grade IB MYP students as measured by the Terra Nova NCE Achievement test on reading, language, and math with students taught from a more Traditional Academic Program (TAP). Wilson's conclusions reported IB MYP seventh grade students demonstrating higher student achievement on all three assessed disciplines. Batson's (2010) study concurred with previous studies on the positive effects of the IB MYP on student achievement.

The CSAP Science assessment is a criterion reference assessment that aligns with the Colorado state academic science standards. Such an assessment specifically measures 5th, 8th, and 10th grade student performance with respect to those standards, rather than to the performance of peers nationally (NCES, 2008). While Colorado NAEP data from 2009 demonstrated fourth grade students scored higher than 24 states and eightth grade students scored higher than 22 states nationally, according to CDE (2010), 39% and 32% ratings respectively are not high enough to demonstrate strength in science performance. Longo (2010) advocated that educators believe that science achievement growth on state assessments is compromised by the use of creative and meaningful pedagogical practices in the classroom, such as inquiry-based instruction. Research suggests such a perception does not prove to be accurate; in fact, the use of inquiry can increase science achievement based on state assessments and motivate students to take more ownership in learning (Longo, 2010).

Teachers vary relative to instructional fidelity and research-based best practice and know little of the correlation to student achievement and effective programs such as the IB. According to Creswell (2008), one purpose of effective research in the educational field is to add to a body of literature as it relates to research-based best practice with written, taught, and assessed curriculum. A research exploration focusing on a particular educational leadership issue contributes to the field of education and student achievement. Collaborative administrative and teacher leadership may make a positive difference on student achievement in science and other disciplines (Dufour, Dufour, Eaker & Many, 2006). The IB Program embraces the practice of collaboration and teacher leadership with curricular decision-making (IBO, 2009). For the purposes of this quasi-experimental study, the focus was on the instructional strategies inherent to one particular program: the IB Program and correlated best practice strategies within the science discipline.

Marzano (2007) posited schools need more research on effective practices and programs to advance student science achievement. Research supports teachers and students becoming more engaged in using critical thinking and becoming more involved in participating in leadership opportunities with curricular decisions (Marzano, 2007). The IB Program encourages such a paradigm through the process of collaborative decision making about curriculum and instructional practices. Student performance scores on standardized tests have increased with the addition of critical thinking as tool in the classroom (Boyer, 1995).

Conclusions

The premise forming the basis of this quasi-experimental study was that the conceptual philosophy behind the IB Program compliments the theories recommended for science theoretical understanding. Research into science instructional practices supports a constructivist methodology and transdisciplinary approach to learning scientific concepts. However, according to Gilbert (2009), science inquiry methodologies can be challenging to implement within the classroom without a framework or staff development support. As shared by Coke (2008), teachers can integrate best practice into instruction and achieve educational mandates.

Summary

Chapter 2 delineates the connection of inquiry and constructivism to the science discipline. Detailed descriptions of the IB PYP and MYP with a connection to the IB science continuum, including graphic views of such perspective, supported the literature review. Chapter 2 maintains current theories and research in science education and proposes a link to how the IB Program supports science achievement.

Chapter 3 reflects the research method and design appropriateness for the study and the study's population, sampling frame, and data collection and analysis procedures. The internal and external validity were also described along with the scope, limitations, and delimitations of the study. Chapter 3 includes a restatement of the purpose of the problem and an elaboration of rationale for research methods and the design appropriateness.

CHAPTER 3: METHOD

The purpose of this quasi-experimental study examined the effect of participation in the IB Program on 5th through 10th grade student achievement in science as measured by the CSAP. This study conducted in a Colorado suburban school district compared IB students and non-IB students using 2006-2011 CSAP results in science. Currently, little prior academic research delineates such influence on student growth of participation in an IB Program that may be aligned with science academic performance or longitudinal growth (Gilliam, 1997).

According to Rothstein and Jacobsen (2006), learning built upon prior knowledge from one grade or course of study to another is effective for longitudinal growth. The quasi-experimental study examined if an inferred correlation exists between the effect of participation in the IB Program and science performance. Archived science CSAP scores formed the basis of analysis from a medium-sized suburban school district of a cohort of 50, 5th through 10th grade students who participated in the IB and a comparison group of 50, 5th through 10th grade students who did not participate in the IB Program over a period of 5 years.

Chapter 3 provides a detailed analysis of the research design techniques and their appropriateness to the study. In addition, an inclusion of the research design and population sampling, the rationale for the study and the validity of the instrument, and data collection procedures and techniques are evident.

Research Method and Design Appropriateness

An effective research design stems from the premise of a consideration of three elements: philosophical assumptions, a procedure to collect data, and specific research methods (Creswell, 2009). A postpositive worldview supported this study with a theory and collection of data that either supported or refuted the hypothesis. Protection against bias and a built-in control for alternative explanations are important elements included to minimize confounding upon internal validity (Creswell, 2008, 2009; Salkind, 2009). A generalization and replication of the study support further research on the hypothesis. A statistical test measures the tendency for two or more variables to differ in a predictable pattern (Creswell, 2008). A study that incorporates predetermined data as the basis of analysis aligns well with a quasi-experimental research design (Creswell, 2008; Salkind, 2009).

This particular quasi-experimental study began by grouping students from a medium-sized suburban school district into two cohorts based upon whether they were in IB or non-IB between 5th through 10th grades in 2006-2011. Next, 50 students were randomly chosen from each of the cohorts to form the sample population. Due to the specific criteria needing to be met for this study, a pure random sample was not possible. The sample groups were demographically alike through SES, gender, and ethnicity and are more likely to represent the population by ensuring that demographic equivalence was evident from the start of the compilation of data. The Levine's Test for Equality of Variances tested the distribution of both groups. The quasi-experimental study supported analysis and prediction of the variance of the dependent variable through use of an Eta Squared η^2 to see how much variance in the dependent variable was being explained by the group (Salkind, 2009). A 2x3 mixed ANOVA was used to discern the differences between the two groups (IB and non-IB) for 5th, 8th, and 10th grade to establish if the program type made a difference and if that difference also occurred over time.

A quasi-experimental design provided statistical data in determining the effect of the treatment (participation in the IB Program) on the experimental group. A 2x3 mixed ANOVA using archived Science CSAP assessment scores formed the basis for analysis. An analysis of variance (ANOVA) is a helpful statistical tool with a quasi-experimental design when a critical variance, such as academic performance between two groups becomes evident (Salkind, 2009).

Research Questions

The research question focused on the statistical significance of the inferred correlation between participation in the IB PYP and MYP Program and science achievement of 5th, 8th, and 10th grade students. The following question guided the quasi-experimental study: Does participation in the IB PYP and MYP Program from 5th through 10th grades make a statistically significant difference in state science performance compared to those students not participating in IB?

Population

Quantitative researchers select a group from a target population that meets the prescribed criteria of the study (Creswell, 2008). A sample of individual students representative of the population formed the basis of a random selection of 50 students who either participated in the IB Program or 50 who did not during the 2006-2011 school years from a specific suburban Colorado School District. Archived CSAP Science scores served as a vehicle to analyze any measurable difference between those students participating in the IB Program and those not.

Sampling Frame

The definition of a sample, as stated by Creswell (2008), includes a representative portion of a population used in making a generalization in a research study. Salkind (2009) mentioned that random assignment of participants in a research group is an effective method to control for variance within the study; however, matching subjects within a group may prove beneficial to ensure equivalence between groups. For the purposes of this study, a sample is a specific random number of students who have participated in the IB Program or not from 5th through 10th grades in a medium-sized suburban school district and took the state science assessment.

This quasi-experimental study had some components of being a stratified convenience sample because the cohort groups stemmed from a set of parameters or stratification: attendance in district from 2006-2011, participation in IB or not, and evidence of 5th, 8th, and 10th grade CSAP science scores. A specific number of randomly chosen students from each cohort group represented a generalized sampling of the population (Salkind, 2009). According to Creswell (2009), a multistage procedure is appropriate when a study includes first identifying the cohort group, assigning a code for each student, and using such a cluster to correlate information between the control group and the treatment group. Once groups formulated; an assumption that each group will progress or regress equally occurs, so that any variance may connect to participation in the IB Program. Babbie (2007) stated that cluster grouping of sample populations is an effective design component when the researcher does not have knowledge of all of the elements of the population.

