International Baccalaureate Assessment Methods

In the following report, Hanover Research provides an overview of the assessment practices, tools, and methods utilized by schools offering International Baccalaureate programs to evaluate the program’s success in achieving its stated goals.
Introduction

This report provides an overview of methods that have been used by school districts to assess implementation of the International Baccalaureate (IB) program and measure its success in achieving its stated objectives. At the core of the IB program is an ambition to achieve a number of goals alongside the rigorous study of academic curriculum areas. These objectives, as stated in the IB Learner Profile and elsewhere, include developing personal responsibility for learning and social service; nurturing critical, independent thinking and inquisitiveness, developing an esprit de corps among the student body and within a school; promoting the involvement of parents and the larger community; and providing an international perspective to students facing a multicultural, globalized future. Such goals are more difficult to quantify than traditional academic achievement. While standardized tests and other assessment methods are well-developed measurements of academic achievement, the IB program’s more nebulous aims may demand innovative or specialized measures to gauge their success.

The report begins by reviewing the International Baccalaureate Organization’s (IBO) suggestions for evaluating program success. Although the IBO’s evaluation guidelines are written as a means to gather evidence in support of the IB program, the guidelines can also serve to measure a program’s lack of success, if that proves to be the case.

Three studies undertaken to measure the efficacy of the IB program by the Academy School District 20 in Colorado Springs, Colorado are then considered. Two are quantitative, measuring the IB program against student achievement on the Colorado Student Assessment Program’s mathematics and reading examinations. The third is qualitative, using interviews, surveys, focus groups, and classroom observations to draw its conclusions.

Following is a review of a study of the implementation of the Primary Years Program (PYP) across multiple schools in Georgia. The Education Policy and Evaluation Center (EPEC) at the University of Georgia undertook the study. Its analysis consists of a broad-based survey followed by interviews, focus groups, and classroom observations at three schools. In addition, the report briefly summarizes a study of IB program efficacy completed by a doctoral student at the University of Nebraska.
International Baccalaureate Organization Recommendations for Program Evaluation

The International Baccalaureate Organization (IBO) presents its recommendations for program evaluation as a means to demonstrate program efficacy. However, the IBO’s evaluation strategies should also be useful in determining if an IB program is not achieving its desired results. The IBO addresses the types of data to be used in assessment, methods for data collection, maintenance of databases, and procedures for sharing data and results in a general way.

The IBO recommends beginning with two big-picture questions: “Why am I collecting data?” and “How will I look at the data?” In other words, clearly establishing what program outcomes are desired and expected, and how they will be measured, is the first step. Once the desired outcomes are established, a plan needs to be created to effectively measure those outcomes. The IBO advises taking an initial baseline measurement, then using the same metrics to measure yearly student progress.

The IBO stresses that there is “no magic formula” for program evaluation, but does recommend some general processes:

- **Identify** the data to evaluate. Data sets could include the number of applicants to the program, test scores, classroom observations, college acceptance rates, as well as attitudes or perceptions about participating in IB.
- **Review** the collected data, such as a group of student surveys. Ensure that the data sets are complete and they cover the areas of interest.
- **Analyze** the data. Look at the information at the micro level (e.g. specific class surveys), and then to compare it at the macro level (e.g. the school or grade level as a whole).
- **Utilize** the collected information. Use the information to advise decisions on appropriate changes. It should provide a reference point to see how well the program is running at a school, and help to adjust its implementation accordingly. Also, share progress with the community to keep it informed and solicit its feedback.

According to the IBO, school and district analyses of program efficacy vary according to school capacity to perform evaluations and school sophistication in

---

2 Ibid.
gathering and manipulating data. The IBO lists indicators schools have collected in order to measure program implementation and efficacy.³

- **Number of applicants to the IB program:** enrollment and participation rates
- **Student performance data:** state assessment scores, GPA, SAT, drop-out rates, graduation rates
- **Alternative assessments:** student portfolios, informal observations
- **Student demographic data:** gender, race or ethnicity
- **Longitudinal studies:** statistical analyses over specific time periods
- **College matriculation rates and quality of the college**
- **College readiness/preparation**
- **Scholarships received by IB students for college**
- **Teacher productivity, unity and support**
- **Testimonials from students, parents, educators or community members:** videos, quotes, interviews

**Community Involvement**

In addition to analyzing easily available metrics such as student performance data, the IBO recommends formal surveys of students, teachers, and other stakeholders, such as parents, to gauge opinions and attitudes. Surveying may allow for measurement of the less quantifiable aspects of the program, such as the development of an international perspective or an increased sense of belonging to and responsibility for the community. Informal surveys or observations may prove to be useful in determining program efficacy. Examples are classroom “walkthroughs” or short, informal questionnaires for parents at the end of a meeting.

