Teacher Collaboration for Mandated Accreditation: A Case Study of Evaluating the International Baccalaureate Diploma Program

by

Ildiko Murrayne Biro

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Curriculum, Teaching and Learning
Ontario Institute for Studies in Education
University of Toronto

© Copyright by Ildiko Murrayne Biro 2019
Abstract

This qualitative case study investigates how teachers work together to complete the evaluation for a mandated accreditation of a program at one school. Specifically, the study describes the participants’ experiences of completing the evaluation, the effects of the evaluation on a school’s program as perceived by the participants, and the future of program evaluation. Qualitative data were collected in the form of document review, participant survey and interviews of evaluation team members.

The evidence from the study shows that the majority of participants find the evaluation of the program transformative for their practices. Five major findings as suggested by the case study can be summarized as follows: (1) participants in this study find that their perceptions of completing the evaluation are influenced by the context of the specific school, their own experience with the given program, the personal relationships that exist between the leadership team and the teachers and among teaching colleagues, and the larger community of stakeholders situated in a specific culture; (2) the freedom to change practices and determine what is needed for the given school is crucial in how participants view the impact of an evaluation; (3) respondents’ perceptions of the intentions and abilities of the internal evaluation leaders may have some effect on how they view their evaluation experiences; (4) participants express the desire to determine what aspect of their program is most useful for them to evaluate, when to gather what data, and how to report their findings to their networks; and (5) participants argue that self-assessment is a regular function of
continuous performance monitoring and a natural part of the learning climate where professional conversations lead to change.

Implications from this study suggest that the context of the evaluation and the individual characteristics of the teachers can have an effect on how they perceive the collaboration required for the evaluation and how much value they place on engaging in these mandated processes. School leadership needs to establish a learning climate that supports the collaborative professional development of teachers. Suggestions for areas of further research are included at the end of the study.
Acknowledgments

This has been an interesting journey and I have many people to thank who made it possible. It was difficult to complete my entire doctoral program while working abroad full time but each and every one of you made me believe that I could finish it.

I would like to thank my supervisor, Doug McDougall, for his continued support from afar. I would not have been able to actually finish my program without your reliable guidance. A grateful thank you to my committee members: Elizabeth Campbell and Ruth Childs. Thanks also to my External Examiner, Cheryl Poth, for the detailed and insightful assessment of my thesis. I would further like to express my appreciation to Jim Hewitt for his participation in my final oral examination. All of your collective suggestions and feedback made this thesis what it is right now.

Of course, none of this would have been possible without the support and kind cooperation of the school and its leadership who allowed me to share this case. I am especially indebted to the three interview participants, Lewis, Joanne, and Bill, who recounted their experiences to me so freely and eloquently. It was a joy to work with you.

My parents have been a constant support my entire life and this endeavor was no different. I thank you for your belief in my abilities.

Above all, I dedicate this work to my husband, Jon, and my children, Shinzan and Nirana. You are the light of my life.
# Table of Contents

## Chapter One: Introduction

1.1 Introduction .............................................. 1 
1.2 Research Context ........................................ 2 
1.3 Purpose of the Study ..................................... 10 
1.4 Statement of the Problem ................................. 10 
1.5 Significance of the Study ................................. 11 
1.6 Background to the Dual Roles of the Researcher ...... 12 
1.7 Plan of the Thesis .......................................... 15 

## Chapter Two: Literature Review

2.1 Introduction .............................................. 17 
2.2 School Improvement Networks and Frameworks Used for Large-Scale Education Reform .............................................. 17 
  2.2.1 School Improvement around the World .............. 21 
  2.2.2 School Improvement in International Schools: The Impact of the IB Diploma Program .............................................. 24 
2.3 Characteristics of Successful Schools .................. 27 
2.4 Obstacles for Educational Change and School Improvement .............................................. 32 
2.5 Collaborative Program Evaluation for Accreditation .............................................. 37 
  2.5.1 What is collaborative evaluation? .................... 39 
  2.5.2 What is internal evaluation and how does it relate to external evaluation? 41 
  2.5.3 What are the pitfalls of using the collaborative evaluation approach for accreditation? 44 
2.6 Conditions for Successful School Evaluation .......... 46 
2.7 The Future of Collaborative Program Evaluation ...... 49 
2.8 The Need for This Study .................................... 54
Chapter Three: Methodology

3.1 Introduction 56
3.2 Research Design 56
  3.2.1 Research Context 57
3.3 Participants 59
  3.3.1 The School 60
  3.3.2 The Interviewees 60
    3.3.2.1 Lewis 61
    3.3.2.2 Joanne 62
    3.3.2.3 Bill 62
3.4 Data Collection 63
  3.4.1 Document Review 63
  3.4.2 Qualitative Survey 65
  3.4.3 Semi-Structured Interviews 67
3.5 Data Analysis 68
3.6 Ethical Considerations 69

Chapter Four: Findings

4.1 Introduction 72
4.2 The School Context 72
  4.2.1 Context of the Diploma Program Evaluation Process 75
  4.2.2 School Characteristics 75
  4.2.3 Timeline of the Diploma Program Evaluation 77
4.3 The Survey 78
  4.3.1 Process 79
  4.3.2 Effect of the Evaluation 80
    4.3.2.1 The Progress 81
    4.3.2.2 The Pushback 81
  4.3.3 Future of DP Evaluation 82
4.4 Interviews with Lewis 83
  4.4.1 Process 83
    4.4.1.1 Systematic Planning of the Evaluation 84
      4.4.1.1.1 Timeline – Synchronization with the MSA Evaluation 84
4.4.1.2 Level of Involvement 85
4.4.1.3 Transparency and Accessibility 86
4.4.1.4 Accountability and Self-Assessment 87
4.4.1.2 Pedagogical Leadership of the Evaluation 88
4.4.1.2.1 Experience Level with DP and Evaluation 88
4.4.1.2.2 Stakeholders in the Evaluation 89
4.4.2 Effect of the Evaluation 90
4.4.2.1 The Progress 90
4.4.2.1.1 Process as Progress 91
4.4.2.1.2 The Professional Development of Staff 92
4.4.2.1.3 Structures and Procedures to Increase Student Achievement 93
4.4.2.2 The Pushback 99
4.4.2.2.1 Lack of Time 101
4.4.2.2.2 Personal Feelings 101
4.4.3 Future of Evaluation 102
4.4.3.1 Future of Evaluation at the School 102
4.4.3.1.1 Continuous Change Model 102
4.4.3.1.2 Synchronization of Evaluations 104
4.4.3.2 Future of Evaluation for the IB 105
4.4.3.2.1 Suggestions for the IB 105
4.4.3.2.2 Feedback on IB Proposed Changes 107

4.5 Interviews with Joanne 109
4.5.1 Process 110
4.5.1.1 Systematic Planning of the Evaluation 110
4.5.1.1.1 Timeline – Synchronization with the MSA Evaluation 111
4.5.1.1.2 Level of Involvement 111
4.5.1.1.3 Accessibility and Institutional Memory 112
4.5.1.1.4 Accountability and Self-Assessment 113
4.5.1.2 Pedagogical Leadership of the Evaluation 113
4.5.1.2.1 Experience Level with Evaluation 114
4.5.1.2.2 Stakeholders in the Evaluation 115
4.5.2 Effect of the Evaluation 115
4.5.2.1 The Progress 116
4.5.2.1.1 Process as Progress 116
4.5.2.1.2 The Professional Development of Staff 117
4.5.2.1.3 Structures and Procedures to Increase Student Achievement 119
4.5.2.2 The Pushback 122
4.5.2.2.1 Lack of Time 123
4.5.2.2.2 Personal Feelings 123
4.5.3 Future of Evaluation 124
4.5.3.1 Future of Evaluation at the School 124
4.5.3.1.1 Continuous Change Model 124
4.5.3.1.2 Synchronization of Evaluations 125
4.5.3.2 Future of Evaluation for the IB 126
4.5.3.2.1 Suggestions for the IB 126
4.5.3.2.2 Feedback on IB Proposed Changes 127
4.6 Interviews with Bill

4.6.1 Process
   4.6.1.1 Systematic Planning of the Evaluation
      4.6.1.1.1 Timeline – Synchronization with the MSA Evaluation
      4.6.1.1.2 Level of Involvement
      4.6.1.1.3 Transparency and Accessibility
      4.6.1.1.4 Accountability and Self-Assessment
   4.6.1.2 Pedagogical Leadership of the Evaluation
      4.6.1.2.1 Experience Level with DP and Evaluation
      4.6.1.2.2 Stakeholders in the Evaluation

4.6.2 Effect of the Evaluation
   4.6.2.1 The Progress
      4.6.2.1.1 Process as Progress
      4.6.2.1.2 The Professional Development of Staff
      4.6.2.1.3 Structures and Procedures to Increase Student Achievement
   4.6.2.2 The Pushback
      4.6.2.2.1 Lack of Time
      4.6.2.2.2 Personal Feelings

4.6.3 Future of Evaluation
   4.6.3.1 Future of Evaluation at the School
      4.6.3.1.1 Continuous Change Model
      4.6.3.1.2 Synchronization of Evaluations
   4.6.3.2 Future of Evaluation for the IB
      4.6.3.2.1 Suggestions for the IB
      4.6.3.2.2 Feedback on IB Proposed Changes

4.7 Summary

Chapter Five: Discussion and Interpretation of Findings

5.1 Introduction

5.2 The Research Questions
   Question 1: How do participants describe their experiences of the IB Diploma Program evaluation?

5.3 Program evaluation?
   5.3.1 The Context of the Evaluation
      5.3.1.1 School Context
      5.3.1.2 Stakeholder Context
   5.3.2 Accountability and Self-Assessment
   5.3.3 Pedagogical Leadership of the Evaluation
   5.3.4 Trust
   5.3.5 Organization of the Evaluation Activities
Question 2: What changes are implemented in the program as a result of completing evaluation? How are these changes perceived by the participants?

5.4

5.4.1 The Value of the Evaluation: The Progress
5.4.1.1 The Continuous Evaluation of Practices
5.4.1.2 The Professional Development of Staff
5.4.1.3 Structures and Procedures to Increase Student Achievement

5.4.2 The Pitfalls of the Evaluation: The Pushback
5.4.2.1 Lack of Time
5.4.2.2 Personal Feelings

Question 3: What are the participants’ perspectives on the future of program evaluation in the Diploma Program at the school level and at the IB level?

5.5

5.5.1 Future of Evaluation at the School
5.5.2 Future of Evaluation for the IB
5.5.2.1 Suggestions for the IB
5.5.2.2 Feedback on IB Proposed Changes

5.6 Major Findings

5.7 Implications of the Study

5.8 Suggestions for Further Research

5.9 Limitations of the Study

References

Appendices

Appendix A: School Consent Letter
Appendix B: Participant Consent Letter
Appendix C: Teacher Feedback on the Process of Completing the DP Evaluation Survey
Appendix D: Interview Guide
Appendix E: IB Diploma Program Evaluation 2012-2014 – Timeline
Appendix F: Collaborative Planning and Reflection Meetings for the Evaluation
Appendix G: Logic Model for the Evaluation of the Extended Essay Supervisor Training Program
## Tables

Table 1: Survey Participants

Table 2: Lewis’s perceptions of structural and procedural changes as a result of the evaluation

Table 3: Joanne’s perceptions of structural and procedural changes as a result of the evaluation

Table 4: Bill’s perceptions of structural and procedural changes as a result of the evaluation

Table 5: Summary of participant suggestions to the IB
Figures

Figure 1: Process of professional development of teaching staff as a result of the evaluation 164

Figure 2: Process of structural and procedural changes in the program as a result of the evaluation 166
Chapter One: Introduction

1.1 Introduction

Schools seeking accreditation by the International Baccalaureate (IB), the Middle States Association of Colleges and Schools (MSA), and the Council of International Schools (CIS), just to mention a few organizations, need to complete self-study questionnaires as the first step of the program evaluation process. Internal evaluators, who are teachers or administrators at the school, are responsible for engaging all stakeholders to participate in the preparation of the self-study that is subsequently submitted to the appropriate regional offices for external review. Evaluation visits usually follow this submission to verify the self-studies and determine the accreditation or reaccreditation of a school’s program (International Baccalaureate Organization, 2010). Teachers are, therefore, important participants in the accreditation process and they collaborate with all stakeholders including parents, administrators, school boards, and national and international accreditation bodies to secure appropriate approval for the programs they are delivering.

School evaluation refers to the evaluation of individual schools as organizations. It generally concentrates on key processes such as teaching and learning, school leadership, educational administration, school environment and the management of human resources. Evaluation also includes an analysis of student outcomes, both for student achievement and for the equity of student results. Other inputs, such as infrastructure, funding and the characteristics of the school staff, are also considered. For the purposes of this study, school evaluation is defined as the evaluation of the following major aspects of the school: the effectiveness of the structures and processes in place at the school, the implementation of national and international educational policies and regulations within the school, the quality of student learning outcomes at the school, and the capacity for the school to improve (OECD, 2013).
In this study, I investigate how teachers work together to complete the evaluation for a mandated accreditation of a program. I chose this particular topic because, through research and personal experiences and speaking with other practitioners, I realize that educators do not have the training or experience necessary to conduct evaluations, and that many stakeholders are not valuing these expectations of the accrediting bodies. I want to discover how the process of meeting this requirement can be better conducted in a school setting and what the effects of the evaluation are on the school’s program. In this chapter, I outline the research context and questions, significance and limitations, and my personal and professional connection to the study.

1.2 Research Context

“The secret of change is to focus all of your energy, not on fighting the old, but on building the new” (Millman, 1984, p. 113). Since teachers and schools are entrusted with preparing generations for the future, “educational change is therefore front and center of all the talk about change in general. Educational reform is indeed pervasive right now” (Hargreaves, 2005, p. 278). Hargreaves (2005) further asserts that “a growing change literature is also helping people understand how teachers and schools cope with educational change, and what sense they make of it” (p. 278).

Organizations in every sector of the economy evaluate their practices to bring about desired changes using appropriate performance management systems on a regular basis. Ritchie and Dale (2000) emphasize that “if a process of continuous improvement is to be sustained and its pace increased, it is essential that an organization monitors … what activities are going well, which have stagnated, what needs to be improved and what is missing” (p. 241). Novicevic et al. (2008) further argue that “the corporate scandals of the 1990s have accentuated the need for organizations to develop codes of conduct and to institute formal ethics training programs designed to facilitate ethical behavior” (p. 1061). As
a result, self-assessment models have been introduced to organizations against which they can compare their practices during self-evaluation and report their progress to stakeholders and external organizations to ensure the fidelity of implementation of the set standards.

Self-assessment is a comprehensive, systematic and regular review of an organization's activities and results. The self-assessment process allows “the organisation to discern clearly its strengths and areas in which improvements can be made and culminates in planned improvement actions which are monitored for progress” (Ritchie & Dale, 2000, p. 241). Ritchie and Dale (2000) describe some immediate benefits of using self-assessment at the organizational level such as developing a common approach to improvement across the company and encouraging employee involvement and ownership. Long-term benefits include improving business results and increasing the ability to meet and exceed customer expectations. Ritchie and Dale (2000) also note the difficulties that organizations experience with the self-assessment process. They point out that the self-assessment process is time consuming, employees do not know where to start, leaders find it difficult to sell the concept to the staff as something other than an add-on to their existing duties, and the inherent problems with establishing and maintaining the self-assessment skills of the assessors. Consequently, organizations need to carefully consider the ways in which they employ self-assessment processes to take advantage of the benefits while avoiding the difficulties.

The need for quality assurance in educational organizations grew out of this general movement and a number of high-quality school self-evaluation instruments have been developed and implemented to support this activity (Schildkamp et al., 2012). Adriaan D. de Groot argued that “Quality needs to be demonstrated” (Blok, Sleegers & Karsten, 2008, p. 379), and he concluded that assessing educational quality needs to be based on what students learn from education. As such, school quality cannot be deduced only from process indicators such as the number of hours of teaching, the social climate of the school or the involvement
of parents in school life. Greater accountability has been advocated for schools as an integral part of much broader school reform initiatives (Blok, Sleegers & Karsten, 2008).

In response to these changing demands of society, schools have been undergoing “restructuring” (Cousins & Earl, 1995, p. 4). One of the major contributing factors to this educational change has been the “professionalization of teaching [which] implies significant reform in what teachers do and think” (Cousins & Earl, 1995, p. 4). Cousins and Earl (1995) emphasize “inquiry-mindedness” in teachers which manifests in their “genuine participation in the determination of school goals and the means adopted to achieve those goals” (p. 4).

When teachers are viewed as professionals, their joint work is reflected not only in collaborative curriculum design, but also in direct participation in non-curricular, managerial or organizational governance processes including school program evaluation. Therefore, participatory program planning and evaluation can be used to direct education reform with collaboration being seen as a reasonable route to desired changes in schools (Cousins & Earl, 1995).

Many authors explore the function of teachers in education reform (Fullan, 2005; Harris & Chrispeel, 2006; Glazer & Peurach, 2013). However, in their book, The Far Side of Educational Reform, Hargreaves and Shirley (2011) state that “even when policy-setting and policy-transporting [governments and other organizations] speak on behalf of teachers, teachers often have little to no voice. Teachers are rarely asked to speak on their own account” (p. 1).

Hargreaves and Shirley (2011) argue that teachers are always the “objects and never the subjects of change” and discuss the “three ways of change” (p. 1). They recount the First Way, which was the Golden Age of teachers’ professional autonomy in the 1960’s and 1970’s. After the Interregnum, or Transition Years, the Second Way brought about what is
described as the Global Education Reform Movement (GERM) in the 1990’s, which has the following five defining characteristics:

- Standardized Teaching and Learning
- A focus on Literacy and Numeracy
- Teaching for Predetermined Results
- Renting Market-oriented Reform Ideas from other systems or sectors
- Test-Based Accountability and Control through continuous monitoring of data

(Hargreaves & Shirley, 2011 p. 4).

The need for the professionalization of teachers grew out of these characteristics. The Third Way introduced increased autonomy for teachers but also brought increased accountability through performance targets. The main idea was that if the government offered professionals more support of finance and resources such as training, coaching and materials, then they will be motivated to meet higher expectations and could be pressured for better results (Hargreaves & Shirley, 2011, p. 6). Hargreaves and Shirley (2011) offer the opinions of researchers about how the educational changes in Ontario and Alberta are examples for educational reforms. These researchers state that “[in these provinces] teachers are not the objects of research but its participants and drivers in a commitment to learning, innovation and self-developed improvement rather than delivery of other people’s policies” (Hargreaves & Shirley, 2011, p. 8). However, the authors point out that there is a need to move beyond the Third Way and push for the Fourth by achieving the following eight reforms:

- Put responsibility before accountability
- Eliminate standardized testing connected to system targets
- Develop and disseminate diagnostic and developmental assessment alternatives
- Abandon the obsession with technology as an end to itself
- Raise quality and standards for all teachers by teachers and with teachers together including becoming partners with administrators and policymakers in developing school curriculum and educational policy not just delivering them
- Blend and interconnect the best of different provincial reforms
- Pursue strategic international coalitions especially among teachers
- Promote public engagement as part of the bone and marrow of profound educational improvement (p. 17).

Hargreaves and Shirley (2011) conclude that “the country needs more than professional peace with or even ‘buy-in’ from its teachers. It needs its teachers and their federations to be co-creators of change on the broadest scale for a strong and just society in the future” (p. 18).

Harris and Chrispeels (2006) similarly note that “school improvement has become increasingly multi-faceted, complex and diverse, embracing research and researchers not traditionally located within the field, and including significantly more overt policy action” (p. 8). Examples of these systemic efforts to enhance individual school improvement also include the No Child Left Behind Act in the US and the National Literacy and Numeracy strategies in England.

The potential roles and responsibilities of local education authorities, school districts, or other intermediate agencies, such as the International Baccalaureate Organization, fall into “several critical leadership categories including setting direction, providing professional development, especially for school leaders, providing data to guide the change process, and marshalling resources to meet needs and ensure equity” (Harris & Chrispeels, 2006, p. 13). This current study is situated in this historical context of educational change and school improvement with a particular focus on international education and the International Baccalaureate.
In the early to mid-twentieth century “international schools had sprung up in different parts of the world to cater to the children of parents working outside their home countries, particularly for United Nations (UN) agencies and outposts” (Hill, 2002, p. 18). These schools sought to organize teachers through conferences and workshops to discuss issues related to international education. The Conference of Internationally-minded Schools (CIS) was formed and supported by the United Nations Educational, Scientific and Cultural Organization (UNESCO), and at a course provided by this organization, an agreement on a definition of international education was reached in 1950. Another organization, the International Schools Association, was founded in Geneva in 1951 for the following purposes:

1. To develop close cooperation between existing international schools by means of regular or occasional consultations on educational or administrative questions;
2. To stimulate, facilitate or carry out research work on educational or administrative questions;
3. To promote the establishment of new international schools; and
4. To publicize aims and principles of international schools. (Hill, 2002, p. 19)

The establishment of an internationally recognized common curriculum and examination emerged from the first two purposes. The transient nature of these internationally educated students posed the problem of the nonexistence of an international program that could be taken by all students and then be accepted around the world by post-secondary institutions. The International School of Geneva brought the predecessors of the IB together to address these issues.

Their work was greatly influenced by Alec Peterson and his curriculum model that he proposed for Atlantic College in 1962 in his capacity as the head of the Department of Educational Studies at the University of Oxford. Later, Peterson became the first Director General of the IB in the mid-1960s (Hill, 2002). In order to establish such a program, common standards for curriculum and examinations needed to be designed that would be adapted by all the schools wishing to participate in the program. Hill (2002) argues that the
IB Diploma Programme came about for ideological, utilitarian and pedagogical reasons and its goals were:

1. To provide a perspective that would promote international understanding, prepare students for world citizenship and promote peace;
2. To provide a school-leaving diploma that would be recognized for university entrance around the world with common curriculum and examinations; and
3. To promote critical thinking skills (rather than an emphasis on encyclopaedic knowledge) via a balanced programme in the humanities, the experiential sciences and experiential learning. (p. 18)

The IB celebrated its 45th anniversary in 2013 and published statistics about its programs in *IB World*, the magazine of the International Baccalaureate. According to this source, the IB provides education for more than a million students in 146 countries. The number of authorized schools in the four IB programs stood at 3,631, and 57,000 teachers were trained by the IB in 2012 alone. By 2017, when the IB reached its 50th birthday, the number of schools has grown to 4,775 schools offering 6,282 programs. As of February 2018, there are 3,182 schools authorized for the IB Diploma Program alone, in 153 countries worldwide (ibo.org). This represents exponential growth in the last five years, especially considering that the number of DP schools in the first five years of the program only changed from 7 schools in 1968 to 10 schools in 1973.

Hill (2012) notes that the IB’s “origins led to its perception as elitist, but it has since been democratised in public (state) schools in quite a number of countries, notably in North America” (p. 341). Indeed, as of July 31, 2013, 1,438 schools offer the three IB programs (Diploma Program, Middle Years Program, and Primary Years Program) in the United States and 330 in Canada, making North America the largest concentration of IB schools in the world (IB World, September 2013). The IB hoped to make its program offerings even more inclusive with the introduction of the fourth IB program in 2012, the IB Career-related Certificate (IBCC). The program has now been officially renamed as the Career-related Program (CP) and it offers a “high quality career-related qualification program for students
aged 16 to 19” who would like to receive vocational training. This program has also proven to be popular and the IB welcomed the 200th CP authorized school in June 2018 (ibo.org).

As a result of interest from public schools looking for a way to reform their own education systems, a program that was originally intended for international schools turned to an “international programme for all schools” (Hill, 2012, p. 344). Hill (2012) lists the following statistics to support this argument:

1. In May 2012, 57% of the 3,393 IB schools in 141 countries were state schools where students pay no tuition.
2. In May 2012, 67.5% of the students who took the IB world examinations (80,576 students) were studying in these schools.
3. In the United States, out of the 1,059 public schools (91% of the US IB schools) that offered one of the three IB programs in July 2010, 247 were eligible as designated Title 1 schools (which indicates 40% of the school’s population being qualified for subsidised meal programs). This represents 23% of the IB state schools in the United States.
4. The IB also introduced accommodations for special needs students to enable them to partake in the examinations at the end of the diploma program. These accommodations include examination papers in Braille or enlarged versions or different formats, and digital copies of the papers for students using Specialist Reading software. (p. 344-345)

Hill (2012) argues that the program can be used for school improvement by helping “to infuse rigor, challenge, critical thinking skills, and much-needed analytical writing skills among [the] student body” (p. 345). Hill (2012) concludes that “by focusing on the dynamic combination of knowledge, skills, attitudes, independent critical and creative thought, and international mindedness, the IB espouses the principle of educating for a life of active, responsible citizenship, both local and global” (p. 358).

In his article, “With World Growing Smaller, IB Gets Big”, Scott J. Cech (2007) also explains why this international program is becoming so popular in the United States. Students choose to enrol in the IB Diploma Program in order to be eligible to apply to any university in the world. Additionally, parents view the program as an opportunity for their children to get an international education that otherwise would not be an option.
Similarly, Resnik (2012) claims that “middle- and higher-status families view international education as a response to economic and job market needs and, therefore, a means for social reproduction or social mobility for their children” (p. 265). This tendency has made it easier for the IB to incorporate into national education systems and become a significant alternative in educational settings to bring about needed change for stagnating programs. International education, through the processes of economic and cultural globalization, satisfies “the increasing demands for educational qualifications that are portable between schools and transferable between education systems, and the spread of global quality standards through quality assurance processes such as accreditation” (Cambridge & Thompson, 2004, p. 164). Resnik (2012) further asserts that “analysis of the IB curriculum has shown that it aims to inculcate in students the skills seen as necessary to compete in the global job market, which may explain the rush of middle- and upper-class families to IB schools” (p. 250). Resnik (2012) finally points out that “the IB was legitimized as a global education actor because of the reputation of the DP as the ‘gold standard’” (p. 257). To maintain this level of excellence for the IB brand, schools are required to go through a rigorous accreditation process followed by mandatory evaluation of their programs every five years. Program evaluation at the IB represents both a formal authorization requirement and a service for IB schools (IBO, 2010). Schools have to complete a guided self-study questionnaire with the participation of all staff members who deliver the program, and submit evidence on how their program is meeting the IB’s standards and practices.

1.3 Purpose of the Study

The purpose of the study is to examine and describe the participants’ reflections on their experiences during the collaborative evaluation practices employed by one specific authorized IB Diploma Program school to complete the guided self-study questionnaire for the mandated five-year evaluation for the International Baccalaureate Organization. The
thesis also seeks to explore the interviewees’ perspectives on the effects of the evaluation. Moreover, participants’ opinions regarding the next steps in program evaluation for the school and their suggestions for the International Baccalaureate for the Diploma Program evaluation are presented.

1.4 Statement of the Problem

This project has three objectives: to describe in rich detail the experiences of participants as they completed the IB Diploma Program evaluation; to examine the changes implemented in the program as a result of the evaluation activities; and to present possible directions for the future of program evaluation of the IB Diploma Program at the school and organizational level.

In light of these objectives, the research questions are:

1. How do participants describe their experiences of the IB Diploma Program evaluation?
2. What changes are implemented in the program as a result of completing the evaluation? How are these changes perceived by the participants?
3. What are the participants’ perspectives on the future of program evaluation in the Diploma Program at the school level and at the IB level?

1.5 Significance of the Study

Collaborative evaluation has grown in popularity, along with participatory, empowerment, and utilization-focused evaluation approaches, since it brings together evaluators and practitioners to exchange ideas about how collaboration can be used as a strategic tool for strengthening evaluation practices (Rodriguez-Campos, 2012). In addition to involving stakeholders, “collaborative evaluators need to design and conduct evaluations that are responsive to the organization's culture surrounding evaluation, staff’s readiness to conduct evaluation, available evaluation resources, and evaluators' and clients' desired level of involvement” (O’Sullivan, 2004, p. 11). These practices will set evaluators on the road to
designing more culturally competent evaluations that engage practitioners in school reform. Since “the number and types of collaborative and participatory evaluation approaches are expanding and the accumulation of evidence supporting the approach is growing” (O’Sullivan, 2012, p. 8), further inquiry into collaborative practices in program evaluation is significant to education research in general.

A school that completed the program evaluation required by the International Baccalaureate Organization serves as the context for the study since it is required that all stakeholders, including administrators, teaching staff, students, and parents, work together to complete the self-study questionnaire every five years. These mandated practices can be conceptualized as engagement in collaborative evaluation. The goal of my research is to inform international curriculum models and their evaluations around the world using the lessons learned from this specific context.

1.6 Background to the Dual Role of the Researcher

I never set out to be an educator but I am glad that I happened on this profession. I never thought I would be living and working in six different countries but I would not change these experiences in diverse locations from North America to East Asia and the Middle East. Career theorist John Krumboltz (2011) calls this planned happenstance: putting yourself where good things are likely to happen and then taking the right actions. Repeatedly, opportunities came my way and I dared to take the road less travelled by.

The first such chance came my way right after finishing my undergraduate degree in the US. I took an English teaching job in East Asia without formal teacher training. I had to design lessons and work together with my colleagues, similarly untrained, to provide an acceptable and effective curriculum for our students. This limited exposure to curriculum design and program planning was followed up by a teaching position with a private school in the Greater Toronto Area that was implementing the International Baccalaureate (IB) Primary
Years Program. This experience inspired my Master’s thesis, which focused on two teachers’ collaborative lives as they were implementing the IB PYP at their school.

After completing my Bachelor of Education degree, my fascination with teacher-teacher interactions continued while I moved to an Ontario public school and spent six years hosting teacher candidates from various pre-service programs. A chance conversation with one such new teacher led me to register with an overseas teachers’ recruitment fair along with my husband, a fellow educator. By now, I was teaching high school courses for several years and my previous exposure to the IB world was an asset while trying to find a job abroad. I was offered the position of IB Diploma Program (DP) Coordinator at an American curriculum school in the Middle East. Thus, I found myself, with children and spouse in tow, leading the Diploma Program for two hundred and twenty students and thirty teachers.

At the time of my arrival in August of 2013, the school was already engaged in the reaccreditation evaluation for the Middle States Association of Colleges and Schools (MSA), and the IB Diploma Program evaluation was also scheduled to be completed by the following April. As the DP Coordinator, it was my job to head the efforts for IB DP evaluation. At the same time, I also started my doctoral studies and it felt natural to engage in research that would complement the demands of my new position. Since I already had an interest in teacher collaboration, it felt right to focus my research on how teachers worked together to complete the mandated evaluation documents.

As the work progressed, I also realized that coordinators, principals, and teachers were woefully unprepared and untrained for the rigors of evaluation requirements. The mandated evaluations were conducted in addition to the regular day-to-day running of the school, and many stakeholders resented the extra time and effort that had to be spent on the paperwork that is associated with the documentation required by the accrediting bodies. Many teachers expressed that they felt that the evaluation was just unnecessary requirements.
to keep our accreditation. I began to wonder what would make this unavoidable experience better for all involved and how we could improve our program for our students as a result of these hoops. This current study grew out of these first musings.

During this time, I also became more active in the larger IB community by presenting our experiences at regional conferences and becoming a workshop leader for the IB teacher-training program, while taking doctoral courses in school improvement and program planning and evaluation. These experiences gave me a broader view of evaluation in school settings and my desire to provide a rich description of an evaluation journey.

Although the idea for the study was born during the completion of the evaluation, the actual study did not take place until the following academic year with the bulk of the data set being collected almost a year after the completion of the evaluation. In essence, the study presents the recollections of the participants of the processes that were used to complete the required documents for the IB Diploma Program evaluation. This represents a unique situation that while I was the Diploma Program Coordinator and it was my mandate to lead the meetings of the entire team of teachers who completed the evaluation, collect the subject groups’ ratings of the standards and practices, and facilitate the efforts of the evaluation team during the preparation of the evaluation, I was also the researcher who was collecting the data from my colleagues as they recounted their experiences with the evaluation. Moreover, many of the changes that were introduced as a result of the examination of the school’s practices during the evaluation had been already implemented for most of the academic year by the time the interview participants commented on them. As such, the participants shared their perceptions on the impact they believed these changes had on the school’s program, including student achievement and teacher efficacy.

It appears that the time that lapsed between the evaluation activities and the actual interviews taking place also helped the interviewees to distance me from the evaluation
activities. All three interview participants mainly referred to me as the Diploma Program Coordinator in their responses and talked about how we worked together referencing me in the third person. This was not by design or agreement between us but it emerged organically during the data collection. It also eased my way during data analysis as I tried to separate my roles as researcher and colleague to follow the participants’ lead and refer to my contributions to the evaluation by using my job title. I also tried to use the participants’ words in direct quotes as much as possible in order to reduce the possibility of introducing my personal biases to the presentation of the data. The only exception to this practice was during the discussion of the work I did with Joanne regarding the supervision of the Extended Essay. Since we worked very closely together on this project, we developed a short-hand of discussing this topic. Therefore, this section during the interview was not clear enough without additional, personal explanation. I asked my interview participants to read the section of the thesis that presented their responses to ensure that they agreed with the way I represented their recollection of the evaluation. They positively perceived my portrayal of their contributions to the study. It was a rewarding experience to work with a group of passionate educators on this project. I was drawn to this topic that I happened upon as a researcher and a practitioner.

1.7 Plan of the Thesis

This thesis consists of five chapters. Chapter One provides an overview of the study including the research context, the research questions, and the significance of the study.

Chapter Two is a review of the existing research conducted on this topic. First, various networks and frameworks used for large-scale education reform in international education are introduced. Specific attention is paid to the International Baccalaureate Diploma Program as a vehicle for school improvement. The second section focuses on collaborative evaluation practices for the purposes of program accreditation. As my case
study investigates teacher collaboration in a mandated accreditation evaluation, the advantages and pitfalls of both internal and external evaluation methods are discussed.

I outline and justify the research methodology used for my study in Chapter Three. Within this section, I further describe the research context of the study and present the characteristics of my participants. I outline the qualitative data sources for study and the method that is used to analyse the data. I discuss the conceptual framework that is used for the presentation of my findings. The ethical considerations of the study are in the conclusion of this chapter.

Chapter Four presents the findings from the case study. Chapter Five integrates the findings from the survey and the three interviewees to answer the research questions introduced in Chapter One and links these findings to the existing literature on program evaluation. I also share some implications of my study and suggest areas for further research.
Chapter Two: Literature Review

2.1 Introduction

In this chapter, I situate my study within the related areas of literature. I first examine the benefits of belonging to large-scale education reform models with specific emphasis on the International Baccalaureate. Within this chapter, I explore the characteristics of successful schools and the obstacles for educational change and school improvement. Next, I discuss participatory evaluation approaches. I further provide an explanation of collaborative evaluation, and how internal and external evaluation models are currently used in tandem. The pitfalls of using the collaborative evaluation approach for accreditation are also addressed. Some conditions for creating a successful school evaluation are also discussed. Finally, some ideas for the improvement of educational change and collaborative program evaluation are introduced.

2.2 School improvement networks and frameworks used for large-scale education reform

Glazer and Peurach (2013) claim that “the development and scale-up of school improvement networks is among the most important educational innovations of the last decade, and current federal, state, and district efforts attempt to use school improvement networks as a mechanism for supporting large-scale change” (p. 676). Glazer and Peurach (2013) offer an analytic framework which “identifies the constituent components of networks and environments, and the relationship between them” (p. 676). Glazer and Peurach’s (2013) central arguments are that “school improvement networks and educational environments exist in mutually dependent relationships and that turbulence in educational environments can have profound implications for the function, structure, and sustainability of school improvement networks” (p. 678). The authors further suggest that school networks can be effective with supporting large-scale school improvement when the network provider develops a detailed
design for instructional and leadership practice that is matched with extensive, practice-based learning opportunities (Glazer & Peurach, 2013). Researchers use the notion of networks to conceptualize organizational forms that support or provide the context for improvement efforts. Formal organizations have “clients linked by codified designs for instruction, leadership, school organization, as well as formal contracts” (Glazer & Peurach, 2013, p. 681). Schools seeking authorization by the International Baccalaureate have to undergo a formal process to verify their adherence to the framework of instruction and school organization that is outlined in the Programme Standards and Practices (IBO, 2005). Schools pay annual membership fees established by a formal contract of authorization and subsequent evaluation. The IB requires teachers to be formally trained at IB-approved professional development workshops, and student achievement is measured by external examinations and external moderation of internally assessed components. Using Glazer and Peurach (2013) definition above, the International Baccalaureate Organization can be conceptualized as a large-scale school improvement network; therefore, the conclusions from related research can be applied in the continued development of the program.

Glazer and Peurach (2013) maintain:

Network providers that work at multiple levels of systems (e.g., schools and districts), that serve clients with weak capabilities, that develop highly elaborated and scaffolded designs, that support implementation schools, and that are committed to continuous improvement and sustainability over time require extraordinary capabilities for program development, training, research, and management. Conversely, for network providers that offer less comprehensive designs, that work with stronger clients, that do not provide practice-based implementation support, or that are less committed to continuous improvement and sustainability, the organizational burden is reduced. (p. 682)

Due to the exponential growth in the number of IB Diploma Program schools over the last decade, the organization now has institutions that vary from 6000 students in Grades 9 to 12 where over 1000 candidates sit for the world examinations in a publicly funded high school that is part of a large school board on one continent to 120 students from Pre-
Kindergarten to Grade 12 schools only examining 3 or 4 candidates in a small not-for-profit international school or in a larger for-profit school where the DP is only taught to a small percentage of the entire student body on another continent. Some schools have offered the DP for several decades now, and some are in their first five-year cycle. Surely more mature programs have different needs than programs that are still in the development stages. Some schools have transient international students but very stable teaching staff. Others, like the school in this case study, educate predominantly local students who stay at the school for their entire primary and secondary education, yet the teacher turnover rate is high. These contexts pose very different approaches to successfully engaging with the requirements of the IB. As such, the on-going support provided by the IB to member schools has to be highly developed, scaffolded, and adaptable to many different contexts.

Glazer and Peura (2013) conclude that “if school improvement networks are to grow as a viable and effective option for large-scale school improvement, they will require a community infrastructure that includes coordinated institutional arrangements, resource endowments, market functions, and proprietary activity” (p. 705). Institutional arrangements include performance standards and accountability assessments to provide legitimacy for the network. Resource endowments ensure funding for the development of standards and assessments, research and evaluation, aligned curricula, and technology support. Market functions are designed to identify and assess external partners with which schools could contract for programs and services. The providers devise technologies, services, and programs that are aligned with the established standards. Leaders in schools associated with large networks like the International Baccalaureate need to be able to manage all of the demands associated with this infrastructure.