Informed Consent

Salkind (2009) noted the critical importance of permission for research studies. Every research project that has human participants must have a document indicating permission to use such data. A permission mandate to access information comes from the district and needed support for collation of data of the private archived information becomes evident. However, these data have anonymity coding. Informed consent from the parents or the students prior to data collection was not necessary because of the double-blind nature of the study. CSAP archived data at the individual district where the test was administered and made known to parents of the participants in the study during the years of 2006-2011 formed the basis for the study. The District of the data origin gave permission for an inquiry of the CSAP data of the specific students that met the specific criteria for the purposes of this study (see Appendix A and B).

Confidentiality

A double-blind codification strategy facilitated the collection of 2006-2011 CSAP data in which the researcher and the district assessment director were unaware of the specific individual student names. The coordinator of data for this study, district assessment director, queried the information about how the subjects of the research were assigned; however, the researcher responsible for collecting data did not have the confidential information and could not provide cues that may impact data results. The district assessment director responsibilities include collation of archived, private data for information and research purposes, such as a retroactive study on science performance. The assessment director provided student numbers aligned with a random set of student scores; therefore, both parties were blind to participation. After 3 years, demolition of the private files of the archived data from this research will take place. There were no FERPA issues with this study because aggregate group data that contains no mention of personally identifiable information about any specific student formed the basis for evaluation.

Geographic Location

A medium-sized suburban school district in Colorado was the source of the data for this research. Two out of 19 elementary schools in the district met the criteria of attendance of students who attended an authorized IB PYP school in 2006 as fifth grade students. Those two schools were matched to two non-IB schools with similar demographic backgrounds. The non-IB schools were matched to the IB schools based upon statistically similar gender, SES as measured by the free and reduced lunch program, and ethnicity data. The middle and high schools that were chosen aligned with those schools that the elementary students articulated to for 6th through 10th grades. At the time of this study, the current district student enrollment was approximately 23,119 students; with a 24.1% minority count at the time of this study. Student demographic make-up included: White-17,541, Asian-853, Hispanic/Latino- 2,730, Black/African American-685, American Indian/Alaskan Native-123, Hawaiian Native or other Pacific Islander-93, and two or more races-1,094. Five out of 19 elementary schools received Title One funds (CDE, 2011). See Table 1.

Data Collection

A study specific medium-sized urban school district provided data from the 2006 CSAP assessments in science for 5th grade students to select randomly the cohort groups for those identified as IB and non-IB students in 5th through 10th grade. Science

Table 1

Demographic Data for IB and Non-IB Schools

Parameter	IB	Non-IB
Gender		
Female	48%	48%
Male	52%	52%
Ethnicity		
White	80%	83%
Non-White	20%	17%
Free and reduced lunch status	4%	3%

Assessment is administered in the state of Colorado in 5th, 8th, and 10th grades (CDE, 2010). A compilation of 2006-2011 CSAP data, using a double blind codification strategy in which the researcher and the district assessment director did not have access to specific student names supporting this study. The individual student science CSAP test score and the Performance Level (PL) was provided for 5th, 8th, and 10th grade randomly chosen IB students and non-IB students for the study (see Appendix D and E). This was important to ensure student confidentiality.

Instrumentation

The instrumentation included in this study was the CSAP science data from 2006-2011. Such data were compiled by the CDE in compliance with the guidelines of the NCLB Act. CDE reports state assessment data on reading, writing, math, and science to all districts in the state of Colorado, the United States Department of Education, holistically by grade level for individual Colorado schools to the media. Assessment reports provided the district and parents of the student individual information about student participation and scores on CSAP. Table 2 shows the CSAP categories and science score ranges for Unsatisfactory, Partially Proficient, Proficient, and Advanced for all 3 testing years.

Table 2

CSAP Science Score Ranges

Performance categories	5th grade 2006	8th grade 2008	10th grade 2010
Unsatisfactory	300-506	300-458	300-468
Partially proficient	507-568	459-506	469-506
Proficient	569-613	507-578	507-580
Advanced	614-900	579-900	581-900

Selection Process of Study Participant Schools

The selection process of study participant groups began with a selection of students who met the prescribed criteria of participating in authorized IB schools and non-IB schools from 2006-2011 and took the CSAP Science test at 5th, 8th, and 10th grade. Relatively even demographic data supported greater external validity. Two cohort groups were formulated from the research population; one group included those students who participated in the IB Program from 5th through 10th grade in a medium size suburban district and the other group is made up of 5th through 10th grade students from the same school district who did not participate in IB. Each group was formed from a random selection of 50 students from each sub-group. A selection of an equal number of participants within the study supports a traditional model of research (Salkind, 2009). This study took under consideration the effect size of the two groups as the IB group was smaller than the non-IB group. According to Litschge, Vaughn, and McCrea (2010), small effect sizes can reflect on a larger population with relative reliability and validity.

Validity and Reliability

According to Trochim (2006), reliability supports the validity of a study to be repeated over time using the same instrument and procedure as used in the original study. The reliability of this study limited itself to the reliability of the established measure. The archived state assessment data formed the basis of analysis and conclusions from this study. Such data elements are those reported to the U.S. Department of Education, CDE, and the local district as a part of evaluating AYP goals required by the federal NCLB Act. Information reported in AYP reports are considered true and accurate under penalty of laws against false reporting.

The IB Program claims to be a program that is appropriate for all students regardless of location or socioeconomic background (IBO, 2009). The degree of generalizability of this study was significant to being able to suggest whether the claim of IB is supported and reliable. The primary research instrument for this study was Science CSAP data results from a particular medium-size suburban school district at the 5th, 8th, and 10th grade levels.

Internal Validity

According to Creswell (2009), internal validity involves those factors that may affect the outcome of an experiment. Such elements can include those involving the participants (history, maturation, regression, selection, and mortality) or those involving the experimental treatment (diffusion, compensatory, resentful demoralization, and compensatory rivalry), or those involving procedures (testing and instruments). Internal validity threats are common in any study involving randomly selected groups (Creswell, 2008). Trochim (2006) posited the following threats to internal validity may exist in any study: testing instrumentation, regression, selection of study participants, mortality and maturation.

A number of patterns existed within this quasi-experimental study because of the nature of maturation of specific experiential events that may cause historical events and family background. Other events may have occurred with students involved in the IB or not which may have affected the science performance. A random selection of participants selected from each subgroup support greater probability of being equally distributed among the two groups and addressing the internal validity type of selection within a specific cohort of students. CSAP science scores presented a view of student achievement as a point of reference and a longitudinal report of performance of 5th through 10th grade students from 2006-2011.

External Validity

External validity supports the degree of generalizability of the study with other populations (Creswell, 2008, 2009). Researchers can support such a goal through drawing upon common elements from each type of setting. This study measured the degree of generalizability of the IB Program as a treatment on science performance using a between group design with a quasi-experimental approach. A retrospective quasi-experimental design uses intact groups as a population for research. According to Neuman (2009), a quasi-experimental design format can provide a researcher a means to test a causal relationship between two or more groups. Creswell (2008) noted that the purpose of an experimental design is to analyze the effect of an action on an outcome.

The parameters of this experiment required students to come from a specific suburban school district in Colorado who participated in IB or not from 5th through 10th grades and participated in the CSAP Science assessment from 2006-2011. Selection, geographic location or setting of study, and history affecting the study may pose a threat to any study involving a comparison between two groups (Creswell, 2008). A quasi-experimental study may experience an element of external validity if a false inference or generalization is made from the data (Creswell, 2008). The process of including a quasi-random approach to the study added another layer of validity. A replication of the study using different levels or settings may support broader generalizations, and therefore, greater validity.

Reliability

According to Creswell (2008), the primary forms of reliability are test-retest, inter-rater reliability, and using a combination of alternate forms and test-retest reliability. For the purposes of this quasi-experimental study, an integration of the test-retest form is evident. Reliability measures the extent that one set of scores from a testing instrument are stable from one testing period to the next (Creswell, 2008). Kiplinger (1998) stated that the CSAP test meets the requirements of reliability with the following indicators: consistently strong relationship of standards and the subject-specific assessment, a strong relationship between standards, historical stability, and consistency between testing periods. CSAP assessments are challenging, developmentally appropriate and provide equal access, unbiased, standardized administration, and make available an accurate measurement of student knowledge (Kiplinger, 2005a).