The IBO recommends a comprehensive approach that combines quantitative data and qualitative data collection with a timeframe of expectations for the program. It gives the following timeframe template for evaluating the PYP or MYP.⁴

---

³ Ibid.

⁴ Ibid.
Measure outputs associated with the IB program: enrollment, test scores etc. and determine how they have changed at three stages.

- Initial outcomes: define expectations for and measure immediate changes among students after an initial exposure to the IB curriculum.
- Intermediate outcomes: define expectations for and measure changes among students after long-term exposure to the IB curriculum.
- Ultimate outcomes: the goals and measurement thereof for students who successfully complete the IB program.

In order to measure outcomes against expectations, the IBO recommends evaluating several general categories. To evaluate the implementation of the program, the IBO recommends benchmarking student performance on IB exams against other IB students in the state, the United States, and globally. It recommends measuring changes in “school culture, student perceptions, PSAT or SAT scores, matriculation rates, enrollment by demographics, number of honors courses taken, and student self efficacy.” In order to appraise less quantifiable changes like school culture, the IBO suggests using questionnaires, individual interviews, and focus groups.

The IBO highlights the methods that two districts have reported using to measure the IB’s implementation and efficacy. Moody Middle School and Henrico High School in Richmond, Virginia reported using three quantitative methods to quantify outcomes: calculating the number of applicants to the program, monitoring the passing rate of students to the world IB average, and quantifying the amount of scholarship dollars graduating students receive from higher education institutions. For qualitative measures of program success, Moody and Henrico used student surveys and informal, anonymous surveys of parents following other events, such as parent coffees.

Park View Education Centre in Bridgewater, Nova Scotia used four methods to evaluate program implementation and efficacy. Park View monitored grades received by students for IB coursework and measured class sizes as metrics of program implementation. To determine program efficacy, the school tracked the number of scholarships received by graduates and assessed student success at the post-secondary level by surveying students who had completed a semester at college or university. Students were asked whether IB coursework had left them well-prepared for higher education. Park View also used wide-ranging interviews with IB graduates to further gauge students’ perspectives on the program’s success.

---

5 Ibid.
6 Ibid.
7 Ibid.
8 Ibid.
Case Studies of International Baccalaureate Program Evaluation

Academy School District 20

Academy School District 20 (ASD 20) in Colorado Springs, Colorado analyzed the impact of the IB curriculum on student achievement in reading and mathematics within the district. The math study followed the MYP and DP programs, grades 5-8 and 8-10 respectively, from 2001-2004 to determine achievement outcomes. The reading study analyzed the PYP, MYP, and DP programs in grades 3-10. The district’s primary research goal was to determine whether the IB program had a discreet, statistically significant impact on achievement while controlling for student background factors such as family, motivation, and prior academic achievement. Outcomes were compared for different grade levels, genders, and against students not participating in the IB program.

In addition, ASD 20 published a qualitative analysis of trends relating to the academic growth of students in 2007. The qualitative segment of the ASD 20 study relied on focus groups, interviews, surveys, and classroom observations to assess the quality of “IB structures, teaching strategies, and student responsiveness.”

Quantitative Analysis

The longitudinal study, divided into two parts, focused on student achievement and growth in reading and mathematics. The same analysis techniques were used on different data sets to examine student improvement in each discipline.

For the portion of its analysis examining mathematics, the study selected its student groups (2001-2004 grades 5-8 and 2002-2004 grades 8-10) based on participation in the 2001-2002 and 2004 Colorado Student Assessment Program (CSAP) math assessment. The CSAP is Colorado’s statewide, standardized assessment test conducted by paper and pencil in a timed environment. The math study also used initial reading status and reading outcomes; reading scores are predictors of math

---

11 The Colorado Department of Education http://www.cde.state.co.us/cdeassessment/documents/csap/csap_frameworks.html
outcomes and growth. For this data, it used scores from the reading portion of the CSAP. Middle school students who did not participate in the IB program or participated for only one year were classified as non-IB students for analytical purposes. Middle school students who participated for two or more years were considered IB students. High school students participating in IB in tenth grade (2003-2004 school year) were considered to be IB students; those who never participated or opted out of the program prior to tenth grade were not considered to be IB students.

The reading achievement study used data from the reading CSAP to benchmark initial proficiency and measure improvement. Specifically, it drew on data from 2000-2004 for students in grades 3-5, 4-8, and 7-10. Although the grade groupings overlap, cohorts did not overlap given that the data was collected over a time period, 2000-2004. Table 1 presents this visually:

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Year/Grade</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002-2004 Grade 3-5</td>
<td>Cohort</td>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2000-2004 Grade 4-8</td>
<td>Cohort</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>2001-2004 Grade 7-10</td>
<td>Cohort</td>
<td></td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

The reading study does not define IB students and non-IB students explicitly, but from the context it seems likely that IB students were those from each cohort who participated in the IB program every year over the period.