Fullan (2009) argues that a new leadership paradigm is emerging that is better suited to leading large-scale system reforms. This new paradigm involves having a broader
directional vision, but it has “humility – listen to others including those with whom you disagree, respect and reconcile differences, unify opposition on a higher ground, identify win-win scenarios, be hopeful and confident no matter what” (Fullan, 2009, p. 109). Similarly, in their key findings of a comprehensive review of the literature, Leithwood, Harris, and Hopkins (2008) present seven claims about successful school leaderships:

1. School leadership is second only to classroom teaching as an influence on pupil learning.
2. Almost all successful leaders draw on the same repertoire of basic leadership practices.
3. The ways in which leaders apply these basic leadership practices – not the practices themselves – demonstrate responsiveness to, rather than dictation by, the contexts in which they work.
4. School leaders improve teaching and learning indirectly and most powerfully through their influence on staff motivation, commitment and working conditions.
5. School leadership has a greater influence on schools and students when it is widely distributed.
6. Some patterns of distribution are more effective than others.
7. A small handful of personal traits explains a high proportion of the variation in leadership effectiveness. (p. 27-28)

The authors identify building vision and setting directions, understanding and developing people, redesigning the organization, and managing the teaching and learning program as key characteristics of the practices of successful leaders. Moreover, schools with the highest level of student achievement reported high level of influence from all sources of leadership. The characteristics of the larger context of the school, school districts or organizations, also influence the efficacy of school leaders. In larger contexts where there is clear focus on student learning and achievement and a commitment to data-based decision-making, school leaders had a better idea about their roles which in turn affected all levels of school leadership and teachers.

Byrk et al. (2010) identify “school leadership as the driver for educational change” (p. 62). After comparing a hundred or so schools that made significant progress to their peer schools that did not, they found four interrelated forces to explain the success of these schools: the professional capacity of the teachers (individually and collectively), school
climate (ensuring safety and orderliness in aid of learning), parent and community ties, and what the researchers call the “instructional guidance system” (instructional practices that engage students in relation to key learning goals) as these affected the students in all classrooms (p. 62). Leaders and larger organizations that embrace these forces have more successful schools. Indeed some researchers argue that practice drives theory better than the other way around (Fullan, 2011, 2014; Kirtman, 2013) which is why large-scale reform models should be informed by the practices of successful schools. Singapore, Hong Kong, Finland, Ontario, Alberta, and England are some areas in the world that are engaged in self-conscious strategy-formulation and implementation looking to what they can learn from their own and others’ practices (Fullan, 2009).

2.2.1. School Improvement around the World

Jones and Tymms (2014) observe that the majority of countries across Europe have “implemented systems of evaluation and accountability in order to ensure quality of education in schools. These systems primarily take the form of school inspections” (p. 316). In their study, the authors highlight the underlying mechanisms that explain how The Office for Standards in Education, Children’s Services and Skills (Ofsted) intends to promote school improvement in England. The key mechanisms that are intended to bring about school improvement are: setting standards, giving feedback, sanctions and rewards, collecting information, and public accountability. Jones and Tymms (2014) point out that “more research is needed [on the use of school inspections for school reform], which focuses on the individual mechanisms of inspection, aimed at the school level” especially in order to establish a causal relationship (p. 327).

Ko et al. (2012) similarly extend the research on leadership and capacity building as a means of school improvement, in the process elaborating on their impact within a non-Western society. The authors argue that “scholars have increasingly sought to understand
how the process of school improvement differs among schools operating in different school levels, conditions, and contexts” (p. 216). To this end, they use Rosenholtz’s (1985) conception of “moving” and “stuck” schools as a framework for thinking about school improvement in Hong Kong secondary schools. Ko et al. (2012) identify the following factors that differentiated moving schools where more school improvement was evident from stuck schools:

1. Resource management of principals and school capacity in terms of professional learning community;
2. Workload of teachers;
3. Alignment, coherence, and structure;
4. And resource capacity. (p. 230-231)

In the key findings, the authors note that variables such as trust, communication, and professional learning communities, which are related to the working relationship among teachers, play a crucial role in school improvement efforts since these are key factors in the development of teacher leadership and distributed leadership. This “highlights the importance of learning-directed professionalism in daily exchanges among teachers… The reciprocal support and learning represented by these communities will enhance the effectiveness of teacher instruction initially and student learning eventually” (Ko et al., 2012, p. 230).

Additionally, Hargreaves (2012) and Sahlberg (2011) explore how the Fourth Way of education change is being realized in Singapore and Finland respectively. They argue that, in countries where there is a national vision for education, public collaboration with educators is encouraged, and teachers feel empowered to make the necessary change, systematic education reform can become a reality. Hargreaves (2012) identifies six pillars of purpose and partnership that characterize this new wave of reforms: (1) an inspiring and inclusive vision, (2) public engagement, (3) the acknowledgement that there can be no achievement without investment, (4) responsible corporate involvement in education, (5) recognition that students should be partners in change and in leadership, and (6) approaches to mindful
teaching and learning. To enable the implementations of these pillars in actual school practices, both authors point out the importance of teachers’ professionalism.

Hargreaves (2012) describes three principles that drive educator professionalism: the availability of high quality teachers, the dynamic role of positive and powerful professional associations and teacher unions, and lively collaborative professional learning communities for teachers. Coherence in the form of sustainable leadership, integrated networks, and differentiation and diversity in approaches is the final catalyst for successful reform.

Harris and Chrispeels (2006) suggest a fifth phase of school reform in the form of networked learning communities and their relationship to district reform initiatives. The authors present challenges and concerns for the future of school improvement under three broad themes: goals and outcomes, capacity building and structures for school improvement.

One of the tensions in goals and outcomes is articulated as an understanding of “how to measure school improvement outcomes [which] needs critical discourse among policy-makers, practitioners, researchers, community and students” (Harris & Chrispeels, 2006, p. 297). As far as capacity building in schools is concerned, “a major practice and research ‘blank spot’ in the school improvement field has been the challenge of transferability and scalability of reform (Harris & Chrispeels, 2006, p. 299). Basically, reform initiatives that work like miracles in one school are not successful in others. Context has been suggested as an important factor in school improvement, “yet [it] remains a relatively under-explored dimension within the field, although it is clear that [context] can be a barrier to growth and expansion of certain programmes” (p. 300). In addition, reform initiatives can be undermined by the “high rates of administrative and teacher turnover, which is often typical in many inner city areas under high stress and in isolated rural communities” (p. 300). Harris and Chrispeels (2006) summarize their recommendations:

More research is needed to fully understand the interplay of legislative mandates and how they support or constrain the development of (a) local capacity and effective
structures needed to respond to and sustain school improvement, (b) safe collaborative learning environments for children and adults, and (c) the retention of competent teachers and administrators. (p. 301)

The authors argue that there are certain conditions for a fifth phase of school improvement for networked learning communities. These conditions include:

1. Decreasing the ranking and comparison of schools currently found in many accountability systems.
2. Ensuring that all schools have the same possibility and opportunity to be part of a network, federation or partnership.
3. Developing ways of discerning the different sets of needs of schools and identifying the various combinations of networks, partnerships or federations that would best meet those needs.
4. Developing the role for local educational authorities. (Harris & Chrispeels, 2006, p. 303)

However, not every region in the world is engaging in successful educational change. For example, Karami Akkary (2014) reports that educational change in the Arab world is often driven by political agendas, is dominated by the top-down approach to change, lacks a culturally grounded knowledge base, broadly neglects planning for implementation, and does not build capacity in those participating in the reforms. The author suggests that Arab nations adopt a transformative view of change, abandon the one-size-fits-all and prepackaged designs of educational reform, and adopt a systemic view while setting goals for reform (Karami Akkary, 2014). Conclusions from these research studies could be used to improve the structures for the school evaluation systems around the world.

2.2.2 School Improvement in International Schools: The Impact of the IB Diploma Program

The IB Diploma Program is designed for students aged 16 to 19. The program is comprised of six groups of subjects (language and literature, language acquisition, individuals and societies, sciences, mathematics, and the arts) and the DP Core, which includes Theory of Knowledge (TOK), Creativity, Activity, Service (CAS), and the Extended Essay (EE). Each course’s requirements contain several internal and external assessments
over the two years of the program. These assessments measure skills such as the ability to analyse and synthesize information, to propose a reasoned, well-supported argument, to present on a topic, and to conduct short and long-term research. Internal assessments are evaluated by the classroom teacher at the student’s school and IB examiners mark the external components of the courses. IB examinations are knowledge-based since students are evaluated on what they have actually studied, and criterion-referenced (IBO, 2019). The variety and quality of assessments are what further separates the International Baccalaureate from other curricula like Advanced Placement (AP) and the A-Levels. The assessments are “designed to inspire the very best kind of teaching” (Mathews & Hill, 2005).

As a result of the rigorous and highly structured and regulated comprehensive curriculum, coupled with the on-going intensive teacher training, the IB DP has become the so-called “Cadillac” of college-prep programs (Gehring, 2001) and a “supertest” capable of transforming a failing education system (Mathews & Hill, 2005). IB DP students are said to outperform students from other education systems with respect to retention, degree completion, and securing the highest-paying jobs upon graduation (IBO, 2015).

Conley (2008) argues that “at the heart of college readiness is the development of the cognitive and metacognitive capabilities of incoming students: analysis, interpretation, precision and accuracy, problem solving, and reasoning” (p. 3). In addition, the author identifies self-management skills, self-awareness, self-control, and intentionality as key elements in student success in higher education (Conley, 2008). These are precisely the capabilities and skills that Conley and Ward (2009) find that the IB Diploma Program develops in its graduates. The view that the IB DP is a premier qualification for university admission is prominently displayed on high school websites, and universities around the world are trying to entice IB students to come to their institutions (Fitzgerald, 2017).
Moreover, research suggests that participating in IB professional development activities can have a positive impact on teacher efficacy. Chadwick, Thier, and Todd (2019) report that the benefits of participating in IB Educator Network (IBEN) learnings, such as becoming IB subject examiners or IB teacher workshop leaders, include: better understanding and preparation for IB assessments, improved subject knowledge and pedagogical skills, access to the latest resources, mastery of IB concepts, and connection to the IB and other IB Educators. Schools whose teachers became involved with IBEN also noted benefits such as developing systematic approaches to collaboration, increased school prestige, possible improvements to student world examination performance, and greater engagement in the IB community.

Schools that implement IBEN learnings provided planning time for teachers and a range of different meetings to promote professional development. Successful schools quickly and efficiently disseminated IBEN information. School leadership also seems to be an important factor in influencing the ability to gain benefits from IBEN participation. Finally, schools with cross-disciplinary collaborative cultures shared information and skills that IB Educators gained from their participating in IBEN. Similarly, initial findings from the ongoing research on the impact of the IB’s professional development suggest a positive indication of the potential for IB teacher training workshops to influence teaching practice, and increase feelings of general self-efficacy, context-specific self-efficacy, and collective self-efficacy of teachers (IBO, 2018).

However, it is prudent to acknowledge that there is scarce empirical research to support these claims from outside the IB (Bunnell, 2008; 2014). Many of the studies were directly commissioned by the International Baccalaureate and were not peer reviewed or published outside of the IB. Indeed, Bunnell (2010) states that it “seems strange that after four decades of IB activity to comment upon the lack of empirical investigation” (p. 359).
Yet, even critics can agree with some of the findings on the impact of the IB Diploma Program. After characterizing *Supertest: How the International Baccalaureate Can Strengthen Our Schools* (Mathews & Hill, 2005) as a “marketing tool”, McKenzie (2005) acknowledges that the IB DP “does provide excellent teacher training, which the AP does not. And it has produced great success stories for students and teachers” (p. 238-239). Similarly, after exploring the perceptions of admissions officers in Canadian universities across 10 provinces, Fitzgerald (2017) finds a similar trend to previous studies with a “highly positive disposition towards the IBDP, particularly when compared to the local (in this case, Canadian provincial) curricula” (p. 94).

Although some evidence has been presented in research that adopting the IB framework improves school performance, particularly in the areas of university acceptance for graduates and professional development for teachers, the program alone cannot ensure that the school will be successful.

### 2.3 Characteristics of successful schools

In order to improve their practices, what characteristics can schools hope to adopt from successful schools? Collaborative leadership has been identified as a mechanism to promote school improvement. Telford (1996) offers a definition of collaborative leadership as transforming leadership in which there are such reforms as “new and more complex expectations for student outcomes; school-leaders able to provide instructional leadership; high expectations by the public for its schools and many associated external pressures of change; a rapidly expanding body of technical know-how concerning instruction; and changing family environments” (p. 27). The author further presents structural, political, human resource, and symbolic elements that have a significant impact on the success of educational reform in a given school. According to this study, successful schools were noted
by “a flat hierarchy, frank and open communication, listening, respecting, and valuing people and empowerment” (Telford, 1996, p. 27).

Collaborative leadership can take into account the social and emotional needs of teachers during the change process and can develop the trust needed to embrace a growth mindset essential for reform. Hargreaves (2005) maintains, “Teachers’ emotional connections to students, and the social and emotional goals they wanted to achieve as they taught those students, shaped and influenced almost everything they did, along with how they responded to changes that affected what they did” (p. 286). Moreover, teachers rated the feeling of freedom in planning and pedagogical decision making regarding teaching strategies as “exceptionally important” (Hargreaves, 2005, p. 292). As a result, Hargreaves (2005) concludes that “teaching cannot be reduced to technical competence or clinical standards. It involves significant emotional labor as well” (p. 293).

Furthermore, in their book Linking Leadership to Student Learning, Leithwood and Louis (2012) conclude that principals who had the greatest impact on student learning in the school focused on instruction – including teacher knowledge, skills, motivation – and on ensuring supportive working conditions, such as time for collaboration. They say that “leadership affects student learning when it is targeted at working relationships, improving instruction, and indirectly, student achievement” (p. 234). Hallinger and Heck (2010) posit that school improvement leadership is highly contextualized. They suggest that the type of leadership that is exercised by the school’s leadership team must be linked to the school’s improvement capacity at any point in time. The authors discuss the notion that “every school is on its own unique improvement trajectory” (Hallinger & Heck, 2010, p. 106). Academic structures, school norms, and on-going organizational processes create both opportunities and restrictions for the leadership in their improvement efforts. Effective leadership for school improvement must be responsive to these contextual characteristics. Hallinger and Heck’s
(2010) findings indicate that “the impact of the school’s culture on leadership was greater than vice versa” (p. 106).

Therefore, it is not surprising that Fullan and Hargreaves (2016) additionally argue that, at the heart of an effective and continuously growing teaching profession, is professional learning and development which is “firmly rooted in a system culture of collaborative professionalism that cultivates individual and collective efficacy” (p. 2). Indeed, in order to achieve sustainable change in a school, stakeholders must identify with the success of the organization and develop a mutual identity, understanding their individual roles and committing to a higher, moral purpose (Fullan, 2010).

Successful reform leaders of schools must have three characteristics: informed optimism, persistence, and willingness to confront behaviour that is incongruent with moral purpose, and the ability to help others uncover their own moral purpose. Effective leaders move forward even when things seem to be stagnant, or worse, falling apart. This cannot happen in isolation so effective relationship building is a crucial leadership competency. Fullan (2010) points out, “if you want to challenge someone to do better, you’d better build a relationship first” (p. 6). Successful school leaders are no longer desk-bound problem solvers; rather they are in the hallways, the classrooms, the library, and the cafeteria, in the middle of the daily ebb and flow of organizational life. Moreover, successful leaders realize that they have a moral obligation to collaborate with and learn from other organizations.

Successful school leaders are now viewed as fulfilling three important roles in their schools: agent of change, system player, and lead learner (Fullan, 2014). As a change agent they move people and organizations forward under difficult conditions. These leaders create a climate for change which is non-judgmental. They recognize that, if you want to change something, you need to name it, model it, and monitor it. A good change process shapes and reshapes good ideas while developing capacities and ownership in the group. The changes are
positioned as voluntary but inevitable, and the process is insistent. Good leaders use the

group to change the group and recognize that the individual will be more influenced by the
culture than the leader. It is essential that leaders use good instructional practices as the
foundation of any change effort. Passion matters but must be earned through actually getting
better at leading change, which is achieved by trial and error through a process of learning.
As Fullan (2014) states, “passion without skill is dangerous” (p. 125). Fullan further argues
that educators and leaders only feel passion emotionally when they are skilled at the work
they are required to do and are actually experiencing success. Additionally, as system players,
principals contribute to and benefit from system improvements. Within a large network of
schools, they balance the needs of their specific organization by balancing autonomy and
cooperation, improvement and innovation, internal accountability and transparency. Their
focus is on diffusion of ideas and on specific hard to solve problems within their own
contexts. These leaders acknowledge that talented schools will improve a weak teacher, that
talented teachers will leave a weak school, and that the strength and sustainability of an
organization is a function of the quality of its lateral relationships which creates the culture at
the school.

Good collaboration reduces bad variation of practices and good variation of practices
produces innovation. Successful schools use the collective expertise of the large networks
that they belong to in order to improve their own practices. Finally, lead learners model
learning and shape the conditions for all to learn. This does not mean that the principal leads
all instructional learning. Rather, the principal works to ensure that intense instructional focus
and continuous learning are at the core work of the school and does this by being “a talent
scout and social engineer, building a culture of learning, tapping others to co-lead, and, well,
basically being a learning leader for all” (Fullan, 2014, p. 90).
Fullan and Hargreaves (2016) advocate for the development of instructional practices that is deeply embedded in a learning culture since it is this culture that drives the quality of professional learning and development which is essential for successful schools. The authors point out that this type of culture is highly contextual, which is why it is futile to try to implement a borrowed system of high performance instead of developing a learning culture suitable for each individual context. Cole (2013) argues that effective professional learning focuses on developing the core attributes of an effective teacher. It enhances teachers’ understanding of the content that they teach and equips them with a range of strategies that enable students to learn that content. This type of professional development is directed towards providing educators with the skills to teach and assess for deep understanding, and to develop students’ metacognitive skills.

In order for professional learning to be effective, it needs to be embedded in or directly related to the work of teaching, grounded in the content of teaching, organized around collaborative problem solving, and integrated into a comprehensive change process. Professional learning should primarily be school-based and school-managed and be focused on improving teaching practice. However, this does not necessarily guarantee that it will impact on teaching practice in ways that will produce school improvement. There needs to be a move away from the traditional practice of professional learning as an isolated event initiated by the individual teacher who attends an externally-provided conference or workshop focused on the acquisition of knowledge. Alternatively, a move is needed towards professional development as a routine practice within the school that involves all teachers with a focus on the implementation of teaching strategies and techniques that make the biggest difference to student learning.

This type of professional learning is promoted within the school by instructional coaches, structured meetings and forums, teaching demonstrations, workshops conducted by
teachers and external experts, and other routine opportunities for formal and informal professional discussions. Moreover, individual teachers’ professional learning practices should contribute to colleagues’ professional development as a matter of common practice, and individual, group and whole-school professional learning should centre around school improvement. Being part of a learning community is not simply about the pursuit of individual learning goals; rather it is also about contributing to the learning and knowledge base of one’s colleagues and the school.

Guskey (2002) also suggests that professional development efforts need to be evaluated in order to determine whether they actually make a difference for student achievement and well-being and individual and collective teacher efficacy. The basic use of evaluation is to investigate the merit or worth of a system which denotes appraisal and judgment by answering such basic questions as, “Is this program or activity achieving its intended results? Is it better than what was done in the past? Is it better than another, competing activity? Is it worth the cost?” (Guskey, 2002, p. 45). The author advises that professional development should be a purposeful endeavour that supports the learning culture for the particular school even if making a simple causal inference between professional development activities and student achievement is impossible due to the complex and variable real-world setting of a school. Fullan and Hargreaves (2016) issue the call to action to “establish a culture of collaborative professionalism in which teachers develop and grow day by day through feedback and joint work in which student and teacher learning and well-being form a mutual feed for the betterment of society” (p. 21).

2.4 Obstacles for educational change and school improvement

The above examples show that there are schools and systems that are successful with the implementation of education reform. So why are there still so many unsuccessful schools when it comes to school improvement? As writer and journalist Arnold Bennett (1913)
expressed, “Any change, even a change for the better, is always accompanied by drawbacks and discomforts” (p. 82). Hargreaves (2005) similarly argues that “one of the most neglected dimensions of educational change is the emotional one. Educational and organizational change are often treated as rational, cognitive processes in pursuit of rational, cognitive ends” (p. 278). Hargreaves (2005) further observes that “emotions are at the heart of teaching … Good teachers aren’t just well-oiled machines … [They] are passionate about ideas, learning and their relationships with students” (p. 278-279). Yet, Hargreaves notes that “even the idea of organizational learning which is on the very cutting edge of change theory, is almost exclusively cerebral in its emphasis…It is as if teachers think and act; but never really feel” (p. 279).

Hargreaves (2005) warns how dangerous it is when emotions are discounted in educational reform since they enter into the change process anyway as “festering resentment [that] undermines and overwhelms rational decision-making” or as poisoned committee work which is a result of “unresolved grudges and grievances” (p. 280). The author emphasizes that “caring occupations like teaching call not only for emotional sensitivity; they also require active emotional labor” (Hargreaves, 2005, p. 280). As such, educational reform needs to recognize that conditions for and demands on teachers’ work need to make it easier for them to do their “emotional work” as well (Hargreaves, 2005, p. 281).

Louis (2007) argues that “in the past several decades, issues of trust have emerged as central to discussions about the future of education” (p. 1). The author examines how trust affects teachers’ willingness to work with innovations introduced by central office administration. Louis (2007) identifies social trust as it is manifested in Western society in two common forms:

- Institutional trust (also referred to as social contract trust), is the expectation of appropriate behavior in organized settings based on the norms of that institution. Parents, for example, generally trust that schools will do their utmost to try and educate and protect their child during school hours.
Relational trust (also referred to as situated trust) is the inevitable result of repeated interactions with others in modern organizations. While personal relationships may be limited, individuals interact repeatedly with the same individuals, which leads to expectations specific to that individual or group. (p. 3)

Louis (2007) summarizes how trust and change are viewed in the school improvement literature by stating that “betrayal of trust between teachers and administrators is common in schools, and undermines improvement efforts” (p. 4). The most common recommendation to combat this lack of trust is that leaders create relational trust by involving subordinates in planning, implementing, and making adjustments in the change as it is carried out. Louis (2007) concludes that “trust is a core resource for improvement” both in elementary and high schools (p. 17).

Teachers’ experiences with change are mainly influenced by relational rather than institutional trust. Based on the results of the study, Louis (2007) offers the following recommendations to administrators as they orchestrate complex change:

1. Combine transformative goals (often associated with leadership) with more attention to transactional considerations (usually categorized as management). Paying attention to daily relationships with teachers inspires confidence in administrators as a person, which in turn provides a foundation for trust in institutional leadership for change.
2. Assess the current level of trust in a building prior to initiating a significant change. If trust is low, trust issues need to be addressed if other organizational improvements are to be introduced on solid ground.
3. Understand that trust cannot be easily separated from expanded teacher empowerment and influence. Teachers are not passive actors in the school, but co-constructors of trust. As active professionals, teachers who feel left out of important decisions will react by withdrawing trust, which then undermines change.
4. Keep [administrators’] fingers on the pulse of trust during a change process, and should not assume that their trustworthy reputation will persist.
5. Pay particular note to embedded cultures of distrust among the staff: If teachers cannot trust each other, they cannot work together effectively to create systemic change.
6. Scrutinize how [administrators’] own behavior and context are interpreted by others. Getting information about how one is being perceived requires developing strategies for getting honest assessments, possibly through techniques like rapid feedback 360 evaluations, in which subordinates and peers provide anonymous assessments of performance.
7. Realize that cutting corners in the change process to speed up planning and implementation is unwise. (p. 17)
Louis’s study points to three variables that may warrant further investigation as significant enhancers of trust: (1) perceived influence over how decisions are made; (2) a sense that decision makers take stakeholder interests into account; and (3) an agreed upon and objective measure of the effects or outcomes of implemented decisions. All three may not need to be present at all times, and the presence of even just one may be regarded by teachers as evidence of trustworthiness. However, the absence of all three appears to breed mistrust. Louis (2007) concludes that “school leaders need to view trust as the bridge that reform must be carried over, but rather than being solid, that bridge is built on changing emotions” (p.20).

Tschannen-Moran (2014) further develops this point when discussing “the control paradox” (p. 224). The author points out that, instead of promoting organizational learning, standardized controls and rigid procedures lead to the breakdown of teacher efficacy, which is precisely needed for the achievement of the very goals of the proposed reforms. School leaders’ attempts to bring about change by using one-size-fits all procedures often backfire because they take away the educators’ professional discretion to act in the best interest of their students.

Although teachers might outwardly comply with the expectations, they will always find ways to sabotage change efforts if they do not feel empowered. The resulting power struggle is counterproductive to the development of the professional orientation of educators, which is needed for sustained change efforts. Tschannen-Moran (2014) finds that the degree of teacher professionalism is related not only to the professional orientation of the school leaders but also to faculty trust. Without trust, there is no capacity building in the educational staff to bring about the desired changes.

Furthermore, Fullan (2014) asserts that, in order for a change process to be successful, sufficient time needs to be allocated for the progress to become a reality. Leaders need to be as assertive with the modifications that they want to introduce as they can get away with. If
the process is too fast, the culture will rebel. However, if the process is too slow, inertia will set in. Principals and program leaders have to manage the changes according to the needs of the individual school.

Fullan (2013) suggests that people take to change when it is intrinsically interesting, it is pursued in a non-judgmental culture, they have some say in its evolution, they are developing ownership with others, and they enjoy doing something worthwhile with peers inside and outside their school. Leaders also need to know how to deal with resistance to change. It is advisable to give people respect before they have earned it, to practice impressive empathy for the teachers who are in the way of change, and to deal firmly with members of the staff who still do not embrace the new ways. Creating emotionally favourable conditions for success will aid the change process since the road to reducing resistance is to increase enthusiasm.

In order to effectively lead change efforts, Fullan (2014) endorses the following actions for principals: challenge the status quo, build trust through clear communication and expectations, create a commonly owned plan for success, focus on the team over self to create sustainability of leadership, have a sense of urgency for sustainable results, commit to continuous improvement of self, and build external networks and partnerships.

In addition, there are program-specific challenges. Despite growing in popularity as a program of choice internationally, there are considerable obstacles for the further spread of the IB Diploma Program. Cech (2007) points out that “the process of becoming IB-authorized and offering IB classes can be expensive and time-consuming, and the research base on IB's efficacy in the United States at this point is thin” (p. 22). Furthermore, some schools decided to “discontinue [their] IB offerings after parents complained that the program allowed less flexibility for students' extracurricular activities than AP” (Cech, 2007, p. 24). Similarly,
Resnik (2012) cites the financial burden that accompanies gaining and maintaining IB accreditation as one of the obstacles for the further spread of the IB.

Moreover, the IB has undergone a major metamorphosis as a result of the dramatic increase in the number of IB schools in the last two decades. Resnik (2012) points out that the IB began as a small community of teachers who often volunteered their efforts to create and develop an innovative university preparatory diploma for international schools. However, the current primary focus of the organization is financial and organizational measures of efficiency and cost-effectiveness, following the advice of international management consultants in the provision of its educational programs and expansion of its services. While the early development of the IB was an action research project in international education that resulted in curriculum innovation through teacher participation, now however, “the need to maintain the ‘same high quality’ of the IB brand globally places constraints on how progressive and innovative the IB can be, thus entailing a certain rigidity in the provision and content of its educational programs” (Resnik, 2012, p. 264).

In the face of such difficulties, it is important to further explore the program evaluation practices of the organization. The IB seems to recognize the need for such reforms. Albright, Atkinson, and Bender (2014) identify the following areas for improvement in the evaluation process for DP schools: differentiated process to provide meaningful context and specific feedback, clearer report format and structure, improved process for schools without an evaluation visit, consultation during the evaluation process and follow up support and services, aligned evaluation for groups of schools, and additional support and training for IB external evaluators.

2.5 Collaborative Program Evaluation for Accreditation

Participatory evaluation emerged to include teachers as professionals in education reform. Cousins and Earl (1995) define participatory evaluation as “applied social research
that involves trained evaluation personnel (or research specialists) and practice-based
decision makers working in partnership” (p. 9). They point out that “participatory evaluation
is best suited to formative evaluation projects that seek to understand innovations (programs)
with the expressed intention of informing and improving their implementation” (p. 9). The
underlying justification for using this approach is “problem-solving in professional work
which is tied to reflection-in-action and reflection-on-action” (p. 10). The term “participant”
in the context of program evaluation refers to “people who have a stake or vested interest in
evaluation findings” (MacLellan-Wright, Patten, dela Cruz, & Flaherty, 2007, p. 101). The
participatory approach can be sensitive and responsive to each unique program in various
stages of program development, geographical locations, and level of stakeholder
involvement. “The benefits of participatory evaluation can be both instrumental (an aid to
enhancing quality, relevance, and uptake) and transformative (leading to individual and
organizational development) … [However,] if participation is not accompanied by an
opportunity for real involvement and influence, it may cause more harm than good”
(MacLellan-Wright, Patten, de la Cruz, & Flaherty, 2007, pp. 102-103). It is hoped that
participation in the evaluation process will motivate organization members to rethink their
professional practice based on what they observe during the evaluation process.

Cousins and Earl (1995) identify five requirements that need to be in place in order
for participatory evaluation to be viable:

1. Evaluation and local applied research must be valued by schools and districts.
2. The administration must provide the time and resources required.
3. School districts need to be committed to organizational learning as a route toward
improvement.
4. Teachers and administrators participating in evaluation activities must be
motivated.
5. It is necessary to assume that staff likely to participate in evaluation activities do
not have sufficient research experience and knowledge to carry out such tasks, but
that they have the ability to learn given appropriate training. (p. 12-13)
The authors conclude that “participatory evaluation holds the promise espoused by advocates of collegial work, it is likely to provide a practical and cost-effective alternative and it appears to offer a distinct approach for schools and school systems wishing to develop organizational learning capacity” (Cousins & Earl, 1995, p. 13). It could be argued that the self-studies required by many international organizations such as the IB, CIS, and MSA are a way to answer the call of researchers to use teachers as co-creators of change at least on the school level of educational reform.

### 2.5.1 What is collaborative evaluation?

Collaborative evaluation is under the umbrella of participant-oriented program evaluation. O’Sullivan (2012) states that “Collaborative Evaluation systematically invites and engages stakeholders in program evaluation planning and implementation” (p. 518). O’Sullivan contrasts “distanced” evaluation approaches, which reject stakeholder participation as evaluation team members, with collaborative evaluation which “assumes that active, on-going engagement between evaluators and program staff, result in stronger evaluation designs, enhanced data collection and analysis, and results that stakeholders understand and use” (p. 518).

In the participatory evaluation approach, stakeholders are engaged in limited aspects of the evaluation process. However, according to O’Sullivan (2004), “the term collaboration implies that people share responsibility for the evaluation, including decision making” (p. 6). Therefore, in collaborative evaluation, rather than participatory evaluation, there is a varying level of involvement that considers the extent to which program staff and other stakeholders should be included as part of the evaluation team. Stakeholders bring important program knowledge and insights to any evaluation activity, and to the degree that they are able and willing, they need to be encouraged to participate in the process. The participation of program staff in the evaluation, however, does not rescind the external evaluator’s fundamental
responsibility for the evaluation. By virtue of training and experience, the external evaluator brings to the collaborative evaluation table a wealth of knowledge about evaluation strategies. O’Sullivan (2004) argues that “with a collaborative approach, the program, participant, and evaluation expertise all need to be honored and encouraged to contribute to the evaluation process as much as is feasibly possible” (p. 7).

O’Sullivan (2004) further points out some of the reasons for collaborative evaluation growing in popularity as an evaluation approach. The experiences of evaluators showed that too often evaluation studies went unread, misunderstood, or unheeded. In some programs, staff did not see the relevance of evaluation questions posed by external evaluators and would not even read the report. Evaluation reports were perceived as technical in nature and beyond the understanding of program staff. “In other instances, program developers new to data-driven decision making decided to ignore evaluation findings because it was easier or did not suit their beliefs. Similar events occurred when evaluators failed to gain the confidence of the program staff” (O’Sullivan, 2004, p. 6).

In recent years, there is also more attention paid to bringing evaluation approaches into more alignment with culturally responsive practices as more and more international programs are dealing with different cultures and their values (Askew, Beverly, & Jay, 2012). Furthermore, there is a shift in education to focus more on accountability, specifically through the expectation of effective use of evaluative-focused efforts about educational program implementation (Corn et al., 2012). Therefore, education leaders need specific guidance and training on how to plan, implement, and use evaluation to critically examine district and school-level initiatives. Collaborative evaluation is a strategy to meet these new experiences and extend the responsive approach.

Logically, if teachers delivering the program are collaboratively involved in the program evaluation, their use and understanding of the findings should increase. In addition
to involving stakeholders, “collaborative evaluators need to design and conduct evaluations that are responsive to the organization's culture surrounding evaluation, staff's readiness to conduct evaluation, available evaluation resources, and evaluators' and clients' desired level of involvement” (O’Sullivan, 2004, p. 11). The argument can be made that, if these organizational needs are met, the utilization of the evaluation can be increased through collaborative practices.

Nonetheless, as a result of the inherent complexities of program evaluation, most organizations will choose an eclectic approach to program evaluation which is analogous to the ‘A la carte method’ or ‘carpenter’s toolbox’, and the specific requirements for schools will depend primarily on the context in which the program operates. Therefore, collaboration most often is just one of the many methods used in a program evaluation. The IB, CIS, and MSA all require internal evaluators to complete a self-study specifically designed for the given program, and even dictate how stakeholders need to be involved in the collaborative process.

2.5.2 What is internal evaluation and how does it relate to external evaluation?

As discussed above, collaborative evaluation can be very beneficial. It can result in the collection of more valid data since stakeholders have knowledge and perspective about the specific context of their program that evaluators do not have. Moreover, if evaluation is a result of a partnership between the stakeholders and the evaluators, the use of the findings will become more likely. This would certainly relate back to the professionalization of teaching for educational change. Therefore, self-evaluation, such as the guided self-study of the IB Diploma Program, can be perceived as a valid, reliable and easy-to-use instrument to commit professionals to external quality assurance.

Internal evaluation is the “monitoring of any aspect of a school’s work by its key stakeholders: its staff, its pupils, its parents. … Internal evaluation is usually seen as
synonymous with self-evaluation” (MacBeath & McGlynn, 2002, p. 14). Self-evaluation is a collective gathering of data. According to MacBeath and McGlynn (2002), self-evaluation “involves teachers and school leaders coming to judgments based on their first-hand knowledge of what is happening in classrooms, workshops and laboratories throughout the school” (p. 15). Whenever existing conditions of the program are “examined with a view to its improvement, internal evaluation is at work, often informally but increasingly in a self-conscious and systematic way” (MacBeath & McGlynn, 2002, p. 16).

On the other hand, “external evaluation is used to mean the review and reporting on a school’s work by people who are not part of the school’s organization. External evaluators may belong to different agencies and come with different mandates. Local authority personnel, inspectors and advisers, have long played a role in reviewing school performance, with varying combinations of audit and support, feedback and advice” (MacBeath & McGlynn, 2002, p. 15).

MacBeath and McGlynn (2002) argue that “an effective system of school evaluation needs to contain elements of both internal and external evaluation” (p. 16). The school is based at the centre of school evaluation, which represents self-evaluation. Self-evaluating schools want to know “How good they are?” and “How do they know?” Secondly, schools are compared at the local level, next at the national level, and finally international evaluations are considered (MacBeath & McGlynn, 2002).

MacBeath and McGlynn (2002) argue that self-evaluation has to become an integral part of teachers’ work since all professionals involved in education benefit from the knowledge of how well they are performing and the extent to which their own judgment agrees with the judgment of others. “More and more [teachers] have to realize that making things better for young people demands a commitment to evaluating what [schools] are offering on a monthly, weekly, and daily basis. Evaluation is becoming a habit rather than an
ad hoc, one-off response to a crisis or unexpected review” (MacBeath & McGlynn, 2002, p. 14).

The case for external evaluation is then made as a “reality check” (MacBeath & McGlynn, 2002, p. 20). The authors acknowledge the presence of “twin dangers: Critics of internal evaluation point out that it can lead to self-deluding while critics of external inspection argue that it leads to a culture of dependency. Both are a valid concern” (MacBeath & McGlynn, 2002, p. 24). In order to reconcile these twin dangers, MacBeath (2008) suggests the development of “internal accountability” as an ideal condition in schools which can respond to the external pressure of performance (p. 396). “Internal accountability, moral and professional, implies openness of dialogue, to the nature of evidence, a form of self-evaluation that is genuinely embedded in teachers’ thinking and day-to-day practice” (MacBeath, 2008, p. 396).

When teachers open their practices up to each other with the intention to learn, rather than to judge, professional trust can develop which enables the creation of learning cultures where leadership and learning meet and senior leadership becomes lead learners (MacBeath, 2008, Fullan, 2014, Fullan & Hargreaves, 2016). Marsh (2015) also situates learning and leadership as a community-wide activity and responsibility. The author encourages valuing learning partnerships with parents and other stakeholders, contextualized within local, national and even international communities, advocates for the expansion of learning beyond the school’s physical environment, and reinforces the centrality of collaboration in a learning culture where a holistic understanding of learning is pursued, supported, evaluated and improved. These practices will increase student achievement, which is the ultimate goal for program evaluation.

Caldwell (2008) suggests that self-management of schools is an appropriate strategy considering that each school has a unique mix of needs, interests, aptitudes and aspirations
which can provide conditions for a positive impact on student learning. In these schools self-evaluation can take root and external review, or inspection, plays a valuable supportive and challenging role.

2.5.3 What are the pitfalls of using the collaborative evaluation approach for accreditation?

The program evaluation requirements for the Diploma Program established by the International Baccalaureate blend internal self-evaluation and external inspection. The IB provides the standards and practices that the school’s program is measured against. The school has to examine their implementation of the 177 practices based on the evidence they collect and judge the implementation of them on a scale from high to low. Decisions are then based on these indicators on how to improve the program. All of the major stakeholders have to be actively engaged in the completion of the guided self-study required by the IB. The stakeholders include the teaching staff, the administrators, and owners of the school, the students, and the parents. However, their participation in some cases is limited by their understanding of the program and also by cultural considerations.

Podsakoff and Organ (1986) also point out “the apparently widespread suspicion that self-report methodology is the soft-underbelly of the organizational literature” (p. 531), despite its ubiquitous use as a form of data collection. It seems that organizational researchers and accreditation bodies do not necessarily like self-reports, but neither can they do without them. Nonetheless, managers and school administrators rely heavily on their own self-report research in employee opinion surveys, program evaluation, and human research planning.

The dilemma is that self-evaluation reports, at best, “provide soft data, perhaps better than mere opinions with no data at all, but vastly inferior to most other kinds of data” (Podsakoff & Organ, 1986, p. 531). Many teachers and administrators might feel that honestly reporting on the school’s practices can endanger their jobs and, therefore, they do
not accurately describe concerns and shortcomings. Owners of the school might not be educators and could be only interested in profits so they might not want to know about the real educational needs that exist and focus on window dressing for the parents who pay the tuition for their children.