The state of Colorado shows a reliability rating for the CSAP tests that appear higher than most norm-referenced tests in reading and writing (Kiplinger, 1998). Historically, the CSAP science assessment has followed the same pattern of performance. Longitudinal CSAP data results demonstrated a statistically valid correlation between individual student science tests between the 5th, 8th, and 10th grades (CDE, 2010). The use of an instrument that is appropriate for the population and has been administered with consistent results over time adds a level of internal content validity to the study results (Creswell, 2008; White, 2005).

Data Analysis

Student data from 2006-2011 CSAP results for Science were analyzed to determine significant differences between IB and non-IB students. Exploratory analysis of the data purposefully contributed to an investigation of the significance of the IB program as a potential contributing factor in increasing student achievement. More specifically, the compared data result supported provision of feedback on whether students in the IB track had a CSAP score mean of significance higher than those in the non-IB track between the testing years of 2006-2011.

SPSS (Statistical Procedures for Social Sciences) software made available statistical information to provide descriptive and inferential results of the transferred collected student achievement data. A 2x3 mixed ANOVA, used to measure the overall effect of program type on test scores provided data analysis for statistical significance. In order to use a 2x3 mixed ANOVA, variances between the groups must be equal. Preliminary findings indicated that the variances between the IB and non-IB groups were equal at 5th, 8th, and 10th grade testing years. A grouping variable supported a method to create stratification for two separate cohorts, such as IB and non-IB and the test variable of the CSAP scores provided a quantitative component.

In addition, configuration of an *F* value with an Eta square η^2 supported analysis as to whether or not the difference was statistically significant between the effect size and the grouping variable. According to Green and Salkind (2011), the effect size is the measure of the extent of the difference between two or more groups. While an implied correlation was evident with this study; a correlation analysis would not have been appropriate as this study did not use a multiple regression analysis to see if a combination of independent variables come together to have an effect size on a dependent variable.

A comparison of longitudinal significance within each group across the three testing periods was conducted using a 2x3 mixed Analysis of Variance (ANOVA). An ANOVA statistical analysis provides data for factoring two or more groups, such as IB and non-IB, whereas the dependent variable differentiates individuals on some quantitative aspect--such as the CSAP Science scores (Salkind, 2009). An Eta Square η^2 analysis demonstrated the effect size of the statistical data, instead of the r-square analysis. Table 2 indicates that ranges for the categories of unsatisfactory, partially proficient, proficient, and advanced varied across the three testing years. Preliminary findings for the Levene's test of equality indicated that the four categories of Performance Levels based on State assignment were equal between the two groups for 5th and 8th grade years, but not for the 10th grade testing year. In order to maintain the integrity of the variance of the original data, each state category was divided equally into two sub-categories. The following eight categories were used for all data analysis: upper and lower unsatisfactory, upper and lower partially proficient, upper and lower proficient, and upper and lower advanced.

Summary

Chapter 3 provided a description of the method of the study used to determine if there was a statistical difference between science achievements of 5th, 8th, and 10th grade students participating in the IB Programs and those not participating in IB. The research method chosen for the study was a quasi-experimental design that examined the differences in archived state science achievement scores. Quantitative procedures in the quasi-experimental study involved collecting, calculating and examining the test results (Creswell, 2009).

A quasi-experimental design supported the data collection of comparing CSAP data of two groups. Group one criteria included: participation in IB from 5th through 10th grades in a medium-sized suburban Colorado school district and receiving a score on the Science CSAP assessment in 5th, 8th, and 10th grades. Group two criteria included: not participating in IB from 5th through 10th grades in a Colorado School District and receiving a score on the Science CSAP assessment in 5th, 8th, and 10th grades. Preliminary analysis indicated that the two groups were demographically similar at the beginning of the study.

CHAPTER 4: RESULTS

Chapter 4 presents the findings from the gathered data (Creswell, 2008). The chapter includes a description of the data collection and procedures, the presentation of descriptive data with reference to Tables 3 and 4 with an explanation of the results of the analysis, according to the research design and method. Additionally, Chapter 4 includes an explanation of the significance of the figures and tables depicting the results of data analysis.

This quasi-experimental study examined if participation in an authorized IB Program supported greater student achievement in science performance based on results from a state standards-based assessment. An analysis of data indicated the degree of generalizability of the IB Program as a treatment on science performance using a between group design with a quasi-experimental approach. Data from this study came from 5th through 10th grade 2006-2011 Science CSAP scores from authorized IB schools and non-IB schools in a suburban Colorado school district. A test-retest reliability procedure provided a reasonably high level of reliability of individual student scores from 5th through 10th grade.

Data Analysis Procedures

Data were collected from IB and non-IB students enrolled in the 5th through 10th grades during the 2006-2011 school years, inserted into an Excel spreadsheet, and transformed into SPSS software for analysis. CSAP data were available in an archive of public record created and maintained by the Colorado Department of Education and the local district involved in the study. Permission was provided from district and all student identities were protected. The data were numbered in the Excel file and a random sample was taken from each group using a random number generator. A random sample of 50 students from the population of 66 IB students formed the basis of the treatment group and a random sample of 50 students from the population of 132 non-IB student formed the control group. A *t*-test was conducted on the fifth grade Science CSAP scores between the students of the two schools that were in the IB program that met the student criteria (t (48) = .052, p = .96) and between the students of the two schools that were never exposed to the IB program that met the study criteria (t (48) = -.337, p = .74) to ensure no significant differences between the students that were being grouped were evident. CSAP science scale scores also were compared between IB and non-IB students at 5th, 8th, and 10th grade testing periods. The results of the data analysis represent a generalized sample of the population.

Through the use of Levene's test of Equality, an F statistic assuming homogeneity of variance was computed indicating that the assumptions of equal variances were met between the two sample groups across the three testing periods using State test scores (F (1, 98) = 2.314, p = .13); (F (1, 98) =.496, p = .48); (F (1, 98) = 3.559, p = .06) respectively across time). The four Colorado state performance levels (unsatisfactory, partially proficient, proficient, and advanced) were tested for equality and to examine data longitudinally, but the 10th grade scores showed to have unequal variances between the two groups. It was determined that the four Colorado state performance levels categories were not adequately reflecting the actual testing scores. Therefore, to maintain the integrity of the variance of the original scores, eight categories were created utilizing a lower and upper range for each of the four state performance levels (lower and upper unsatisfactory, partially proficient, proficient, and advanced). The Levene's test for equality of variances was not found to be violated across the three testing periods using the adjusted performance levels ((F (1, 98) = 1.760, p = .19); (F (1, 98) = .031, p = .86); (F (1, 98) = 3.070, p = .08) respectively across time).

Demographics

IB School data and non-IB school data were collected from schools that fit the criteria for the sample study. The samples included CSAP Science Assessment scores for the same 5th through 10th grade students in a suburban Colorado school district for the 2006-2011 school years. There was no significant difference in the mean fifth grade Science CSAP scores between the students at authorized IB schools in the district that met the criteria of the study (t(48) = .052, p = .96). Likewise, there was no significant difference between the non-IB schools in the district that met the study criteria (t (48) = -.337, p = .74). Furthermore, analysis showed that there was no significant difference in gender (t (98) = .800, p =.43), free/reduce lunch compared to full pay (t (98) = -.581, p = .56), and Minority status (Whites compared to non-Whites) (t (98) = .479, p = .63) between the two groups at the fifth grade testing year.

Findings

The quantitative data analysis compared significant differences in the science achievement scores of 5th through 10th grade students from the experimental group (n = 50), who received science instruction through participation in the IB PYP and IB MYP program, and the control group (n = 50), who received instruction using traditional science instructional methods. According to Creswell (2009), a p value that exceeds the predetermined alpha level indicates statistical significance. A 2x3 mixed ANOVA was performed to access overall effects of time and program type on CSAP adjusted performance level scores. Tables provided a method for displaying the findings because Creswell (2008) maintained that tables and figures are widely used to display data results. Table 3 provides a comparative view of the mean and standard deviations of the original CSAP Science scores for the experimental group: IB and the control group: non-IB students from the testing years of 2006, 2009, 2011.

Table 3

Means and Standard Deviations of IB to Non-IB Original CSAP Science Scores During the Testing Years of 2006, 2009, and 2011

	5th grade 2006		8th grade 2009		10th grade 2011	
	M	*SD*	*M*	*SD*	*M*	*SD*
IB science scores	600.28	39.147	551.20	44.202	543.28	41.007
Non-IB science scores	581.32	45.256	528.46	47.315	523.08	58.707

Research Question 1

Does participation in the IB Program from 5th through 10th grades make a statistically significant difference in state science test performance of students participating in IB compared to those students not participating in IB?