As a result, self-evaluation bias, among other variables, has been identified as a reason why self-reporting might not be considered as reliable data. There is a serious risk of “dramaturgical compliance”, window dressing, which strongly jeopardizes the legitimacy of the accreditation process (Kemenade & Hardjono, 2010, p. 257). In plain terms, school leadership members might not report the actual state of affairs on self-evaluations, and this window dressing might be committed unintentionally to protect themselves. Daigneault (2014) also reports that “relying only on self-reported data could generate potential social desirability bias in that participants might be tempted to overestimate their level and participation [in] and use [of the evaluation]” (p. 178). Therefore, as Novicevic et al. (2008) conclude, “it is important that team members and their coordinating managers experience extensive training based on their differential propensity for unethical decision making, and that they also receive training in other potential decision-making biases” (p. 1085).

In their article, “A critique of the use of self-evaluation in a compulsory accreditation system”, Kemenade and Hardjono (2010) suggest that writing a self-evaluation report should not be part of a compulsory external quality assurance system. They argue that it is not a reliable instrument for control. Self-evaluation should only be used for internal quality management, where it can be a powerful instrument for improvement. Kemenade and Hardjono (2010) offer the following conclusion regarding the use of self-evaluation in accreditation: “In an accreditation system that is compulsory and that has serious consequences, professionals will be careful in writing down the truth, expressing their
weaknesses and showing their vulnerability. That is a serious drawback, especially in education” (p 265). Therefore, external evaluation should be used for accreditation purposes.

The IB conducts evaluation visits to address these concerns and verify the contents of the self-evaluations. Ten percent of authorized Diploma Program schools are selected for external evaluation visits. Through the use self-evaluation and inspection visits the IB practices quality assurance and accountability in the IB Diploma Program. Research suggests that self-evaluation seems to be more useful for internal assessment that is undertaken voluntarily for the purpose of performance monitoring as opposed to being required as part of a mandatory accreditation process. However, current practices by most organizations do not follow what the research is showing to be true: most organizations will “dress up” their programs to protect their own interests.

Yet “dressing up” programs can undoubtedly make them better if the practices that are changed for the accreditation process actually become part of the regular delivery of the program. As the IB (2010) points out, “Programme evaluation is both a requirement and a service provided by the IB Organization to IB World Schools. The aim is for the IB to ensure on a regular basis that the standards and practices of the programmes are being maintained” (p. 1). It is a requirement that the school has to perform but it is a service as well since practices that need to be changed or updated are scrutinized both at the school level and at the IB level. Ultimately, even “dramaturgical compliance” can bring about needed reform.

2.6 Conditions for Successful School Evaluation

Around the world, there is a growing interest in using school evaluation as a key means to identify areas of improvement at the school and at the system level (Altrichter & Kemethofer, 2015). Consequently, school leaders are tasked with ensuring that the mandated evaluation activities result in higher school performance and student learning. Many action plans require specific learning targets to be set with an often daunting schedule to reach these
set goals. However, the connection between school evaluation and its intended outcomes is not so clear-cut, and there is limited evidence to suggest that school evaluation activities always lead to increased school performance and student learning outcomes (Eddy-Spicer et al., 2016; Ehren et al., 2014; Ehren, Altrichter, McNamara & O’Hara, 2013).

Although there is a lack of data in the literature to establish a causal relationship between self-evaluation and student achievement, evidence is emerging that completing self-evaluation activities may contribute to changes in school practices (Mason & Calnin, 2018). Nevertheless, many of these studies rely on school self-perceptions of change and effectiveness (Schildkamp et al., 2009; Ehren et al., 2014; Meuret & Morliax, 2003), rather than externally reported data as indicators of change. Meuret and Morliax (2003) find that self-evaluation processes, when conducted as participatory exercises, are positively related to school perceptions of their own effectiveness and their capacity to improve. Similarly, Gustafsson et al. (2013) also find strong, positive linkages between the quality of school self-evaluations and the frequency with which schools adopt improvement activities. Longitudinal studies of schools that conduct self-evaluation have also shown small improvements in teachers’ attendance at professional development, teachers’ perceptions of school functioning, and teachers’ level of achievement orientation after conducting self-evaluation processes (Schildkamp et al., 2009). One study suggests that schools with more advanced self-evaluation systems also tended to have a higher quality of teaching and learning processes than those with less advanced self-evaluation practices (Hofman, Dihkstra & Hofman, 2009).

How then can school leaders still use the evaluation requirements to attempt to positively impact the school’s program? In their review of empirical literature on school evaluation, Mason and Calnin (2018) promote “a model of school evaluation that combines external accountability with strong self-evaluation and on-going support” (p. 4). The authors
argue that eight common factors can be identified that support a relationship between school
evaluation activities and school performance:

1. Adopting differentiated models: Although data are inconclusive on the ‘ideal’
frequency of school inspection, Ehren et al (2014) found that the impacts of school
inspection on school and teaching conditions tend to last two years. However,
differentiated inspection models (i.e., those that target potentially failing schools with
more frequent visits) have been shown to have the greatest effect on school and
teaching conditions (Ehren et al., 2014).

2. Embedding external accountability: Higher stakes evaluation models (i.e., those
where there are consequences that flow from evaluation findings) have a greater effect
on the school’s adoption of improvement activities (Altrichter & Kemethofer, 2015).
However, schools within higher stakes inspection models also experience higher rates
of negative outcomes, and there is little evidence that high stakes models positively
impact student learning (Ehren et al., 2014).

3. Establishing public reporting: Similar to accountability pressure, schools whose
inspection results are shared publicly—at the individual school level—tend to report
greater improvements in school and teaching conditions after inspection visits (Ehren
et al., 2014). Additionally, when parents are involved in the performance monitoring
process and receive public report cards, school performance and outcomes tend to
improve (Eddy-Spicer et al., 2016; Barr et al., 2012).

4. Providing ongoing support and capacity building: Ensuring that schools have the
capacity to undertake—and to use—evaluation findings is essential if the potential
link between school evaluation and school improvement is to be realised
(Schildkamp, Visscher & Luyten, 2009; Schildkamp, VanNoof, van Petegem &
Visscher, 2012). When school staff receive training, toolkits and guidance on how to
interpret and use evaluation results, the impacts on school management are greater
(Lassibille et al., 2010). Additionally, when parents build the capacity to interpret
performance monitoring results, service delivery and learning outcomes tend to
improve (Eddy-Spicer et al., 2016).

5. Embedding external advisors/facilitators: Schools which conduct self-evaluation with
the support of Advisory and Research teams (i.e., individuals with expertise in both
educational effectiveness and evaluation) have positive effects on student
achievement (Demetriou & Kyriakides, 2012; Antoniou & Kyriakides, 2011).
Ongoing, collaborative support from these Advisory and Research teams is also
critical for the longevity of these effects (Antoniou & Kyriakides, 2011).

6. Building learning cultures: When school leaders place value in understanding
performance rather than simply ‘looking good’, the connection between school
evaluation and school performance is enhanced (Eddy-Spicer et al., 2016). Schools
categorised as being “learning organisations” (i.e., organisations where teachers and
leaders work collaboratively, agree on improvement strategies and accept
responsibility for learning outcomes) demonstrate higher academic performance
(Verscio et al., 2008).

7. Providing instructional feedback: Feedback from school inspections tends to provide
limited guidance around how schools can improve their performance and what
constitute effective ameliorative strategies (Visscher & Coe, 2003). Consequently,
few schools engage with findings from school inspection reports (Gaertner & Pant,
2011). Feedback that offers guidance on how to respond to areas of weakness tends to
support greater change.
8. Ensuring participation and autonomy: When decisions about what to measure as part of school evaluation involve participatory processes in which school communities select indicators relevant to their needs and interests, gains in student learning appear to be greater (Barr et al., 2012). (Mason & Calnin, 2018, p. 4-5)

As of September 2017, the IB’s Schools Division has been engaged in the redesign of its program evaluation to enhance the current self-study and visit model with approaches and procedures that improve the experience and results for schools while making the best use of IB resources to provide on-going support to IB schools. A cross-divisional redesign group has been charged with leading discussions regarding the new evaluation processes that will reflect the IB’s updated standards and practices coming into effect in 2020 (IBO, 2017). A combination of relevant research, as presented in this study, and practitioner experiences will aid the IB in this important work. This dissertation reports participants’ experiences during the evaluation of the Diploma Program at the school, the changes that the evaluation introduced to the program and their perceived impact, and the future direction the participants suggest for IB program evaluation. As such, the findings of this study can inform the design of the IB evaluation as it unfolds.

2.7 The Future of Collaborative Program Evaluation

What should be the direction for this ever-growing field in program evaluation? Cousins, Whitmore, and Shulha (2012) acknowledge the “evident growth in the popularity and credibility approached to inquiry” (p. 8). However, the authors find that some developments in the field are not as productive as they hoped them to be, especially the use of program logic models to represent collaborative evaluation theory and compartmentalizing collaborative approaches (Cousins et al., 2012, p. 8). They argue that “collaborative inquiry in evaluation is about the relationships between trained evaluation specialists and nonevaluator stakeholders (i.e., members of the program community, intended program beneficiaries, or other persons with an interest in the program) and that practice should, in the first place, be sensitive to stakeholder interests and context, and it should be principle-driven”
Gargani (2013) encourages theorists to “more prominently feature collaborative approaches to developing communities of practice that already exist” (p. 87) and use logic models in a way that would allow practitioners to learn from them.

Alpert and Bechar (2007) also conclude that “there is a major advantage to collaboration between teachers and researchers for academia. …teachers have practical knowledge that is valuable for understanding educational phenomena” (p. 252). They claim that “learning while doing is still the best way to acquire the skill of evaluation” (p. 253).

Cousins, Whitmore, and Shulha (2012) further state:

The program context is the ever-present filter through which subsequent activities and decisions flow. Understanding the context in which we work is central to what we do, why we do what we do and how, or the methods we use. As a context changes so will the appropriateness of decisions around the purpose and form of collaboration. (p. 16)

Riffert (2005) similarly maintains that “the school is the essential education unit” (p. 244). As such, the author suggests the creation of a “school specific measurement tool” as an instrument of self- or internal evaluation as opposed to the one-size-fits-all approach of many mandated evaluations (p. 245). Riffert argues that such a flexible approach, despite the seeming dilemma, is possible to take into account the unique contexts of each school. Indeed, the evaluation process itself can be very time-consuming and teaching staff can feel like the evaluation is never-ending.

When is an evaluation finished? Cronbach (1980) suggests that schools should aim for good enough evaluations. This idea simply implies that the evaluation is good enough if it can help the school staff make decisions about the program in some way. It is not good enough if it either fails to do so or actually leads the team to making wrong decisions. What is good enough is a judgment call and highly contextual.

Datnow (2011) revisits collaboration and contrived collegiality in the age of data-driven accountability. She posits that we need to consider the impact of recent educational
reform models on the professional lives of teachers. While it is likely that sharing more information with teachers about teaching and learning at the school will inform educators’ practices, it does take additional time and increases teachers’ work load. The density of information flow that is related to a mandated evaluation can be overwhelming for teachers. School leaders have to acknowledge that “using data to guide instruction is time intensive work” (Datnow, 2011, p. 157). Additionally, teachers need to be reassured that they have the flexibility to actually make changes in their classrooms based on the data.

It is all well-intentioned to prepare action plans as part of the evaluation documents, but if practitioners do not have the resources or time to teach in new ways then the action plans are worthless. Moreover, leaders should encourage teachers to provide feedback on the effectiveness of the data they are asked to use and the tools that are used to get said data. As Fullan and Hargreaves (2016) point out, collaboration without an appropriately established culture of learning is overrated.

Moreover, the evaluation of practices needs to be viewed as a change process and principals and program administrators as change leaders (Fullan, 2010, 2011, 2014). A good change process shapes and reshapes good ideas while developing capacities and ownership in a group of educators. Good leaders use the group to change the group. They reduce the fear of change and create a nonjudgmental climate by making it acceptable to fail while trying something new. Successful principals talk the walk – they can articulate what they are doing and why they are doing it. In collaborative cultures, it is essential to develop leadership in others by acknowledging that today’s junior leaders will become tomorrow’s leaders, by encouraging many people to look at the problem at hand to develop the best solutions, and by admitting that no one is indispensible, not even the principal. In these environments, school leaders need to become nuanced in order to go beyond surface-level change and tackle complex challenges like evaluations with depth and clarity. Most recently, Fullan (2018)
advocates for three habits of nuance in change leadership: joint determination, adaptability, and culture-based accountability. With the aid of these habits, evaluation leaders can develop processes that are appropriate for the given context of the school.

Sahlberg (2010) rethinks accountability in a knowledge society by trying to combat “the conditions of distrust and blame” suggested by Tschannen-Moran (2007, p. 100). The author proposes that there are two change forces present in education reform that are more contradictory than complimentary. Sahlberg (2010) argues that “accountability policies have trapped teachers in a dilemma between schooling for social capital and moral purpose with student-centred pedagogy on one side, and efficiency-driven education with teacher-centred instruction and achievement on the other” (p. 49). The solution lies in intelligent accountability that uses a wide variety of data to ascertain the genuine strengths and weaknesses of a given school in meeting its goals. This type of intelligent accountability is only possible if educational change efforts focus on building trust and collective responsibility in schools and their communities and on allowing autonomy for schools to decide their own teaching arrangements and how to measure their own achievement against goals they set themselves.

Instead of the competition that is a result of school rankings, school leaders should be cooperating and networking with other schools. Educational institutions should cultivate attitudes, skills and cultures that are needed in collaborative learning environments. Students and teachers should feel safe to take risks and explore the unknown in order to be creative and innovative in their approaches. Evaluation processes, therefore, need to keep pace with the demands of a knowledge society.

Cousins, Whitmore, and Shulha (2012) offer criticism of the recent developments in collaborative evaluation practices and hope to start “ongoing constructive discussion and dialogue” (p. 19) that will result in a “specific set of principle-development strategies” (p. 18)
for the field. These discussions are necessary in order to address the growing demand for culturally-sensitive program evaluation that is needed for the ever-growing international education community. Fetterman et al. (2014) offer their own criticism of Cousins et al.’s conclusions. These authors’ advocate for the distinguishing of stakeholder involvement approaches as follows: collaborative evaluators are in charge of the evaluation but they “create ongoing engagement between evaluators and stakeholders”, participatory evaluators jointly share control of the evaluation, while empowerment evaluations view program staff members, participants, and community members as in control of the evaluation (p. 145). In light of this type of delineation of approaches, internal and external evaluators need to be clear regarding the type of evaluation in which the school is engaging.

Shulha et al. (2016) introduce a new set of principles to guide collaborative approaches to evaluation, which they base on the experiences of evaluators who have engaged in collaborative evaluations in a wide variety of settings and contexts. The authors propose the following integrated set of principles: clarify motivation for collaboration, foster meaningful relationships, develop a shared understanding of the program, promote appropriate participatory processes, monitor and respond to the resource availability, monitor evaluation progress and quality, promote evaluative thinking, and follow through to realize use of evaluation findings. Shulha et al. (2016) expect these interconnected principles to further evolve as evaluators learn more about collaborative approaches in different contexts.

Therefore, in the ever-changing landscape of evaluation approaches, constant updating of the self-study questionnaires and processes for the IB programs is necessary in order to appreciate the complex, complicated, or sometimes even chaotic contexts in which teachers as professionals are delivering and evaluating their programs. This study explored how one specific school navigated the requirements of the mandated accreditation evaluation of the IB Diploma Program.
2.8 The Need for This Study

Large-scale school improvement networks such as the International Baccalaureate Organization are a mechanism for ushering in educational reform. At the heart of educational change is a new leadership paradigm, which focuses on creating a learning culture in schools that supports the collaborative professional development of teachers (Byrk et al., 2010; Fullan & Hargreaves, 2016; Glazer & Peurach, 2013). For these cultures to be successful, leadership teams need to take into account the emotive nature of teaching and develop trust within their schools, and they have to provide more autonomy and time for teachers to improve their practices (Hargreaves, 2005; Louis, 2007; Tschannen-Moran, 2014).

In order to evaluate the effectiveness of such programs and to ensure the participation of teachers as professionals in the evaluation processes, collaborative evaluation has emerged, which includes program stakeholders as an integral part of the evaluation process (O’Sullivan, 2004, 2012). Most organizations use a combination of internal and external evaluation, which combines the monitoring of a school’s systems by its key stakeholders such as its teachers, students, and parents as a form of self-evaluation, and the reviewing of the school’s practices by agencies for inspection or authorization purposes (MacBeath; 2008; MacBeath & McGlynn, 2002). Self-reporting bias and dramaturgical compliance have been identified as possible pitfalls of these processes (Kemenade & Hardjono, 2010; Podaskoff & Organ, 1986). Several conditions for successful school evaluations have been identified in the research literature that school leaders can employ in an attempt to ensure positive impact on school performance and student achievement (Mason & Calnin, 2018).

Researchers promote intelligent accountability that fosters learning cultures as a way to ensure school improvement that results in better student achievement and well-being. To ensure the creation of such school environments, program planning and evaluation processes need to be viewed as change processes and school leaders as change leaders. Nuanced
leadership attributes are necessary in order to serve the specific needs of a given school within its unique context. Evaluation tools and requirements will need to adapt to the contextual nature of evaluation processes as opposed to the current one-size-fits-all approach (Fullan, 2018; Fullan & Hargreaves; Riffert, 2005; Sahlberg, 2010).
Chapter Three: Methodology

3.1 Introduction

In order to address the research questions, I used a qualitative design for this project. Specifically, I employed the case study design as I focused on the participants’ perceptions and recollections of the evaluation of the Diploma Program at one authorized IB Diploma Program school that took place during the 2013-2014 academic year.

The project was exploratory and uncovered processes used by teachers during the evaluation that affected the practices implemented in the IB Diploma Program at the school. A qualitative approach was chosen to gain a deeper, more authentic and descriptive perspective on the issues being investigated (Denzin & Lincoln, 2000). Moreover, a qualitative design was selected to allow for a more open-ended approach to data collection methods, in which the participants within the case could choose how they would articulate their responses (Creswell, 2012). Since my study focused on providing a rich description of the evaluation at the school, only a minimal portion of the survey used a quantitative component to collect background information about the teachers answering it.

3.2 Research Design

Creswell and Guetterman (2019) define a case study as “an in-depth exploration of a bounded system (e.g., activity, event, process, or individuals)” over time through detailed, in-depth data collection involving multiple sources of information rich in context (p. 477). The systems in case studies are bound by time and place, and it is the program, event, activity or individuals that are being studied. Multiple sources of information are used which include observations, interviews, audio-visual material, documents and reports. The context of the case involves situating the case within its setting, which may be a physical setting or the social, historical, or economic setting of the case. The type of analysis of the data can be a holistic analysis of the entire case or an embedded analysis of a specific aspect of the case.
The investigator narrates the study through techniques such as a chronology of major events followed by an up-close or a detailed perspective about a few incidents. In the final interpretive phase, the researcher reports the lessons learned from the case.

Bogdan and Biklen (2007) describe qualitative research in education. They identify five features of qualitative research that are present in qualitative studies to varying degrees, with participant observation and in-depth interview studies being exemplary:

1. **Naturalistic**: The researcher enters and spends considerable time in the school and the researcher’s insight is the key instrument for analysis.
2. **Descriptive Data**: The written results of the research contain quotations from the data to illustrate and substantiate the presentation.
3. **Concern with Process**: The researcher is concerned with process rather than simply with outcomes or products.
4. **Inductive**: The data is analyzed inductively. The researcher does not search out data or evidence to prove or disprove a theory before entering the study; rather, the patterns are built as the individual instances are grouped together.
5. **Meaning**: The researcher is interested in how different people make sense of their lives. Therefore, participant perspectives are paramount. (p. 4-8)

Data collected and analysed in earlier parts of the study were more explorative and key issues, recurrent events, and activities that surfaced became topics for later parts of the study. This earlier data guided both data sources collected and the data analysis. The core themes were then further examined through the lens of already existing research, and I looked for additional examples in the document review, the survey responses, and subsequent interviews.

### 3.2.1 Research Context

The IB Diploma Program (DP) is an academically challenging and balanced program of education with final examinations that prepare students, aged 16 to 19, for success at university and life beyond. It has been designed to address the intellectual, social, emotional and physical well-being of students. The program has gained recognition and respect from the world’s leading universities (IBO, 2009). Schools seeking authorization to offer the IB Diploma Program go through a lengthy process to meet the external requirements of the IB.
After receiving authorization, the school is expected to evaluate its program every five years (van Loo, 2015). The aim for program evaluation is for the IB to ensure on a regular basis that the standards and practices of the program are being maintained (IBO, 2010). Program evaluation is both a requirement and a service provided by the IB to IB World Schools. The IB views program evaluation as a service to schools since it supports school self-reflection and on-going school improvement. The current process has its roots in a regional initiative developed in the late 1990s that was expanded to the other regions with the publication of the first *Programme Standards and Practices* in 2005.

In 2010, the IB underwent a global restructuring and the design of the evaluation process became the responsibility of the new Global School Services department. This department developed a process to be used across the three newly established regions with the intent of producing more consistency in experience and outcome for schools. The process was based on the Programme Standards and Practices updated in 2010 and introduced evaluation visits for select Diploma Program schools (IBO, 2017).

The *Programme Standards and Practices* (IBO, 2005) is the foundational document used by schools and the IB to ensure quality and fidelity in the implementation of its programs in IB World Schools. It is expected that the school makes a commitment towards meeting all the standards, practices and program requirements. This process allows the IB to work closely with the schools in their on-going development of the programs. The evaluation does not seek to appraise or assess individual teachers or students. It is a process of formal reflection involving all stakeholders within the school community (IBO, 2010).

Within this process, the school is expected:

- to determine its own assessment of the implementation of the program, according to the *Programme Standards and Practices* and program requirements,
- and to identify major achievements during the period under review and to identify practices that need further development (IBO, 2010).

To this end, the school completes and submits a self-study of the Diploma Program as required by the IB. Upon receipt of the self-evaluation, the IB assigns external evaluators to the school. During this external evaluation, the IB is expected:

- to analyze and evaluate the school’s implementation of the program, according to the *Programme Standards and Practices* and program requirements,

- to commend schools on practices that address the *Programme Standards and Practices* in ways that solve challenges faced by the school and/or show outstanding implementation,

- to provide guidance on enhancing the implementation of the program in the school,

- and to point out areas within a school’s practice that, if not addressed immediately, will jeopardize the integrity of the program and thus the school’s entitlement to be considered an IB World School (IBO, 2010).

The specifics regarding the process of the evaluation at the school are provided in Chapter 4.

### 3.3 Participants

The project is a case study involving one school after the completion of the mandated five-year evaluation of the IB Diploma Program. All 30 teachers and two administrators (the High School Principal and Assistant Principal) who taught and supervised the Diploma Program during the evaluation were invited to complete a survey designed by the school after the completion of the evaluation. These teachers and administrators were chosen to be invited to complete the survey because they were required by the IB to participate in the completion of the self-study.
As the researcher, I asked for volunteers from the practitioners who were required to participate in the evaluation since these were the teachers and administrators who had the knowledge and experience necessary to engage in the research project. Although all 32 teachers and administrators were invited to be part of the interviews, only three staff members volunteered to participate further in the study. One participant was the High School Principal and two participants were Heads of Departments who also taught in the Diploma Program at the school. The participants were part of the evaluation team that collaborated in the completion of the self-study on a volunteer basis. The evaluation team consisted of the Diploma Program Coordinator, the High School Principal, one representative from each of the six subject groups (Language A, Language B, Humanities, Sciences, Mathematics, Arts), one Creativity, Activity, and Service (CAS) advisor, and one Theory of Knowledge (TOK) teacher.

3.3.1 The School

A private, proprietary IB school in the Middle East was the site of the study. The characteristics of the school are further described in the findings in Chapter 4. I served as the Diploma Program Coordinator during the evaluation of the school’s Diploma Program. This close connection to the case study and its possible ramifications was explained to the Superintendent of the school prior to the commencement of the study and was included in the School Consent Letter (Appendix A).

3.3.2 The Interviewees

Three staff members of the school volunteered to participate in the semi-structured interviews after the completion of the evaluation and their contributions to the survey. None of these staff members reported to me as the Diploma Program Coordinator and it was made clear to them that any responses given to the interview questions would not be used in any
professional evaluations for their positions. This was reiterated in the Participant Consent Letter (Appendix B). Pseudonyms were used for the three interviewees.

3.3.2.1 Lewis

Lewis was the High School Principal during the evaluation. He was in his third year at the school and his first year as principal. He previously served as Head of the Physical Education Department and then as Assistant High School Principal. The interviews happened after the completion and submission of the evaluation during Lewis’s second year as principal of the high school.

During the first interview with Lewis on February 14, 2015, he explained that he had been an administrator for thirteen years at the time of the study. Before starting out in international education and moving to the school, he trained and worked in a large publically-funded school board in a metropolitan centre in Canada. Lewis had not taught Diploma Program courses and prior to becoming assistant principal, he had not have any experience with the DP.

He completed IB DP training for administrators once he took on the post of assistant principal. Lewis identified collaborative leadership as being an important aspect of his practice. He had previous experience with evaluating schools according to Ministry of Education requirements at the school board in Canada, and the MSA accreditation pre-dating the DP evaluation at the school. Lewis also participated as a team member in an accreditation visit to a school in a different part of the Middle East. He noted that these experiences prepared him for being an integral part of the Diploma Program evaluation at the school.

As the High School Principal, Lewis was my line manager in my role as the Diploma Program Coordinator. Consequently, we worked closely together on all aspects of the evaluation from determining what the requirements were to developing a plan of action and timeline for the evaluation activities and preparing the final version of the self-study for
submission. We met weekly as pedagogical leaders for the Diploma Program at our school, and the evaluation was a large part of our discussions. Lewis expressed that while he did not lead the evaluation efforts, he was very heavily involved in all activities that resulted in the finished self-study and related documents.

### 3.3.2.2 Joanne

Joanne was the Head of the English Department at the school. She was in her fourth year at the school and her first year as head of department during the evaluation. She maintained the same position during the interviews. Joanne previously taught Diploma Program English courses at the school and maintained her DP courses after her appointment to this leadership role. According to her responses during our first interview on February 14, 2015, Joanne’s training and earlier teaching experiences also took place in a large publically-funded school board in a metropolitan area of Canada. This school was her first international teaching post and she was in her twelfth year of teaching at the time of the study.

Joanne recounted that she was part of the MSA evaluation efforts completed by the school in the previous academic year. She was part of the Facilities Committee; however, she did not feel that she had an “integral or integrated” role and felt quite disconnected from the process of the MSA evaluation (Joanne interview, February 14, 2015). Yet, she acknowledged that working on the MSA evaluation provided her with institutional knowledge that became useful during the DP evaluation. Indeed, in addition to volunteering to be a member of the evaluation team who worked on all three sections of the self-study, she also completed Chart 3-Facilities since it was a natural fit after her work on the MSA accreditation. As such, Joanne was an important contributor to the DP evaluation activities.

### 3.3.2.3 Bill

Bill was a teacher in the Humanities Department during the evaluation and was promoted to become the head of the same department by the time the interviews took place.
He was in his second year at the school during the evaluation. During his first interview on February 15, 2015, Bill stated that he was in his fifth year of teaching after completing his teacher training in a university on the west coast of the United States. This school was his second international teaching experience after spending two years in South-East Asia.

Previously, Bill experienced program evaluation in an alternative-program school in the US, which was conducted in order to maintain grant funding. He did not participate in the MSA evaluation as a leader but Bill believed that the internal evaluators did a good job to keep teachers informed regarding the reaccreditation process and he felt ready to join the DP evaluation team as the representative for the Group 3 Humanities teachers. Together with the then head of his department, Bill also led the group’s completion of Section C-Curriculum. In addition, he volunteered to participate in the writing of Section A-Philosophy of the self-study. As such, Bill was a very active participant in the DP evaluation.

3.4 Data Collection

My study took place during the 2014-2015 academic year after the DP evaluation was submitted to the IB in April 2014 and the response was received from the IB in May 2014. The participant interviews were conducted from February to May of 2015, inclusive. In accordance with the above guidelines for conducting case study research, data was collected from the following three sources: 1) document review; 2) a qualitative survey; and 3) interviews of evaluation team members. The main data set was collected from three semi-structured interviews with Lewis, Joanne, and Bill with their consent. The remaining data came from the document review and the responses to the qualitative survey.

3.4.1 Document Review

A document review of all of the required materials that were submitted to the IB for the purposes of the evaluation in April 2014 and the corresponding response from the IB external evaluators received in May 2014 was conducted at the same time that the participant
interviews were taking place in the spring of 2015. These documents included the completed self-study and its corresponding charts and all supplementary papers required by the IB. Specifically, the following documents were examined as part of the data set: the school’s assessment policy, academic honesty policy, admissions policy, special education policy, and language policy; the description of the self-study process including the timeline of the evaluation activities; the scheduling of DP courses with examples of student and teacher timetables; the calendar of school deadlines for the submission of internal and external assessment components over the two years of the program; the description of the process of the supervision of the Extended Essay; the charts that detail the organization of teaching time for each of the courses, the qualifications and IB-approved training of all teaching staff, the improvements of the school’s facilities to deliver the program, and the budget for the implementation for the Diploma Program; the school’s promotional materials that explain how the philosophy of the IB DP is delivered by the school; and the teacher and parent-student handbooks that explain the specific practices used by the school. The participants referenced all of these documents in their responses to the survey and the interview questions. As such, I needed to be familiar with the contents of the documents. All of these documents were located on the school’s Atlas Rubicon website which was accessible by all staff members at the school. Since I was working at the school, I had direct access to these documents. Once data collection started for the study, the superintendent gave me permission to use these documents.

Since the project was conducted after the completion of the self-study, the main focus was on the recollection of the participants of the evaluation activities that happened to complete the required documentation for the IB. As such, the original meeting agendas and minutes and parent and student survey responses that were collected by the evaluation team to complete the self-study questionnaire were not part of the data collected for this study.
However, participants, at times, reference these materials in their responses as they recount their experiences with the completion of the evaluation documents. This is in line with Bogdan and Biklen’s (2007) statement that “documents clearly fit the criteria of using rich data in description but to what extent the researcher uses them in a manner that is naturalist, inductive, and concerned with the process of meaning construction” depends on each individual case (p. 64).

3.4.2 Qualitative Survey

An electronic qualitative survey was administered to all teachers and administrators who participated in the process of completing the self-study. The survey was prepared as per the school’s regular requirements to collect feedback from teachers and administrators at the end of a school evaluation for the purposes of improving the processes for the next evaluation cycle. The questions for the survey were developed by the members of the evaluation team. The survey was sent out to the teachers by me in my capacity as the Diploma Program Coordinator as per the usual processes employed by the school. The survey was completed by participants during April 28 and May 15, 2014. All 30 teachers and two administrators (the High School Principal and Assistant Principal) who taught and supervised the Diploma Program were invited to complete the survey. These teachers and administrators were required by the IB to participate in the completion of the self-study. As the Diploma Program Coordinator, I also answered the survey questions to provide feedback to the school on the evaluation.

At the close of the survey, 16 out of the 32 invited staff members volunteered to participate by providing anonymous answers to the survey questions. The teacher survey is presented in Appendix C. The anonymous responses were shared with members of the evaluation team and the Superintendent at the close of the survey in May 2014. All three
interview participants were members of the evaluation team. In this capacity, they had access to all of the answers provided to the survey questions.

Once I started to design this project and prepared for the ethical review, it became apparent that the survey previously completed for the school could provide valuable insights for the study. Since many of the original participants had already left the school where the study took place, it would have been difficult to recreate the responses to a new survey. Therefore, at the start of the study in February 2015, the Superintendent gave me written permission to use the responses from the school’s survey as data for this project. Thus, the study presents a secondary analysis of a pre-study qualitative survey. However, the responses from the survey provide a rich description of the evaluation of Diploma Program at the school, which adds an important element to the study.

The Superintendent’s consent was deemed sufficient on behalf of the participants as the responses were given anonymously and no personal information was connected to their answers. In order to try to separate my role as researcher and my role as the Diploma Program Coordinator as much as possible, I asked for clarification from the interviewees when coding and analyzing the data generated by the survey since my answers were already part of the aggregated data and it was not possible to remove them from the data set. Neither the interview participants nor I had access to the identity of the staff members who answered the survey questions and we could not connect specific answers to specific respondents.

During the first interview, each of the three interview participants also reported that they had answered the survey during the original administration of it in the spring of 2014. The Superintendent of the school also provided feedback on the full rough copy of the thesis in order to ensure that the participants’ responses were truthfully reported.
3.4.3 Semi-Structured Interviews

Since the data collection took place after the completion of the evaluation, direct observation of the evaluation activities was not possible. Therefore, the bulk of the data was collected using qualitative interviews. “A qualitative interview occurs when researchers ask one or more participants general, open-ended questions and record their answers. The researcher then transcribes and types the data into a computer file for analysis” (Creswell & Guetterman, 2019, p. 217). Open-ended questions allow the participants to best voice their experiences without any constraints from the perspectives of the researcher. Open-ended responses similarly enable the participants to create options for responding. Qualitative interviews can provide useful information when direct observation of the participants is not available and they also permit participants to describe detailed personal information. However, the information can be filtered through the views of the interviewer if the researcher summarizes the participant’s views in the report (Creswell & Guetterman, 2019).

In order to counteract this disadvantage of using interviews, I tried to use direct quotations of the interviewees’ actual words as much as possible instead of summarising information in the findings.

Each of the three interviewees was asked to participate in three hour-long semi-structured interview sessions, which were audio taped. The first interview was informed by the teacher feedback survey to address the participants’ perceptions of the process of completing the evaluation. During the second interview session, participants identified the effects of the evaluation on the Diploma Program at the school as they saw them. The third session explored the future of program evaluation at the school and IB level. The participants engaged in a “future perfect” reflection: What, in their opinion, was the next step for the DP both at the school and at the IB level in terms of evaluation? The Interview Guide for all three of the sessions can be found in Appendix D.
All interviews were audio taped and transcribed. Each interview was transcribed verbatim and participants had the opportunity to review transcripts of their interviews if they wished. Participants also had the choice to add their own comments, clarify and elaborate on their ideas if they wished. None of the teachers decided to exercise this option.

3.5 Data Analysis

The process of data analysis was done iteratively and concurrently with the processes of data collection and report writing, as these three steps are interconnected (Creswell, 2008). Audio recording and transcribing of interviews, interview logs, memo writing in study journal, analytic files, and development of coding throughout the process was used during data gathering. Situational and contextual analysis was used to develop patterns for coding.

The audiotaped interviews were transcribed and reviewed for accuracy. The transcripts were coded manually without the aid of a computer program since I was only dealing with one survey and three interview participants and the data collected was straightforward enough to classify by hand, guided by the inductive approach (Thomas, 2006). The inductive approach to data analysis, a common method for qualitative data, requires careful readings of the data in the identification of key concepts, themes, and models (Thomas, 2006). In initial coding, an openness to explore possible theoretical explanations is important (Charmaz, 2006). I carefully read through all of the documents, survey responses, and initial interview transcripts to develop a general sense of the key ideas to develop some possible categories and over-arching themes. I asked the participants to provide feedback on this initial analysis of the data during subsequent interviews.

I drew upon previously existing literature about school evaluation for which potential themes could be created and began coding through the use of “coding categories” (Bogdan & Biklen, 2007, p. 173). Apart from setting/context codes that helped me to organize the data sources, codes were mainly created to fit within categories of “definition of the situation” (p.
“perspectives held by subjects” (p. 175), “process” (p. 176), and “relationships and social structures” (p. 177) codes. Within these coding categories, I created more specific codes based on evaluation and supplemented this with specific recurring themes for the participants.

Once all the data was coded, I identified the major themes that emerged from the data using the three main objectives of the study as a guide. With these categories in mind, all the data, including the transcripts, observation notes and documentation provided by the participants, were examined for evidence that supported or contradicted these themes. I conducted within-case analysis for the three participants (Merriam, 1998). Within-case analysis was applied to develop a strong sense of each participant’s perception of the evaluation of the Diploma Program at the school. Similarities and differences were also noted between the participants, which generated more complete conclusions and observations at the end of the project.

3.6 Ethical Considerations

My study was reviewed by the Office of Research Ethics of the University of Toronto and was granted approval for completion. Pseudonyms for the participants and the school were used in this study to ensure confidentiality. The Superintendent and the participants agreed to participate in the study by signing a consent letter (see Appendices A and B) and they were informed of their right to withdraw from the study at any time. The Superintendent also gave me written consent to use the results of the electronic survey that had been conducted prior to the commencement of the study. As per the established protocols of the school, the Superintendent could give permission for the study to be completed at the institution and then inform the Board of Owners. The Diploma Program was part of the High School, and the High School Principal was one of the interviewees who gave his consent for the research to be conducted with his participation.
All of the survey responses were provided anonymously and voluntarily. Since this study was designed to examine the participants’ perspectives on collaborative practices during a mandated accreditation, it was imperative that I asked for clarification from the interview participants throughout the collection, analysis, and interpretation of the data. All of the interviews were audio taped and transcribed, and I specifically asked the teachers about informal discussions during these formal interviews. In this way, I could ensure that I only used quotes from the transcripts for accurate representation of their perceptions.

As Bogdan and Biklen (2007) point out, the researcher’s insight is the key instrument for analysis in qualitative research. Therefore, it is important for a qualitative researcher to protect his or her findings from bias. I engaged in the procedures to monitor my “subjective lenses” (Glesne, 1999, p. 109) and ensure that this study presents its findings in an objective manner, which truly represents the opinions and actions of the participants. These procedures included triangulation of the data sources, audio-taping interviews, and asking for clarification of information from the participants while collecting data and finalizing findings and conclusions.

Since I was the Diploma Program Coordinator at the school during both the process of the evaluation and the subsequent data collection for this project, I was particularly sensitive to the possible ethical considerations of conducting “backyard research” (Trowler, 2011; Zulfikar, 2014). Endogenous research was beneficial for this project since I had better access to both the data and the participants, I was able to produce results that would be actually used by the teachers and administrators of this school and the IB in general, and it was a practical and easier way to collect data. In essence, I felt that I was in the right position to offer a rich description of the evaluation efforts at my school with the help of the participants.

However, I had to be very conscious of the most important problem with completing the research at my own institution: the problem of institutional and personal anonymity for
both the school and the participants (Trowler, 2011). This aspect of backyard research was, therefore, incorporated into the informed consent letters for the Superintendent of the school and the teachers. Actually, the fact that I was part of the leadership team at the school assisted with gaining access to the site since there was precedence of doctoral research being conducted by previous administrators employed by the school. Secondly, when selecting the participants, I ensured that there were no ethical issues around disparities in power; none of the participants were directly reporting to me and it was made clear that the data set was not going to be used in any assessment of their performance at the school.
Chapter Four: Findings

4.1 Introduction

This chapter describes the evaluation of the Diploma Program that took place during the 2013-2014 academic year. First, the specific characteristics of the school are explored including the process and timeline of the creation of the documents for the evaluation. Then the survey results are discussed followed by the information provided by the three participants in their individual interviews. Each of these areas of data collection is examined in light of the three main domains of inquiry: process, effects of the evaluation, and the future of DP evaluation.

4.2 The School Context

The founder of the school sought to create a school of high calibre with a disciplined, yet relaxed atmosphere that would provide opportunities for local and expatriate children to gain access to the world's best universities. The school opened as a private, proprietary organization in 1991 and during the first year, twenty-five teachers and 300 students overcame a shortage of textbooks and classroom supplies, inadequate facilities, and a skeleton curriculum. Since the passing of the school’s founder in 1995, four owners act as the governing board of the school with the educational side of the organization being managed by professional staff. After the school opened, the school developed rapidly, adding staff and students and in October 1994, the school became fully accredited by the Middle States Association of Colleges and Schools (MSA).

In 1995, the school was authorized by the International Baccalaureate Organization (IB) to offer the Diploma Program (DP) as an option for Grade 11 and 12 students in addition to the American high school diploma. In 1993, in response to the MSA self-study, a strategic plan was created outlining significant changes to curricular programs and facilities that would accommodate a growing population of students. In 2005, the school was authorized to offer
the IB Middle Years Programme (MYP) in Grades 6 to 10 and in the following year, the IB Primary Years Programme (PYP) was implemented in the Elementary School. These programs, along with a focus on strengthening students’ literacy and critical thinking skills, led to a greater number of students choosing to study in the Diploma Program in Grades 11 and 12.

At the time of the DP evaluation, the school was the only fully authorized IB World School in the country offering all three IB programs under one roof. An American-based written curriculum is used in most subjects. Along with core subjects, elementary teachers also offer specialist instruction and support in Arabic, Art, Music, Religion, Exploration, Physical Education, and Technology. Once students complete the PYP in Elementary School, they move to the Middle School for Grades 6 to 8, where they embark on their MYP education. Students receive specialist instruction in Language Arts, Math, Science, Humanities, and Physical Education, and have the opportunity to take electives such as French, Fine Arts, Technology, Religion, and Arabic.

The High School works within the framework of the MYP (Grades 9 and 10) and the DP (Grades 11 and 12). Students have the option whether or not to take the full Diploma Program, or a mix of some DP classes and other non-DP classes. As such, some of the graduates of the school are eligible for the IB Diploma, but all are eligible for an American High School Diploma at the conclusion of their program in Grade 12. High School students can select courses from eight subject areas. The subject areas offered to High School Students are: Language A, Language B, Math, Science, Humanities, Fine Arts, Technology, Physical Education, and Religion. For the purposes of this study, only the Diploma Program at the school will be considered in detail.

In order to prepare students for the rigor of the Diploma Program, the school recognizes the importance of high academic expectations from the earliest grades. A School
Improvement Plan (SIP) was adopted as a result of the curriculum review conducted for the Middle States Association of Colleges and Schools (MSA) reaccreditation process in 2012-2013.

The first two objectives from the MSA SIP focused on building capacity for improvement of students’ literacy and numeracy skills. The need for vertical and horizontal alignment within the school’s curriculum was also identified as an important aspect in preparing the students to become IB Diploma graduates. Through the curriculum review process, the Common Core was selected. Consequently, the school was preparing to implement the Common Core State Standards for Language Arts and Mathematics curricula for Grades Pre-Kindergarten to Grade 12 during the DP evaluation process. It was the school’s belief that these initiatives would better prepare students to enter the Diploma Program in Grade 11.

The High School leadership team promotes the Diploma Program to students and parents in a variety of ways. In Grade 10, presentations are made to the students and parents about the Diploma Program starting at the early September Welcome to High School night. The Diploma Program Coordinator, DP teachers, and current DP students are available to provide information and answer questions about the program. It has proven to be a powerful tool to have students who are currently enrolled in the program to act as ambassadors for the merits of the DP. Although the target audience for this presentation is Grade 10 students and their parents, all high school students and parents are invited. Many Grade 11 new DP 1 students and Grade 9 students attend with their parents.

The Diploma Program Coordinator is available during the Parent Conferences in November to answer any questions regarding the program. In January of the Grade 10 academic year, the Diploma Program Coordinator visits all Grade 10 English classes to prepare students for course selections. At the same time, parent meetings are held to inform
parents and guardians about the specific benefits of choosing the DP path. The school also invites parents of alumni and alumni to attend these parent information evenings to share their experiences with the DP. The Diploma Program Coordinator and the Guidance Counsellors are also available for individual parent and student meetings to assist families with making the correct scheduling decisions. Additionally, the Diploma Program Coordinator and the Guidance Counsellors attend the Parent Conferences in April to clarify any questions regarding DP registration for the upcoming academic year.

4.2.1 Context of the Diploma Program Evaluation Process

The school was involved in two separate yet related program evaluations during the 2012-2013 and 2013-2014 academic years. The entire school community from grades Pre-Kindergarten to Grade 12 was part of the Middle States Association of Colleges and Schools reaccreditation process from November 2012 to October 2013, which culminated in a visit by the accreditation team in November 2013. At the same time, the Diploma Program offered in Grades 11 and 12 was undergoing evaluation for the International Baccalaureate Organization from April 2013 to April 2014.

Both evaluations used meetings, surveys, and document studies to gather the necessary information for the completion of the two self-studies. The DP Coordinator and the High School Principal used information generated for the MSA accreditation as part of the DP self-study when it was appropriate. In preparation for the completion of the self-study, the DP Coordinator attended the Category 3: Evaluating Your DP IB training workshop in September 2013.

4.2.2 School Characteristics

The students are children of professional and skilled technical local and expatriate community members. The school is organized into three divisions: Elementary School for Kindergarten to Grade 5, Middle School for Grades 6 through 8, and High School for Grades
9 to 12. Each of the three divisions is headed by a principal and an assistant principal. They report directly to the Superintendent who is responsible for the all of the educational operations of the school. The Superintendent reports to the Board of Owners.

At the time of the DP evaluation, there were 1,826 students enrolled at the school from Pre-Kindergarten to Grade 12. The High School had 401 students. Out of this, 87 students attended Grade 11. The Class of 2014 had 84 graduates. While the majority of the families were from Middle Eastern countries, there were also students from North America and other parts of the world. A total of 12 nationalities were represented at the school in the Diploma Program during the 2013-2014 school year, including several Middle Eastern countries and American, Canadian, Spanish, Indian, Belgian, Chinese, South African, and Australian students. Eighty-four percent of the Grade 11 and 12 students had Arabic as a first language. Only 9% had English as a mother tongue. However, many of the students had been at the school since pre-school and, therefore, had English as the language of instruction most of their academic careers. During the evaluation of the program, 94% of the Grade 12 students were taking at least one DP course, while 74% of the Grade 11 students were taking at least one DP course. However, 41% of the Grade 11 students were enrolled in the full Diploma Program and only 32% of the Grade 12 students were attempting the IB Diploma.

The teachers, with the exception of Arabic and Islamic Studies teachers, held valid North American teaching credentials. More than 80% of the entire school staff held a BA degree or higher. The quality of the staff was considered excellent. There were 30 teachers who were involved in the delivery of the Diploma Program during the 2013-2014 academic year and the evaluation of the program. All of these teachers held teacher certifications and had received appropriate IB Diploma Program training. Fifty-three percent of the DP teachers were American and 27% were Canadian. Other teachers came from Jordan, Lebanon, Ireland, Tunisia, and Australia. All teachers had great command of the English language.
4.2.3 Timeline of the Diploma Program Evaluation

The timeline presented for the completion of the Diploma Program self-study was officially accepted by the Administrative Council at the December 9, 2013, Administrative Council meeting (Appendix E). The Superintendent and the DP Coordinator subsequently presented the timeline for the completion of the evaluation to the Board of Owners of the school.

In accordance with the timeline, the Diploma Program Coordinator scheduled meetings with all stakeholders to participate in the completion of the self-study. The High School Principal and the Diploma Program Coordinator met once a week to discuss issues related to the Diploma Program and the completion of the evaluation was a regular topic. All DP Teacher Meetings touched on principles and practices that were required to be addressed for the evaluation. One of the DP teacher meetings in November was devoted to the creation of the school-based descriptors that was used by all stakeholders throughout the examination of the DP principles and practices.

In addition, teachers met in their subject groups, as a group for Theory of Knowledge (TOK) teachers, and as a group for Creativity, Activity, Service (CAS) advisors to complete Section C - Curriculum of the evaluation and to choose a representative from each group for the Evaluation Team. The Diploma Program Coordinator then scheduled meetings with the Evaluation Team to complete the combined school report for Section C - Curriculum and Chart 5 – Overview Levels of Achievement for the Standards on Section C. Heads of Departments also analysed results from the previous five years for each of the subject groups, which were then compiled by the Diploma Program Coordinator. The Diploma Program Coordinator sent out surveys to both parents and students in the Diploma Program so they could make their voices heard in the self-study.
The High School Principal and the Diploma Program Coordinator reviewed the feedback provided by the parents and the students and prepared the statements and quotes for the evaluation on their behalf. Once all required documents were gathered and the recommendations from the self-study were compiled, the Evaluation Team prepared Chart 7, the Action Plan. The completed self-study was then presented to the Superintendent who approved the self-study for submission as a representative of the Board of Owners in accordance with school policy.

All of the meetings that were conducted as part of the preparation of the MSA self-study and the DP Evaluation have been documented for transparency and institutional memory purposes. A complete list of the types of meetings conducted to implement the Diploma Program at the school is presented in Appendix F.

All agendas, action plans and supporting documentation used were available on the school’s Atlas Rubicon website, the curriculum planning platform, to all teachers who contributed to the completion of the self-study. In addition, all documents generated for the self-studies were stored on the school’s info drive, which was available to all teachers at the school. Parent meetings were conducted to share the action plans with the parent community after the completion of the self-study.

4.3 The Survey

The 16 volunteers who responded to the survey had different levels of involvement in the completion of the evaluation. One participant could choose several of the options depending on how many roles they fulfilled during the evaluation process. Based on their responses, the participants can be categorized as shown in table 1.

The IB requires that all teachers participate in the completion of Section C – Curriculum of the self-study. For all other sections, teachers volunteered to be part of the evaluation team and could choose how much they wanted to be involved. Indeed, 94% of
<table>
<thead>
<tr>
<th>Level of Participation</th>
<th>Number of Respondents</th>
<th>Percentage of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Contributor – Section C (departments)</td>
<td>12</td>
<td>75%</td>
</tr>
<tr>
<td>Evaluation Team Member – Section C (whole school)</td>
<td>5</td>
<td>31%</td>
</tr>
<tr>
<td>Evaluation Team Member – Sections A and B</td>
<td>1</td>
<td>6%</td>
</tr>
<tr>
<td>Evaluation Team Member – Charts 1-7</td>
<td>2</td>
<td>13%</td>
</tr>
<tr>
<td>Diploma Program Coordinator</td>
<td>1</td>
<td>6%</td>
</tr>
<tr>
<td>High School Administrator</td>
<td>2</td>
<td>13%</td>
</tr>
<tr>
<td>Parent of Diploma Program Student at the school</td>
<td>1</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 1: Survey Participants

Survey respondents agreed that they could choose their level of involvement in the preparation of the self-study, while only one respondent strongly disagreed.

4.3.1 Process

Eighty-one percent of survey participants agreed that they were aware of the entire process that was used for the preparation of the self-study. Similarly, 88% felt that they were included in the evaluation process and 94% reported that stakeholders were appropriately represented in the preparation of the self-study during a transparent process. Thirteen participants reported that they believed that there was competent pedagogical leadership for the Diploma Program at the school, and 11 respondents at least skimmed through all of the evaluation documents with 4 of them reading all documents submitted to the IB carefully.

In the general comments section regarding the process of completing the Diploma Program Evaluation, some participants expressed their positive experiences thusly:

Despite the amount of paperwork involved, the process was straightforward. Our MSA Evaluation prepared us well for the Diploma Evaluation.

Transparent and easy to follow and understand.

Having worked in this school’s DP program for several years, I feel the self-study accurately reflects the current reality.
The DP Evaluation was a collaborative process, involving all DP staff. The evaluation was extremely reflective and has brought many positive changes to the DP Program at the school. (Survey responses)

Teachers appreciated that they could choose their level of involvement in the evaluation while they still felt that the evaluation activities were transparent enough that they felt a part of the process.

On the other hand, 1 participant out of the 16 respondents consistently chose to strongly disagree to every question in the survey. Regarding the process of the self-study completion, this teacher wrote:

Frustrating, annoying, time wasting. I was lied to about the real object behind the process. I was told it was for the “students”. In reality, it was for the [DP Coordinator] pursuing her personal goals at the expense of [the school], its students and reputation. The lack of transparency, blatant lies, and creation of a personal fiefdom is appallingly pompous. (Survey response)

Two other participants also expressed their views on the process:

The self-study was loaded on top of the other tasks we were assigned to do in the DP and MYP. Its relevance to our students seemed limited. The relevance of the many documents we were expected to read to our area of expertise, seemed limited. Nor was there enough time allotted to read them. Many of the questions were difficult to answer honestly and authentically because we lacked information.

Not very close to the process, at least not as much as I would have liked. (Survey responses)

While most teachers viewed the process of the evaluation as a positive experience, some were clearly frustrated and annoyed by this requirement.

4.3.2 Effect of the Evaluation

Thirteen respondents agreed or strongly agreed that the documents submitted to the IB provided an accurate and honest depiction of the Diploma Program at the school. Most of the respondents agreed that completing the evaluation made the DP at the school better. These teachers also reported that they developed a better understanding of DP standards, assessment, teaching and learning practices, and the supports required for the DP core after completing the evaluation.
4.3.2.1 The Progress

Teachers described the progress that they observed as a result of the evaluation efforts:

The self-study forced us to reflect on our current practices and make positive and collaborative changes. The process brought on many deep level discussions which forced us to think about our current practices.

It was helpful to go through the experience a second time and see the changes or improvement made by the school since the last evaluation. The review of the assessment policy was very needed and helpful.

When people are more involved there would be more sense of responsibility and willingness.

I believe the integration and focus on teacher responses is essential to the overall team effort of teaching and learning DP. It allows for transparent forms of open dialogue and an understanding of the vision of DP at the school. I am a firm believer of the DP program and my philosophies of teaching are similar to the philosophies of the DP; therefore, I was engaged and involved in the self-study process and keen to respond to the issues addressed. Having classroom teachers involved, and transparent meetings, is essential to reach an authentic action plan.

It allowed us to look critically at our practices and examine the difference between our program on paper and in practice. There are gaps that we have to address.

Many positive aspects - a good school continually self-evaluates and documents. This will help the school immensely. (Survey responses)

Reflection on current practices and the impetus to change what was not working was most often mentioned as a positive effect of the evaluation.

4.3.2.2 The Pushback

While most teachers appreciated the positive changes the evaluation brought to the program, some respondents reported different effects of the evaluation:

As with many other aspects of professional development at the school, there simply is not enough time allotted to do the work asked thoroughly and professionally on top of the many other teaching tasks and teaching preparation tasks we are assigned to do in both DP and MYP. This is not necessarily the fault of the school; we do not have enough school days in the year to accomplish what the IB wants us to accomplish in teaching. It follows that the other support tasks are also artificially limited. So the process is rushed and superficial.

I honestly and truly feel that most paperwork is simply a time-waster. The sheer volume of this means that most of it will only be occasionally referenced, so I wonder about its usefulness-to-time-spent ratio. (Survey responses)
Indeed during the study of the documents that the school completed for the various evaluations during this period, the Superintendent of the school addressed this very issue:

If [the evaluation] leads to systematic, institutional change and growth, we will succeed. The targets we have set for measuring improvements in student performance in mathematics and literacy are impressive. The copious data on which they are based is weighty. They will, however, remain words on paper if they do not provoke thoughtful action on our part. It is imperative, therefore, that each of us knows our strategic objectives, goals and actions and makes them guiding documents in our daily planning. (Superintendent’s Notes, November 6, 2013)

Moreover, there was significant pushback from one participant regarding the effect of the evaluation, which, at times, took on a personal tone against the DP leadership at the school:

[The DP evaluation caused] incalculable damage to the school, students, teachers, and parents as a result of pursuing a self-serving, dictatorial agenda as opposed to a truly collaborative, thoughtful, open, and worldly process. … Please find another line of work to avoid pervasive, long-lasting academic damage to the school, students and teachers. Also, have the decency, morals, and character to include all the responses to your surveys, versus taking only those that stroke your ego and support your agenda. (Survey response)

Lack of time provided for the evaluation and personal feelings towards the leadership of the Diploma Program at the school prove to be the most significant sources for pushback against the process according to survey respondents.

4.3.3 Future of DP Evaluation

Eighty-two percent of survey participants agreed or strongly agreed that the action plan submitted to the IB provides adequate resources for the future of the Diploma Program. One participant suggested to “have DP, MYP, PYP, and MSA [evaluations] in a closer time frame so that duplicate work isn’t done by different groups merely 6 to 7 months apart” (Survey response). Another respondent asked for “more time set aside to do the work” (Survey response). An additional recommendation was to “periodically visit the evaluation report so we keep [the] goals for our program in mind” (Survey response). Finally, one
participant felt that the evaluation was a real team effort and that “[the school] will feel the positive impact on the work completed for many years” (Survey response).

However, one teacher had a very different view of the future of evaluation at the school and continued to express criticism of the DP leadership:

Be willing to hear some point of view other than that which supports your [opinion]. Stop trying to make things "politically appropriate" and massage teachers’, students’ and owners’ egos to make it appear as though you are doing something of value for the school, when, in fact, the object is to be self-serving. (Survey responses)

This example shows that personal feelings towards the internal evaluators can significantly affect an individual teacher’s experiences with the evaluation process.

4.4 Interviews with Lewis

Lewis served as the High School Principal during the evaluation of the Diploma Program in the 2013-2014 academic year. The previous academic year, Lewis was the Assistant Principal of the High School for one year. During this time, he was an integral part of the team who prepared the self-study for the Middle States Association of Schools and Colleges (MSA) reaccreditation requirements. He led the committee tasked with evaluating the MSA standards related to school improvement planning and as such collaborated on the school’s strategic plan.

Lewis and I worked very closely on all aspects of the Diploma Program evaluation process. It was my first year as a Diploma Program Coordinator and my first year working at the school. This coincided with Lewis’s first year as High School Principal. Lewis also mentioned that there was communication to the Superintendent of the school throughout the evaluation process to keep him abreast of our efforts and to ensure that we were meeting his expectations.

4.4.1 Process

In this section, I describe Lewis’s perceptions of the process of the evaluation of the Diploma Program at the school. Lewis asserts that the experience of going through the
accreditation process was “extremely positive … I think it provided an opportunity to bring staff together to work on something in common” (Lewis interview, February 14, 2015). According to his perceptions, the evaluation helps to put structures in place not just in the Diploma Program but in the entire school. Lewis also values the leadership opportunities that an evaluation can open up for staff members across the organization. Overall, Lewis expresses that the evaluation was really beneficial for the school.

4.4.1.1 Systematic Planning of the Evaluation

Lewis recounts that the evaluation process unfolded according to a set plan. The timeline and the sequence of evaluation activities were established by the Diploma Program Coordinator in conjunction with Lewis as the High School Principal and the Superintendent. The plan established each stage of the evaluation from beginning to end to ensure that all stakeholders were appropriately involved. Responsibilities were divided amongst the three leaders to communicate with the different groups throughout the academic year and to determine how the plans would go from one stage to the next.

There were multiple meetings with various groups such as the Administrative Council, DP teachers, and the various evaluation teams during which different parts of the evaluation were discussed and these pieces were consolidated as needed. Most importantly, Lewis believes that the weekly meetings with the Diploma Program Coordinator kept him apprised of where the program evaluation was within the timeline as it developed over the course of the academic year.

4.4.1.1.1 Timeline – Synchronization with the MSA evaluation

Lewis acknowledges that the documents developed for the MSA accreditation nicely align with the Diploma Program evaluation and that it was helpful that he had a thorough knowledge of the MSA self-study since this experience enabled him to seamlessly move to a new self-study. He states:
[Due to] the fact that we had collected a lot of data to support and involve a lot of staff in the MSA accreditation process through surveys and their work on committees, we had a lot of the information, which prepared us for the beginning of the [DP] evaluation. I think that was key because both the Diploma Coordinator and the High School Principal were new, so I think that it was [good] to have that as a starting point. But I also think it was good that we were kind of doing this together and working through it as a first step to my high school principalship and the diploma coordinator position. Actually, I think that ended up being really helpful, that we were starting fresh and moving from that direction. (Lewis interview, February 14, 2015)

Lewis appreciates the fact that different documents did not have to be duplicated for the two different evaluations in which the school was engaged at roughly the same time. This allowed for a better use of time for all the stakeholders involved.

4.4.1.2 Level of Involvement

Lewis states that, while he did not lead the DP evaluation, he was an integral part of all evaluation activities as appropriate for his role as High School Principal. The Diploma Program Coordinator, who was the leader of the evaluation, continuously met with Lewis to keep him informed about where the teams were in the timeline that had been previously developed. He further explains:

We still do meet because we constantly evaluate the program, whether it is a formal evaluation or not, but we meet an hour every week to go over different aspects of the Diploma Program. So, that continued throughout the [evaluation] process and still continues and that is where we reviewed the different aspects of the evaluation during that time. I also did attend the Diploma Program teacher meetings where evaluation was a big part of those as well. So, I am in attendance during those meetings. There were meetings that I was not involved in, but those were more team meetings, where different departments were leading their teams through different aspects of the evaluation, but we reviewed that, though, together. So any part of [the evaluation] that came out of [the other meetings] was reviewed either during our weekly meeting or during a subsequent meeting (Lewis interview, February 14, 2015).

Lewis points out the continued meetings regarding the evaluation of the program even after the mandated accreditation was finished. The meeting schedule that was established for the evaluation process has become a significant part of the structures for the program and the High School at the school.
Lewis also relates that other stakeholders were an important factor in the evaluation activities. He specifically emphasizes the contributions of the DP teachers and the parents of DP students. Lewis believes that one of the best parts of the evaluation was the communication and keeping everyone informed and involved in the process. He recounts that DP teachers could volunteer to be on the evaluation team and some members of staff were strategically asked to lead certain parts of the evaluation based on their expertise or role in the school so that the right people could fill those roles. Lewis believes that the evaluation was an inclusive process representing the teachers delivering the Diploma Program equitably.

**4.4.1.1.3 Transparency and Accessibility**

Lewis attributes the success of communicating with DP teachers regarding the evaluation to the accessibility of all documents generated for the evaluation on Atlas, the school’s curriculum planning platform, and the transparency of all meetings and subsequent decisions being recorded in meeting minutes and action plans which were also shared out to all teachers. He asserts:

[The evaluation documents] should be accessible by everyone on staff through Atlas because we can search and find and go on to anything that is put on Atlas. So it should be accessible to staff if they want to go in and see it. As everything that we do now is based on our accreditation, I think that is extremely important, just to be able to go back and take a look at the history because those pieces are sometimes lost. So even though the process is great, if we do not capture it we can't remember what made it great. (Lewis interview, February 14, 2015)

Lewis further explains that the process of the evaluation was transparent since the activities associated with the accreditation were embedded into various aspects of the time teachers spent together during their collaborative times and the meetings that were held. Since so many people were involved in the evaluation, it just became an organic part of all discussion about the Diploma Program. As he put it, “To me, it was just the way we went about doing things” (Lewis interview, February 14, 2015).
Lewis believes that this is the only way that the changes implemented by the evaluation can be sustained: schools document the discussions that staff have about the program, record the changes needed and how they are actualized, capture the reasoning behind the decisions, and, finally, they share these documents with the entire staff to serve as institutional memory. This way, even in a context with a highly transient staff, systematic planning can be realized.

**4.4.1.1.4 Accountability and Self-Assessment**

Lewis argues that, in order to hold all stakeholders in the evaluation process accountable, everything needs to be highly organized and accessible. With so many teachers involved, it is essential that the evaluation documents were visible on Atlas so people could go in and take a look at what happened at what time. This also enables leaders to go back and review what decisions were made and why. Lewis claims that this type of systematic planning will also be useful the next time an evaluation happens at the school. He asserts, “The process we went through for the Diploma Program [evaluation] should be a model of what we use [next time]. I mean, the way it is documented and the way it is planned and how we went about doing it I think should be used as a model” (Lewis interview, February 14, 2015).

This type of transparency also ensures that the evaluation is collaborative and not prepared in isolation by leadership and encourages accountability. Lewis maintains that an evaluation run according to the model used at the school can help to avoid window dressing. He states that “for the most part, self-evaluation is actually the way to go” (Lewis interview, February 14, 2015). Lewis believes that staff members are pretty harsh when they are giving their thoughts on their own program because they are genuinely trying to make their practices better. Lewis asserts that the information and feedback received during the preparation of the self-study was excellent and it was used to improve the school’s program.
4.4.1.2 Pedagogical Leadership of the Evaluation

Lewis claims that the documents submitted to the IB as part of the mandated accreditation were an honest depiction of the state of affairs at the school. Lewis perceives the fact that the pedagogical leadership for the Diploma Program was new as a positive aspect of the experience:

We did not have any preconceived notions on how things were supposed to work. … We did not think it should work this way or should not work this way, so that we had no bias, I do not believe we did. Our goal was to make the program better and so we took all that information and moved forward from there. I think it is a great entry plan as far as moving into a new position to go through an evaluation because you've reviewed the program with the entire staff, it is collaborative, and then you end up developing an action plan and a starting point to move forward. So, I actually think it was extremely helpful. (Lewis interview, February 14, 2015)

Lewis states that “I am a collaborative leader, so it is not about me, it is about the team” (Lewis interview, February 14, 2015). As High School Principal, he is responsible for more than just the Diploma Program, which only includes students and courses in Grades 11 and 12. As a result, he acknowledges that the Diploma Program Coordinator is largely in charge of the pedagogical leadership of the program. However, there is significant amount of distributed leadership on staff. Lewis believes that the job of pedagogical leaders is to move staff forward.

He states that the fact that eight staff members became IB subject examiners during the process of the evaluation speaks to the environment of teaching and learning at the school and proves that newness of the leadership did not have a negative effect on the program. Instead, Lewis claims that it brought a fresh look to the DP at the school.

4.4.1.2.1 Experience Level with DP and Evaluation

While teachers could choose their level of involvement in the evaluation process since they were all given the choice to participate, Lewis recognizes that some newer teachers would have chosen not to get as involved as they were making the transition to a new school, a new program, and a new culture. However, since attending DP teacher meetings was
mandatory, these staff members would still be receiving information about accreditation and learn more about the program. Many staff members also did not have any previous experience with accreditation. These factors might have contributed to self-assessment ratings being lower on some standards and practices than the expert rating the school received from the IB external evaluators. Lewis observes that, within the school, the team always wants to improve. He states, “I think we have established a culture of continuous learning here, which means we are never there. So if we are doing a survey or something and we are checking that we are in the highest band that means we are done. Of course we are never done” (Lewis interview, February 14, 2015).

Moreover, Lewis points out that external evaluators are exposed to many different schools’ evaluations and different ways in which the program is implemented while teachers in the context of just a single school, especially if they are new to the program, are fairly isolated in their practices and experiences. Lewis wonders if IB external evaluators use the criterion reference model to assess schools. As a next step to improve his own understanding of evaluations, Lewis wishes to see other schools’ evaluations because those would be great to share.

4.4.1.2.2 Stakeholders in the Evaluation

In addition to the heavy involvement of Diploma Program teachers, parents of DP students and the students themselves are part of the stakeholder group that is included in the evaluation of the program. Lewis states that, as the father of a student in the Diploma Program, he filled out the parent surveys and expressed his opinions about the program using that lens. Lewis reports that great attempts were made to get parents of the program to participate in the evaluation, but most of the parents who answered this call were also teachers at the school, some of them not necessarily teachers in the Diploma Program.
Lewis acknowledges that the school received some good information from teacher parents but he would have liked to see more parents from the community provide more feedback as well. Lewis opines that “everyone is busy so I think [not participating in the evaluation of the program] is a thing of trust. A lot of parents trust us that what we are doing is good for their children so they do not feel they need to be part of [the evaluation]” (Lewis interview, February 14, 2015). However, Lewis is grateful for the excellent feedback provided by the students on an on-going basis. Since they are the end users of all of the teachers’ efforts, Lewis states that it is crucial to get the students’ opinions on any of the changes made to the program. Students participated in surveys and focus group discussions about different elements of the accreditation and then after the changes were actually implemented.

4.4.2 Effect of the Evaluation

This section explores Lewis’s responses regarding the effects of the evaluation on the Diploma Program. Lewis views the effects of the evaluation very positively. He expresses that just because the formal self-study that was required to be submitted to the IB is completed does not mean that evaluation is finished. Lewis reports that teachers and program leaders constantly evaluate the program. As the High School Principal, Lewis focuses on the entire program and, at times, the entire High School when he is making his observations about the effects of the evaluation.

4.4.2.1 The Progress

Lewis finds that, as a result of this continuous evaluation set in motion by the evaluations mandated by the MSA and the IB, the staff are constantly looking at how they can get better. He states that the program is continually monitored as the teachers receive feedback from students and parents, as the structures set in place are annually reviewed, as more and better-gathered data is used to guide decisions, and as the teachers and leaders
reflect on their growing understanding of the program. Lewis attributes the development of this relentlessly evaluating learning climate to the evaluation process of the Diploma Program.

4.4.2.1.1 Process as Progress

Lewis says, “[The evaluation] was not about getting it done, it was about learning as we were moving through it” (Lewis interview, March 19, 2015). Lewis argues that the simple act of going through the evaluation activities resulted in needed changes in the program. He elaborates:

We were improving as we were going through the evaluation, so I just think it has been two years of steady action from the part of our Diploma Program. I would say our Diploma Program is probably better than what some of the information [of the self-study] shows from the beginning as we started because of what we found when we were doing the self-study. We continue to make changes to our program for the betterment of the staff and students. I think the evaluation had a big part in that, though. (Lewis interview, February 14, 2015)

Regarding the timing of the implementation of these changes, Lewis states:

Because as we were going through [the evaluation] and we were looking at things and uncovering things that we wanted to change, we weren't going to wait until the end of the evaluation to make changes, so we started to implement changes before that. So when we were getting information and we were looking at ways to implement change, we were doing it as we were moving through, so I think that is the most important thing. I say we are a lot better now than we were before we started this process. (Lewis interview, February 14, 2015)

Furthermore, Lewis opines that the evaluation gave staff the impetus to find solutions for problems that were years in the making. He recalls the discussions about the scheduling of DP courses:

We did a lot of think tank type things. I mean we really worked hard to try to see how we can get the proper hours for HL courses and reduce the hours that we were having in the SL courses. It has been years of thinking about it but we finally just sat down to come up with a solution. It is a very complex timetable we have now. But only complex looking from the scheduling aspect, when looked at from the student aspect, it is very easy. It is not complex at all. (Lewis interview, March 19, 2015)

Lewis credits the evaluation process with facilitating the implementation of desired changes.

At the time of the interviews, Lewis envisions a school that is going to be under
constant evaluation. Bringing in ideas from the large network of schools, the staff will be in a constant state of refinement and taking a look at their practices as a matter of course not because of an evaluation. Lewis reflects, “We have a lot of what needs to be in place in place now; I think we’ll be able to move forward a lot easier” (Lewis interview, March 19, 2015).

4.4.2.1.2 The Professional Development of Staff

Lewis suggests that the evaluation of Diploma Program practices prompted staff to examine not just the practices of the program but their own practices as educators. He argues that this type of reflection is on-going within the whole team who delivers the program. The conversations that elicit changes are still happening. He elaborates:

It was good for people to examine their practices [for the evaluation]. I think we changed our assessment practices significantly based on looking at our program and I think that is continued to have an impact on our everyday delivery of the program. So, yes, we did this evaluation, but I do not feel we have stopped. ... We are continuing to use the same model we used in evaluation, to continue to refine our practices and reflect on what we are doing. It has put us in a mode of being reflective practitioners and we have continued that. It has created a culture in the school that we have sustained after our evaluation was completed. It has changed the way people look at things. (Lewis interview, February 14, 2015)

In addition, the program leadership realized that IB DP subject training workshops were not enough to strengthen the overall program. The DP Core, Extended Essay and CAS, have not been considered for training purposes. Lewis credits the Diploma Program Coordinator with driving instructional leadership within the core and developing leadership in other individuals to support the delivery of the Extended Essay and CAS. One teacher was chosen to attend IB training for supervising the Extended Essay and another participated in CAS training as a direct result of the evaluation.

Moreover, the changes happening in the Diploma Program in Grades 11 and 12 triggered the examination of processes in the rest of the High School. Lewis says:

I think we do a lot of professional development within our Diploma Program. What I do think though is that the focus of our whole [high] school professional development on [rigorous] assessment came out of the Diploma Program. So by examining what we needed within the Diploma Program, [we realized] we needed to build more than
just two years of rigor within the school. So we have been looking at a resource and work through as a staff about developing rigorous instruction and assessment as a whole staff. So I think we are starting to build structures in Grades 9 and 10 and our focus is on different aspects that are guided by our results in the Diploma Program. … So I think whether you are a Diploma Program teacher or not, you are working towards developing instructional and assessment practices that support the end goal. And that is to have the students be successful within the Diploma Program. (Lewis interview, March 19, 2015)

Thus, Lewis believes that, in order to have a better foundation for the DP, the whole school’s professional development plan has been informed by the DP evaluation.

4.4.2.1.3 Structures and Procedures to Increase Student Achievement

Lewis lists many positive structural and procedural changes that he attributes to the completion of the DP accreditation process. Lewis states:

By going through all the documents [required by the IB for the evaluation] and developing them, we built a better understanding of the program that we are doing at the school. … We gained a lot of knowledge as a team through the development of those documents. (Lewis interview, March 19, 2015)

Lewis acknowledges that, because the IB requires the submission of certain supporting documents such policies and handbooks, the school had to update or create these documents. However, he finds that, while teachers were engaged in the creation of these documents, they examined their practices and made changes. He recalls:

It is one thing changing a document and another thing changing a practice. I think it kind of worked both ways. We changed our practice, which lead us to examine the documents as well but we also, through developing the documents, found out we need to change some of our practices. (Lewis interview, March 19, 2015)

Lewis believes that the biggest change in the program as a result of the evaluation is a shift to move towards transparency about program practices by making all of these documents accessible to all staff, parents, and students. Through the communication of all of the policies and procedures of the program, Lewis suggests that there is a shared understanding of the program by its stakeholders. The structural and procedural changes and their perceived effects that Lewis describes can be summarized in table 2.
Lewis states that a lot of the growth that the school has seen in its program comes from the structural changes that were implemented as a result of the evaluation. Lewis maintains that the changes happened because the staff wanted to better support the students and use resources like time in the program better. These structural changes also then change the thought process of the staff. The school introduced many standardized processes across the program which means more teacher accountability. For example, Lewis states that, by

<table>
<thead>
<tr>
<th>Change to the Diploma Program</th>
<th>Effects Perceived by Lewis</th>
</tr>
</thead>
</table>
| 1. Communicating program specific information to teachers, students, and parents:  
- Diploma Program Handbook for Students  
- Diploma Program Handbook for Teachers  
- CAS Handbook  
- Extended Essay Student Handbook  
- Extended Essay Supervisor Handbook  
- Atlas resource page for DP teachers  
- Parent and student information section on school blog  
- Parent information evenings | 1. Sharing understanding of the program by all stakeholders  
2. Increasing transparency and accessibility of all program procedures  
3. Establishing organizational capacity to capture institutional memory and yearly review and update of all documents |
| 2. Scheduling the Diploma Program:  
- Providing appropriate timings for HL and SL courses including A block and H block HL classes and SL tutorials  
- Building collaborative department meeting times into teacher schedules | 1. Meeting IB requirements for the teaching of HL and SL courses  
2. Providing appropriate supports for students to improve academic achievement while reducing stress  
3. Increasing collaboration time for teachers to grow their capacity to deliver the program |
| 3. Updating all assessment policies for the Diploma Program:  
- Aligning of grades with predicted grades and IB grades  
- Moderating IB world exams scripts  
- Updating all achievement charts  
- Regulating homework, retake, and formative assessment procedures  
- Establishing use of GradeBook for recording marks | 1. Using data-guided decision making process to reach conclusions about program practices  
2. Standardizing all Diploma Program practices  
3. Increasing teacher confidence in grading and reporting practices |

Table 2: Lewis’s perceptions of structural and procedural changes as a result of the evaluation.
clarifying expectations and making all processes transparent, teachers report that they have
more confidence in their grading and, consequently, in their communication with students
and their parents. Teachers know how their marks are generated whereas before there was
ambivalence in assessment practices. Lewis particularly emphasizes importance of the
increased alignment between the school’s internal percentage marks on report cards, the
predicted grades that are sent to universities for admission consideration, and the final IB
marks that are based on mainly externally assessed components. Lewis states:

We have made giant steps with regards to alignment [of grading] practices and I think
this is linked to [more] collaborative time [in the schedule] and the [mandated]
m Moderation [of assessments]. And we have really brought our [school] grades and
practices in alignment with the IB. So, if our students are getting 85%, we know it is
going to be 5 [on the IB grade scale]. It is very much aligned, we feel very confident.
We have made major changes in our practices. Just on our report [cards] that went out
last week, the predicted grades are there and we could not do that before. We could do
that now because we are confident. We are confident in our practices, we are
confident in our alignment and that comes from examination of the [assessment] data.
A key part of the evaluation is taking a look at our [assessment] data. What did our
report card marks say? What did our predicted grades say and what did our world
[examination] scores say and whether they are in alignment or not. And we have
made major gains in there to the point where we can put that out there [to the parents
and the students]. … It has been a lot of education with the students as well as the
teachers [about the new assessment practices]. But we have fewer students asking
about things like predicted grades because they know that whatever’s on the report
card is really what we think they are going to get on the world exams. I just think by
being transparent, by talking about it, by being open with it, we really haven’t had
problems like we had in the past where students were wondering what their predicted
grade was. Well, they do not wonder anymore, they know whatever they are getting in
class is what they are going to get on their predicted grade [that we send to
universities and to the IB]. Students know where they are and we are keeping them
abreast of the situation with regards to their achievement level (Lewis interview,
March 19, 2015)

Lewis further recalls that one of the key changes for the program was the scheduling
of Diploma Program courses. These changes in the scheduling of Standard Level (SL) and
Higher Level (HL) courses enable students and teachers to maximize instructional time as
appropriate for the level of the courses. Lewis explains that the High School operates on an
eight-day cycle of eight rotating instructional blocks, noted as A to H Blocks, with each
assigned to a course. There are six 60-minute blocks per day. As per IB requirements,
Diploma Program students take six DP courses: three courses at Standard Level (150 instructional hours) and three courses at Higher Level (240 instructional hours). In addition, Theory of Knowledge (TOK) is scheduled as an SL course, and one Supervised Study period is offered during H Block. H Block serves as a time to complete the Extended Essay (EE) requirement, to plan and document Creativity, Activity and Service (CAS) activities, to meet with EE Supervisors and CAS advisors, and to schedule additional HL instructional hours.