Hypotheses

$H1_a$: Students from 5th through 10th grade who participated in IB PYP and IB MYP will indicate a statistical difference in science student achievement in CSAP compared to those students who did not participate in IB PYP and IB MYP.

$H1_0$: Students in 5th through 10th grade who participated in IB PYP and IB MYP will indicate no difference in science student achievement in CSAP compared to those students who did not participate in IB PYP and IB MYP.

A 2x3 mixed ANOVA was used to analyze the between-subjects effect of program type. There was a significant main effect (F (1, 98) = 7.783, p = .006) with IB students scoring higher on the Science CSAP performance levels than the non-IB students regardless of the time factor. As shown in Table 4, the mean performance levels were higher for the IB students than for the non-IB students for the 2006, 5th grade test, the 2009, 8th grade test, and the 2011, 10th grade test. When the partial Eta squared η^2 was calculated (partial η^2 = .07), the analysis determined a significant relationship indicating approximately seven percent of the variance in the testing scores was due to whether students attended an IB or non-IB school.

Table 4

Means and Standard Deviations of IB to Non-IB Science Performance Levels During the Testing Years of 2006, 2009, and 2011

	5th grade 2006		8th grade 2009		10th grade 2011	
	M	*SD*	*M*	*SD*	*M*	*SD*
IB science scores Performance levels (n = 50)	5.80	1.340	5.68	1.347	5.46	1.281
Non-IB science scores Performance levels (n = 50)	5.06	1.490	5.02	1.421	4.68	1.558

Note: Categories of the adjusted performance levels are: 1 and 2 = Unsatisfactory, 3 and 4 = Partially proficient, 5 and 6=Proficient, and 7 and 8 = Advanced.

In addition, a 2x3 mixed ANOVA was used to examine the longitudinal effects of time on Science scores as well as the interaction effect of time and program type (IB vs. non-IB). The partial Eta square η^2 was calculated to show the overall main effect of time on Science scores as well as any interaction effects on scores. The 2x3 mixed ANOVA showed a significant overall time effect (F (2, 196) = 8.208, p = .000) indicating that change did occur over time regardless of the program type. The reported partial η^2 indicated that approximately 8% of the variation in the testing scores was due to the time factor η^2 = .08. In addition, there was no significant interaction between time and program (F (2, 196) =.214, p = .807) indicating that although there was change over time, the change was consistent for both IB and non-IB groups.

On the whole, IB students scored higher on the Science CSAP performance level scores than the non-IB students (main effect of program type: (F (1, 98) = 7.783, p = .006)), and the main effect of time decreased performance level scores for both IB and non-IB students longitudinally F (2, 196) =8.208, p = .000). The interaction between program type and time failed to reach statistical significance, F (2, 196) = .214, p = .807, indicating there was a change over time for both groups but the change occurred in a similar pattern.

Summary

This quasi-experimental study was designed to determine if participation in an authorized IB Program supported greater student achievement in science performance based on results from a state standards-based assessment. This study measured the degree of generalizability of the IB Program as a treatment on science performance using a between group design with a quasi-experimental approach. The study used a 2x3 mixed ANOVA on state assessment scores, categorized into adjusted performance level scores, to assess statistical differences and effect size on science performance of 5th through 10th grade IB and non-IB student science CSAP scores from 2006-2011. The analysis determined a positive statistical difference between IB and non-IB test scores with approximately six percent of the variance in the overall testing scores being due to whether students attended an IB or non-IB school.

The research study analysis supported accepting the hypothesis for Research Question 1 and rejecting the null hypothesis. The results concluded that the students who participated in the IB Program showed statistically significant higher science achievement performance level scores at the 5th, 8th, and 10th grade levels as measured by CSAP, in comparison to non-IB students. Additionally, when analyzing the within subject effects, there were significant decreases in performance level across time for both IB and non-IB students indicating that increased exposure to IB does not appear to have increasing effects on performance scores although the IB students consistently scored higher than the non-IB students. Furthermore, the interaction effect between time and program type was shown to be non-significant. Chapter 5 presents a summary of the conclusions and implications of the study, especially pertinent to educational leaders with suggestions for future research.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

Curriculum leaders have to redefine what it means to be a knowledgeable, successful person in the first quarter of the 21st century (Partnership for 21st Century Skills, 2011; Wagner, 2008). The National Governors Association (2008) stated that a world class education includes a global measurement of state-level education performance in science, mathematics, and literacy by examining student acquisition of skills and concepts to compete nationally and internationally in a 21st century economy. Standardized testing may provide only a limited quantifiable estimate of what students know as measured by high-stakes proficiency testing. Boykin and Noguera (2011) shared that comprehensive curriculum models that systemically meet the needs of the whole child have proven successful in increasing student achievement.

A need for effective educational practices, particularly in reading, math, and science, are critical as schools dedicate efforts to meeting the NCLB requirements (AAAS, 2005). Educators are called upon to ensure that students can compete in a 21st century international economy as measured by student achievement on state, national, and international assessments (Tarc, 2010). According to Kaniuka (2009), administrators and educators can make more informed decisions about the implementation of school reform within their district or schools with the support of research-based best practices, models, and curricular frameworks. The International Baccalaureate program provides the kind of framework that supports student achievement, progressive educational reform, and meets the needs of the 21st century learner (Tarc, 2010).

The purpose of this quantitative quasi-experimental study was to determine the relationship between participation in IB and science academic achievement scores. The research methods employed included: (a) defining the variables (measures), (b) collecting data, (c) constructing statistical hypotheses, (d) calculating *t*-tests defining the association among the variables, and (e) conducting 2x3 mixed ANOVA and Eta Square analysis to assess the contribution of demographic variables to any correlation between IB and student achievement. This study explored the effectiveness of participation in the IB Program on science performance as measured by the CSAP. Current experts in science education support the need for further study of the relationship between a student's understanding of science concepts and proficiency in standardized testing programs (Myers & Botting, 2008).

The CSAP Science assessment from 2006-2011 provided a composite of science scores for the study. The study compared the differences in science achievement scores of the experimental group (students who participated in IB schools), who received instruction through a transdisciplinary approach using inquiry-based instruction and the Conceptual-based Backwards Design Model and the control group (students from non-IB schools), who received instruction using traditional science teaching methods. The results indicated that IB students outperformed their non-IB counterparts at 5th, 8th, and 10th grade. The implication of the study analysis results indicated that IB did make a statistically significant difference in science performance. However, a focus on alignment between research-based science pedagogical practices and standardized tests may be prudent to provide a clearer picture of student science performance as it relates to an indicator of quality and future success in a globally competitive world.

Chapter 5 includes a review of the conclusions from the literature review, along with recommendations to educational leaders and suggestions for future research. In addition, conclusions from the research question and findings from the data analysis support a synthesis of the research study.

Limitations

The study's first limitation was the possibility of the experimental and control groups' population experiencing different teaching styles over the course of the study. The possibility exists that some of the non-IB students received instruction through a constructivist model. In addition, other variables, including the schools' methods of instructional strategies, and student experiences, and understanding of science concepts may have caused a variance in results. This study on the effect of IB participation on science performance was limited in its generalization to a defined population of Colorado elementary students who were selected from a quasi-experimental group of students who met specific criteria.

Other variables may have been present during instruction between the 5th, 8th, and 10th grade tests that influenced the results of this study. Research indicates that elementary classroom teachers and secondary science teachers teach to the test and still use methods from the past (Wagner, 2008). IB teachers may be developing effective abilities to use IB strategies in science instruction at the 5th, 8th, and 10th grades. A constructivist approach may have not been implemented within the science classroom at all three levels. Test formatting may have been different at each of the three test levels. Variables that may have affected the drop in scores between 5th and 8th, and 8th and 10th grade IB and non-IB groups may have been behavioral, such as a lack of understanding of the importance of the science test and student motivation. IB methodologies may not be aligned with the format or framework of the state CSAP assessment.

This quasi-experimental study had the limitation of a pure random selection of IB students not being possible as students had to meet the requirement of participating in an authorized IB Program in the same district from 5th through 10th grade. The teaching experience and gender of the teaching staff was not considered in this study. The exclusive use of the CSAP is an element of this study to measure growth in science. Student science performance was limited to the results of only the CSAP test at 5th, 8th, and 10th grade with this study, which cannot provide every essential piece of information necessary to evaluate school science curricula and programming.

A possibility exists that students in non-IB schools may have been influenced by teachers using similar strategies, such as inquiry-based instruction, effective science instructional strategies, and groupings of students who include a variety of socio-economic levels and academic levels. Another possibility may be that non-IB schools used a pull-out program as a structure for teaching science. One advantage to the structure of this quasi-experimental design was the subjects within the time of the study brought the same demographic background and experiences to the study situation. According to Salkind (2009), such a factor minimizes the intra-individual variability within the study. Student group makeup was predetermined through the use of retrospective data; therefore, no risk of students dropping out of the study was evident.

Conclusions

The quantitative data analysis compared differences in the science achievement scores of fifth grade students from the experimental group (n = 50), which received instruction using IB pedagogy and the control group (n = 50), which received instruction using traditional differentiated instructional methods.

Conclusions that may be drawn from this analysis are that the statistical difference in scores may be contributed to participation in the IB Program at the PYP and MYP levels. The drop in scores from 5th through 10th grades may be the result of research-based science pedagogical practices not being systemically integrated into science instruction at the 8th and 10th grade levels.

Hypothesis 1

Fifth through 10th grade science scores of students who participated in IB will indicate a significant difference in science student achievement in CSAP compared to those students who did not participate in IB PYP and IB MYP.

The research study analysis supported accepting the hypothesis for Hypothesis 1 and rejecting the null hypothesis. The results concluded that the students who participated in the IB Program showed statistically significant improvement in their science achievement scores at the 5th, 8th, and 10th grade levels as measured by CSAP, in comparison to non-IB students. In conclusion, positive statistical significance was evident when comparing 5th through 10th grade science student achievement of IB students and the 5th through 10th grade science achievement of non-IB students.

Statistical results indicated no statistical variance between the sources of schools used for the student groups that formed the basis for the study. Upon determination, a 2x3 mixed ANOVA was used to analyze for statistical significance between the experimental group and the control group. A statistical analysis did indicate a significant difference between IB students and non-IB students at the 5th, 8th, and 10th grade levels as measured by the science CSAP tests from 2006-2011, resulting in the rejection of the null hypothesis.

Variables that may have affected the results of the research might include teacher skills, state science test incompatibility with teaching approach between 5th and 10th grades and student test taking skills. The study used only one testing instrument to investigate significant differences between the groups, and the results were limited to outcomes measured by the CSAP. Additional or alternative measurements could determine if participation in the IB supported student science achievement. An analysis using a larger sample size and multiple assessments could add to a body of evidence providing more validity to the study.

Implications

The purpose of this study analyzed the differences in science performance on Science CSAP of students who participated in an IB Program compared to those students who learned science from a more traditional approach. Based on the data analysis, two conclusions were drawn. First, learning within an IB framework supported student achievement in science from 5th through 10th grade when comparing IB students with non-IB students who are demographically alike. Second, longitudinally, the variance between 5th through 10th grades in science achievement measured by the state science assessment supported a statistically significant variance of student achievement.

Implications to Leadership in the Organization

Such findings indicate that the IB Program may make a difference in science student achievement involving a larger sample size comparison and measuring from a different norm-based assessment. Further investigation into the specific written, taught, and assessed science curriculum models used in each school in the district may provide more information about best practices in the area of science performance. Articulation between schools (IB and non-IB) and levels, both horizontally and vertically may inform and guide future science curriculum decisions within the district.

Implications to Leadership Globally

The IB Program is an international curriculum framework implemented worldwide as a successful program of study for 3 to 18 year old students. In addition to literacy, science, and mathematics continue to be disciplines placed in high regard as subjects that support life-time student success. According to Wagner (2008), students need critical thinking and inquiry skills in a global economy. The results from this study may provide information to global leaders as they consider the cost effectiveness of initiating or continuing implementation and support of such an educational program.

Recommendations

Results of the current study indicate a significant difference between the IB and non-IB student science CSAP Scores, implying that participation in the IB Program at the primary years and middle years levels contributes to a significant difference in student performance. The current research was limited by a sample size of 50 and a focus on only science performance as measured by the Science CSAP assessment. More research on the effect of participation in IB PYP and MYP on science performance is encouraged through the use of a larger sample size and with a more demographically and social-economically, diverse group of students.

Recommendations for Educational Leaders

The results of this study provide educational leaders with a possible instructional curriculum framework that supports student achievement in science performance and lifelong learning skills that support success in a global economy. Instructional leaders must make critical decisions about the programs that best support student achievement based on valid research within a High stakes political educational environment. With the implementation of a program, such as IB, teachers support student construction of conceptual understanding that brings not only lifelong skills and knowledge but also successful achievement on assessments. By examining the inferred correlation of participation in the IB Program with state science assessment performance through a quasi-experimental research study, school administrators, and educators may gain information about the integration of the IB Program, support of student achievement, and long-term growth in science within a school culture and specific framework.

Recommendations for Future Research

Salkind (2009) noted that inconsistencies of characteristics between groups within a study may pose a threat to internal validity. Teachers may vary from the IB design treatment implementation causing a possible internal validity threat. The generalizations from the study sample refer only to grade five through grade ten students from a suburban medium-sized school district. Broader generalizations may warrant a replication of the study to provide a more comprehensive response to the impact of IB on science performance.

The current research was limited by a sample size of 50 IB students and 50 non-IB students and a focus in particular on science achievement as measured by the CSAP assessment. Further research may present findings of the use of IB strategies and participation in the IB Program longitudinally as it relates specifically to science achievement and a larger sample size of students. A quasi-experimental research design supported the analysis of the current study. A mixed-method study may prove helpful in drawing out specifically those research-based instructional strategies inherent to the IB Program that may lend them to successful student science achievement.

A meta-analysis of research on the effect of IB on all disciplines may provide a body of knowledge to guide future instructional decisions. Future studies including assessments that provide increased global validity, such as ACT, SAT, and PISA with 10th through 12th grade students may indicate a more in-depth analysis of program differences. Continued research should occur to monitor the longitudinal effects of the IB PYP and IB MYP program on student achievement. Limited research has taken place on the articulation between the three levels of the IB Program. School districts may benefit from further studies with different demographic populations and different baseline assessment scores to explore the effectiveness of the IB Program at all three levels on student achievement in all commonly assessed disciplines.

Based on past and current research of the importance of critical thinking, additional recommendations are to evaluate educational policies and the format of state assessments. Research indicates a need for students to acquire critical thinking skills for them to survive in a 21st century global marketplace (Paul & Elder, 2006; Partnership for 21st Century Skills, 2011; Wagner, 2008; Zhao, 2009). The implementation of critical thinking in the K-12 education system may help students increase test scores and achieve lifelong success in the 21st century (Mansilla & Gardner, 2008). The IB Program philosophy supports such a premise. An investigation into the connection of IB to critical thinking from a qualitative stance may benefit educational leaders as curricular decisions involve a focus on increased student achievement.

Summary

The findings of this research study and the evidence of lack of any previous research on this topic indicate a need for further studies that go into more depth to determine whether IB methodologies make a significant difference in science performance and produce resulting conceptual understanding and higher test scores. This study on the effect of IB participation on science performance used a quasi-experimental design to determine whether participation in IB led to a difference in science achievement of 5th through 10th graders as measured by CSAP. The study compared the differences in the mean of the CSAP science achievement scores of the experimental group (students participating in IB from 5th through 10th grades), and the control group (students participating in a non-IB instructional model using traditional science teaching methods from 5th through 10th grade). The analysis of data concluded that although scores declined between 5th, 8th, and 10th grades with IB and non-IB students, a statistical difference was indicated at each level between the two groups: IB and non-IB in the area of science performance as measured by the CSAP assessment.

Based on the research study, two conclusions were drawn. Students participating in IB had a higher mean on student science performance at 5th, 8th, and 10th grades when compared to demographically aligned non-IB students. In addition, when comparing the science mean score between 5th and 10th grade, IB students demonstrated a positive statistical difference longitudinally compared to non-IB students. According to

White (2005), longitudinal data can provide information about program performance and the effect on student achievement. Educational leaders, teachers, and administrators may benefit from information about research-based pedagogical strategies that support increased science performance to meet the needs of diverse student populations and sustain student success in the 21st century.