There had been several attempts to change the scheduling for SL and HL courses prior to the evaluation. A scheduling committee comprised of the High School Principal, the Middle School Principal, the previous Diploma Program Coordinator, and the High School Guidance Counsellors was formed to address the recommendation from the previous IB evaluation regarding the scheduling of HL and SL courses. The main issue identified in the evaluation report was that the regular High School schedule allows for 210 hours of instructions for both SL and HL courses over the course of the two DP years. HL courses are also scheduled extra instructional hours during H block to meet the 240 hour requirement. However, this schedule can possibly result in work overload in SL courses and not enough time to complete HL requirements possibly causing students unwarranted stress. The scheduling committee proposed several versions of schedules for changing the length of the teaching blocks and moving from the eight-day cycle to a two or ten-day cycle. However, upon further consideration, none of the proposed changes could work with the larger structure of the school. The main stumbling blocks were the fact that most DP teachers also teach in the IB Middle Years Program (MYP) where eight subjects are required therefore eight periods work well, the fact that the High School shares school facilities like the canteen and sport areas with the Middle and Elementary schools whose recess time needed to be taken into consideration, and the fact that nutritional breaks need to respect prayer times to accommodate Muslim students. As a result, the only change regarding the bell schedule that
was implemented for the 2013-2014 school year was the separation of a longer lunch period into two shorter breaks: a nutritional break after the first two periods of the day, followed by a lunch break after the fourth period. This way, students could have a balanced school day. As part of the DP evaluation, the evaluation team proposed several additional changes within the existing bell schedule to avoid work overload in SL classes, reduce TOK teaching hours, and provide more instructional hours for all HL classes. These changes were gradually introduced with all of them being fully implemented by the start of the 2014-2015 academic year.

Lewis explains that although the schedule allows for 210 instructional hours for SL courses, according to the new scheduling guidelines, teachers design the SL courses to be completed in 150 hours over the two years. This arrangement gives students one period of tutorial support per eight-day cycle for each standard level course. This means that all homework and class work should be done in class: out-of-class work for SL students should be minimal. Most homework time for students should be taken up with the HL classes. At the same time, teachers design the HL courses to be completed in 240 hours over the two years. Therefore, students are scheduled one additional period of instruction per eight-day cycle for each higher level course during H Block. Most homework time for students should be taken up with the HL classes.

To reduce the number of instructional hours for TOK, this course is now scheduled for the second semester of the DP 1 year and the first semester of the DP 2 year to reduce workload for students at the beginning and the end of the program. All TOK classes are scheduled for A Block. During semester one of the DP 1 year and semester two of the DP 2 year, A Blocks will be used as an additional Supervised Study blocks. This arrangement allows DP 1 students to adjust to the demands of the program focusing on the six subjects prior to starting TOK. Similarly, at the end of the program, students have more time to
prepare for the world exams having completed TOK requirements in the first semester. As an added benefit, additional HL classes can be scheduled during A block to ensure that the HL courses are meeting the 240-hour requirement. Moreover, since A and H blocks follow each other on Days 2 and 6, teachers have the opportunity to schedule longer labs or assessments during the school day for the first time instead of requiring students to come to school on the weekend. Lewis states:

It has been years of thinking about the schedule but we finally sat down to come up with a solution. It is a very complex timetable we have now, but only complex looking from the scheduling aspect. When looked at from the students' aspect, it is very easy. It is not complex at all. We provide time through supervised study to get extra HL hours and we reduced the amount of content they were given in SL courses. We can build in time during the day for students to get some of their work done so that they can have a more balanced life at home. These changes made the program more accessible for more students. (Lewis interview, March 19, 2019)

Lewis demonstrates how changes in the teachers' schedules allowed the groups to develop better teaching and assessment practices which in turn provide more supports for the students.

Lewis summarizes his perception of the impact that these changes have had on the Diploma Program at the school:

We scheduled the entire department to have one block off. For example, the English department does not teach during C block, so they are all off working in groups. Every department has a different block now. I think that has been key. However, having time [to collaborate] without a focus wouldn't help. But having collaborative time and having moderation built in that time and the expectation that we have to moderate, to use later, has been instrumental. This guided us to ordering back student work from the IB, I think that was the initial step to say, you know what? We should get our [student work samples] back from the IB and take a look and moderate those to see where we are. I think that was our digging deeper to examine our practices.

We further comments that teacher timetables were also changed as a result of the examination of DP practices. Scheduled collaborative time was built into the High School timetable for each subject department. Lewis explains:

It is great to have a network within the school now. We have gone outside and we have got the examining if they weren't. So, we know that that is a true mark. (Lewis interview, March 19, 2015)
We have a lot of procedures we put in place, we have a lot of policies we put in place. We have a lot of documents we put in place but really the measure of whether those are really doing their job is our [school and IB] results. It was great to see the significant improvement in our alignment of our results. Our [world exam] results did improve. The school average increased from 30 to 32 points. But I think it was the increase in the alignment of the school and IB results that were the most exciting for me. … Moreover, part of our plan was to move as many kids into DP courses as possible. I think with the success we are having and with the structures we have put in place that is one area where we have shown significant growth. Our numbers are increasing within the program. We are providing more time and more supports for students in a more nurturing environment. (Lewis interview, March 19, 2015)

Overall, Lewis concludes, that the evaluation of the Diploma Program continues to improve teaching and learning.

4.4.2.2 The Pushback

Lewis claims that there were no negative aspects of the evaluation. He states, “We either recognize the great things we are doing and celebrate, or recognize some things that we need to improve and we work towards improving them, so either way we win. So I do not think there is any negative aspect at all” (Lewis interview, February 14, 2015). However, Lewis acknowledges that different attitudes towards the evaluation can stem from change. He states:

We made a major overhaul of the program. …I have a child within the program and I have to say that everything we have done has supported and nurtured the students within the Diploma Program. So with regards to comments that [the changes are] not helping students I would absolutely say 100% is incorrect. And I can speak to it as a parent and I think that my child in the program has received a tremendous amount of support. … Support looks different [to different people]. Having deadlines is support, [the program being] planned over two years is support, getting extra time within the day to work on your school work is support. Really, being hard on kids sometimes to make sure they get things in is support because we know that things build up and they'll get overwhelmed and that is support. Changing practices of teachers is support. … Our teachers are better teachers now than they were before. (Lewis interview, March 19, 2015)

In response to the negative experiences some survey participants expressed, Lewis says:

The [teachers] who understood that we need to change are better, the ones that did not probably felt that they knew what they were supposed to be doing and took offence to the fact that we asked them to change what they were doing. And change is hard so I understand why a staff member might make negative comments. Nonetheless, it is very disappointing on two aspects because I do not believe that to be true, the
statement. On the second I do not want staff to feel that way. I think [the evaluation] was open to all staff to be a part of the process. And really as teachers we decide what we get out of something. If you decide that you are going to learn from this process you will and in the end those comments that are positive people have embraced the change, wanting to get better. Some people do not necessarily feel that way. It is a hard thing to examine your own practice and say, you know, maybe I am not doing things the way I should be doing. We have had success, our results are better. We are more aligned, our parents know more, our students know more, our teachers are more aware. I just can't say the things that we have done have been negative or are hurting anyone in the least. (Lewis interview, March 19, 2015)

Lewis also recognizes that it is natural that some pushback will happen as a result of a major overhaul of a program. Yet he insists that, if the changes are in the best interest of the students and the majority of the staff, and as a pedagogical leader, it is his responsibility to ensure that the program gets better. He shares:

We have high expectations here, we do a lot. There is a lot of work, there is no doubt about it. … There are expectations of how [teachers] write a report card for example, there are expectations of what's included on the exam, and those are all expectations that we hold staff to. We were doing different things before [the evaluation]. I do not think we were serving our students though. It is hard to go through that [change]. I think now that we are there and now that staff knows the expectation and it is there I do not think staff would feel [a negative] way. But kids need consistency so by building in that accountability of staff and saying this is how we do things the kids are getting that consistency and in the end that is the most important thing. So I would say that transparency and consistency are what have come out of the [evaluation]. And high expectations, I mean we want to consider ourselves a school that does extremely well and we need to have high expectations to get there. (Lewis interview, March 19, 2015)

Finally, Lewis reiterates that the fact that both the High School Principal and the Diploma Program Coordinator were new to the program during the completion of the evaluation had nothing to do with pushback from teachers who have been long time participants in the program at the school. He concludes:

Good instructional practices are good instructional practices, good assessment is good assessment. … To me it is about leading change. We have great staff here so we lead those staff to make change, it was not just [the leadership] doing it. (Lewis interview, March 19, 2015)

Lewis believes that the collaborative leadership model that he follows distributes ownership of the changes to the program. The majority of the staff agrees with this assessment.
4.4.2.2.1 Lack of Time

The completion of the evaluation activities is an extra requirement in addition to teachers’ regular teaching duties. Lewis admits that some teachers might not get involved more with the evaluation simply because they did not feel they had enough time to give to the process. As Lewis reflects on what the leadership could have done to make teachers feel more part of the evaluation process, he says:

The only thing I could think of would be we could have released teachers [from teaching to work on the evaluation]. I do not know if we actually could have, but we could have explored releasing teachers for chunks of time within a day to see if it would have maybe got other teachers involved that maybe did not necessarily feel they had the time. That is a catch-22 though, so we are taking them out of classes and then they are not there to teach, so I am not sure, but, I mean, that is the only thing I could think of. I think we kept everyone fairly involved in the process. (Lewis interview, February 14, 2015)

4.4.2.2 Personal Feelings

Lewis believes that the majority of teachers had a great experience during the evaluation. He states that the changes were not done to teachers and students, rather the changes were developed with them so that the team could have agency in how the process worked. He points to evidence that teacher confidence and efficacy are on the rise at the school, and world exam results and student well-being increased. He reflects:

The big evaluation for us is at the end of the year when it is all done and students have finished and teachers have finished within that year. And we have graduated diploma students and they come back. I think that is when we get the best feedback. When they are outside of the school for a year and they say, you know what? The best thing for me is when you said I had to have this by this time. And you were right, I was a five [for my IB grade]. So, I think really getting information within the heat of change sometimes is difficult. Sometimes it does not give staff enough time to reflect. I think different sampling over the course of the year might get a better idea. … So even though we have great structure they will continue to change and grow. Because we are at the point where we can look at the small pieces [of our program] and refine them and add to them. I mean we already have plans for next year on how we are going to grow and how we are going to change things. (Lewis interview, March 19, 2015)
Ultimately, Lewis advocates for continuous reflection on practices which results in constant change. His challenge is to lead his team through these processes with getting maximum results for all stakeholders.

4.4.3 Future of Evaluation

For Lewis, the future of program evaluation focuses on teaching and learning: How can the programs be better for students and how can the evaluation process help with this improvement? Lewis claims, “How [the evaluation looks] is not the most important thing because that is not what this is about. [Evaluation] is about improving teaching and learning for student achievement” (Lewis interview, April 22, 215). In Lewis’s opinion, anything that makes the process more able to help with this fundamental goal is an improvement of the evaluation system.

4.4.3.1 Future of Evaluation at the School

At the school level, Lewis believes that the MSA and DP evaluations should serve as a model for an increase in organizational capacity in the form of continuous performance monitoring of the school’s programs. He argues that this constant reflection will make the fulfillment of formal accreditation requirements much easier in the context of the highly transient international staff of the school.

4.4.3.1.1 Continuous Change Model

Lewis states, “The best part of the evaluation is not the evaluation itself but what it did for us and what we are going to do now up until the next evaluation” (Lewis interview, February 14, 2015). The action plan was a new requirement for the DP evaluation self-study for the school. However, Lewis argues that putting the action plan into practice is for the improvement of the program not a fulfillment of a compliance requirement. At the same time, the whole school had to produce a school improvement plan for the MSA reaccreditation process. Lewis comments that all of these plans need to coalesce. He says,
All those pieces need to work together. So, the PYP plan, the MYP plan, the diploma plan and the school improvement plan should work together and everything we do should be focused on our graduating diploma students and building from their needs from when they graduate. (Lewis interview, March 19, 2015)

Lewis believes that synchronization of all program action plans and making them available to all stakeholders has been a significant improvement in the school’s organizational capacity and transparent planning ethos.

Lewis further expresses his ideas for maintaining the reflective state regarding teaching and learning practices introduced by the evaluation of the Diploma Program. He elaborates:

Although we actually have a plan developed coming out of the evaluation, and the evaluation being successful, allowing us to identify reflective pieces that we want to work on, I think what I would like to do at the school is to take pieces of the program evaluation and embed them into our regular practices. So we are constantly looking at the different pieces we do for evaluation, doing them at a smaller level as we progress through each year. So we never really get out of [evaluation]; it becomes less of a one-off type thing. We are constantly in a reflective pattern where we are reviewing the specific documents and looking at them from the eyes of an evaluator to see what we can do to improve and what direction we need to go in. I think that is always the way we need to look at it. So I would like to be in a constant state of reflection and not necessarily always just thinking about that one year where we are going through a formal evaluation. So I would like to always be prepared to go through a formal evaluation so it is just about pulling the pieces together to show somebody else how we are meeting our targets and improving the program. (Lewis interview, April 22, 2015)

In order to achieve this continuous model of evaluation, Lewis advocates for the ongoing collection of data to allow year-on-year comparison of results in order to guide decisions and build institutional memory. To determine what data to collect and how, Lewis looks to school-specific measuring tools to complete yearly review of the program which can also serve as preparation for the external review on a five-year basis. Lewis argues that this is the best approach due to the high turnover of leaders and teachers. He summarizes:

This way, when we get to our formal evaluation again, it is there. The documents are there and it is better because we tweak them each year and we have gone through the review process and we never left it. So, it is not going to be this gigantic monster at the end. It is just like if you constantly work on a report each month and then by the
time the report is due, you are just putting the pieces together. But if you wait until the end, yeah it can be pretty daunting. (Lewis interview, April 22, 2015)

Indeed, none of the interview participants remained at the school until the next evaluation of the program but the systems put into place served the new team well.

4.4.3.1.2 Synchronization of Evaluations

At the end of the MSA and DP evaluations, the school planned to take advantage of the newly available continuum evaluation for all three IB programs at the school. This means that, during the 2016-2017 academic year, the PYP, MYP, and the DP programs all participated in the preparation of a continuum self-study and school visit. Moreover, this coincided with the mid-term review of the MSA reaccreditation process. Lewis believes that the synchronization of the formal evaluations required by the school’s accrediting bodies will streamline processes for the staff completing the evaluations. He mentions that the already prepared Kindergarten to Grade 12 policy documents that were compiled for the DP evaluation set the right tone for a continuum evaluation. He encourages the creation of standing committees to continue the work of the evaluation team and suggests that these groups perform the yearly reviews of the school’s practices. He states:

I think we had excellent teacher teams working on the evaluation and going through that process. They are disbanded. I think our goal next would be continuing to develop teacher teams as we move through so not disbanding them. We did not formally say, okay you are done, walk away now. It is just as part of the process they just ceased to work on something at that point. We should leave those teams in place, adding new members, having focus groups so we have an extended team that would meet and discuss at certain times. … Because if we create those teams that are constantly working on that, then evaluation is nothing, I mean we are going to have those teams, they are already working on different pieces and then they are contributing to the evaluation because it is just something we do. (Lewis interview, April 22, 2015)

Lewis’s words illustrate his belief that the continued work of standing committees would enable the school to meet the challenges of external evaluations easily while improving the structures supporting instruction at the school.
4.4.3.2 Future of Evaluation for the IB

This section will explore Lewis’s suggestions for the IB on how to improve the evaluation requirements schools have to complete in order to increase the likelihood that the process is used for improvement and not just compliance. In addition, Lewis also provides feedback on the changes that the IB plans for the evaluation of its programs.

4.4.3.2.1 Suggestions for the IB

Lewis suggests that the IB changes the format of the external evaluation mandated for authorized schools to match the school’s needs for a continuous evaluation model he envisions. As opposed to approaching the accreditation as a one-time event occurring every five years, he argues for yearly information to be submitted to the IB that track data to prove that the school is functioning within set parameters. At the five-year point, he advocates for schools to be able to choose which practices in their specific context they want to explore in more depth after the IB confirms that they meet minimum compliance requirements. He explains:

This way we can focus on the program from a different lens I guess and not take our time filling out paperwork. The most important thing to me is examining our practices and how it impacts kids and instruction and achievement. If we [submitted statistical data] every year and just updated with number of students and that and that was every year and then the IB always has that information and it is at their fingertips all the time. So instead of waiting five years, there is central information that we have that can be housed separately and updated each year rather than waiting. I mean that is my thing because then the focus is what I think the focus should be on, not on collecting trivial information. It is important, do not get me wrong because they need to know our context. But I do not want most of the time being directed towards that. I want to get down into student achievement, delivery of the program and instruction. (Lewis interview, April 22, 2015)

Lewis also relates that not having a school visit detracted from the possible benefits of the evaluation. He says:

I think that not having a person come to see our program that we are really proud of and the work we have done, the improvements we have made [was a missed opportunity]. We did get feedback although it was not in person that we did a good job and we had very minor things that we needed to take a look at as far as next steps.
But I really think actually having a person onsite and feel and be here and be able to speak to our teachers [is important]. I think the feeling that the teachers had about filling in the paperwork and all the work would have been different because they would have felt more validated by actually having somebody in the building talking to them face to face. (Lewis interview, April 22, 2015)

He further comments that the feedback the school received from the IB lacked depth in relation to the amount of work the preparation of the self-study represented to the staff collectively. Lewis suggests that schools with evaluation visits should receive a different feedback report than schools without visits. He also recommends that schools be given the opportunity to present their self-studies using a variety of modalities. Lewis thinks that teachers preparing presentations that are videotaped and submitted to the IB would be more meaningful. Since it is the teachers who deliver the program, they could demonstrate how they are living the DP at the school as opposed to having a static self-study completed by the DP Coordinator on the teachers’ behalf.

Lewis applauds that school visits have become available for DP schools and are part of the continuum evaluation requirements. He would like to see the professional opportunities for teachers expand in this regard. He says, “I’ve been on a visiting team for accreditations through the Council of Independent Schools. And I think that the professional development you get by being part of that is probably better than any professional development I have had” (Lewis interview, April 22, 2015). Lewis argues that the pool of school visiting team members (SVTM) should be opened up to allow more leaders and teachers to participate. Moreover, there should be an option for teachers to participate as observers in the evaluation in order to learn more about the process but also about the program. Compared to other accreditation bodies, Lewis finds the ranks of SVTMs at the IB too closed off. He says:

You need strength [and experience] on your [visiting] team but I think it is an opportunity to have new people each time and make sure that they have that experience. And I just think it would be a phenomenal learning experience for people. Like could not the IB just open it up so that people could apply and they could be chosen more easily? The teachers wouldn’t require the training, they wouldn’t have the same kind of responsibility but they would still be part of the experience? … The
IB could provide different types of training. I mean it does not have to be on the same level. Like when I went through at different visits [with CIS and MSA], I did some online training so I had a general sense of what it would look like. I mean our instructional knowledge, that is there. I mean we have those pieces. I think there are a lot of teachers out there that definitely could be part of this. And I think if we put together small training online and teachers are able to read the information the schools develop before the visit, teachers would get so much out of that. (Lewis interview, April 22, 2015)

The current training available for DP coordinators only focuses on ensuring that leaders know how to meet the compliance pieces of the mandated evaluation. However, Lewis approaches the school visits not as inspections but as an invaluable opportunity for school improvement.

Finally, Lewis would urge the IB to see the evaluation as a chance to celebrate successes both at each individual school’s level and globally. He states, “We got [the teachers] moving and the more success they get, the bigger the snowball gets as we move through evaluations” (Lewis, April 22, 2015). He shares the experience of the science team at the school who decided to completely change the G4 project as a result of the evaluation. The teachers were so happy with the alteration of the program that they submitted an article for consideration to the IB World magazine. Their submission was rejected because the magazine had already published something similar. After this experience, no other team attempted to share their triumphs with the larger IB community. Lewis wonders if there is a better platform for celebration of achievements on a larger scale to encourage continuous improvement in the IB network of schools.

4.4.3.2.2 Feedback on IB Proposed Changes

The IB recognizes that, while the number of IB schools has grown exponentially in the last two decades, the evaluation processes for its programs have not been updated to meet the needs of a changing international organization. The IB has to better balance the need to protect its brand with providing personalized services to its members. At the October 2014 Regional Conference in Rome, key members of the school services department outlined the
proposed direction for the future of program evaluation at the IB (Albright, Atkinson & Bender, 2014). During the final interview, I asked the three participants what their feedback is to the proposed changes to the mandated accreditation process as presented at the conference.

Lewis believes that he would be in favour of synchronizing evaluation self-studies and school visits between different accrediting bodies, like the IB arrangement that exists with the Council of International Schools (CIS), as long as the documents and expectations are perfectly aligned. He states:

I would have to look at it but it sounds like a structure that would be useful. … My only fear in that is we are going to spread ourselves very thin because someone has to be on [all the different groups for the evaluation like] facilities. And I would hate to pull someone out of a program that they want to focus on with regards to instruction and learning. That is my only fear. Does it make sense? Absolutely. Would it be easier? I think it would be. (Lewis interview, April 22, 2015)

Lewis is also supportive of any technology that would make the preparation of the self-study easier and less annoying. He says:

We want the time being put into quality discussion about the program. And I think if we are given a platform to work on that is already structured and formatted is great. I look back at the amount of time we wasted with paper or even the accreditation process trying to format it. That is time that we do not need to waste. (Lewis interview, April 22, 2015)

Lewis is particularly excited about the potential of virtual evaluation visits piloted by the IB. Heralded as an alternative to on-site school visits, virtual visits conducted via Skype can be especially useful to reduce the cost of evaluation for the schools who are solely responsible for the travel and accommodation expenses of the visiting team in addition to the evaluation fee levied by the IB.

Additionally, virtual visits have been used in politically unstable areas with travel advisories that made it impossible for visiting teams to personally conduct the evaluation. However, Lewis sees further opportunities to be exploited from virtual accreditation visits. He claims:
I think it would be great. Really, if you put together a virtual [school visit] or a video of your evaluation and the process at a school, it could be used for so much more. New teachers coming in they could watch it. You could put it on your school website. Parents could see it and get a better understanding of what’s going on in the classrooms without actually having to walk in the classes during the day and interrupting. I just think it presents a great opportunity for sharing outside the evaluation process. It could be used as a teaching tool; it could be used for so much more. And schools could share with each other or the IB could post it and share with each other [for professional development purposes]. (Lewis, April 22, 2015)

Lewis’s words further demonstrate how he sees the time and effort put into preparing evaluation materials as useful tools for further development of stakeholder understanding of the program.

Finally, Lewis supports providing schools with options in how they present their self-studies and what they focus on in their specific contexts after they have fulfilled the minimum compliance requirements. He reiterates:

I think there needs to be some must do’s. I think you should have to demonstrate some things but [the IB] should give us various ways to demonstrate that. So whether it is students or parents, I think they should [be giving] various models. The evaluation is all about the information, getting the information and then having the visit. I do not think that matters as long as they get the information or we can give them the information. So that would be one way to differentiate. Because I know we had trouble in certain areas as well. You do not want to spend your time running around trying to get so many students, like that is difficult. And in the end, you do not get the sample size that you need and really is that data valid anyway? I think having some options [would be beneficial]. I mean obviously some things you have to do but then you could provide that data in various ways so give some options for schools.

(Lewis interview, April 22, 2015)

Lewis believes that providing more differentiated approaches suited to the particular context of the school will result in more ownership of and engagement in the evaluation process. Correspondingly, Lewis restates that in-depth, contextually appropriate feedback on the evaluation in turn will sustain a meaningfully reflective culture at the school and develop leadership within the school which will further improve the program for the students.

4.5 Interviews with Joanne

Joanne was in her first year as the Head of the English Department during the evaluation of the Diploma Program during 2013-2014. This was a position that she continued
while I interviewed her for this study. She led a group of experienced English teachers and I worked closely with her on the curriculum and delivery of the Grade 11 DP English Language and Literature course which was in its second year of implementation during the evaluation process. Joanne had minimally participated in the creation of the accreditation self-study for the Middle States Association of Schools and Colleges prior to becoming a more integral member of the evaluation team for the IB DP evaluation.

4.5.1 Process

Joanne’s overall feelings regarding the process of the evaluation of the Diploma Program reflect that she believes that it was a worthwhile exercise. She reports that she learned more about the program and what the IB’s expectations are for a solid program. Most of all, she appreciates that the preparation of the self-study made her more reflective both as a teacher and as a Head of Department. The evaluation made her stop and take a look at the implementation of the English Literature course and the Language and Literature course to ensure that IB standards are met and that students are appropriately supported. She reiterates, “A moment of self-reflection and department-reflection was really important to us” (Joanne interview, February 14, 2015). Joanne further elaborates:

Ultimately, I think about the students and what impacts them the most and they can make do with no wireless internet, they can make do with a cramped classroom, ultimately though it is how we are running our courses that make a difference to them. (Joanne interview, February 14, 2015)

4.5.1.1 Systematic Planning of the Evaluation

Joanne states that the organization of the accreditation activities was efficient. She says:

I remember when it was finished and I felt shocked because I think that the process went quite smoothly. It seemed very organized and efficient and I think [the leadership] used us when they required input and feedback and it did not take up a lot of my time, like a huge investment of my time. I think it was a fair amount of work but nothing that was, you know, exhausting. (Joanne interview, February 14, 2015).
Joanne believes that the organization of the evaluation activities can make a difference in how the amount of work required is perceived by teachers.

4.5.1.1.1 Timeline – Synchronization with the MSA evaluation

Joanne acknowledges that using materials that had already been produced for the MSA evaluation streamlined the process for the preparation of the DP self-study. She says, “I do not see why we [should] reinvent the wheel. If the information is there already, I think we might as well just use it rather than having to redo everything” (Joanne interview, February 14, 2015). Personally, Joanne indicates that she simply cut and pasted the information she had gathered about changes in the school’s facilities from the MSA report she prepared into the chart required by the IB. She already knew that the data was accurate since it was used for another accreditation. This type of synchronization of the two evaluations made the process easier for the participants.

4.5.1.1.2 Level of Involvement

Joanne reports that a fair amount of the information used to complete the DP evaluation was compiled by teachers and Heads of Department within the DP community not just by the Diploma Program Coordinator. This made her feel that she had more of a stake in the evaluation. Joanne indicates that she could choose how she wanted to participate in the evaluation beyond what she was required to do as a result of her leadership responsibilities.

She volunteered to complete the chart used for the facilities report since it was a natural fit for her after the MSA evaluation. She appreciates that she could work on this alone because it made sense for her at the time. Regarding working in a committee, she states that she does not mind it “if it works properly and everyone has an equal share in the work, but I find that sometimes working in committees the work is not evenly distributed, so I rather just start off, particularly for [the facilities chart] because I would had experience in it, doing it myself” (Joanne interview, February 14, 2019). Joanne also explains that she felt that she had
more institutional knowledge to contribute to the evaluation after spending four years at the school than many of her colleagues. She jokes that even after a relatively short stay at the school she is considered a veteran in a context where there is a high turnover of teachers and administrators.

Joanne further elaborates that, as a member of the leadership team, she put the evaluation close to the top of her priority list, only outranked by teaching her courses and supporting her fellow teachers in her department. She feels that the evaluation is important because she has seen the growth in the program and views acknowledging the successes of the team as crucial. Joanne opines that all teachers in her team had the opportunity to be a part of the evaluation but sees that some teachers would not prioritize the evaluation over the needs of their students. She even contrasts her own heavy involvement in the DP evaluation with her very marginal contributions to the MSA accreditations. Joanne reflects, “I was quite disconnected from the [MSA] process. I completed a report, edited it, received feedback, reviewed it, revised it and resubmitted. … I think I was just more a namesake than anything else” (Joanne interview, February 14, 2015). The self-study is on top of teachers’ regular duties which take precedence unless they see a value to the evaluation. Joanne perceives the improvements made to the program as a result of the DP evaluation as a compelling reason to significantly contribute to the collaboration required for its completion.

4.5.1.1.3 Accessibility and Institutional Memory

Joanne finds that documenting the evaluation process on the school’s curriculum planning platform (Atlas) and storing all the materials there are steps in the right direction. She relates:

I think [storing evaluation documents on Atlas] is fantastic because I think a critique particular of many programs or many schools is that there's a lack of institutional memory. Teachers change, teachers move and I think that having something on Atlas that is there forever is extremely important. The other thing about Atlas is everyone has access to it and I think that that provides people at any time with the information
that they need and they feel a part of [the evaluation] because they can access it. (Joanne interview, February 14, 2015)

In the context of the school’s very transient international staff, maintaining structures and materials can be chaotic at times. In Joanne’s five years at the school, she worked with three different High School Principals and three different Diploma Program Coordinators. Given these circumstances, it is understandable that she is eager to have all efforts for the improvement of the program properly documented in the hopes of sustaining favorable changes for the students of the school.

4.5.1.1.4 Accountability and Self-Assessment

Joanne advocates for self-assessment as an important part of improving a program. She identifies looking at a school’s strengths and weaknesses as a way to grow. She believes that educators are intrinsically motivated to get better and have a strong program. Therefore, self-evaluation of and reflection on practices are a valid way to determine how to change the program for the better. Joanne points out that there are negative aspects to an evaluation as well. She does not believe that a perfect evaluation exists; schools will always need better facilities, teachers will always ask for more time, and improvement is never done.

However, Joanne indicates that having a fair assessment of the school’s practices, like she thinks the school completed, will lead to improved teaching and learning approaches. Ultimately, Joanne says, “the IB gets to decide whether we get reaccredited, I am sure there are checks and balances to make sure that we had a sound evaluation even though it was self-assessment” (Joanne interview, February 14, 2015). Her words express faith in the system accrediting bodies use to ensure that schools deliver the programs as they should, thus, being accountable for their practices.

4.5.1.2 Pedagogical Leadership of the Evaluation

Joanne indicates that pedagogical leadership for the Diploma Program at the school is shared by teachers and leaders. She acknowledges that the school constantly trains teachers
so there is appropriate subject-specific leadership in the departments. She also argues that both the Diploma Program Coordinator and the High School Principal have sound knowledge of the Diploma Program’s requirements. She recalls, “I do not think that it was in anyone’s mind that the Diploma Program Coordinator and the High School Principal were new. They had the knowledge to lead the evaluation without us having any doubt [about their leadership]” (Joanne interview, February 14, 2015).

4.5.1.2.1 Experience Level with Evaluation

When asked about staff underrating practices in the self-study compared to the expert rating of the IB readers of the school’s evaluation, Joanne says that, when teachers self-reflect, they are harder on themselves than external evaluators would be. She acknowledges that individual teachers have a narrow view of the program that might only apply to their specific course or department. The school’s ratings apply to the whole school and external evaluators have an even wider perspective on many school’s realizations of the IB DP standards and practices. According to Joanne, this can result in teachers being negative or harsh in their assessment of the program.

Joanne further elaborates when she mentions the importance of allowing teachers to use different ways to express their opinions about the program. The current self-study requires a choice of four levels for each of the practices. Joanne wonders “if teachers had to describe [their ratings] using words rather than numbers if the results would be different” (Joanne interview, February 14, 2015). She feels that many people feel uncomfortable with an excellent rating because they do not like to toot their own horn. Moreover, in a very international setting with many nationalities represented, cultural norms can dictate how teachers approach the evaluation. Joanne relates that Arabic teachers behave differently in committee settings than North American or European teachers. Evaluations need to become more culturally sensitive and allow for diverse ways to capture data from all stakeholders.
4.5.1.2.2 Stakeholders in the Evaluation

Joanne shows particular interest in having more representation on the evaluation from parents and students. She was aware that parents and students were invited to give feedback in surveys and meetings but was concerned that the surveys were only available in English. Joanne indicates that many of the parents do not have English as their first language and this contextual reality would limit their ability to express their opinions regarding the ways in which the school delivers the Diploma Program. This deprives the self-study of an important lens.

In a similar vein, Joanne advocates for developing a common vocabulary for all teachers when it comes to the standards that they are discussing. In an international setting, leaders cannot assume that educational terms are used the same way by teachers trained in different parts of the world. Joanne recalls conversations about the meaning of “reflection” and “collaboration” and how they are conceptualized by members of her department. Their answers varied based on their experiences and even their cultures. Agreed-upon definitions of such seminal words can have a profound effect on how the evaluation proceeds.

Moreover, some Arabic teachers in the program felt uncomfortable expressing their opinions because of their limited English skills. Culturally and pedagogically sound clarification of what the teachers are actually evaluating will yield more reliable data for schools to use to improve their practices.

4.5.2 Effect of the Evaluation

This section examines Joanne’s perceptions of the effect of the evaluation on the Diploma Program. Overall, Joanne views the evaluation of the Diploma Program as a reflective process that helped her to see that the school had a strong program already that needed adjustments to make it better in order to support students more.
4.5.2.1 The Progress

As a teacher, Joanne’s approach to the changes made in the program is naturally more focused on the specific teaching and learning practices. She says:

As a result of the evaluation, our program, I would say, was more streamlined. A lot of processes, policies, practices, were streamlined across departments, and I think prior, everyone had their own practices and different methods of doing things. And I think that the process of evaluation and discussing things related to the evaluation helped us to come to an agreement on what our policies and practices should be. (Joanne interview, March 14, 2015)

4.5.2.1.1 Process as Progress

Joanne maintains that not the evaluation itself but the reflection during the process made the program better. She elaborates:

[The evaluation] was a great thing to do for reflection, for us to stop and look at how far we have come and how we still need to progress. I think that our day-to-day actions make the program better, I do not think a survey or the evaluation necessarily did, but I think that moment of reflecting really helped us to think about where we have to go from here, give us some direction in general. (Joanne interview, February 14, 2015)

Joanne reports that the changes proposed by the evaluation team were put into practice immediately upon a decision being reached. She states, “The discussions were brought up during the evaluation but we did not wait for our feedback from the IB to implement the changes” (Joanne interview, March 14, 2015). The only exceptions to this immediate action were the scheduling changes and the assessment policy changes for the Grade 12 courses.

Firstly, the team could not change the school’s schedule in the middle of the academic year but all of the recommendations from the evaluation were implemented in the timetabling for the subsequent 2014-2015 academic year. Secondly, due to the nature of Diploma Program courses running over the period of two academic years, the evaluation team ‘grandfathered in’ the assessment practices used during the graduating class’s Grade 11 year. Teachers felt that it would not have been in the best interest of the students to change practices mid-program. When possible, all the changes were implemented for the Grade 11
students and also communicated to the incoming Grade 10 students and their parents. Joanne reiterates, “We saw the changes immediately. Something would be discussed, and then the change would be made. There was not a delay, and I think that helps with teachers with feeling some kind of agency in the process, because their input is being listened to and immediately acted upon” (Joanne interview, March 14, 2015). The freedom to act on their decisions immediately increased the teachers’ feeling of efficacy.

4.5.2.1.2 The Professional Development of Staff

Joanne appreciates the professional development opportunities provided by the school in the Diploma Program. She believes that the strength of the program is based on the strengths of the individual teachers. She reports that the teachers in her department all received subject-specific IB training prior to commencing teaching their courses. Joanne again emphasizes the reflection on specific practices required by the evaluation because the process made the program clearer to the teachers. She shares that her team had to come up with rationales behind what they were doing and look at the weaknesses of their practices.

As a direct result of the evaluation, the Diploma Program Coordinator requested IB world examination scripts of students back from the IB. These scripts contain examiner comments and marks. Joanne recalls that her team believed that the moderated marking exercise using the IB scripts was the most practical professional development that they ever had since they were using their own students’ work and the examiner marks to calibrate their assessment practices. Joanne agrees that reaching out to the larger IB community can be a powerful tool to increase teacher efficacy.

She reports that she became an IB examiner for the English Language and Literature course as a form of professional development. She recounts:

I undertook examining for Language and Literature and written tasks because I wanted to improve my teaching practice. I felt that our students were weak in preparing their written tasks, and I wanted to see how it could support the practices. Examining for written tasks, when I look at samples from other schools, our school’s
doing really well in terms of supporting our students, preparing written tasks. And I have learned several things about how I can better support my students in the process. (Joanne interview, March 14, 2015)

Joanne recalls that around the same time as the evaluation of the Diploma Program was happening at the school, the number of DP examiners on staff increased from two to ten. She states:

I think, really, going through the evaluation process helped the teachers feel stronger as DP teachers. They feel like they know the program pretty well because they were forced, in some ways, and volunteered in other ways, to step back and take a look at practices and acknowledge the good things that we are doing. … that is one good thing that came out of it is we actually do have a really strong DP program, and I think that helped boost teachers’ confidence to go on and examine. (Joanne interview, March 14, 2015)

Joanne further says that being an examiner gives teachers an opportunity to have more access to the DP world when they can see what other schools are doing and assess where the school fits.

Moreover, Joanne cites the systematic examination of the school’s world examination records as one of the pillars for data-guided decision making in the Diploma Program. She suggests that the evaluation introduced new ways to look at data at the school. She uses the changes made in the supervision of the Extended Essay to illustrate this point. Based on the analysis of Extended Essay grades, teachers wanted to become better supervisors in order to support the students more. Joanne recalls the solution for this problem:

I was the person who was sent for [IB] training for supervising the Extended Essay. I think it opened up my eyes about a lot of things – how we are running our Extended essay program, I’ll say. I do not think we were scaffolding students enough through the process. We did not have deadlines. We did not have in-term checks to make sure that they were on the right track. I think the approach from supervisors was too hands-off. We were trained, but I do not think teachers were trained to supports students properly. There was a lack of consistency in terms of how our students were going through the research process. Every student was receiving a different experience from the Extended Essay. (Joanne interview, March 14, 2015)

After her IB training, Joanne was a central member of the team that produced a new Extended Essay Handbook for students and she prepared a separate one for supervisors. A
new training program for Extended Essay supervisors was also devised with significant assistance from Joanne. Her exposure to other Extended Essay programs within the IB network of schools while at IB training enabled her to bring best practices to students and teachers. As Joanne mentioned above, supervisors were trained for the Extended Essay but the training was ineffective. When the new training program was put into action, we wondered how we would know if we were actually making any difference.