REFERENCES

American Association for the Advancement of Science. (2005, February 28). Strategies for closing the 'performance gap' in science and mathematics. *ScienceDaily*. Retrieved from http://www.sciencedaily.com/releases/2005/02/050224105720.htm

American Association for the Advancement of Science. (2006). *Project 2061*. Retrieved from http://www.project2061.org/publications/2061Connections/2005/2005-11c.htm

Andain, I., & Murphy, G. (2008). *Lifelong learners: Challenges for education in the 21st century.* Cardiff, Wales: International Baccalaureate Organization.

Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, CA: Wadsworth/Thomson

Batson, B. (2010). *The effects of the International Baccalaureate Middle Years Program on student achievement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 746476437).

Boyer, E. L. (1995). *The basic school: A community for learning*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Boykin, A. W., & Noguera, P. (2011). *Creating the opportunity to learn: Moving from research to practice to close the achievement gap.* Alexandria, VA: Association for Supervision and Curriculum Development.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.

Brooks, J. G., & Brooks, M. G. (2001). *In search of understanding: The case for constructivist classrooms.* Upper Saddle River, NJ: Prentice Hall.

Bunnell, K. (2009). The International Baccalaureate in the USA and the emerging 'culture war.' *Discourse: Studies in the Cultural Politics of Education, 30*(1), 61-72. doi:10.1080/01596300802643090

Bybee, R. W. (2008). Scientific literacy, environmental issues, and PISA 2006: The 2008 Paul F-Brandwein Lecture. *Journal of Science Education Technology, 17*(1), 566-585. doi: 10.1007/s10956-008-9124-4.

Coke, P. (2008). Uniting the disparate: connecting best practice to educational mandates. *English Journal*, *97*(5), 28-33. Retrieved from www.ncte.org/journals/ej/issues/v97-5

Colorado Department of Education. (2010). *School accountability reports.* Retrieved from http://reportcard.cde.state.co.us/reportcard/CommandHandler.jsp

Creswell, J.W. (2008). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches (*3rd ed.). Thousand Oaks, CA: Sage.

Davey, I. (2011). *Learners without borders: A curriculum for global citizenship*. Cardiff, Wales: International Baccalaureate Organization.

Davidson, S., & Carber, S. (2009). *Taking the PYP forward*. Suffolk, England: A John Catt Publication.

Dewey, J. (1938). *Experience and education*. New York, NY: Touchstone.

DuFour, R., Dufour, R., Eaker, R., & Many, T. (2006). *Learning by doing: A handbook for professional learning communities at work.* Bloomington, IN: Solution Tree.

Eick, C. J., & Reed, C. J. (2002). What makes an inquiry oriented science teacher? The influence of learning histories on student teacher role identity and practice. *Science Teacher Education, 86*, 401-416. doi:10.1002/sce.10020

Fisher, D., & Frey, N. (2008). *Better learning through structured teaching: A framework for the gradual release of responsibility.* Alexandria, VA: Association for Supervision and Curriculum Development.

Getchell, L. (2010). *Effects of International Baccalaureate Primary Years Programme on teacher philosophy, perceptions of efficacy, and outlook on education* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 250764463).

Gewertz, C. (2008). Assessing 21st century skills won't be easy, paper says [Electronic version]. *Education Week, 28*(12), 8. Retrieved from http://nepc.colorado.edu/news/200811/us-measuring-21st-century-skills-will-be-complex

Gilbert, A. (2009). Utilizing science philosophy statements to facilitate K-3 teacher candidates' development of inquiry-based science practice. Early Childhood Education Journal, 36(5), 431-438. doi:10.1007/s10643-009-0302-7

Gilliam, L. H. (1997). The impact of the International Baccalaureate Program on school change (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 304411722).

Gredler, M. E. (2005). *Learning and instruction: Theory into practice (*5th ed., pp 84-85. Upper Saddle River, NJ: Prentice Hall.

Green, S. B., & Salkind, N. J. (2011). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Hancock, T. (2009). Academic performance: A case study of mathematics and science educators from rural Washington high schools (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 305130039).

Hanover Research Council. (2010). *Research without limits.*. Retrieved from www.hanoverresearch.com/

Hassard, J. (2011). PISA test results through the lens of poverty. *The Art of Teaching Science*. Retrieved from http://www.artofteachingscience.org/2011/01/05/pisa-test-results-uncovering-the-effects-of-poverty/

Hayes-Jacobs, H. (2010). *Curriculum 21: Essential education for a changing world.* Alexandria, VA: Association for Supervision and Curriculum Development.

Headrick, V. (2009). The effect of non-traditional science instruction in the middle grades (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 305151064).

Hermann, R., & Miranda, R. (2010). A template for open inquiry: using questions to encourage and support inquiry in earth and space science. *Science Teacher, 77*(8), 26-30. Retrieved from http://highbeam.com/doc/1G1-240864370.html

Hill, I. (2008). BAC to the future: The International Baccalaureate. *Phi Delta Kappan*, *3*(3), 3-18.

Hung, M. (2009). *Achieving science, math, and reading literacy for all: The role of inquiry-based science instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3370970).

International Baccalaureate Organization. (2009). *Making the PYP happen.* Retrieved from www.ibo.org

International Baccalaureate Organization. (2010). *The Primary Years Programme as a model of transdisciplinary learning.* Retrieved from www. ibo.org

International Baccalaureate Organization. (2011). *The IB Programme continuum of International education: Science across the IB continuum.* Retrieved from www.ibo.org

Jerome, D. (2010). *Exploring the relationship between the use of affect in science Instruction and the pressures of a high-stakes testing environment* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 518983003).

Kaniuka, T. (2009). NCLB, school-based instructional policy and decision making: A proposed research agenda. *College Student Journal, 43*(4). Retrieved from http://news-business.vlex.com/vid/nclb-instructional-proposed-agenda-68231887

Kiplinger, V. (1998). *Assessing the CSAP: How can we assess the quality of large scale standards-based assessment?* Retrieved from http://www.cde.state.co.us/ cdeassess/documents/csap/1998/asbrek98.pdf

Kiplinger, V. (2005a). *Effects of IB participation on mathematics achievement and growth, 2000-2004, Academy School District Twenty.* Retrieved from http://www.ibo.org/programmes/research/resources/math/mathematics.cfm

Kiplinger, V. (2005b). *Effects of IB participation on reading achievement and growth, 2000-2004. Academy School District Twenty.* Retrieved from http://www.ibo.org/ programmes/research/resources/reading/reading.cfm

Kuhlthau, C., Maniotes, L., & Caspari, A. (2007). *Guided Inquiry: Learning in the 21st Century*. Westpoint, CT: Libraries Unlimited.

Lamanauskas V. (2010). Integrated science education in the context of constructivism. theory: Some important issues. *Problems of Education in the 21st Century*, *25*(5-9). Retrieved from EBSCOhost database.

Lauer, S. R., & Yodanis, C. L. (2004). A tool for teaching with an international perspective. *Teaching Sociology*, *32*(3), 304. doi:10.1177/0092055X0403200305

Lee, O., & Luykx, A. (2005). Dilemmas in scaling up innovations in science instruction nonmainstream elementary students. *American Educational Research Journal, 42*(5), 411-438. doi:10.3102/00028312042003411

Lindfors, J. W. (1999). *Children's inquiry: using language to make sense of the world.* New York, NY: Teacher's College Press.

Litschge, C. M., Vaughn, M. G., & McCrea. C. (2010). The empirical status of treatments for children and youth with conduct problems: An overview of meta-analytic studies. *Research on Social Work Practice, 20*(1), 22.

Longo, C. (2010). Fostering creativity or teaching to the test? Implications of state testing on the delivery of science instruction. *The Clearing House, 83*(2), 54-57. doi:10.1080/00098650903505399

Longo, C. (2011). Designing inquiry-oriented science lab activities: Teachers can create inquiry-oriented science lab activities that make real-world connections. *Middle School Journal*, *43*(1), 6-15. Retrieved from http://www.amle.org/Publications/MiddleSchoolJournal/Articles/September2011/Article3/tabid/2458/Default.aspx

Lord, T., & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *American Biology Teacher*, *68*(6), 342-345. doi:10.1662/0002-7685(2006)68

Mansilla, V. B., & Gardner, H. (2008). Disciplining the mind. *Educational Leadership, 65*(5), 14-19.

Marx, R., & Harris, C. (2006). No Child Left Behind and science education: Opportunities, challenges, and risks. *The Elementary Science Journal, 106*(5), 467. doi:10.1086/505441

Marzano, R. (2007). *The art and science of teaching: A comprehensive framework for effective instruction.* Alexandria, VA: Association for Supervision and Curriculum Development.

Myers, L., & Botting, N. (2008). Literacy in the mainstream inner-city school: Its relationship to spoken language. *Child Language Teaching and Therapy, 24*(1), 99-214.

Minner, D., Levy. A., & Century, J. (2010). Inquiry-based science instruction--what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474-496. doi:10.1002/tea.20347

Morse, T. (2007). *Academy District 20: IB and non-IB student comparison 2007 CSAP.* Unpublished raw data.

Muller, D., Judd, C. M., Yzerbyt, V. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*, 853-863. Available from APA PsychNet Direct database.

National Center for Educational Statistics. (2008). *State education reforms*. Retrieved from http://nces.ed.gov/programs/statereform/tab2_5.asp

National Center for Educational Statistics. (2010). *What does the NAEP Science Assessment measure?* Retrieved from http://nces.ed.gov/nationsreportcard/science/whatmeasure.asp

National Commission on Mathematics and Science Teaching for the 21st Century. (2007). *Before it's too late: A report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century*. Retrieved from http://www2.ed.gov/inits/Math/glenn/toc.html

National Governors Association, Council of Chief State School Officers. (2008) *Joint statement on reauthorization of the No Child Left Behind Act (NCLB).* Retrieved from http://www.stand.org/Document.Doc?id=590

National Science Foundation. (2005). Inquiry: Thoughts, views, and strategies for the K-5 classroom. *Foundations*. Retrieved from http://www.nsf.gov/pubs/2000/nsf99148/htmstart.htm

Neuman, W. L. (2009). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Boston, MA: Pearson.

Oliva, P. F. (2008). *Developing the curriculum* (7th ed.). Needham Heights, MA: Allyn & Bacon.

Osbourne, J. (2008). *Best practices in quantitative methods*. Thousand Oaks, CA: Sage.

Partnership for 21st Century Skills. (2011). *Framework for 21st century learning*. Retrieved from http://www.p21.org/index.php?option=com_content&task=view&id=254&Itemid=119

Paul, R., & Elder, L. (2006). *Critical thinking; Tools for taking charge of your learning and life* (3rd ed.). Boston, MA: Pearson.

Poelzer, G. H., & Feldhusen, J. F. (2006). An empirical study of the achievement of International Baccalaureate students in biology, chemistry, and physics in Alberta. *Journal of Secondary Gifted Education, 8*(1), 28-40. Retrieved from ERIC database. (EJ540982)

Powell, K. C., & Kalina, C. J. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education, 130*(2), 241-250.

Prensky, M. (2011). *The reformers are leaving our schools in the 20th century*. Retrieved from http://www.scribd.com/doc/50443236/Prensky-The-Reformers-Are-Leaving-Our-Schools-in-the-20th-Century-please-distribute-freely#

Prince, M., & Felder, R. (2007). The many faces of inductive teaching and learning. *Journal of College Science Teaching, 34, 14-20.*

Pritchett, M. (2005). *Removing labels and breaking boundaries.* Retrieved from http://www.convergemag.com/story.php?catid=231&storyid=103612

Richards, L. (2011). *Characteristics of a successful employee for a Fortune 500 company*. Retrieved from http://www.ehow.com/list_6848145_characteristics-employee-fortune-500-company.html

Rothstein, R., & Jacobsen, R. (2006). Goals of education. *Phi Delta Kappan. 88*(4), 264-272. Retrieved from http://www.senate.state.tx.us/75r/Senate/commit/c835/handouts08/0714-Richard-Rothstein-2.pdf

Rothstein, R., Wilder, T. & Jacobsen, R. (2007). Educating the whole child: Balance in the balance. *Educational Leadership, 64*(8), 8-14. Retrieved from http://www.ascd.org/publications/educational-leadership/may07/vol64/num08/Balance-in-the-Balance.aspx

Salkind, N. J. (2007). *Statistics for people who (think they) hate statistics (The Excel Edition).* Thousand Oaks, CA: Sage.

Salkind, N. J. (2009). *Exploring research (*7th ed.). Upper Saddle River, NJ: Pearson Education.

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T., & Lee,Y. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching, 44*(10), 1436-1460. doi: 10.1002/tea.20212

Schunk, D. (2004). *Learning theories: An educational perspective* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Scott, T. J., & O'Sullivan, M. K. (2005). Analyzing student search strategies: Making a case for integrating information literacy skills into the curriculum. *Teacher Librarian. 33*(1), 21-25.

Silva, E. (2009). Measuring skills for 21st century learning. *Phi Delta Kappan, 90*(9), 630-634. Retrieved EBSCOhost databse.

Tan, L., & Bibby, Y. (2010). *PYP and MYP student performance on the international schools'assessment*. Australian Council for Educational Research. Retrieved from http://www.ibo.org/research/programmevalidation/documents/2010ISASummaryReportFinalwebsiteversion.pdf

Tarc, P. (2010). *Global Dreams, enduring understandings: International Baccalaureate in a changing world.* New York, NY: Peter Lang Publishing.

Trochim, W. (2006). *Research methods knowledge base*. Retrieved from http://www.socialresearchmethods.net/KB.intaval.htm

University of Innsbruck. (2011, March 2). *Hands-on learning turns children's minds on to science*. *ScienceDaily*. Retrieved from http://www.sciencedaily.com/releases/2011/03/110302075953.htm

U.S. Department of Education. (2006). *Strengthening teaching*. Retrieved from http://www.ed.gov/teaching

Van Oord, L. (2007). To westernize the nations: An analysis of the International Baccalaureate's philosophy of education. *Cambridge Journal of Education, 37*(3), 375-390. doi:10.1080/03057640701546680

Wagner, T. (2008). *The global achievement gap: why even our best schools don't teach the new survival skills--and what we can do about it*. New York. NY: Basic Books.

Walker, G. (2011). *The changing face of international education: Challenges for the IB*. Cardiff, Wales: International Baccalaureate Organization.

White, S. H. (2005). *Beyond the numbers: Making data work for teachers and school leaders.* Englewood, CO: Advance Learning Press.

Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.

Wilson, J. C. R. (2007). *The effect of a founding International Baccalaureate Middle Years Programme on participating seventh grade students' achievement, behavior, extra-curricular involvement, and perceptions of life skills* (Doctoral dissertation). Available from ProQuest Dissertations and Theses databe. (No. AAT 3277634).

Zabel, D. (2004). A reaction to information literacy and higher education. *The Journal of Academic Librarianship, 30*(1), 17-21. doi:10.1016/j.jal.200311.003

Zhao, Y. (2009). *Catching up or leading the way: American education in the age of globalization.* Alexandria, VA: Association for Supervision and Curriculum Development.

APPENDIX A: PERMISSION TO USE PREMISES

UNIVERSITY OF PHOENIX

PERMISSION TO USE PREMISES, NAME, AND/OR SUBJECTS
(Facility, Organization, University, Institution, or Association)

Academy District 20

Name of Facility, Organization, University, Institution, or Association

Check any that apply:

☒ I hereby authorize Margaret Irene Healer, student of University of Phoenix, to use the premises (facility identified below) to conduct a study entitled: A quasi-experimental study of the effect of IB on science performance.

☒ I hereby authorize Margaret Irene Healer, student of University of Phoenix, to use archived 2005-2010 District CSAP data for the study entitled: A quasi-experimental study of the effect of IB on science performance.

☒ I hereby authorize Margaret Irene Healer, student of University of Phoenix, to use the name of the facility, organization, university, institution, or association identified above when publishing results from the study entitled: A quasi-experimental study of the effect of IB on science performance.

Susan Field — 11/15/2011

Signature — Date

Susan Field

Name

Executive Director for Learning Services

Title

1110 Chapel Hills Drive Colorado Springs, CO 80920

1

APPENDIX B: DISTRICT PERMISSION LETTER

Academy District Twenty

Dr. Mark Hatchell, *Superintendent of Schools*

Education and Administration Center
1110 Chapel Hills Drive, Colorado Springs, CO 80920-3923
Website: www.asd20.org

Phone: 719-234-1200
Fax: 719-234-1299

November 18, 2011

Dear Margaret,

I'm pleased to inform you that your research study entitled, *A Quasi-Experimental Study of the Effect of IB on Science Performanc* has been approved by the Academy District Twenty Internal Review Board. You may begin working with the Director for Assessment as soon as you receive approval from the university IRB.