Fresh out of a doctoral course where I learned more about program planning and evaluation, I proposed to the team that we could create our own measurement tool to see if the new program was successful. Joanne and the team worked with me to put together a logic model that would provide us with information regarding the usefulness of all the new practices we put in place for teachers and students to improve our program for the Extended Essay. The logic model can be found in Appendix G. The logic model demonstrates the team’s commitment to continuous performance monitoring. The data collected using the logic model was subsequently used to fine-tune and refine practices for the Extended Essay supervision at the school.

4.5.2.1.3 Structures and Procedures to Increase Student Achievement

Joanne sees proceduralizing long discussions about practices as one of the positive effects of the evaluation. Joanne reports that prior to the evaluation there was inconsistency across departments regarding how the program was interpreted. Based on the subject matter and even the background of the teacher, courses had different expectations. For example, Joanne recounts that teachers, at their discretion, discarded marks as they pleased using their professional judgment. As a result of the evaluation, using IB policy guides to adjust practices, all teachers need to follow the Best Fit approach when marking.
Joanne reflects that teachers’ interpretations of IB practices actually conflicted with requirements so it was a worthwhile effort to consult IB documents during the evaluation. Using similar processes, many of the program’s elements were altered. Joanne reports the following structural and procedural changes and their perceived effects after the evaluation of the Diploma Program in table 3. Joanne believes that these changes ensure that her students receive the best practices. However, she acknowledges:

It probably took some people longer to accept change, because their practices had been based on tradition rather than what was correct or what was from the [IB] guide or what was best for the students. So I think some people were slower to accept changes, but I think now, they would acknowledge the consistency across departments, across subject areas. I would hope. (Joanne interview, March 14, 2015)

Joanne further advocates for consistency among the different subjects as well. She says:

I think it is really important for students to see that there is this continuity across the program rather than, I think, prior. We had a lot of students saying, okay, well, in science, our teachers do this; in math, our teachers do this; in English, our teachers do this. But it is good for them. I think [consistency] gives them more peace of mind and I think their expectations are clear to them when there’s consistency with the achievement charts. (Joanne interview, March 14, 2015)

Moreover, while Joanne believes that it was important for teachers to be involved in the decision making during the evaluation, she does not feel that teachers are losing agency now that many of these practices are standardized and more top-down as opposed to being at teachers’ discretion. She notes:

I think that what you do day-to-day in your classroom and your backward planning and how you design assessments is where your agency comes in. But I think certain things have to be consistent, and there should be procedures and policies put in place. I am sorry, consistency is important, and ultimately, the students are doing assessments that are mandated. So we need to make sure they are properly prepared, and some things have to be mandated as well. I do not think it is a lack of agency. I think the agency they put in to design these procedures and policies is the important part. (Joanne interview, March 14, 2015)

Joanne further makes a case for practices being tailored to the needs of the specific school when she mentions the scheduling for a balanced school day that provides a nutritional break in the morning for students. She argues that, in the context of the school where local
<table>
<thead>
<tr>
<th>Change to the Diploma Program</th>
<th>Effects Perceived by Joanne</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Developing program specific information for teachers and students:</td>
<td>1. Providing more structures for students to support them in the Diploma Program</td>
</tr>
<tr>
<td>- Diploma Program Handbook for Students</td>
<td>2. Proceduralizing discussions about the program</td>
</tr>
<tr>
<td>- Diploma Program Handbook for Teachers</td>
<td></td>
</tr>
<tr>
<td>- CAS Handbook</td>
<td></td>
</tr>
<tr>
<td>- Extended Essay Student Handbook</td>
<td></td>
</tr>
<tr>
<td>- Extended Essay Supervisor Handbook</td>
<td></td>
</tr>
<tr>
<td>- Atlas resource page for DP teachers</td>
<td></td>
</tr>
<tr>
<td>5. Scheduling the Diploma Program:</td>
<td>1. Meeting IB requirements for the teaching of HL and SL courses</td>
</tr>
<tr>
<td>- Providing appropriate timings for HL and SL courses including A block and H block HL classes and SL tutorials</td>
<td>2. Providing appropriate supports for students to improve academic achievement while reducing the material covered</td>
</tr>
<tr>
<td>- Reducing TOK scheduling through A block semestered approach</td>
<td>3. Increasing collaboration time for teachers to grow their capacity to deliver the program</td>
</tr>
<tr>
<td>- Building collaborative department meeting times into teacher schedules</td>
<td></td>
</tr>
<tr>
<td>- Providing more professional development time on early release days</td>
<td></td>
</tr>
<tr>
<td>6. Updating all assessment policies for the Diploma Program:</td>
<td>1. Using data-guided decision making process to reach conclusions about program practices</td>
</tr>
<tr>
<td>- Aligning of the school grades with predicted grades and IB grades</td>
<td>2. Streamlining all Diploma Program practices</td>
</tr>
<tr>
<td>- Moderating IB world exams scripts</td>
<td>3. Introducing consistent practices among departments</td>
</tr>
<tr>
<td>- Updating all achievement charts</td>
<td>4. Increasing teacher confidence in grading and reporting practices</td>
</tr>
<tr>
<td>- Regulating homework, retake, and formative assessment procedures</td>
<td>5. Providing more communication about assessment to parents and students</td>
</tr>
<tr>
<td>- Establishing use of GradeBook for recording marks</td>
<td>6. Spreading out internal assessments over the course of the two years of the program to reduce stress on students</td>
</tr>
<tr>
<td>- Differentiating between academic achievement and learning skills</td>
<td>7. Building more teacher accountability into the program</td>
</tr>
<tr>
<td>- Providing descriptive feedback to students</td>
<td></td>
</tr>
<tr>
<td>- Building skills through formative assessment to increase performance on summative tasks</td>
<td></td>
</tr>
<tr>
<td>- Moderating all summative assessments</td>
<td></td>
</tr>
<tr>
<td>- Determining report card grades using the Best Fit approach</td>
<td></td>
</tr>
<tr>
<td>- Establishing common vocabulary to explain assessment practices</td>
<td></td>
</tr>
<tr>
<td>- Establishing firm due dates calendar</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Joanne’s perceptions of structural and procedural changes as a result of the evaluation.
families live in a night culture, it is important to give a chance to eat in the morning to sustain them until lunch time. Furthermore, she emphasizes the importance of easy accessibility of all of the DP documents on Atlas as the hallmark of a transparent program. She says:

I think putting everything in one place gives us something common to look at, and unites us in some way, because we have all of our documents in one place, and we are sharing resources, and I think that helps with continuity from one teacher to another. We have high turnover here. So I think having something like Atlas really helps with turnover and making sure that things are best practice and not prior practice. (Joanne interview, March 14, 2015)

Joanne’s words show her understanding that meaningful improvement for a program is contextual and culturally sensitive.

4.5.2.2 The Pushback

Overall Joanne feels that the evaluation of the Diploma Programs was a positive experience for her that improved the program significantly and provided her with opportunities for reflection and professional development. However, she acknowledges that others might have a different lens on the evaluation activities. She states:

I think a lot of people have an unrealistic expectation of the program itself and I think that maybe not enough experience in the program, too. But I think every DP program across the world has issues or weakness and I think that if they were to go teach at another school they would find weakness there as well; maybe in different areas. I also think that it is human nature; there are people that tend to be more negative. We will find that negativity anywhere, I do not think it is necessarily just here. (Joanne interview, February 14, 2015)

Joanne also recognizes that there are a lot of things that are out of teachers’ or even leaders’ and owners’ control such as facilities, access to certain resources, censorship, and the local Ministry of Education requirements. Joanne points out the many times North American teachers have unrealistic or culturally inappropriate expectations and remain uninformed about the local and national context in which the school is situated. Their inexperience with these contextual realities can create a conflict when they engage in evaluation activities.
4.5.2.2.1 Lack of Time

Joanne claims that one reason why some teachers did not fully engage with the evaluation process is lack of time. She says that the primary function for teachers at the school is to prepare lessons for students and they choose to spend their finite time and effort on instruction. Moreover, Joanne points out that teachers who are new to the school or the program might feel disconnected from the evaluation in other ways as well. She elaborates:

I talked earlier how teacher turnover here is an issue and I think that the less time you've spent here, the less you feel involved in the program. And so you may be teaching in the DP program but you may not feel like you need to worry about its organization overall or care about its organization overall. … You are here for two years; you are here for three years, you leave. You have your DP qualification, that is all. … After five years, I am a veteran teacher [here]. (Joanne interview, February 14, 2015)

Joanne finds that lack of time is not limited to time provided for the preparation of the evaluation. The short time that teachers spend at the school also affects how invested they feel in the success of the Diploma Program. Hence, the context of a transient teacher population has varied consequences for the process of the evaluation: new teachers need more time to get up to speed on teaching the DP and might lack time for the evaluation and they do not spend enough time at the school to get invested in its evaluation.

4.5.2.2.2 Personal Feelings

Joanne mainly sees negative reactions to the changes in the program as a function of the increased accountability that was brought to teachers as a result of the evaluation. This means a lot more work for teachers. She explains:

I mean, there’s always resistance to change. I think those teachers are quite short-sighted if they feel like that, but I think the [accountability] is for the betterment of the program, but also to better support students, so I mean, if they have an issue with it, too bad. Really. … [The negative comment] sounds like a teacher who’s been in the program for a long time and who is used to a more laissez-faire approach to the program, and, I think, probably a lack of rules, perhaps – deadlines, guidelines – and maybe a lack of work, because with deadlines – I mean, this is all work for a teacher. But that is to support students so ultimately, there shouldn’t be anything wrong with it. But I think it seems like a teacher who has been here a while and is averse to change. But that does not mean that someone’s ego or lack of work was called into
question…. Change is hard and some people are short-sighted. They are personally impacted; they are not looking at the larger picture. (Joanne interview, March 14, 2015)

Joanne’s words show that pushback from some teachers who are averse to change should not discourage staff from making changes to a program in order to improve practices to make students more successful. Joanne concludes, “It is for the greater good” (Joanne interview, March 14, 2015).

4.5.3 Future of Evaluation

Joanne observes that, as the school moves forward after the evaluation, there is “more of a drive towards planning, and long-term planning, not just short-term just for the year, about where we are going in general with our school and with the Diploma Program” (Joanne interview, April 30, 2015). She welcomes this change and encourages any practices that make it easier to improve the students’ and teachers’ experiences at the school.

4.5.3.1 Future of Evaluation at the School

Joanne’s main recommendation for the future of the Diploma Program is to increase the number of students at the school who attempt the DP. She believes that, by including the students and parents in program planning and evaluation and adjusting the practices of the program to the specific local context of the school, the enrollment in the program can grow. She would like to see more face-to-face and focus group discussions with all members of the community to ensure that the DP will continue to thrive at the school.

4.5.3.1.1 Continuous Change Model

Joanne points to the action plan that was developed as a part to of the evaluation as an important factor in the future of the program at the school. This is a roadmap for the staff that they can follow to implement all of the recommendations made by the teachers and by the IB. Joanne says, “Some of the things [in the action plan] have actually happened so it is good to see growth and I think, again, the fact that we are involved in the process also helps us want
to make [the changes] happen because we had that role” (Joanne interview, February 14, 2015). However, she warns that the next evaluation should not be put off until it is required again by the IB in five years’ time. She insists:

[Measurement of practices] should be annual. We know, based on our last evaluation, that it does take time to implement change. But the change has to be consistent, it has to be sustainable, so this needs to be done every year. Trying to incorporate new voices, for example if we are doing a promotional brochure, incorporating new voices every time there’s a graduating class and making it relevant to students. Stakeholders always change. Our community is highly transient so the stakeholders are always changing, the population of the stakeholders is always changing, and the demographics are always changing so I think this needs to be a consistent effort. (Joanne interview, April 30, 2015)

Joanne continues to emphasize the importance of school-specific and continuous evaluation in order to serve the students best.

4.5.3.1.2 Synchronization of Evaluations

Joanne welcomes the idea of a continuum evaluation for the three IB programs at the school. She believes that amalgamating the information required for the entire program will streamline the process for the staff. To Joanne, it makes sense to provide the same information at one time instead of the previous system of three different timelines for the three programs, even if this schedule brings the mandated evaluation of the DP forward by two years. She states:

[Continuum evaluation] is a good thing because I think a lot of the information we provide for all of the evaluations, including the MSA evaluation, is similar. There are some unique aspects, which is fine, but I like the idea of we are all working on the same thing at the same time rather than being evaluated every year for a different program, including our MSA accreditation. We could always be in evaluation mode. It is almost like standardized testing. We are working towards an external goal or teaching to the task rather than focusing on building our community and other things. So I think this alignment gives us time to make changes and have a five year cycle, after of course this next evaluation because this is quite soon. (Joanne interview, April 30, 2015)

Joanne is not worried about staff pushback caused by the shortened evaluation cycle for the Diploma Program. She elaborates:
To be honest, I think that the turnover here allows for fresh voices and new teachers, and so I do not think there will be that much pushback. I do not know how many teachers who went through the last MSA evaluation and the DP evaluation will still be here. So I do not think it will be negative. I mean I do not think evaluations are ever something people want to do, but knowing that they have to be done, I am sure there won’t be any negativity towards that. (Joanne interview, April 30, 2015)

The transient nature of teachers at the school would be alleviated by more frequent evaluation of practices to promote deeper knowledge of the program and more ownership of the changes proposed.

4.5.3.2 Future of Evaluation for the IB

Joanne continues to encourage the accrediting bodies to reduce the redundancy of information required from the schools. She believes that using technology to simplify processes and allow teachers to use their time more efficiently will increase teacher participation in the accreditation processes.

4.5.3.2.1 Suggestions for the IB

Joanne suggests that synchronized and continuum evaluations can serve the schools in a more efficient way. She further argues for similar context school visits: school visit team members who are familiar with the particular setting of the school for more personalized feedback. She urges the IB to refresh the IBEN members and upscale the experience of IB educators, including school visiting team members. Joanne compares the different models used by MSA and the IB thusly:

It is important to have fresh people on the team. And I really like the MSA model. … The MSA evaluators are chosen from MSA participant schools, or accredited schools, so those teachers and administrators have a view of what it is like to be working in those communities and a more realistic picture of how quickly change happens or it does not. (Joanne interview, April 30, 2015)

Using this model for external evaluators would also ensure that the feedback provided by the IB to school is more detailed and personalized. Joanne argues that, using a drop down menu
for the IB report to the self-study, does not give enough descriptive feedback to staff to feel further inspired by the accreditation process.

4.5.3.2.2 Feedback on IB Proposed Changes

In her feedback on the proposed IB changes to the evaluation, Joanne opines that the plan to include more concrete evidence about instructional and collaborative practices should encourage staff to base their rating of practices on more than just mere opinions. She admits that this will create more paperwork for the teachers but it will result in a more consciously reflective teaching environment.

Joanne finds the idea of virtual evaluation visits great. She muses, “I am surprised that they do not change the system actually overall rather than sending people which can be quite costly” (Joanne interview, April 30, 2015). Admittedly, the IB programs are fairly expensive to run and one of the criticisms of the IB is that it needs to be more inclusive; not every school can afford to have an evaluation team to come to their school, so they could do it virtually.

However, Joanne also sees a challenge to this process. She worries about schools putting on the show for the virtual evaluation. Joanne says, “It would be a lot easier via virtual evaluation than if people were actually here, because you can’t really hide … you can hide things on a Skype phone call which you could not do [if the evaluators are present]” (Joanne interview, April 30, 2015). Joanne would suggest verification visits as part of the process to ensure that the school is not just window dressing their program. Joanne recognizes that producing some of the materials for the virtual visit or as a part of the self-study would be beneficial for the school in other ways as well. Schools could use the evidence that they collected in video form for training purposes of new staff, or as promotional material for prospective students and parents.
Finally, Joanne thinks that differentiating evaluation requirements based on the length of time that the school had authorization from the IB would benefit the schools. Brand new schools have very different needs from schools where the DP has been offered for decades. Joanne states:

Looking at an established program versus a newer program is that you can have benchmarks for where you should be in your programming, like if you are still working on developing an assessment policy then that is something that a new school or a new program should be doing, not a school that is 15 years old. … So those benchmarks will definitely help with making sure your program is measuring up to where it should be. (Joanne, April 30, 2015)

Joanne suggests that some basic requirements should be established for all schools but the different needs of the school have to be met through differentiation.

4.6 Interviews with Bill

Bill was a teacher in the Humanities Department during the evaluation of the Diploma Program during the 2013-2014 academic year. He became the Head of the Humanities Department by the time the interviews were recorded for this project. In preparation for his upcoming leadership role, he was approached by the High School Principal and the Diploma Program Coordinator to work closely with the outgoing Head of Department as the Humanities teachers were working on the evaluation together. Bill and I worked closely together as he transitioned into his first leadership role and participated in his first accreditation process at the school.

4.6.1 Process

Bill shares that he found many positive aspects in engaging in the process of completing the self-study of the Diploma Program both on a personal level and on a greater school level. He reports that he could reflect on his understanding of the Diploma Program since previously he taught more courses in the Middle Years Program. He especially learned more about the Diploma Program core requirements and even found identifying what he did
not know valuable so that he could address the gaps in his knowledge. Bill says, “I find that I want to be the best teacher that I can be and the best well-informed teacher to better support my students. And I learned a lot from this process” (Bill interview, February 15, 2015). Bill’s focus during the evaluation was on how he could personally improve his practices to provide better instruction to his students.

4.6.1.1 Systematic Planning of the Evaluation

Bill identifies organization as an important characteristic of the evaluation of the Diploma Program. He admits, “There is nothing more that I hate than being a part of something and not knowing what I am doing and why I am doing it” (Bill interview, February 15, 2015). He asserts that the process of accreditation was a well-organized and productive experience for him because it was clear what the team was doing and why. He recalls the first meeting of the evaluation team where the plans for the evaluation were laid out step-by-step by the DP leadership. The timeline and documents from the IB were shared with all members of the team that were made available for later reference as needed. He feels he was well-informed and that he could give his opinions that were authentically heard in the evaluation process. Bill says, “If you are going to be invited to give your time and someone wants to hear your voice, it is nice that I felt that my voice was heard” (Bill interview, February 15, 2015).

4.6.1.1.1 Timeline – Synchronization with the MSA evaluation

Bill did not sit on any of the committees charged with completing the MSA accreditation. He reports that he answered the staff surveys that the evaluation team sent out looking for staff feedback and attending general staff meetings where the different groups updated the teachers regarding the evaluation. He feels that the groups did a good job of keeping the process of reaccreditation for the MSA very transparent at a larger, institutional
scale. These processes prepared Bill to take a more active role in the subsequent DP evaluation that was of more personal importance to him.

### 4.6.1.1.2 Level of Involvement

Bill recalls that he was approached to become a part of the evaluation team after he was selected to be the Head of Humanities. During discussions with the DP leadership and the Head of Department at the time, he agreed that it would be beneficial to him and the department to take on some of the responsibilities of the headship earlier. Bill and the outgoing Head of Department led the group evaluation for the Humanities teachers together and Bill represented the team for the evaluation committee.

In addition, Bill volunteered to be a part of the subgroup that completed the section of the self-study rating the philosophy standards and practices of the program. Bill says, “I chose that section because I felt most personally connected with it and interested in it. … I thought it would be a more authentic approach because I had an intrinsic interest and I would be more engaged with that” (Bill interview, February 15, 2015). He states that he felt a special affinity towards exploring how the mission statement of the school is actualized to provide a rigorous but balanced education for the students while still nurturing them so that they can be successful in the Diploma Program.

As a result of these experiences, Bill felt highly included in the preparation of the self-study. He believes that his involvement was appropriate for the role he had in the evaluation team and he had a choice to say how much above and beyond the basic requirements he wanted to go. Bill states that there was a shared responsibility for the completion of the self-study and all related documents. Bill acknowledges that his priority was teaching his classes and the evaluation was added to his regular instructional duties. However, he states that he could choose to engage only in the parts of the evaluation that he was particularly interested
in and passionate about which provided enough motivation and personal connection to balance the extra effort required for the accreditation activities.

4.6.1.1.3 Transparency and Accessibility

Bill sees the availability of all evaluation documents on Atlas, the school’s curriculum planning platform, as a crucial ingredient for the smooth running of the evaluation. He argues:

I would say that [having the documents on Atlas] would be beneficial both proactively and retroactively. So, going into a meeting that is a Diploma Program meeting or if it was a committee meeting, the agenda was posted what was going to be happening, what we would be addressing was all there. But also it is still easy to access if I need to revisit it or work deeper to understand what we are actually talking about. The documentation is still accessible in one spot. Not getting buried in email, not getting buried in the mountains of papers that can engulf teachers at points in time. (Bill interview, February 15, 2015)

Bill elaborates that having free access to the documents also developed a feeling of shared responsibility for the process and shared ownership of the self-study. He expresses:

I think the fact that they were there and accessible and could be accessed at the leisure of anyone participating [in the evaluation] is wonderful versus a coordinator saying well, I am keeping all the documents. That just can be a road block in someone’s day. But if you can access those [documents] on a prep period at school, or if you are at home and you are curious about something, you know, I think that the accessibility of those documents through Atlas is wonderful. It makes the process very transparent and accessible at the leisure of the person doing the accessing. (Bill interview, February 15, 2015)

4.6.1.1.4 Accountability and Self-Assessment

Bill believes that the self-study submitted to the IB accurately reflects the Diploma Program at the school as it functioned during the 2013-2014 academic year. He reports that there was open dialogue and in-depth conversation regarding the ratings of the standards and practices in order to come to an agreement to say that the documents were a true assessment of the program not just window dressing. Bill states:

Everyone came with the documentation and they were ready and had their voices heard before a decision was made collectively as a group. … I think that some of the discussions became more heated than others and I am not sure everyone would agree
with the assessment that every single mark truly reflects their opinion. But I think it was to a pretty high degree. (Bill interview, February 15, 2015)

Bill continues to say that “self-assessment can be tricky but it is also necessary and I think that it just needs to be set up well with parameters and expectations” (Bill interview, February 15, 2015). He acknowledges that there need to be checks along the way in order to hold individuals and institutions accountable so that their self-reflections are true in nature. He suggests that appropriate evidence needs to be used to underpin the conclusions of self-studies. He calls for a close understanding of the expectations of the accrediting organizations by the internal evaluation leaders and the external readers who will put the final stamp on the school’s self-study since it is their perceptions that count as evidence of a school’s practices.

Bill says:

The people that they have within the positions that are part of the committee for the assessment can hopefully keep things legitimate in nature and the assessment becomes legitimate and fair within it. … And I guess I am thinking more that the individuals that are working for the school, that are doing the assessing, are properly trained practitioners and that they know what they are talking about when they are giving that assessment. And that they are educators, I guess I would worry that the opinions could get wild if the person does not know what they are talking about. Like if their opinion isn’t rooted in educational theory and in educational practice, but it is based on what they think and what they feel because that is just something that they’ve cooked up then that can be scary. (Bill interview, February 15, 2015)

Bill states that it is important to get the perspectives of teachers who have been at the school for a longer time in order to capture institutional memory and document how the program has evolved. He further insists that the conversations about the program should not go down the road of bashing the school. He argues that the evaluation of the Diploma Program at the school did not become negative because it was structured well. He recalls:

We had clear tasks of what we were doing and how we were doing it every time that we met. And the DP Coordinator was clear this is what we are doing this time and this is how we are going to do it and this is why we are going to do it versus some - I can't even describe it because it did not happen, like some sort of loosey-goosey like let's just talk about what we are doing. And then that is when the spiral happens, when there's no road map to where we are headed. That is when teachers start to tell their stories, oh yeah one time - and then we just start to [go off topic], teachers [laughs]. (Bill interview, February 15, 2015)
In Bill’s opinion, the structured organization of the evaluation meetings ensured that teachers were accountable throughout the process.

**4.6.1.2 Pedagogical Leadership of the Evaluation**

Bill states that there was competent pedagogical leadership for the Diploma Program during the evaluation. He says that good educational leadership is not necessarily program specific so the fact that the Diploma Program Coordinator and the High School Principal were new to the school did not affect how he viewed their ability to lead the program. In addition to the educational training and experience of the leaders, he also takes into account that he has witnessed their experience in an educational setting which leads him to trust the leaders and the process that they set forth for the evaluation of the program. He further shares:

> I see these people that have invested themselves in education, invested themselves in educational leadership, in the development and the betterment of schools and of young people. And that they prove themselves on a daily basis, but I guess on a grander scale as well, to be trustworthy drivers of this ship. … In the time leading up to the [evaluation] in all my interactions with them, they had proven themselves to be competent, knowledgeable, caring educators who were interested truly in creating a legitimate process and legitimate educational system within our independent school. (Bill interview, February 15, 2015)

Bill finds that the relationships that the leaders built with him and others on the teaching staff were rooted in sound instructional practices which led to an authentic evaluation of the program.

As a result of this ethical foundation created jointly by the leaders and the teachers, Bill sees no reason to suspect dramaturgical compliance in the self-study prepared for the IB. He posits that the assessment of the school’s program was realistic as evidenced by the students who graduate from the school with a strong educational foundation which enables them to be successful at the IB world examinations and gaining admission to great universities. Bill concludes, “If a school says it is doing wonderful but students are failing the
world exams, not getting in a school or they are dropping out [of university] quickly, then that for me would become more of a question about the process [of the evaluation]” (Bill interview, February 15, 2015).

4.6.1.2.1 Experience Level with DP and Evaluation

Bill attributes the differences found between the ratings of certain practices between internal versus external evaluators to the personal feelings of the teachers at the school. He says, “I think [the low ratings] reflect that it was a group of individuals that was interested in taking what they thought was a critical look at their practices and the practices of the school … I do not want to say we are our own worst enemies, but we are our own worst critics” (Bill interview, February 15, 2015). Bill hopes that teachers are always looking to make themselves better and make the program better for their students. There is always a need for improvement.

According to Bill, the low ratings can also show that teachers took the evaluation seriously and did not want to give themselves an easy pass. Moreover, he further questions the rating system used for the self-study. The highest rating required that the practice was in evidence explicitly and consistently. Given the complexity and relatively large size of the program, Bill argues that it was difficult to say that all practices were present in such a degree across the entirety of the DP at the school. As a result, Bill wonders that the experience level of different teachers both with the DP and with evaluation gave way to too many different interpretations of how the practices were actualized at the school resulting in varying perceptions of the program and its evaluation.

4.6.1.2.2 Stakeholders in the Evaluation

Bill recounts in detail the consensus building that was used by teachers to first arrive at their ratings of practices in their departments and then to amalgamate these results for the final school ratings which were reported to the IB. He recalls that each practice was discussed
individually and during the conversation each member provided evidence for their opinions. If there were outliers, the specific reasons of why a certain teacher or group did not agree with the larger group was recorded and also included in the final self-study. Bill maintains that the Diploma Program Coordinator acted as a recorder and not a participant during these discussions and that the diverse answers were truthfully communicated to the accrediting body. Bill believes that this process ensured that teachers and pedagogical leadership were authentically represented in the findings of the self-study.

Parents and students of the Diploma Program represent other important stakeholders in the program. Bill reports:

I think the outreach [to parents] was great. But also I remember a feeling of frustration especially I think from the parents’ side of being able to most effectively get that feedback. So, it is like the opportunity was there but I know it is not a missed opportunity but it was a battle to get a large survey of parents to contribute to the information that went into the final product. (Bill interview, February 15, 2015)

Bill says that the cultural context of the school is probably the main reason behind the lack of involvement from the parents. Bill says:

There's not a culture of high parent involvement on a regular basis. … I think we have a bit of disengaged parents. That school is something for their children and if the students are failing then that is when they swoop in but besides that we are going to leave it to the teachers and the professionals at the school and that is kind of it. … And I think lastly with that like a by-product of that disengagement from parents is that parents maybe aren’t confident in their understanding of the programs that their students are involved in. They understand that when their student does not receive a mark that they think their student is capable of but the understanding of the programs and the requirements and the nuances of the program, they maybe do not feel confident to speak to it. (Bill interview, February 15, 2015)

Bill explains that the cultural and societal context of the school has a significant effect on how much stakeholder contribution a school can expect during their program evaluation.

4.6.2 Effect of the Evaluation

In general, Bill feels that the evaluation had a positive effect on the Diploma Program at the school. He ponders:
I guess my belief, and maybe I am naive or idealistic, is that I do not really understand how [the evaluation] could not make the program better. Because for a variety of reasons. Number one, it can make us better since within the committee someone is represented from every single group so we are able to reflect on not just what we are doing but what we are not doing well and that can pave the opportunity for those people to dialogue and discuss. But also to represent that group and have a greater discussion with the people that are on the teams within those groups to say, ‘oh my gosh in assessment group or in this reaccreditation committee we talked about this’. The perception is that as a group or as a school we are not doing this well and that can open up conversations and conversations can be the beginning of change. I think that it made us stronger by assessing the practices that we had been putting in place and being able to evaluate what we are doing well, the progress we have made but also where we want to go. I just feel like any time that you can be reflective on the practices that you can do and you can be reflective in a true way, like there’s no way that you can’t be better. (Bill interview, February 15, 2015)

On a personal level, Bill shares:

I feel like I learned a lot, especially being newer to the program. Like my vision on the program was very focused on making my classroom within my group and my subject area being successful. But I feel like my understanding of the program as a whole was opened up much more by hearing all that is required of the students and required of the school and in other groups as it was shared out from other group members in the committee. It was just eye opening to me and it helped me be a better DP teacher because I better understood the program as a whole. And I think a lot of that came from the information getting that was just part of the sharing process. (Bill interview, February 15, 2015)

Bill’s overall perception of the accreditation process is that it set up the program for continuous improvement.

**4.6.2.1 The Progress**

As a teacher during the evaluation, Bill’s main focus is the improvement of teaching and learning approaches. Accordingly, he finds that “the strength of the program comes from the strength of the teachers” (Bill interview, March 15, 2014). Bill sees self-reflection and the changes in practices, both short and long-term, as the immediate results of the evaluation. In particular, he believes that pulling procedures together into seminal documents helps teachers as there is high turnover rate for leaders and teachers at the school. Bill attributes all of these alterations to the program to the evaluation process. Bill states, “From my perception, the program at our school absolutely, 100% has been changed because of the self-reflecting
reaccreditation process. And I would say that it continues to change and it continues to really only improve because of it” (Bill interview, March 15, 2015).

**4.6.2.1.1 Process as Progress**

Bill indicates that the process of going through the evaluation not just the evaluation itself brought about needed changes in the program. He says:

I notice that change began immediately after the completion of [the evaluation]. … I would say that there’s a potential to change, at least started brewing even in the midst of the process itself. Like the process, for sure, spurred change. Not from a negative aspect, but I think from a – how can we make ourselves better? Not like we are messing all this up, but like let’s make ourselves a better and stronger program now. (Bill interview, March 15, 2015)

Moreover, Bill finds it essential that teachers felt empowered to implement the changes. He elaborates, “We were given the freedom and the venue to give a critical look at what we were doing and to say are we assessing students in the way that is best for them and that supports teaching and learning” (Bill interview, February 15, 2015). Bill concludes that he sees reflection on teaching practices and their alteration as needed a primary function of teaching and learning.

**4.6.2.1.2 The Professional Development of Staff**

Bill suggests that one of the direct positive effects of the evaluation is the professional development of the educational staff at the school. He differentiates between two areas of professional growth at the school during and after the accreditation process: impactful reflection and professional training. According to Bill, learning activities were a result of reflective conversations that informed data-guided decision making and the systematic changes that were discovered during the evaluation process.

In turn, once the training of teachers was completed, they shared their expertise with the rest of the staff and further improved the practices for the program. The two processes feed into each other. Bill illustrates this point with the following example regarding the changes made for the training of Extended Essay supervisors:
One of the Heads of Department went to the Extended Essay [IB] training, came back, and worked with the Diploma Program Coordinator to use the expertise that she had taken from the training along with the expertise of the coordinator to create both Extended Essay Handbooks [one for students, the other for supervisors]. They complement the ideas and the requirements of the IB, but then also they were able to pull in outside knowledge as well from two women with extensive experience in education. So it was nice that there was clarity from the IB, but also providing just general best practices in education so as to support students that could give clarity to teachers and students in the process. (Bill interview, March 15, 2015)

Furthermore, Bill recalls how the conversations during the evaluation made an impact in classrooms to ensure that IB and DP leadership expectations were met. He uses the example of his own subject, IB History, when he explains this idea:

I’ll use IB history. I am not even sure if IB history was being taught in the way that the IB envisioned it happening [prior to the evaluation]. Once we assessed what was happening through our self-reflection, we are kind of going, ‘Whoa.’ We have some serious expectations and guidelines that the IB gives upon to us. And because of that reflection and saying some of us are doing it really well, some of us are doing it okay, and some of us aren’t doing it at all, we made changes. … We said, “If we are going to do [the needed changes], we as a team need to do that.” And that is everyone across all subject areas. And I saw a direct change because of the self-assessment into the team that is the diploma teachers. And the expectation became clear from administration and the DP Coordinator down that these are your expectations in terms of assessment practices which would then lead into teaching practices building a stronger program. (Bill interview, March 15, 2015)

Moreover, Bill realized that he personally wanted to engage in additional professional development. He states, “I felt being reflective for myself like, oh my gosh, I need to make sure I am doing my best to support the Extended Essay and TOK (Theory of Knowledge) because those are integral for the students’ overall success” (Bill interview, February 15, 2015). After conversations about TOK during the evaluation, Bill decided that he needed to know more about the subject and signed up for external IB training for TOK. In this way, Bill acknowledges that the evaluation directly affected his personal choices for his own professional development.
4.6.2.1.3 Structures and Procedures to Increase Student Achievement

Bill emphasizes the importance of standardizing procedures across the subject groups and clarifying expectations in all aspects of the program as the most important improvement caused by the evaluation. These changes made the program more accessible and transparent for all stakeholders which is crucial for Bill. Bill observes the following structural and procedural changes and their perceived effects on the Diploma Program in table 3.

Bill suggests that the structural changes brought clearer expectations for all teachers which made the program delivery more seamless. A year after the evaluation, Bill believes that the program can now be refined as part of a continuous improvement process. In this way, the program is becoming more proactive as opposed to reactive. This means that, instead of dealing with problems that come up in the program, staff are now anticipating problems and putting structures in place to prevent the issues from happening in the first place.

Increased accountability for teaching staff and oversight of the processes by program leadership are a significant element of this shift. Bill reflects that these changes enabled teachers to bring more rigor into the program and allowed student results to increase while reducing students’ stress which is their moral imperative.

Finally, Bill argues that the structural changes that the evaluation precipitated give students and teachers “a gift of time” (Bill interview, March 15, 2015). Bill says that the new schedule allows for more instructional time for students, which enables him to differentiate his instruction and provide better learning for his students. At the same time, teachers have more collaborative time together to plan and moderate which is incredibly beneficial for the teachers. He concludes:

I think that the program has gotten better in terms of scheduling for both students and for teachers. Like for students, everyone is given the gift of time. For HL students,
they are given more time to grapple with some of the extra HL topics that they have to. For SL students, they’ve been given the gift of time to grapple with some of the complex skills which they have to apply in their assessments. Teachers are given the time of being able to sit across from their colleagues face to face and plan and talk. (Bill interview, March 15, 2015)

<table>
<thead>
<tr>
<th>Change to the Diploma Program</th>
<th>Effects Perceived by Bill</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Developing program specific information to teachers, students, and parents:</td>
<td>1. Increasing transparency, accessibility, and clarity of all program procedures</td>
</tr>
<tr>
<td>-Diploma Program Handbook for Students</td>
<td>2. Making it easier for new teachers to the program to locate resources</td>
</tr>
<tr>
<td>-Diploma Program Handbook for Teachers</td>
<td>3. Making the program more proactive as opposed to reactive</td>
</tr>
<tr>
<td>-CAS Handbook</td>
<td></td>
</tr>
<tr>
<td>-Extended Essay Student Handbook</td>
<td></td>
</tr>
<tr>
<td>-Extended Essay Supervisor Handbook</td>
<td></td>
</tr>
<tr>
<td>-Atlas resource page for DP teachers</td>
<td></td>
</tr>
<tr>
<td>2. Scheduling the Diploma Program:</td>
<td>1. Meeting IB requirements for the teaching of HL and SL courses</td>
</tr>
<tr>
<td>-Providing appropriate timings for HL and SL courses including A block and H block HL classes and SL tutorials</td>
<td>2. Providing appropriate supports for students to improve academic achievement</td>
</tr>
<tr>
<td>-Reducing TOK scheduling through A block semestered approach</td>
<td>3. Providing more time to differentiate teaching approaches</td>
</tr>
<tr>
<td>-Building collaborative department meeting times into teacher schedules</td>
<td>4. Increasing collaboration time for teachers to grow their capacity to deliver the program</td>
</tr>
<tr>
<td>3. Updating all assessment policies for the Diploma Program:</td>
<td></td>
</tr>
<tr>
<td>-Aligning of the school grades with predicted grades and IB grades</td>
<td>1. Using data-guided decision making process to reach conclusions about program practices</td>
</tr>
<tr>
<td>-Moderating IB world exams scripts</td>
<td>2. Centralizing all Diploma Program practices</td>
</tr>
<tr>
<td>-Updating all achievement charts</td>
<td>3. Introducing rigorous assessment</td>
</tr>
<tr>
<td>-Regulating homework, retake, and formative assessment procedures</td>
<td>4. Increasing teacher confidence in grading and reporting practices</td>
</tr>
<tr>
<td>-Establishing use of GradeBook for recording marks</td>
<td>5. Providing more communication about assessment to parents and students</td>
</tr>
<tr>
<td>-Making assessments more rigorous</td>
<td>6. Spreading out internal assessments over the course of the two years of the program to reduce stress on students</td>
</tr>
<tr>
<td>-Removing non-IB, “fluff” assessments</td>
<td>7. Developing an understanding of the whole program not just subject specific matters</td>
</tr>
<tr>
<td>-Building skills through formative assessment to increase performance on summative tasks</td>
<td>8. Building more teacher accountability into the program</td>
</tr>
<tr>
<td>-Moderating all summative assessments</td>
<td></td>
</tr>
<tr>
<td>-Establishing firm due dates calendar</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Bill’s perceptions of structural and procedural changes as a result of the evaluation.

In light of these observations, Bill summarizes his experiences with the evaluation process thusly, “I would say we were authentic and critical within our reflection and it is been great to
see that what we criticize ourselves on or what we said we wanted to do better, we are doing better, if not trying” (Bill interview, March 15, 2015).

4.6.2.2 The Pushback

Bill personally finds the process of the evaluation rewarding and the effects of the evaluation positively transformative. In response to some comments expressed by other staff members, Bill muses:

My professional opinion is [that teachers are negative about the changes] because they did not take the time to ask the questions, to educate themselves on the literature that was provided as to why we are doing this. I think it is easy to jump to conclusions when you are ignorant because you have chosen not to educate yourself about a process. So you did not want to read the emails or the explanation of the process of what was happening and you chose not to do that. You can always still ask the questions, but these teachers made the conscious choice to not do that which allowed them to come up with their own conclusion. And I also would say – I mean maybe this is mean – I would say that person is self-centered because they made the automatic assumption that this program, this requirement that the IB does for us, makes us, as a school do, is about an individual and is not about the students. So they are also then assuming that the IB has no interest in the benefit of our students and that they would make us go through this process for something other than the benefit of our students. The ignorance in which this person lives and their choice in not wanting to educate themselves would be the reason for their negativity. I think that it is absurd. (Bill interview, March 15, 2015)

Bill’s words show that he believes that the evaluation was about the students and how the program could serve them better not about the individuals leading the accreditation efforts.