If you have any questions please contact me at 234-1308.

Sincerely,

Susan Field

Susan Field, Ph.D.
Executive Director for Learning Services
Academy School District Twenty

"The mission of Academy School District 20 is to educate every student in a safe and nurturing environment and to provide comprehensive, challenging curricular and extracurricular opportunities that meet the unique needs of every individual by expanding interests, enhancing abilities, and equipping every student with the knowledge, skills, and character essential to being a responsible citizen of our community, our nation, and the world."

APPENDIX C: PERMISSION TO USE IB QUOTES AND IMAGES

From: rhiannon.george@xxxxx.xxx
To: peggygoodhealer@xxxxxx.xxx
Date: Mon, 8 Aug 2011 08:46:02 -0400
Subject: RE: Permission

Hi Peggy,

You may use the quotes and images as part of your dissertation provided that the IB is properly credited. If at some point in the future you wish to publish your dissertation publically you will then need to apply for a license.
Kind Regards
Rhiannon

	Rhiannon George, Paralegal Associate Route des Morillons 15, CH-1218 Grand-Saconnex, Switzerland Tel: | Fax: +41 22 791 02 77| Web: www.ibo.org “International Baccalaureate” is a registered trademark of the International Baccalaureate Organization.

APPENDIX D: IB SCHOOL RAW DATA

Student	5th Test	5th PL	8th Test	8th PL	10th Test	10th PL
1.	503.00	1.00	433.00	1.00	432.00	1.00
2.	512.00	2.00	484.00	2.00	456.00	1.00
3.	530.00	2.00	468.00	2.00	491.00	2.00
4.	548.00	2.00	480.00	2.00	493.00	2.00
5.	631.00	4.00	546.00	3.00	495.00	2.00
6.	597.00	3.00	533.00	3.00	497.00	2.00
7.	546.00	2.00	471.00	2.00	498.00	2.00
8.	554.00	2.00	521.00	3.00	501.00	2.00
9.	576.00	3.00	502.00	2.00	509.00	3.00
10.	561.00	2.00	506.00	2.00	516.00	3.00
11.	600.00	3.00	549.00	3.00	522.00	3.00
12.	598.00	3.00	536.00	3.00	525.00	3.00
13.	600.00	3.00	541.00	3.00	527.00	3.00
14.	630.00	4.00	543.00	3.00	527.00	3.00
15.	570.00	3.00	553.00	3.00	528.00	3.00
16.	543.00	2.00	488.00	2.00	531.00	3.00
17.	552.00	2.00	540.00	3.00	534.00	3.00
18.	639.00	4.00	550.00	3.00	534.00	3.00
19.	584.00	3.00	544.00	3.00	537.00	3.00
20.	582.00	3.00	582.00	4.00	538.00	3.00
21.	569.00	3.00	543.00	3.00	546.00	3.00
22.	624.00	4.00	552.00	3.00	547.00	3.00
23.	580.00	3.00	498.00	2.00	550.00	3.00
24.	590.00	3.00	546.00	3.00	552.00	3.00
25.	629.00	4.00	561.00	3.00	552.00	3.00
26.	587.00	3.00	576.00	3.00	552.00	3.00
27.	614.00	4.00	598.00	4.00	554.00	3.00
28.	669.00	4.00	593.00	4.00	556.00	3.00
29.	605.00	3.00	555.00	3.00	558.00	3.00
30.	615.00	4.00	543.00	3.00	559.00	3.00
31.	604.00	3.00	580.00	4.00	562.00	3.00
32.	596.00	3.00	591.00	4.00	565.00	3.00
33.	638.00	4.00	587.00	4.00	566.00	3.00
34.	621.00	4.00	587.00	4.00	572.00	3.00
35.	616.00	4.00	598.00	4.00	572.00	3.00
36.	644.00	4.00	582.00	4.00	583.00	4.00
37.	644.00	4.00	622.00	4.00	583.00	4.00
38.	621.00	4.00	569.00	3.00	589.00	4.00
39.	676.00	4.00	622.00	4.00	589.00	4.00
40.	625.00	4.00	571.00	3.00	591.00	4.00
41.	622.00	4.00	593.00	4.00	592.00	4.00

42.	629.00	4.00	627.00	4.00	599.00	4.00
43.	682.00	4.00	616.00	4.00	603.00	4.00
44.	633.00	4.00	573.00	3.00	620.00	4.00
45.	628.00	4.00	584.00	4.00	607.00	4.00
46.	611.00	3.00	585.00	4.00	569.00	3.00
47.	612.00	3.00	519.00	3.00	553.00	3.00
48.	569.00	3.00	481.00	2.00	439.00	1.00
49.	587.00	3.00	563.00	3.00	565.00	3.00
50.	618.00	4.00	575.00	3.00	528.00	3.00

APPENDIX E: NON-IB SCHOOL RAW DATA

Student	5th Test	5th PL	8th Test	8th PL	10th Test	10th PL
1.	554.00	2.00	514.00	3.00	521.00	3.00
2.	592.00	3.00	540.00	3.00	533.00	3.00
3.	561.00	2.00	533.00	3.00	458.00	1.00
4.	532.00	2.00	510.00	3.00	487.00	2.00
5.	575.00	3.00	523.00	3.00	499.00	2.00
6.	644.00	4.00	606.00	4.00	582.00	4.00
7.	585.00	3.00	558.00	3.00	557.00	3.00
8.	642.00	4.00	551.00	3.00	539.00	3.00
9.	652.00	4.00	587.00	4.00	641.00	4.00
10.	639.00	4.00	600.00	4.00	562.00	3.00
11.	568.00	2.00	506.00	2.00	484.00	2.00
12.	563.00	2.00	483.00	2.00	476.00	2.00
13.	531.00	2.00	426.00	1.00	355.00	1.00
14.	543.00	2.00	484.00	2.00	497.00	2.00
15.	595.00	3.00	506.00	2.00	534.00	3.00
16.	639.00	4.00	600.00	4.00	562.00	3.00
17.	627.00	4.00	579.00	4.00	619.00	4.00
18.	544.00	2.00	523.00	3.00	479.00	2.00
19.	544.00	2.00	447.00	2.00	501.00	2.00
20.	561.00	2.00	533.00	3.00	458.00	1.00
21.	544.00	2.00	523.00	3.00	479.00	2.00
22.	544.00	2.00	447.00	2.00	501.00	2.00
23.	542.00	2.00	486.00	2.00	474.00	2.00
24.	517.00	2.00	483.00	2.00	470.00	2.00
25.	606.00	3.00	577.00	3.00	567.00	3.00
26.	550.00	2.00	501.00	2.00	507.00	3.00
27.	514.00	2.00	426.00	1.00	420.00	1.00
28.	541.00	2.00	488.00	2.00	497.00	2.00
29.	575.00	3.00	504.00	2.00	513.00	3.00
30.	636.00	4.00	589.00	4.00	623.00	4.00
31.	523.00	2.00	474.00	2.00	500.00	2.00
32.	623.00	4.00	605.00	4.00	649.00	4.00
33.	609.00	3.00	554.00	3.00	556.00	3.00
34.	618.00	4.00	569.00	3.00	565.00	3.00
35.	644.00	4.00	512.00	3.00	498.00	2.00
36.	617.00	4.00	582.00	4.00	551.00	3.00
37.	593.00	3.00	558.00	3.00	532.00	3.00
38.	667.00	4.00	556.00	3.00	544.00	3.00
39.	591.00	3.00	517.00	3.00	510.00	3.00
40.	560.00	2.00	508.00	3.00	492.00	2.00
41.	570.00	3.00	526.00	3.00	528.00	3.00

42.	515.00	2.00	510.00	3.00	578.00	3.00
43.	514.00	2.00	426.00	1.00	420.00	1.00
44.	573.00	3.00	526.00	3.00	523.00	3.00
45.	667.00	4.00	556.00	3.00	544.00	3.00
46.	575.00	3.00	513.00	3.00	542.00	3.00
47.	683.00	4.00	655.00	4.00	667.00	4.00
48.	600.00	3.00	553.00	3.00	547.00	3.00
49.	517.00	2.00	483.00	2.00	470.00	2.00
50.	564.00	2.00	550.00	3.00	519.00	3.00