4.6.2.2.1 Lack of Time

Bill acknowledges that some teachers felt that they perhaps did not volunteer to participate in the evaluation committee because they were not given release time for this activity. Since their priority was mainly their classrooms and preparation, they did not want to spend the time on the evaluation. In addition, the school has a high turnover rate, so some teachers might feel that they did not have enough experience about all parts of the program to contribute. Bill finds this a true yet ironic state of affairs at the school. He relates that as a relatively new teacher to the program, he had only taught the DP for one year when he was approached to participate in the evaluation team, he especially found the accreditation
process useful to build and increase his understanding of the program. Bill wonders if teachers could provide more narrative feedback as the evaluation process is unfolding to combat the lack of time argument against participating in the mandated self-study preparation more actively.

### 4.6.2.2.2 Personal Feelings

Bill admits that changing teaching practices is hard. He states:

It is easy to do what you’ve always done, especially when it is wrong, right? Yeah. I think [some teachers are] living in their own bubble and do not want to be challenged. …I think that those people that are resistant to it, like change can be hard but you have to understand why you are doing it. And these teachers clearly show zero interest in even working to understand the why of the process which is ironic because that is where we started with the Extended Essay as to why we do this. And I guess the hard part for me too – and maybe this an irrelevant rant too – like to me, we are in education for kids. We are. And this person does not sound like they care about the kids at all. They sound like they care about themselves and they care about getting pissed off at other adults. And that is – they have forgotten the kids. To me in no way whatsoever did they say, “This process is bad because it is bad for kids.” They said, “This process is bad because I do not like another adult.” Like it is not about the kids at all. That is messed up. (Bill interview, March 15, 2015)

Bill states that criticism is healthy if it is framed within the aim of the evaluation which, in his opinion, about improvement. Bill concludes:

Sometimes pushback is healthy. We need that. You should question is this, especially in education, is this best for kids? Pushback is healthy because hopefully it makes the process the best that it can be. But I guess I would challenge that saying the pushback here had nothing to do with the kids whatsoever. It had to do with someone’s personal – not vendetta – or just like personal thoughts about another adult in the building not about kids. (Bill interview, March 15, 2015)

Bill advocates for always keeping the students’ needs in the centre of evaluation and not personal feelings about the leadership.

### 4.6.3 Future of Evaluation

After going through the evaluation process, Bill understands the dual nature of accreditation: it is a requirement but also a service. Bill explains:

[Evaluation] is like tough love. Like, we love ya. This isn’t easy, but we are doing this because we care about you. You know, I think the IB does care. It is a business. But I do think they care about the schools to which they give their stamp of approval. They
are interested in maintaining their reputation as well, and their brand. I think because of that, it is a requirement, but it is this tough love. We are going to make you do this because we want it to be a transformative experience for you and the staff to serve the IB program and to serve the students. So even though we are forced, it is for our own good. (Bill interview, April 22, 2015)

This section explores Bill’s opinions on how the school and the IB can improve future evaluations to ensure that stakeholders, including the IB, conceptualize accreditation as a service.

4.6.3.1 Future of Evaluation at the School

Bill argues that having a self-driven action plan will enable the school to continue on a path of continuous improvement. He believes that any practices that can make the lives of teachers easier to engage in evaluation will ensure that the legacy of self-reflection and learning culture will endure even within the context of a highly transient teaching staff.

4.6.3.1.1 Continuous Change Model

Bill reiterates that the evaluation is all about supporting teachers so they, in turn, can support students. The action plans generated for the DP and the whole school are based on the recommendations of the staff and the accrediting bodies provide the guidance for the future to make these plans a reality. Bill argues that making the action plans public to all stakeholders builds more transparency and accountability in the school, which, hopefully, will translate into thoughtful action by the teachers. Bill advocates for a continuous process of evaluation of practices in order to ensure that the action plans do not remain just words on paper. He says:

On-going evaluation in the lead-up to when the IB evaluation has to happen formally again [is ideal]. The school has gone through the process once, and the coordinators have gone through the process, some staff members but definitely the coordinator knows what the IB will be looking for years in advance and all. Ideally, I think it’d be pretty awesome to track the data and get feedback over a multi-year period that could then be applied to the evaluation process, when it needs to be done formally. I think that would be cool to get on-going data, especially with our institution and the turnover rate we have. So [the evaluation] wouldn’t be a snapshot of who was [at the school] at that point in time, but if someone maybe had been here for a couple of
years, that they can give insight – but if they are not here during the evaluation year, that their voice is still heard. (Bill interview, April 22, 2015)

This process would also ensure that staff changes at the school would not be as jarring for the accreditation teams since the model of continuous evaluation would become a part of the school’s institutional memory and increase organizational capacity.

Bill muses what the reactions of the teachers will be the next time they are faced with the evaluation:

I wonder if teachers would be like, no problem. Here we go again. We have done this. We can easily do this again. Some of the leadership that has been really integral in those evaluations will still be here, and that will be nice that they can lead us through. Some will be gone, but I think that they are the legacy, and the documentation that they left might actually leave a bit of an air of confidence among staff, amongst the eye rolls. (Bill interview, April 22, 2015)

Bill’s words show that even under ideal conditions he would expect some resistance to the mandated accreditation requirements but he is hopeful that the majority of the staff could take up the mantle of evaluation with ease.

4.6.3.1.2 Synchronization of Evaluations

Bill is supportive of the school undergoing continuum evaluation for all three IB programs during the 2016-2017 academic year. He is especially excited by the prospect of the school visit. He states:

[The continuum school visit] is a great opportunity, to be honest. It adds a level of accountability knowing that a real individual from the larger organization is going to physically be within the building. Because you can’t hide at that point in time. For me, it is cool, because it is accountability, but it is meaningful accountability. At that point in time, we should be able to showcase the good things that we are doing as opposed to, like, a gotcha. It’d be nice to be able to present what we are doing, and we know when they are coming, versus it may feel different than like checking a box, checking a box, checking a box, and sending the paperwork off. (Bill interview, April 22, 2015)

Bill’s words illustrate the journey from compliance in self-study documents and inspection visits to showcasing improvement as the focus of the accreditation.
4.6.3.2 Future of Evaluation for the IB

In this section, Bill’s suggestions for the IB and for the improvement of the DP evaluation are discussed. Moreover, Bill’s reactions to the IB proposed changes to the accreditation process are presented.

4.6.3.2.1 Suggestions for the IB

Bill encourages the IB to improve the services provided to member schools. For example, while Bill acknowledges that the report from the IB was appreciated, he posits that more personalized feedback to schools is needed. He also believes that school visits are an essential part of the evaluation process. He says:

I appreciate the feedback that we got from the IB, and it is nice that they can give specific feedback about what we are doing well or things that we need to address, but I do not know. There’s always that question, I guess, in your mind, is the person sitting behind a computer, are they copying and pasting those [comments], or are they taking five minutes to scan over [the self-study]? Because we put in a lot of work on this end within the school. I trust the IB and what they have to do but that face-to-face contact and knowing that that person has experienced the school culture and sat down with the stakeholders involved to get a true sense [of the program] to give real feedback, it feels to be more of an authentic process. Understandably not doable on a totally regular basis with the hundreds, if not thousands of schools around the world, but I think it is good to allow schools that opportunity. (Bill interview, April 22, 2015)

Moreover, Bill suggests that the IB should allow schools to report data for the evaluation in different ways. He explains:

There’s not much choice in how we present the information [for the self-study]. Clearly, the IB has standards and benchmarks that they are judging us on and that is fine. But there’s one mould and one way to go about approaching it. We gather data, there’s a lot of typing and checking of boxes, and the boxes were still limited within themselves. As an institution, we were limited in how we present the work that we are doing. And maybe we feel that we are doing something well, but when we look at the scale that is presented to us it does not come across the same way. Or maybe there’s not the opportunity for us to necessarily showcase, using a variety of media, the work that we are doing here at all. It is only via someone’s articulation through writing. And that could be a pain. (Bill interview, April 22, 2015)

Bill grapples with the idea that some practices are difficult to accurately capture in a pen and paper format, which might present an inaccurate picture of the program at the school.
Furthermore, he points out the limitation of a four-level rating system, which can also cause problems. He identifies a major flaw in the process of producing the self-study in the only prescribed way:

If you have someone who’s responsible for [writing the self-study] and they are not the best writer, or I did not even think about schools around the world, in a lot of IB schools you have people that aren’t first language English speakers or native speakers, and that could be a challenge for them if they are tasked with communicating the good work, that could be miscommunicated easily. So I guess the lack of choice in how information [for the self-study] is presented is a concern for me. (Bill interview, April 22, 2015)

Bill advocates for culturally sensitive evaluations that allow different modalities to present the school’s data for IB consideration.

Finally, Bill explores the role of IB Educators, school visiting team members, workshop leaders and readers who serve as external evaluators for the IB. Bill admits:

My experience with the IB is primarily through trainers [IB workshop leaders]. But many of the trainers that you come across have been within the system for a long time, and it can be a mixed bag of experienced educators in the system for a long time who are very up and current with what’s happening in education now, but I definitely have had IB trainers that have not been in the education system for a long time, and are so out of touch as to what’s actually happening in school now. They remember what was happening in school when they were at a great international school 20 years ago. But what I find is that them being at a top tier school does not mean that they are top tier educators. It does not keep them relevant anymore. There are people at their schools that are keeping those schools afloat and world-class, but just because they’ve got 20 years of experience, and they happen to be at a top tier school, does not qualify them by default. I think that upscaling and just additional support for [IB Educators] of course is needed. It should be a requirement. Honestly, I would say if these people are evaluators, they should, on a yearly basis, be evaluated themselves. Do they know what they should be looking for? I mean, they should be given that support as they go out into the field that, hey, here are trends in education, or here’s what we are looking for, as the IB in the year 2015, not IB in 2010 or 2005. (Bill interview, April 22, 2015)

Bill further presents the argument that IB external evaluators should have IB experience. He says, “If you know program evaluation, that is good. But if you do not really know what it means to be an IB teacher in an IB program, how can you really share real knowledge? It is not going to be as meaningful for the actual practitioners” (Bill interview,
April 22, 2015). Bill shows again that the context of the school is really important for teachers and the evaluation process.

### 4.6.3.2.2 Feedback on IB Proposed Changes

Bill appreciates the streamlined processes proposed by the synchronization of the evaluations by different accrediting bodies. He states:

You are killing two birds with one stone. That is awesome. [Evaluations] are time-consuming things, and they are important to do, but at the same time, they should not be all-consuming for the individuals or the leadership that is tasked with doing it, because in the end, those individuals, their job is still to function within the school and to implement programs that are still best for kids. I understand that reflectiveness and evaluations are good for school systems, which should be still good for kids, but we also have got to live in the day-to-day. So if a process can be streamlined, that is awesome. Like that is phenomenal. (Bill interview, April 22, 2015)

Bill is also keen to use technology to simply processes. He is excited by the prospect of having a collaborative digital platform for the self-study. According to Bill the previous static system had so many illogical steps. A better system of a living document that could be used by many people simultaneously would facilitate more collaboration and transparency. Moreover, Bill is intrigued by utilizing virtual school visits. He says:

It is a novel idea. I think I still would err towards having a real live person there, but I would say, this is a wonderful alternative, for many of the reasons, whether it’d be a safety issue, or whether it’d be cost prohibitive. It is expensive to fly someone somewhere; this still provides someone a live insight into what’s happening in the school. And again, you can’t hide. It is like when we say, like, oh man, I would love to be a fly on the wall, this, to me, is almost like you are a fly on the wall, which is actually kind of interesting within itself. It may be less intrusive. It could be awkward and there always could probably be technological issues that could arise within this, but to me, it seems like a brilliant idea. It allows real conversations to happen in real time, as well. (Bill interview, April 22, 2015)

Finally, Bill reiterates the use of different modalities to capture data throughout the year to prove that certain practices are happening at the school. He imagines that these video recordings or presentations can be used for more than just evidence of compliance for the self-study. Bill believes that these snippets of school life can be used for instructional coaching opportunities at the school and, if shared by the IB, on the global level. Bill hopes
that teachers would take the initiative to improve their practices and that would benefit their students. Bill argues that this could lead to the development of a more engaged global learning community.

4.7 Summary

The case of the evaluation of the Diploma Program provides a rich description of the participants’ experiences of the evaluation of the school’s program, the effects of the evaluation on the program as perceived by the participants, and suggestions for the future of evaluation both at the school and at IB schools around the globe. The majority of participants in the study find the evaluation of the Diploma Program transformative.

Lewis believes that the program better serves students now and he maintains that teachers owe it to their students to constantly look for the betterment of practices. Internationally, he states that educators owe it to each other to improve each other’s schools and vicariously enjoy successes. Joanne celebrates the process for synchronizing the efforts of stakeholders to improve the program and move beyond compliance. Bill claims that, if the self-examination of an accreditation is done meaningfully, it can transform the school. He cautions teachers to keep personal feelings about other adults in the school out of the evaluation so that they can focus on what is important: student achievement and well-being. Overall, the majority of respondents report a positive experience that made the program stronger, developed them as teachers and leaders, and enabled them to support their students better.
Chapter 5: Discussion and Interpretation of Findings

5.1 Introduction

In this chapter, I am going to answer my research questions and explore how the case of the school answers those questions by describing some common themes that emerged across the survey responses and the interviews given by the three participants. Then I summarize the main findings and present some recommendations and suggestions for further research.

5.2 The Research Questions

My thesis focuses on the research questions posed in Chapter One. Those questions were:

1. How do the participants describe their experiences of the IB Diploma Program evaluation?
2. What changes are implemented in the program as a result of completing the evaluation? How are these changes perceived by the participants?
3. What are the participants’ perspectives on the future of program evaluation in the Diploma Program at the school level and at the IB level?

I will examine each research question based on the findings from the case study of the International Baccalaureate Diploma Program evaluation of a school in the Middle East.

5.3 Question 1: How do the participants describe their experiences of the IB Diploma Program evaluation?

There have been many changes to the professional lives of teachers in the age of accountability (Cousins & Earl, 1995; Hargreaves, 2005; Hargreaves & Shirley, 2011). Participatory evaluations have become increasingly more popular to engage practitioners in the process of accreditation of their programs (O’Sullivan, 2004; Rodriguez-Campos, 2012). For my study, I used the case of one school going through the mandated evaluation of their
IB Diploma Program. This evaluation requires all stakeholders to collaborate on the preparation of the self-study which is then submitted for external evaluation. The survey respondents and the three interviewed participants provided a rich description of their experiences of the evaluation and their perception of the evaluation activities.

From the document review, the survey responses and the participant interviews, we see that the unique circumstances of the school including the timing of the evaluation and the size and complexity of the program, the personal characteristics of the stakeholders and the relationships between the participants of the evaluations all provide the background for the evaluation process of a school’s program. Moreover, the pedagogical leadership of the program and the trust that the staff and parents have in said leadership can also determine their level of involvement in the accreditation activities.

The more personally and intrinsically motivated teachers are to contribute to the evaluation of the program, the more likely they will positively perceive the changes implemented as a result of the evaluation. These findings are consistent with previous research conducted by Bryk et al. (2010) that suggests that people will take to change if it is intrinsically interesting, it is pursued in a non-judgmental culture, they have some say in its evolution, and they feel ownership of the process.

5.3.1 The Context of the Evaluation

Hallinger and Heck (2010) suggest that the academic structures, school norms, and on-going organizational processes create the environment for any improvement efforts such as a program’s evaluation. Therefore, school improvement leadership is highly contextualized. Cousins, Whitmore, and Shulha (2012) further argue that the context of the program is the constant lens through which all program decisions, including evaluations, have to be examined. Understanding the context of the program which is being evaluated is central to what teachers do, why we do what we do and how, and how and why we propose to
change what we do. I use these ideas to first examine the specific structural and personal contexts that surrounded the evaluation of the Diploma Program at the school during the 2013-2014 academic year.

5.3.1.1 School Context

The respondents to the survey and the interview participants continuously reference the specific characteristics of the Diploma Program. Indeed, the IB does not require every aspect of the program to be actualized in exactly the same way for every single school. The external evaluation performed by the DP be can be conceptualized as the IB acting like a good shepherd: ensuring that all schools stay within certain boundaries to maintain authorization yet allowing for necessary variations among the large number of schools with unique institutional characteristics.

The DP many times coexists with other programs in the school that have their own specific requirements. For example, the High School offers credits in Grades 9 to 12 that are accredited by the Middle State Association of Schools and Colleges. The IB DP courses need to also fulfill the requirements for the credits granted towards the American High School Diploma issued by the school.

As a result, the specific way the DP is implemented in this context might be very different from how another school realizes the program according to their circumstances. The processes that foster the development of a learning culture that allows teachers to adapt their practices to the individual school’s specific needs then also have to be highly contextual. These findings are consistent with Hallinger and Heck’s (2010) conclusions that report that it is futile to try to implement a borrowed system of a high performance instead of developing a learning culture suitable for each individual school context. All participants agreed that the evaluation of the Diploma Program resulted in changes in the program. Most participants, including Lewis, Joanne, and Bill, viewed these changes and needed improvement for the
program very positively while one survey respondent lamented the alterations to the program in the specific setting of the school.

According to Cronbach (1980), schools should aim for good enough not perfect evaluations. Good enough evaluations provide the school’s staff with enough information to make decisions about the program rather than looking for a definite end for the accreditation of the program. What is good enough for the specific school is a judgment call and highly contextual. Lewis, Joanne and Bill agree that evaluation is never finished: it is an on-going conversation about teaching and learning practices with a conscious eye to improvement. This approach can seem never-ending and time-consuming. However, if an appropriate climate for change is created in the school then the evaluation of practices becomes a natural part of what teachers do.

The interview responses support Leithwood, Harris, and Hopkins’s (2008) assertions that successful leaders reduce the fear of change and create a non-judgemental climate using appropriate processes for the given context of the school. Bill advocates for sharing the reasons behind the evaluation with all stakeholders. He argues that, if the focus is on improvement and the process is efficiently organized, teachers can see the value of the evaluation activities and will honestly and conscientiously participate in the preparation of the self-study.

Within the specific context surrounding the evaluation of the DP at the school, the timing for the IB evaluation following closely after the reaccreditation of the whole-school programs for the Middle States Association of Schools and Colleges seems significant for all respondents. For better or for worse, the teachers were engaged in two separate yet highly interconnected evaluation processes at roughly the same time. Lewis and Joanne agreed that serving on committees for the MSA evaluation helped them during the DP evaluation by
providing them with documents that could be repurposed from one evaluation to another and giving them experience with being engaged in an accreditation.

Bill was aware of the MSA evaluation efforts even if he only provided information in staff surveys. It can be argued that ‘evaluation was in the air’ at this time at the school, which created a receptive climate for implementing needed changes in the Diploma Program since a critical look at standards and practices had been performed on a larger scale for the MSA. This is congruent with the findings of MacBeath (2008), Caldwell (2008), and Marsh (2015), who conclude that self-evaluation of schools is an appropriate strategy to address each school’s unique set of needs, interests, aptitudes and aspirations to have a positive impact on student learning. Synchronization of evaluation efforts for external review can play a supportive role in self-assessment by using resources in an efficient way.

5.3.1.2 Stakeholder Context

Teachers who answered the survey and the three interview participants all have their unique characteristics as well which serve as the context for the evaluation of the program. They differ in age, nationality and cultural background, number of years in the classroom or in a leadership role, number of years spent at the school, familiarity with the Diploma Program, and previous experience with program evaluation. O’Sullivan (2004) makes the case for collaborative evaluation by pointing out that program stakeholders bring invaluable knowledge and insight to any evaluation activity and they need to be encouraged to participate in the process to the degree that they are able and willing. This is in line with the observations Lewis, Joanne, and Bill make regarding the level of involvement of stakeholders in the DP evaluation.

Lewis and Joanne indicate that, by virtue of their respective roles as leaders in the High School, they felt that they had a significant role in the evaluation yet still could make personal choices regarding the exact nature of their participation. Lewis acknowledges that
the Diploma Program Coordinator led the evaluation as it is required by the IB. However, he recognizes that, as the High School Principal, he was a part of all decision making and planning at all stages of the process. In her capacity as Head of the English Department, Joanne felt responsible for contributing to the evaluation as a way to support both her teachers and her students. In addition, she claims that, as a teacher with institutional knowledge of the school’s Diploma Program, she could provide historical data otherwise unavailable due to the high turnover rate of teachers and leaders.

Bill states that, although he was approached to be a part of the evaluation committee in preparation for his upcoming leadership role, he was still able to personalize his level of participation. He says that he picked parts of the evaluation activities towards which he had a special affinity. This way, he could really engage with the evaluation because he was personally motivated.

Some survey respondents did not feel that they were appropriately involved in the evaluation processes. Lewis and Bill attribute this to lack of experience with evaluations, while Joanne believes that some teachers new to the program and the school might prioritize preparing for classes over contributing to the evaluation efforts. Overall, most respondents believe that they were appropriately involved in the evaluation as befitting their own comfort level and preferences. This is important to note since O’Sullivan (2004) observes that, if the collaborative evaluation is responsive to the specific organizational needs of the institution, then the utilization of the evaluation can be increased through collaborative practices.

This brings about the question regarding the processes followed during the evaluation of the DP at the school: were the evaluation activities truly collaborative? As Cole (2013) and Kirtman (2013) assert, collaboration without an appropriately established culture of learning is overrated. Lewis, Joanne, and Bill all speak about the process in which they engaged as a truly collaborative exercise. Bill recalls the way that consensus was built during the rating of
the specific practices for the self-study; not all teachers had to totally agree with every item yet decisions about changing the program still could be reached. Many of the survey respondents also believe that a culture of learning to benefit the students in the DP at the school was present during the evaluation.

However, some survey respondents question the sheer amount of paperwork and the time to usefulness ratio of all of the external requirements posed by the IB in order to continue a school’s authorization. Datnow (2011) agrees that in the age of data-driven accountability, compiling and using data to guide instruction is time sensitive work. Moreover, one survey respondent had a very personal reaction to all evaluation activities based on a personal dislike of the Diploma Program Coordinator. Many of the comments made by this teacher were not within the realm of rational decision-making. This echoes the findings of Hargreaves (2005), which warn that it is dangerous to discount the emotive nature of teaching and pursue educational reform as a purely rational endeavour. The author finds that festering personal resentment can lead to poisoned committee work and unresolved grievances which was clearly the case for this particular teacher.

Furthermore, the unique make-up of an international school’s teaching team is also a contextual factor. Joanne and Bill suggest that common vocabulary is needed in order to meaningfully discuss educational terminology for teachers with different professional training background. What the term ‘collaboration’ means to a North American teacher might be very different from how an Arabic teacher conceptualizes it. Joanne indicates that some teachers might not feel that they can meaningfully contribute to the evaluation because of their limited English skills. She advocates for more culturally sensitive evaluations to capture all stakeholders’ data. Askew, Beverly, and Jay’s (2012) findings suggest the same. Karami Akkary (2014) further argues that educational change in the Arab world in particular is often driven by political agendas and is mainly dominated by the top-down approach. Therefore,
teachers who only have experiences with this type of school inspection will require additional training before being asked to participate in a collaborative evaluation process so they can confidently and meaningfully contribute to the evaluation of the program.

In addition to teaching staff, the IB also requires parents and students to participate in the preparation of the self-study. Lewis, Joanne and Bill show concern regarding the relatively low participation of parents in the evaluation. All three interview participants agreed that the opportunity for parents to contribute to the evaluation was extended by the program leadership. Lewis indicates that teachers with children in the program, like himself, had a voice in the self-study; however, he acknowledges that parents, especially of the local students, usually do not participate in school life. Lewis sees this as a function of the trust that culturally parents place in the professional teaching staff to make program decisions on their behalf to improve the education of their children.

Bill also mentions that the parents at the school are disengaged. He indicates that the only time parents really come to school is if their child is not getting the grades that the parents believe they should. He adds that parents might not feel confident about their knowledge about the program so they do not contribute to the evaluation. Joanne elaborates that the surveys for the evaluation were available in English only and many of the parents of local students do not have English as their first language. These findings further advocate for more culturally appropriate evaluations to include the school’s entire community.

5.3.2 Accountability and Self-Assessment

Most of the survey respondents and Lewis, Joanne, and Bill indicate that the self-study and required supporting documentation submitted to the IB in April 2014 honestly depict the practices used in the Diploma Program at the time of submission. Lewis reflects that the self-study was prepared collaboratively with stakeholders engaged in every step and not in isolation by the DP Coordinator, which ensures accountability within the school
community. Joanne shares her belief that teachers are intrinsically motivated to provide a better program for their students and as a result can make decisions about the program’s delivery based on self-assessment. Bill suggests that a well-organized process for self-assessment rooted in sound educational theory and practice yields the best results for the improvement of teaching and learning. These comments are in agreement with the type of internal accountability that MacBeath (2008) describes as part of internal evaluation.

Furthermore, the three interviewees acknowledge that they feel collective ownership of the self-study: in Lewis’s case in its entirety and Joanne’s and Bill’s cases the parts that they chose to work on. They all report that the findings of the evaluation are implemented in the program since they came out of a collaborative conversation on how to improve the program to better support their students. O’Sullivan (2004) similarly concludes that it is logical to assume that, if the teachers who are actually delivering the program are collaboratively involved in the evaluation of the program, their use and understanding of the findings should increase.

In addition to internal accountability, all three interviewed participants find that checks and balances from the IB are an important aspect of the evaluation of the program. Ultimately, they all acknowledge that the external accrediting bodies decide whether a program is authorized or not. Lewis, Joanne, and Bill elaborate that external evaluators also have a broader view of the program and more experience with how the DP is conceptualized in many different contexts. Consequently, they appreciate the feedback and expert ratings on the school’s practices that they receive from a large network of schools like the IB. This echoes Caldwell’s (2008) findings on how external review of self-evaluations can play a valuable role in accreditation. Marsh (2015) further encourages schools to value partnerships with local, national and international communities to expand their learning communities beyond the physical environment of the school. Sahlberg (2010) congruently promotes
intelligent accountability according to which schools cooperate and network with each other to collectively expand their knowledge base in the interest of offering the best possible programs for their students.

5.3.3 Pedagogical Leadership of the Evaluation

Most of the participants in the study report that they believe that there was competent pedagogical leadership of the Diploma Program at the time of the evaluation. One significant outlier response will be further explored in the section about the pitfalls of the evaluation process. Lewis views himself as a collaborative leader and states that leadership is not about him as a leader but about the team as a whole. Joanne says that the DP leadership is able to provide instructional leadership since they have sound knowledge of the DP requirements. Bill finds that the relationships that exist between leaders and teachers are built on sound educational theory and practice. These observations are in line with Telford’s (1996) definition of collaborative leadership which the author identifies as transformative for schools since leaders should be able to provide instructional leadership.

Research by Bryk et al. (2010) suggests that successful leaders are now viewed as the drivers for educational change. The interviewees express their perception that the DP leadership developed a good change process that shaped the collective ideas of the group while developing capacities in the group and ownership of the evaluation. These conditions then enabled the team to implement the changes suggested by the evaluation process.

5.3.4 Trust

Louis (2007) argues that relational trust is at the heart of school improvement. Teachers’ experiences with change are more influenced by relational trust in their leadership than institutional trust especially in a context with a highly transient international staff. This seems to be the case during the evaluation of the DP during which both the High School Principal and the DP Coordinator were new to the school along with many teachers.
delivering the program. Lewis perceived this as a positive aspect of the evaluation since neither he nor the DP Coordinator had any preconceived notions on how practices should work at the school.

Bill disregards the length of time at the school as a basis for successful pedagogical leadership. He says that he developed the trust with the new leaders as he witnessed their abilities during day-to-day interactions. He also mentions that he has trust in the school as an institution that graduates capable students as evidenced by their results on the world exams and in universities. However, one survey respondent strongly disagrees with these opinions and demonstrates how the perceived betrayal of trust between a teacher and a leader can undermine improvement efforts. As Tschannen-Moran (2014) asserts, without trust, there is no capacity building in the faculty to bring about the desired changes.

5.3.5 Organization of the Evaluation Activities

Leithwood, Harris, and Hopkins (2008) indicate that successful program leaders can articulate what they are doing and why they are doing it. They encourage many people to look at the problems at hand to develop the best solutions. The three interviewed participants agreed that this could only happen during the evaluation of the Diploma Program because of the systematic planning of the evaluation activities. The structures set out in the timeline of events, the transparent and accessible storage of all evaluation materials, and the scheduled meetings all contributed to the participants experiencing a climate at the school that channelled the evaluation efforts appropriately. This served as a basis for the entire evaluation process.

Lewis recollects that the evaluation of the Diploma Program unfolded according to a set plan. There was a clear external submission deadline that had to be met so the tasks were divided up and a clear delineation of responsibilities was put in place. Joanne even remarks that she was shocked at the end of the process since it went so efficiently and smoothly that
she did not feel overwhelmed. Bill similarly shares that the process of the accreditation was a productive experience for him because it was clear what the teachers were doing and why. All three interviewees report that having the access to all evaluation materials on Atlas increased the transparency and accessibility of the accreditation process to them.

Moreover, participants acknowledge that, keeping in mind the main reason behind the evaluation was to improve practices in order to deliver a better program for students, focused their efforts on the evaluation and did not allow meetings to become avenues for complaining about the school. The importance of focusing on student achievement as the driving force behind school leadership is consistent with the findings of Leithwood and Lewis (2012).

5.4 Question 2: What changes are implemented in the program as a result of completing the evaluation? How are these changes perceived by the participants?

All of the respondents to the survey and the three interviewees agree that the evaluation changed the Diploma Program at the school. How these changes are perceived by the participants differs significantly. Lewis, Joanne, Bill and the majority of the 16 staff members who completed the survey view the effect of the evaluation positively. They value the creation of a constantly evaluating professional learning climate that continues to fine-tune structures and procedures in the program to further improve the service provided to students and parents.

However, some members of the teaching staff question the time spent on the accreditation activities. One survey participant further claims that the changes heralded by the evaluation caused incalculable damage to the school and the students. The wide range of reactions to the evaluation efforts seems to agree with Hargreaves’s (2005) assertion that educational reform cannot discount the emotions that surround the act of teaching. Moreover, internal evaluators need to safeguard against self-evaluation bias and ensure that they are not risking “dramaturgical compliance” (Kemenade & Hardjono, 2010).
5.4.1 The Value of the Evaluation: The Progress

The majority of survey respondents and the three interview participants perceive the development of a collaborative professional learning climate at the school as the foundation for the progress that happened in the program as a direct effect of going through the evaluation. The teachers argue that the process of going through the evaluation itself resulted in desired changes since the staff and the leadership team had the autonomy to alter practices immediately after identifying needs and devising solutions for them. Participants also indicate that the reflection forced by the evaluation resulted in increased professional development for Diploma Program teachers and leaders.

These new understandings about the program, in turn, allowed staff to make significant structural and procedural adjustments to how the DP is delivered at the school. These findings are congruent with the work of Cole (2013) and Kirtman (2013), who advocate for the development of instructional practices that is deeply imbedded in learning cultures in order to improve the quality of professional learning at a school.

5.4.1.1 The Continuous Evaluation of Practices

Participants who have a positive attitude towards the evaluation process emphasize the continuation of evaluation after the final submission of the self-study and its related documentation to the IB for external review. These teachers admit that they were initially forced to reflect on their practices and look outside their own subject area and critically examine what the program was doing as a whole. They also acknowledge that changing traditional practices is hard and increases teacher workload. Moreover, the survey respondents and the three interview participants recognize that the standardization of practices and the correspondingly higher level of teacher accountability might appear as a form of increased control to some teachers.
However, the majority of staff believes that the continuous improvement of the program benefits the students and raises their academic results while reducing their stress. Lewis emphasizes the importance of communicating program specifics to all stakeholders at the school to make the school’s commitment to the practices of the program transparent and accessible by all. Joanne advocates for the long discussions of the evaluation becoming proceduralized. Bill claims that the program has become more proactive as opposed to reactive as a result of the evaluation and now can anticipate issues as opposed to just trying to solve problems as they crop up. As Bill points out, conversations can be the start of change, and further discussions can sustain an environment where change naturally happens.

Additionally, the three interviewees centre their comments about continuous improvement around their responsibility to the students to provide them with the best possible education. This aspiration for improvement is never-ending; thus, performance monitoring needs to be on-going as well. In order to support such continuous evaluation of practices, program staff need to be able to ascertain if the programs that they have put in place are effective or not between formal program evaluations. Joanne and Bill both talk about the new Extended Essay Supervisor Training program that was implemented at the school to address the low performance of students on the Extended Essay. The staff needed to find out if the new training program was more effective than the previous one. They wanted to see the results long before the next evaluation required by the IB and the full self-study or even its section on the Extended Essay was not appropriate for this task. The team worked on creating a logic model to evaluate the effectiveness of the new processes for the supervision of the Extended Essay. These actions are following Guskey’s (2002) suggestion that professional development programs need to be evaluated for their effectiveness and changes made as needed. Riffert (2005) advocates for the use of school-specific measurement tools in order to customize evaluations to the specific need of the school, which is what team
members needed in the case of the Extended Essay Supervisor Training Program. Finally, the logic model developed at the school is an actual application of Gargani’s (2013) suggestion that practitioners can learn from theorists. All of these efforts contributed to the on-going development of the Diploma Program.

5.4.1.2 The Professional Development of Staff

The conditions and processes as perceived by the interview participants that enabled program staff to create a professional learning climate at the school during the evaluation of the Diploma Program can be summarized in figure 1.

Cole (2013) argues that effective teacher development focuses on growing the core attributes of an effective teacher. Professional learning enhances teachers’ understanding of the program that they teach so they in turn can better serve their students. Bill echoes this sentiment when he states that the strength of the program comes from the strength of its teachers. The conversations about the DP standards and practices that were required by the IB in order to complete the self-study resulted in impactful reflections which uncovered what the teachers were already doing well and what the program was lacking.

According to the three interview participants, a systematic look at world exam results, predicted grades, and report card data also enabled some data-guided decisions to be made about professional development opportunities for the staff. Some of the teachers, like Joanne, were sent on IB external training to fill a void in the expertise of the staff as in the case of hitherto missing training for Extended Essay supervisors. Other teachers personally decided to sign up for external IB training. Bill reported that he took a course on Theory of Knowledge as a direct consequence of the reflection he utilized during the evaluation.

Additionally, the number of IB Subject Examiners increased from two before the
Evaluation to ten after the evaluation. More teachers were looking at the collective practices of IB teachers around the world to improve their own instruction. Moreover, the school also introduced new professional development opportunities for teachers: the moderation of the IB scripts and the Extended Essay Supervisor Training Program are just two examples. These professional learning opportunities increased the individual and collective teacher efficacy at the school and in turn improved the instruction of the Diploma Program at the school. These findings are consistent with conclusions reached by Chadwick, Thier, and Todd (2019) in their study of teacher participation in the IB Educator Network.

The majority of the participants agree that the Diploma Program leadership created an environment that was suitable for a learning culture to take root in. Byrk et al. (2010) similarly argue that it is essential that leaders foster the professional capacity of teachers to make their schools more successful. Similarly, Leithwood and Louis (2012) encourage
program leaders to focus on improving instruction through professional development and providing time for collaboration so teachers can meaningfully reflect on practices. Hargreaves (2012) and Sahlberg (2011) advocate for staff conditions in schools that empower them to implement changes in a program. Joanne and Bill both state that these desirable conditions existed during the evaluation of the Diploma Program. The interviewees report that they did not have to wait for the IB’s response to implement the proposed changes: when possible, the alterations in practices were immediately deployed. This autonomy encourages teachers and leaders to be agile and constantly respond to the changing needs of students by making changes in structures and procedures of the Diploma Program at the school.

5.4.1.3 Structures and Procedures to Increase Student Achievement

The three interview participants report numerous structural and procedural changes in the Diploma Program during and after the accreditation process of the Diploma Program. As High School Principal, Lewis focused more on the overall structure of the program and the accountability of teachers towards students and parents enhanced by the increased transparency and accessibility of the program. On the other hand, Joanne and Bill concentrated on the specific teaching and learning approaches that were altered after the careful examination of the program’s practices. The teachers reflect that standardization of practices across the whole program and the introduction of consistent application of procedures were the most significant benefits of the evaluation. A summary of the process of making structural and procedural changes in the Diploma Program as perceived by the interviewees is contained in figure 3.

Lewis, Joanne and Bill agree that the development of the professional learning climate illustrated in figure 2 allowed the growth of intelligent internal teacher accountability to thrive at the school as advocated by MacBeath (2008) and Sahlberg (2010). This climate fostered the process captured in figure 2. According to the participants’ perception of the
evaluation activities, as teachers started to better understand the program through reflection and training both at the school and at IB events, they were more able to choose and use data to make decisions about the program which were implemented in a timely manner.
At the same time, the clarity of expectations and the standardization of procedures across the program coupled with better instructional and assessment practices increased the communication teachers provided for students and parents in the form of descriptive feedback. This, in turn, increased the students’ achievement particularly on the world exams but also reduced the stress students’ experience, especially since their time management was better scaffolded due to the changes in scheduling. Lewis provides an account of his recollection of the changes made in the schedule and how he perceives the impact these changes had on student achievement and well-being.

This improvement validated the changes in the eyes of the interview participants. According to Lewis, Joanne, and Bill, the DP teachers saw better alignment between their own grades given to students and the IB grades students achieved. At the IB level, more of the teachers at the school felt confident enough to apply to become IB subject examiners and they had the necessary skills to qualify to mark external assessments and world examinations.

The interconnectedness of the two processes is exemplified by the changes made in the supervision of the Extended Essay. As described by Lewis, Joanne and Bill, the initial reflection on the supervision of the Extended Essay led to an analysis of IB results. This resulted in the external IB training of one supervisor who accessed the collective experience present at the training event. Subsequently, the timeline of the Extended Essay, the training of the supervisors, and the type of workshops the students received were all altered.

These procedures were captured in two handbooks: one for students and one for supervisors. The handbooks were made available to teachers, students, and parents to increase transparency and consistency in the program and to promote teacher and student accountability. Finally, in the spirit of constant evaluation, the team created a logic model to evaluate the effectiveness of the new Extended Essay training program. Throughout these events, leaders, teachers, and students sought and provided feedback and developed working
relationships. This process is consistent with the factors for school improvement Ko et al. (2012) identified in successful schools, the conditions suggested by Leithwood and Louis (2012) for successful school leadership, and the common factors that appear to support the relationship between school evaluation activities and school performance as summarized by Mason and Calnin (2018).

Establishing a causal relationship between the evaluation activities and student achievement using externally reported data as an indicator of change is beyond the scope of this study. However, the self-perceptions of change and effectiveness reported by Lewis, Joanne, and Bill are in line with evidence presented in the findings of Schildkamp et al. (2009), Ehren et al. (2014), Meuret and Morliax (2003), and Gustafsson et al. (2013), which suggests that there may be some capacity for self-evaluation to shape school teaching and learning practices.

5.4.2 The Pitfalls of the Evaluation: The Pushback

While the majority of participants value the evaluation efforts of the Diploma Program at the school, a small minority find the changes to the program damaging to the school and its students. These teachers believe that the process was not transparent nor was it inclusive. One particular survey respondent expresses very negative personal feelings towards the Diploma Program Coordinator.

The three interview participants suggest possible reasons for the negative emotions that some staff members anonymously shared in the survey. Lewis feels that change is hard in any setting and some pushback is to be expected any time there is a major overhaul of a program. Joanne opines that it is human nature to be negative and to try to change even practices that are beyond the control of the program’s leadership. Finally, Bill suggests that some pushback is healthy if it is questioning how a particular change is beneficial for the students but not a function of a specific teacher’s dislike of a certain leader. Hallinger and
Heck (2010) suggest that there will always be a certain number of resisters in any staff. Leaders need to have nuanced leadership skills in order to handle difficult situations in different contexts.

**5.4.2.1 Lack of Time**

Survey respondents and interviewees all suggested that lack of time devoted to the evaluation processes was a downside of an otherwise well-run effort. Lewis admits that as the principal, he should have considered providing release time to teachers to engage in the evaluation activities. Joanne and Bill both suggest that some teachers, especially ones new to the program or to the school, would prioritize teaching responsibilities over significant contribution to the evaluation process. This is consistent with the findings of Hargreaves (2005) who suggests that, in order for collaboration to work properly, sufficient time needs to be provided in the teachers’ timetable.

Joanne further claims that some teachers had not been employed at the school for very long and do not feel invested in the betterment of the program. However, the longer time these teachers spend at the school, the more likely they can become too entrenched in practices and develop an investment in the school. Program leadership has to maintain a delicate balance in order to engage all staff members in the evaluation proceedings.

**5.4.2.2 Personal Feelings**

Lewis, Joanne and Bill do not find it appropriate to let negative personal feelings stand in the way of needed reform of the program that benefits the students. However, Lewis and Joanne explain that the changes made in the program do require more accountability from teachers, which is more work for them. All three interview participants believe that, in the past, teachers could decide most of their practices even if they were not in the best interest of the students or went against IB requirements. The standardization of practices could hurt the professional pride of some teachers as suggested by Joanne. This can cause someone to
lash out against the Diploma Program Coordinator who was the leader of the evaluation.

Tschannen-Moran’s (2014) research on the control paradox reaches similar conclusions. Additionally, Louis (2007) warns leaders that any perceived betrayal of trust can undermine improvement efforts. Finally, Hargreaves (2005) concludes that any change, even change for the better, will always be accompanied by drawbacks and discomforts.

**5.5 Question 3: What are the participants’ perceptions on the future of the Diploma Program at the school level and at the IB level?**

The majority of survey respondents and the three interviewed participants agree that evaluation can be a transformative experience for the school if it results in a genuinely reflective teaching and learning environment that creates a climate for the meaningful professional development of teachers. According to the participants of the case study, the future of program evaluation lies with developing processes that focus on the creation of optimal conditions for this culture of learning and improvement through internal and external feedback while, at the same time, meeting the compliance requirements for the accrediting bodies. This finding is consistent with Cole’s (2013) and Kirtman’s (2013) call to action to establish a culture of collaborative professionalism that supports teacher efficacy and student achievement for the betterment of society.

**5.5.1 Future of Evaluation at the School**

Lewis, Joanne and Bill agree that any practice that makes it easier for teachers to engage in evaluation processes will lead to better actualization of the action plan that is a mandated part of the evaluation. Since the action plan contains all of the recommendations for improvement and step-by-step plans for implementation, it is an essential part of the new planning ethos at the school.

The interviewees see the evaluation of the Diploma Program during the 2013-2014 academic year as a model for subsequent accreditation efforts and a significant factor in the
institutional memory and organizational capacity of the school. They advocate for the establishment of continuous evaluation in the form of yearly reviews of the documents produced for the IB self-study, the examination of year-to-year data for comparison, and a change in how the formal evaluation is treated at the school. Instead of looking at mandated evaluations as a stand-alone, distinct moment, the participants envision it as a natural part of what teachers do at the school since all of the requirements are met through the yearly review and now they just need to be externally communicated to the IB. This is congruent with the findings of MacBeath (2008), who identifies internal accountability as the ideal condition for schools to respond to external measurements of performance. The interplay between internal and external review proposed by the participants is also reminiscent of Caldwell’s (2008) assertions regarding self-managing schools which states that once self-evaluation takes root, external review plays a supporting role.

The participants welcome the opportunity to engage in a continuum evaluation for the three IB programs at the school since it will streamline processes and reduce the redundancy of preparing the same information over and over again based on different evaluation timelines. This change will enable stakeholders to use time more efficiently and might encourage more participation in the accreditation efforts.

5.5.2 Future of Evaluation for the IB

According to Glazer and Peurach (2013), large-scale networks of schools like the International Baccalaureate can be powerful agents for school reform. Hill (2012) further argues that the Diploma Program can be a vehicle for school improvement by helping to inject rigor into the curriculum. Indeed, many schools introduce the DP hoping for such improvements. However, the participants find that many of the evaluation practices used by the IB need updating and upscaling to reach the true potential of what it means to be a part of this network.
5.5.2.1 Suggestions for the IB

Lewis, Joanne, and Bill engaged in a future perfect exercise to truly reflect on what an IB school would need from the organization to benefit fully from the long and arduous process of the evaluation. A summary of suggestions for the improvement of the evaluation of the Diploma Program at the IB level is presented in table 5.

<table>
<thead>
<tr>
<th>Suggestion for Improvement</th>
<th>Participants’ Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reimagine the relationship between internal and external evaluation requirements.</td>
<td>1(a). Collect yearly demographic and statistical data about each school to allow for year-on-year comparison and regular check in. 1(b). Focus on specific school-selected practices in depth every five years.</td>
</tr>
<tr>
<td>2. Provide differentiated evaluation services for schools.</td>
<td>2(a). Prepare personalized and in-depth feedback for schools, which is differentiated based on whether school visit is part of the DP evaluation. 2(b). Allow different modalities for the submission of the self-study, for example presentations or video submissions. 2(c). Acknowledge that many IB schools around the world have staff whose first language is not one of the three official languages of the IB (English, Spanish, French) by making the evaluation culturally sensitive.</td>
</tr>
<tr>
<td>3. Increase opportunities for teachers to become school visiting team members (SVTM).</td>
<td>3(a). Refresh the SVTM pool to allow for visitors who have recent experiences in different school contexts. 3(b). Allow teachers to participate in school visits as observers to increase their professional knowledge of evaluation. 3(c). Share school self-studies, presentations, video submissions, and virtual evaluations globally with schools for training purposes.</td>
</tr>
<tr>
<td>4. Celebrate successes.</td>
<td>4. Provide an easily accessible and IB-moderated platform for schools to share best practices and IB feedback.</td>
</tr>
</tbody>
</table>

Table 5: Summary of participant suggestions to the IB.

Overall, the three interviewed participants agree that teachers need in-depth, personalized feedback from the IB through an up-to-date, differentiated evaluation process. They also yearn for the development of a global community of learners to share the burdens and celebrate the successes of the accreditation process. These findings are consistent with the Heads’ Council Brief (2017) that explains the need for providing more contextual
services for IB member schools in order to maintain the high expectations associated with IB brand.

5.5.2.2 Feedback on IB Proposed Changes

Lewis, Joanne, and Bill remained mainly optimistic and positive about the proposed changes outlined in the presentation by Albright, Atkinson, and Bender (2014). The participants reiterate their support for any practices that streamline the processes required, reduce redundancy, aid collaboration and transparency, involve all stakeholders in a culturally sensitive way, avoid wasting time, address the particular context of the program specifically, provide descriptive and in-depth feedback including suggestions for further improvement, showcase best practices, celebrate successes, share program improvements on a global scale within the network of schools, reduce cost, and above all have the students’ best interest at heart.

These observations echo findings by Resnik (2012), who suggests that the IB needs to make improvements in order to maintain its reputation as a gold standard in international education and provide personalized, inclusive and affordable services to its members. The same conclusion is reached by Cech (2007), who claims that more flexibility in the Diploma Program is needed in order to serve students better. The IB seems to be listening to this call to action to bring the evaluation of its programs into the 21st century (IBO, 2017).

5.6 Major Findings

This study investigated how the participants describe their experiences of a mandated accreditation process at an IB Diploma Program school. The effects of the evaluation on the school’s practices and the future of collaborative evaluation at the school and the larger network of schools are also examined. There are five major findings for this study relating to the use of collaborative evaluation for mandated accreditation:
1. Participants in this study find that their perceptions of completing the evaluation are influenced by the context of the specific school, their own experience with the given program, the personal relationships that exist between the leadership team and the teachers and among teaching colleagues, and the larger community of stakeholders situated in a specific culture. Participants who report positive involvement with the evaluation processes believe that there is systematic organization of the activities related to the accreditation. They also state that the leadership team creates a suitable climate for internal evaluation as a learning opportunity to counteract the negative connotations of external verification of the school’s self-assessment.

These participants view the evaluation process as a necessary reflection to bring about changes in the program to increase teacher effectiveness and student achievement. For example, Lewis, Joanne, and Bill all state that, as result of the systematic organization of the evaluation activities, they feel that it is time well-spent to examine the school’s practices on a regular basis.

2. The freedom to change practices and determine what is needed for the given school is crucial in how participants view the impact of an evaluation. Participants view good accountability as a function of reliable data that are collected during the evaluation process which are then used as a strategy for school improvement. These practices require a non-judgmental environment that depends on wide-spread transparency and accessibility to all collected data. The more control the teachers feel regarding their own participation in the evaluation, the more positive view of the evaluation they have. The more control over the process that an individual teacher attributes to the leadership team of the program, the less they are likely to have productive engagement with the accreditation process.
Teachers express a desire to keep the improvement of the school’s practices as the priority of the evaluation and shift away from an emphasis on compliance with the accrediting bodies’ external standards. Participants also stress the importance of involving the entire school community in the preparation of the evaluation documents through culturally appropriate communication with stakeholders. Joanne and Bill both voice the concern that parents of DP students at the school are difficult to engage in school improvement activities despite the efforts of the leadership team. Lewis, as a father with a child in the Diploma Program, also wishes that more parents and students participated in evaluation activities.

The particular society in which the school is located might not provide a context where parents and students feel comfortable to question teaching practices. Furthermore, participants mention lack of time and the high turnover rate of educators as possible reasons why in this particular school setting even many teachers choose to only superficially engage with the evaluation of the Diploma Program. As a result, participants believe that the accountability that accrediting bodies expect from schools has to be in line with the context of the given institution.

3. Respondents’ perceptions of the intentions and abilities of the internal evaluation leaders may have some effect on how they view their evaluation experiences. Competent pedagogical leadership of a program during the evaluation process is essential for creating a climate for change. Principals and program leaders need to engage in appropriate professional development, firstly, to be able to establish the necessary environment for change, and, secondly, to be able to set up the evaluation as a vehicle for school improvement leading to higher teacher efficacy and better student achievement and well-being which should be at the heart of all school reform.
Internal evaluators need to be prepared for different responses from stakeholders and apply sensitivity when engaged in the evaluation process.

Leaders cannot discount the importance of relationship and trust building that serves as a foundation of collaborative cultures. As such, the emotive nature of the teaching profession has to be taken into account as leaders organize an evaluation. It has to be acknowledged that change is hard and it takes time to appropriately move a school forward. Leaders need to learn together with their teams how to change their particular school. This requires the development of a nuanced leadership repertoire as opposed to a specific leadership style. This way program leadership can be agile and adaptable to the unique circumstances at the given school engaged in evaluation. Reaching out to networks of schools going through the same evaluation processes can help internal evaluators develop better practices to get the most out of the evaluation requirements.

4. Participants express the desire to determine what aspect of their program is most useful for them to evaluate, when to gather what data, and how to report their findings to their networks. They argue that differentiation of evaluations helps to address the idiosyncratic nature of their school’s context. The current one-size-fits-all evaluation models do not meet the individual needs of the very different schools that make up large school networks like the International Baccalaureate. Appropriate training enables program leadership members to take the necessary steps towards more differentiated and individualized evaluation models.

5. Participants argue that self-assessment is a regular function of continuous performance monitoring and a natural part of the learning climate where professional conversations lead to change. According to Lewis, Joanne, and Bill, the true value of an evaluation process for a school crystalizes when members of the school community
recognize evaluation as a way to develop their own continuous school improvement model through on-going performance monitoring. They view this as an element of a healthy and productive learning culture rather than an isolated event only brought on by the periodic requirements of mandated accreditation. The school then can take the conclusions of its self-study and its action plan can go from words on paper to thoughtful action. Program leaders who tap into the motivation and passion of the teachers at the school can connect the evaluation to an improvement in their classroom practices which will increase student achievement. This way the school can avoid the pitfall of dramaturgical compliance and only attempt to dress up its program for external evaluators.

5.7 Implications of the Study

This study describes the recollections of the participants’ experiences while they were completing a collaborative evaluation of a school’s program, the effects that the evaluation can have on the program and suggests some possible avenues for the future of program evaluation. The context of the evaluation and the individual characteristics of the teachers can have an effect on how they perceive the collaboration required for the evaluation and how much value they place on engaging in these mandated processes. Indeed, one of the important implications of this study lies in its methodology. Gaining access to a private, proprietary institution in the Middle East in order to closely examine a high-stakes process such as the evaluation of one of its programs provided a unique context to explore. However, this case study is only a snapshot of the school’s program evaluation practices as seen through the lens of one particular set of participants at one particular time. It needs to be noted that only North American teachers volunteered to be interview participants to share their recollections about their experiences almost a year after the conclusion of the evaluation of the Diploma Program.
Research suggests that successful schools and their leaders take into account the emotive nature of teaching and actively develop trust within their community of stakeholders. School leadership needs to establish a learning climate that supports the collaborative professional development of teachers (Bryk et al., 2010). Moreover, teachers need to be provided with the freedom and time to improve their practices and implement changes in their classrooms (Hargreaves, 2005; Louis, 2007; Tschannen-Moran, 2014). This study shows that the complexity of a school’s context can have a noticeable difference in how stakeholders are involved in the accreditation process and how the effects of the evaluation are perceived. Intelligent accountability has been promoted as a way to ensure that learning cultures are at the core of school improvement (Sahlberg, 2010).

From this study, we see that pedagogical leadership of the program being evaluated is an important factor in how teachers perceive the value of the evaluation. Joanne and Bill believe that the school’s Diploma Program is managed by competent leaders. This plays a role in how positively they view the changes introduced as a result of the evaluation. The negative personal feelings of one survey respondent towards the DP leadership have the exact opposite result. Research suggests that principals and program coordinators should become change agents and lead learners, not just mere managers (Bryk et al., 2010).

The current case study mainly reports on the internal processes of evaluation undertaken by the school. The school had very limited interaction with the external evaluators for the self-study completed in the 2013-2014 academic year since an evaluation visit did not follow the submission of the required documents. The DP leadership and teachers only received written feedback from the IB external evaluators in the form of a report sent to the school in May 2014. Research shows that large scale reform models like the International Baccalaureate can be successful in bringing about school improvement (Glazer & Peurach, 2013). This study only focused on the actual evaluation process at one school, which had a
long history of offering the Diploma Program. It would be interesting to discover the
different structures that support schools before, during, and after accreditation and subsequent
evaluation. The IB continuously expends the repertoire of services that are available to its
member schools. Some recent examples include the School Enhancement Services, the IB
Leadership Certificates, the IB World Schools Department and various post-authorization
services offered to schools. However, the effectiveness and appropriateness of these
initiatives are yet to be studied in detail.

5.8 Suggestions for Further Research

Although there is much to be learned from the case of the school evaluating its
Diploma Program, further research is needed to develop a better understanding of how
mandated accreditation can be used as a vehicle for school improvement. More research is
needed on the relationship between mandated accreditation processes, which predominantly
focus on compliance, and self-assessment that is free of judgment and allows for failure. It is,
therefore, worthwhile to explore how schools and accreditation bodies balance the dual needs
of ensuring compliance with set standards and practices with the need to authentically
improve programs. Furthermore, how can large networks of schools differentiate evaluation
measurement tools while still maintaining overall expectations? In order to make evaluations
culturally and contextually sensitive, school specific measurement tools have been proposed
for schools to monitor their own progress (Riffert, 2005). An exploration of how logic models
created by practitioners on a larger scale could shed more light on this issue.

It is also advisable to further discover the ways in which we view the role of internal
evaluators. If the self-studies leaders are required to complete only focus on compliance, how
can the DP Coordinator’s role change from manager to leader? Additionally, how do we train
internal evaluators as pedagogical leaders who view themselves as change agents? And what
type of professional development of teachers is needed to ensure that they are able to develop the school’s vision and school improvement plan that is using data reliably?

Moreover, it would be enlightening to explore how the relatively new requirement of school visits in the DP is changing the ways in which external IB evaluators are perceived. This study shows that evaluation is highly contextual. In light of this finding, it would be prudent to further discover how the IB trains and selects school visit team members and leaders and the readers of the evaluation self-studies. An exploration of whether organizing the evaluation teams by knowledge of the context of the school’s evaluation instead of the current practice of regional teams would be beneficial to the schools that belong to this network.

Furthermore, an examination of the type of responses an accrediting body sends to a school could provide more insight into how much buy-in there is from practitioners to actually turn words on paper into thoughtful action. Does the use of a drop-down menu to fill out reports reinforce the “ticking boxes” feeling of the evaluation as opposed to receiving useful feedback? Are external evaluators who are independent consultants and perceived as out of practical classroom experience for years viewed by teachers as not in touch with the realities of what they face every day? What type of professional development of external evaluators would be useful to be able to create culturally and contextually sensitive evaluations that are actually impactful in a DP school?

Finally, this study focuses on one particular school in one geographic region undergoing evaluation mandated by one specific program. While the case of evaluating the IB Diploma Program of the school provides a rich description of collaborative evaluation in this particular context, it is important to explore whether or not the findings of this study would be generalizable to other schools and their contexts. As program planning and evaluation are here to stay as a part of the professional lives of educators, further research in
this area can increase our understanding of how we can make the most use of these mandated processes to improve the efficacy of teachers in order to better serve our students.

5.9 Limitations of the Study

This study examines the practices used for mandated program evaluation of one specific program in one specific school. The findings are based on only one case study of one school with detailed descriptions provided by only three participants who implemented the IB Diploma Program; therefore, they are not generalizable to other programs and their teachers in other locations. The characteristics of the school and the personal experiences of each individual teacher will undoubtedly affect how teachers collaborate during the evaluation of the given program. Moreover, the study is limited to retrospective recollections of the evaluation processes by the participants following a large gap of time after the evaluation itself. Additionally, there was only opportunity for a secondary analysis of the qualitative survey previously developed by the school. Nonetheless, the study provides a rich description of this school and these teachers’ experiences and perspectives that could provide insights for schools and teachers completing program evaluations.

Furthermore, this project only presents the participants’ self-perceptions of the changes implemented by the school and the possible impact of the evaluation activities. It is beyond the scope of this study to make a claim for causal relationships between self-assessment of school practices and increased student achievement and teacher efficacy.
References


Eddy-Spicer, D., Ehren, M., Bangpan, M., Khatwa, M., & Perrone, F. (2016). Under what conditions do inspection, monitoring and assessment improve system efficiency, service delivery and learning outcomes for the poorest and most marginalised? A realist synthesis of school accountability in low-and middle-income countries. *Social Science Research Unit, UCL Institute of Education, University College London*. Available at: https://assets.publishing.service.gov.uk/media/5851931140f0b60e4c0000bd/SR_Q4_Final_Draft_for_Publication.pdf.


based on interim reports prepared by Centre for Program Evaluation, The University of Melbourne. Bethesda, MD.


Appendix A – School Consent Letter

Dear ________________,

I am seeking permission to conduct a research study about the mandated evaluation of the International Baccalaureate Diploma Program (IB DP) at your school. This research is conducted as part of the requirements for completing the Doctor of Philosophy Degree at the Ontario Institute for Studies in Education of the University of Toronto. The following outlines the study itself and information about the school’s participation.

The study is entitled: **Teacher Collaboration for Mandated Accreditation: A Case Study of Evaluating the International Baccalaureate Diploma Program**

The purpose of the study is to examine and describe the collaborative evaluation practices employed by one specific authorized IB Diploma Program school to complete the guided self-study questionnaire for the mandated five-year evaluation for the International Baccalaureate Organization and to investigate the effects of the evaluation.

**What are the benefits for the school?**
By participating in this study, the school will have the opportunity to examine the different ways stakeholders participated in the evaluation of the Diploma Program and to investigate the effects of the evaluation.

**What risks are there for the school in participating in this study?**
There are no known risks to participating in the study. All the raw data collected will be kept in confidence and the school will not be identified. However, since this is a case study about one school conducted by the Diploma Program Coordinator of the school, there is a possibility that the school might be identified as a participant in the study in published material due to the small sample size, mainly by people who work in IB schools.

**Research Consent:**
I have read and understood the conditions under which the school and its teachers will participate in this study and give Ildiko Murrayne Biro permission to conduct the research at our school.

Name: ___________________, Superintendent

Signature: ___________________________ Date: ___________________________

If you have any questions or concerns about this study, please contact the investigators:

Principal Investigator
Ildiko Murrayne Biro

Project Supervisor
Dr. Douglas E. McDougall
Appendix B – Participant Consent Letter

Dear Participant,

You are invited to participate in a research study about the mandated evaluation of the International Baccalaureate Diploma Program (IB DP). This research is conducted as part of the requirements for completing the Doctor of Philosophy Degree at the Ontario Institute for Studies in Education of the University of Toronto. I have been working with your Superintendent and owners/board of directors to use your school as the site for the study and I am now looking for volunteers who have served on the evaluation team for the completion of the required documents for the IB Diploma Program to participate in the project. The following outlines the study itself and information about your participation.

The study is entitled: Teacher Collaboration for Mandated Accreditation: A Case Study of Evaluating the International Baccalaureate Diploma Program

The purpose of the study is to examine and describe the collaborative evaluation practices employed by one specific authorized IB Diploma Program school to complete the guided self-study questionnaire for the mandated five-year evaluation for the International Baccalaureate Organization and to investigate the effects of the evaluation.

A Brief Overview: If you decide to participate in the study, you will need to commit to spending about 10-15 hours over a five-week period with the researcher. There will be three one-hour interviews. There will also be a need for “reflective time” between interviews to prepare for the next session.

Your involvement in the study will require you to do the following:

1. Be involved in three hour-long interviews which will be audio taped. The questions will focus on how you participated in the evaluation of the Diploma Program and what you perceive to be the effects of the evaluation on the program at your school.

2. Share documentation with the researcher (meeting minutes, staff handbook, teachers’ timetables, etc..) that helps you to talk about how you believe the evaluation changed the Diploma Program at your school. These documents will be used to support information you will provide in the interviews.

What are the benefits for you?
By participating in this study, you will have the opportunity to examine the different ways you participated in the evaluation of the Diploma Program at your school and to investigate the effects of the evaluation. As a pedagogical leader at your school, this will increase your knowledge of the program which in turn will benefit students and teachers at your school.

What risks are there for you in participating in this study?
There are no known risks to participating in the study. Only you, the researcher, and the thesis supervisor will have access to the data collected. All the raw data will be kept in confidence and you will not be identified by name in the study, nor will your school be
identified. The data will not be available to the administration of your school and will not be used to evaluate your performance as part of any school or system evaluation. You will have access to all the raw data collected about you. All the raw data collected during the study will be secured in a locked file and/or secure computer and will be shredded/deleted at the conclusion of the study. However, you need to be aware that there will be only three participants in the study. This small number creates the possibility that you might be identified as a participant in the study in published material due to the small sample size, mainly by people who work at the school.

Your participation in this research study requires a commitment from you but you may, at any time, withdraw from the study by simply indicating to the researcher your intention to withdraw. No evaluative judgment will be made about you if you choose to withdraw from the study. All raw data connected to your participation will be immediately destroyed.

Research Consent:
I have read and understood the conditions under which I will participate in this study and give my consent to be a participant.

Name: ___________________________________
Signature: ______________________        Date: ______________________

If you have any questions or concerns about this study, please contact the investigators:

Principal Investigator  Project Supervisor
Ildiko Murrayne Biro        Dr. Douglas E. McDougall
Appendix C - Teacher Feedback on the Process of Completing the DP Evaluation Survey

Dear Participant,
Please take a few moments to complete this feedback survey regarding the process of preparing the DP Evaluation self-study during the 2013-2014 academic year. Please find the DP Evaluation Documents on Atlas.
Please contact me at _________________________ if you have any questions or concerns.
Thank you so much in advance!
Kind Regards,
Ildiko Murray

I participated in the preparation of the self-study as a (select as many options as appropriate)*Required

- [ ] teacher contributor - Section C (departments)
- [ ] evaluation team member - Section C (whole school)
- [ ] evaluation team member - Sections A and B
- [ ] evaluation team member - Charts 1-7
- [ ] Diploma Program Coordinator
- [ ] High School Administrator
- [ ] parent of a Diploma Program student at the school

I was aware of the entire process that was used for the preparation of the self-study for the DP evaluation.*Required

1 2 3 4 5
strongly disagree  ☐ ☐ ☐ ☐ ☐ strongly agree

I felt included in the preparation of the self-study.*Required

1 2 3 4 5
strongly disagree  ☐ ☐ ☐ ☐ ☐ strongly agree

I could choose my level of involvement in the preparation of the self-study.*Required

1 2 3 4 5
strongly disagree  ☐ ☐ ☐ ☐ ☐ strongly agree

I have reviewed the documents submitted to the IB located on Atlas (please choose one).*Required

- [ ] I read all of the documents very carefully.
- [ ] I read some of the documents carefully.
- [ ] I skimmed through some of the documents.
• □ I skimmed through only the self-study.
• □ I quickly looked at the list of documents but I trust that accurate information was uploaded.
• □ I have not opened any of the documents.

I believe that the submitted documents provide an accurate and honest depiction of the Diploma Program at the school.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree

I believe that all stakeholders, including parents and students, were appropriately represented in the preparation of the self-study.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree

I believe that the process of completing the evaluation was transparent.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree

In general, I believe that completing the evaluation made the Diploma Program at the school better.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree

I have a better understanding of the DP standards and practices as a result of completing the evaluation.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree

I have a better understanding of assessment practices in the DP after completing the evaluation.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree

I have a better understanding of teaching and learning practices in the DP after completing the evaluation.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree

I have a better understanding of how to support the DP core after completing the evaluation.*Required

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
</table>
| strong disagree | 〇 | 〇 | 〇 | 〇 | 〇 | strongly agree
I feel there is competent pedagogical leadership for the Diploma Program at the school.*Required

1 2 3 4 5

strongly disagree [ ] [ ] [ ] [ ] [ ] strongly agree

I believe that the action plan submitted provides adequate resources for the future of the Diploma Program at the school.*Required

1 2 3 4 5

strongly disagree [ ] [ ] [ ] [ ] [ ] strongly agree

Please answer the following questions in paragraph format:

In general, please describe your experiences with completing the Diploma Program evaluation.

What did you perceive as a positive aspect of engaging in the process of completing the self-study for Diploma Program evaluation?

What did you perceive as a negative aspect of engaging in the process of completing the self-study for the Diploma Program evaluation?

What would you suggest that we change for next time?

Are there any other comments that you would like to make?
Appendix D – Interview Guide

Dear Participants,

Thank you again for volunteering to participate in this research study. As we are preparing for the first interview sessions, please find the DP Evaluation Documents on Atlas for your reference. As we discussed, the first session will focus on the process of the evaluation (please see the interview guide below – the survey questions are attached again for easy access). The second session will cover the effects of the evaluation on the program and the third interview will explore the future of program evaluation at the school level and at the IB level.

Interview Guide

Session One

Purpose:
- To conduct an introductory semi-structured interview,
- To get to know the participants,
- To obtain information about the participant’s views on the process of completing the DP evaluation.

Part A: Personal Characteristics of the Participant

Q1: How long have you been a teacher/head of department/administrator?
Q2: What kind of teacher education did you receive (DP training included)?
Q3: How long have you been with the school (and your current position)?
Q4: How would you describe yourself as a teacher/head of department/administrator?

Part B: The Process of Completing the Evaluation

Q1: Have you had previous experience with program evaluation?
Q2: How did you become involved with the DP evaluation?
Q3: See survey questions
Q4: What could have been done to make teachers feel more part of the process?

Reflection for Next Session:

Today we talked about the process of completing the evaluation. For the next session, I would like you to think about the effects of the evaluation on the Diploma Program at the school. If you have any documents (meeting minutes, staff handbook, teachers’ timetables, etc.) that would help you with this, please bring them with you for the next session. Please let me know if you have any further questions or concerns.

Session Two

Purpose:
- To conduct an introductory semi-structured interview,
- To obtain information about the participant’s views on the effects of completing the DP evaluation.
Effects of Evaluation

Q1: What do you perceive as the effects of the evaluation process?
Q2: Were there any changes to the program as a result of the evaluation process?
Q3: How was the program shaped specifically by the evaluation during the process?
Q4: What are the actual implementation plans outlined in the recommendations and in the action plan after the evaluation?

Reflection for Next Session:

Today we talked about the effects of completing the evaluation. For the next session, I would like you to think about the future of program evaluation of the Diploma Program at the school and at the IB level. If you have any documents that would help you with this, please bring them with you for the next session.
Please let me know if you have any further questions or concerns.

Session Three

Purpose:
☐ To conduct an introductory semi-structured interview,
☐ To obtain information about the participant’s views on the future of program evaluation of the DP at the school and at the IB level.

The Future of DP Evaluation at the School Level and at the IB Level

Q1: How should the school proceed after the evaluation?
Q2: What should be the areas of focus for improvement over the academic years prior to the next evaluation?
Q3: Are you aware of any changes in the IB evaluation requirements initiated by the IB?
<table>
<thead>
<tr>
<th>Action</th>
<th>Who</th>
<th>When</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA Accreditation Planning Team meetings</td>
<td>MSA Team</td>
<td>November 2012-October 2013</td>
</tr>
<tr>
<td>Secondary Management Informed</td>
<td>DPC</td>
<td>August 2013</td>
</tr>
<tr>
<td>Timeline drawn up</td>
<td>DPC</td>
<td>October 2013</td>
</tr>
<tr>
<td>Faculty informed</td>
<td>DPC</td>
<td>August/September 13</td>
</tr>
<tr>
<td>DP staff informed in detail and documents given</td>
<td>DPC</td>
<td>November 2013</td>
</tr>
<tr>
<td>Groups meet to discuss section C during DP meeting</td>
<td>Dept./Groups</td>
<td>November-December 13</td>
</tr>
<tr>
<td>Send out 'invites' to parents, board, students, staff</td>
<td>DPC/HSP/SI</td>
<td>November 2013</td>
</tr>
<tr>
<td>Information meeting and distribution of documents to volunteers</td>
<td>DPC</td>
<td>November 13 (end)</td>
</tr>
<tr>
<td>7 representatives meet to discuss overview of section C</td>
<td>DPC + 7</td>
<td>January 2014</td>
</tr>
<tr>
<td>Round table A-C (split up sessions) (Times to be determined once volunteers are in place)</td>
<td>Reps. From Board/parents/students/staff</td>
<td>Throughout February 2014</td>
</tr>
<tr>
<td>Collate school-wide documents</td>
<td>DPC/HSP</td>
<td>Early March 14</td>
</tr>
<tr>
<td>Collate department documents</td>
<td>DPC + 7</td>
<td>Early March 14</td>
</tr>
<tr>
<td>Scan documents</td>
<td>DPC/IT</td>
<td>End of March 14</td>
</tr>
<tr>
<td>Redistribute section C for final review</td>
<td>Dept.</td>
<td>March 2014 (end)</td>
</tr>
<tr>
<td>Principal's summary</td>
<td>HSP</td>
<td>April 2014 (early)</td>
</tr>
<tr>
<td>Give final to Superintendent/admin.</td>
<td>DPC</td>
<td>April 2014 (early)</td>
</tr>
<tr>
<td>Superintendent’s comments</td>
<td>SI</td>
<td>April 2014 (early-mid)</td>
</tr>
<tr>
<td>Respond to admin. feedback</td>
<td>DPC</td>
<td>April 2014 (mid)</td>
</tr>
<tr>
<td>Scan and send to …</td>
<td>DPC/IT</td>
<td>15/04/14</td>
</tr>
<tr>
<td>IBAEM</td>
<td>DPC</td>
<td>15/04/14</td>
</tr>
<tr>
<td>IBAEM deadline</td>
<td>DPC</td>
<td>15/04/14</td>
</tr>
</tbody>
</table>

DPC: Diploma Program Coordinator  HSP: High School Principal  SI: Superintendent
# Appendix F - Collaborative Planning and Reflection Meetings for the Evaluation

<table>
<thead>
<tr>
<th>Name of Meeting</th>
<th>Who attends</th>
<th>Frequency of meeting</th>
<th>Objectives</th>
</tr>
</thead>
</table>
| Diploma Program Teacher Meeting    | The DPC leads the meeting for all DP subject teachers, CAS advisors, TOK teachers, the High School Principal and the High School Assistant Principal | Once a month (usually the first or second Sunday of the month after school)       | -discussion about the day-to-day running of the program including deadlines for IB components and preparation of the students for the world exams  
- professional development regarding assessment and teaching practices  
- professional development for the inclusion of research skills and TOK skills in all subjects  
- the analysis of exam results data to guide our efforts to improve our program |
| Diploma Program Leadership Meeting (Admin) | The Diploma Program Coordinator, the High School Principal and the High School Assistant Principal | Once a week (regularly scheduled for 8AM every Thursday)  
*Note:* The DPC and the HS Principal are in daily contact regarding the day-to-day running of the program as needed. | --discussion about the day-to-day running of the program including deadlines for IB components, the preparation of students for the world exams, teacher, student, and parent concerns, and budgeting for the program  
- planning for all DP related training, meetings and professional development  
- preparing for meetings with the Superintendent and Administrative Council  
- the analysis of our enrolment data to guide our efforts to grow our program |
<p>| Extended Essay Supervisor Training | The DPC leads two one-hour training workshops for EE Supervisors          | Two one-hour workshops scheduled in March of the DP1 year                      | - introduction to the Extended Essay and any changes to the requirements, the role of the supervisor in the program |</p>
<table>
<thead>
<tr>
<th>Name of Meeting</th>
<th>Who attends</th>
<th>Frequency of meeting</th>
<th>Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP Evaluation Team Meeting</td>
<td>The DPC leads all DP teachers, HS Principal and Assistant Principal, the Evaluation team, and Superintendent (Administrative Council) – attendance dependent on task</td>
<td>During the preparation of the DP Self-Study scheduled as needed</td>
<td>-preparation of all the documents required for the DP Self-Study</td>
</tr>
<tr>
<td>(Evaluation Year only)</td>
<td></td>
<td></td>
<td>Extended Essay (Part One)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-using the assessment criteria of the Extended Essay, the preparation of the supervisor report and completing of the Viva Voce (Part Two)</td>
</tr>
<tr>
<td>Committee Meetings</td>
<td>The K-12 Curriculum Coordinator and the four IB coordinators (two for PYP and one each for MYP and DP)</td>
<td>Several meetings in the Fall of the Academic Year</td>
<td>-review and update all procedures for the assessment and language instruction policies</td>
</tr>
<tr>
<td>(Assessment, Language Instruction, Curriculum)</td>
<td></td>
<td></td>
<td>-ensure vertical alignment of AIS curriculum</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-assist the implementation of the Common Core curriculum guidelines</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-ensure successful transition of students through the three IB programs</td>
</tr>
<tr>
<td>Scheduling Committee Meetings</td>
<td>High School Principal, the Middle School Principal, the Diploma Program Coordinator, and the High School Guidance Counsellors</td>
<td>As needed during the academic year</td>
<td>-examine the scheduling of SL and HL courses for the Diploma Program</td>
</tr>
<tr>
<td>Student Services Meeting</td>
<td>High School Principal and Assistant Principal, the two High School Guidance Counsellors are in regular attendance. The DPC attends as needed to be included in decisions regarding students in DP courses.</td>
<td>Once a week (regularly scheduled for 1:30PM every Thursday)</td>
<td>-discussion about at-risk students and development of plan of intervention (concerns regarding academics, attendance, and deportment)</td>
</tr>
<tr>
<td>Name of Meeting</td>
<td>Who attends</td>
<td>Frequency of meeting</td>
<td>Objectives</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Divisional/Departmental Meetings</td>
<td>All High School staff.</td>
<td>Once a week (regularly scheduled for 3PM on Tuesday)</td>
<td>-discussions about AIS program delivery, sometimes specific to the DP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rotating schedule for the year for meetings involving teachers in:</td>
<td>-professional development, sometimes specific to the DP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Divisional meeting (all high school teachers)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Department meetings (Grades 6-12 subject groups)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-All school (K-12 staff)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix G – Logic Model for the Evaluation of the Extended Essay Supervisor Training Program

Logic Model for the Evaluation of the Extended Essay Supervisor Training Program

**Inputs**
- Data analysis of extended essay results for Diploma Program
- Diploma Program Coordinator as Extended Essay Coordinator
- Partnership with University Library to conduct higher level research
- Librarians as support for conducting research

**Activities**
- Development of the Extended Essay Supervisor Training Program Materials
- Two one-hour training sessions for all extended essay supervisors led by the Diploma Program Coordinator; first session focuses on requirements and processes of the extended essay, second session focuses on assessment of the extended essay
- Dissemination of all new training materials including the Extended Essay Supervisor Handbook
- Organization and facilitation of students' visit to University library to conduct research
- Input survey for both supervisors and students regarding their experiences with the completion of the extended essay

**Outputs**
- Number of supervisors trained
- Number of students affected by supervisor training
- Number of full diploma students who receive D or above on the extended essay – data from World Exams results

**Outcomes**

**Short Term**
- All students complete the extended essay by the deadline

**Medium Term**
- All students receive at least a D on the extended essay
- Increase in number of As and Bs on the extended essay

**Long Term**
- Better alignment between predicted grades and actual grades for the extended essay
- More positive attitude towards the completion of the extended essay from students
- Higher level of satisfaction with the process of completing the extended essay both from students and supervisors
- The reputation of the school is improved due to higher average scores and diploma passing rates