ASD 20 used several mathematical models and statistical methods to analyze the CSAP data for the selected student groups: descriptive statistics, bivariate correlations, stepwise multiple linear regressions, and hierarchical linear models (HLM). Descriptive statistics included measures such as frequency distributions, crosstabulation tables, and t-tests to compare variables such as math and/or reading scores, length of time in the IB program, language background, gender, minority, poverty status (defined by eligibility for free or reduced cost lunches), and other demographics.

The study used bivariate correlations and stepwise multiple linear regressions to identify factors impacting achievement on the CSAP in reading or math. Average

---

12 Kiplinger V. October 2005. Op. Cit. 2. According to the study, “Much prior research has shown that the subject area performance having the highest correlation with mathematics achievement is reading. Therefore, CSAP reading scores are included as predictors in the models of mathematics performance.”


gain in scores was the dependent variable, and the above variables functioned as the independent variables to determine their predictive power and interaction. The study then utilized empirical, nonparametric growth plots of achievement of randomly selected IB and non-IB students to determine the parametric model fit for the data. Finally, HLM models were employed to evaluate the extent of the IB program’s effect on student achievement.\textsuperscript{15}

\textit{Qualitative Analysis}

As mentioned above, the qualitative study incorporated personal interviews, focus groups, surveys, and classroom observations. These tools were intentionally used in that order to be “stepwise and complementary.”\textsuperscript{16} Interviews were used to develop and hone the organization of the focus groups to follow. General themes emerging from the interviews became the foundation for questions posed to the focus groups. The interviewees were experts on IB programs within the district and nationwide IB experts. The interviewer focused on questions of “IB structures, teaching strategies, and student responsiveness.”\textsuperscript{17} Specific questions posed during the interview process were not published in the paper.

The study formed six focus groups comprised of various stakeholders: IB students, non-IB students, IB teachers, non-IB teachers, IB assistant principals, and non-IB assistant principals. Student participants were nominated by teachers, and participating teachers were nominated by principals. Both students and teachers were nominated based on high achievement. The study chose assistant principals from various schools to ensure balanced viewpoints. However, only one IB assistant principal was obtained for that focus group. Other focus groups had up to 18 participants with an average of 10. Each group met for approximately an hour. Table 2 compiles questions posed to each focus group. Each question comprised the starting point for a discussion, and was subsequently elaborated upon.

<table>
<thead>
<tr>
<th>Focus Group</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>IB Students</td>
<td>Did you choose to participate in IB? If so, why? If not, who made the choice and why?</td>
</tr>
<tr>
<td></td>
<td>What aspect(s)/part(s) of your classes do you find the most interesting? Least interesting? Like the most? Like the least? Why?</td>
</tr>
<tr>
<td></td>
<td>Tell me about one class in which you are particularly successful. Why do you think you do so well in this class?</td>
</tr>
<tr>
<td></td>
<td>Tell me about one of your best teachers/a teacher you learn most from.</td>
</tr>
</tbody>
</table>

\textsuperscript{17} Ibid.
\textsuperscript{18} Ibid. 9-10.
<table>
<thead>
<tr>
<th>Focus Group</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Why did you choose her/him?</td>
</tr>
<tr>
<td></td>
<td>What parts of your classes do you find most helpful? Not very helpful?</td>
</tr>
<tr>
<td></td>
<td>Do you think that your classmates play a role in your learning? Why or why not?</td>
</tr>
<tr>
<td>IB Teachers</td>
<td>What are the characteristics of a successful IB student?</td>
</tr>
<tr>
<td></td>
<td>Why do IB students tend to be so successful?</td>
</tr>
<tr>
<td></td>
<td>Tell me about a successful critical inquiry/lesson. What made it so successful?</td>
</tr>
<tr>
<td></td>
<td>What role does inquiry play in your classroom? What is it about inquiry-based teaching and learning that is so important? Be specific.</td>
</tr>
<tr>
<td></td>
<td>Describe the typical student in your classes/IB. Why do you think certain students engage in IB while others don’t?</td>
</tr>
<tr>
<td></td>
<td>Tell me about professional learning communities. How are they incorporated into your practice?</td>
</tr>
<tr>
<td></td>
<td>What are essential questions and big ideas and how are they incorporated into your lessons? Do they play a role in student learning and academic achievement? How?</td>
</tr>
<tr>
<td>IB Principals</td>
<td>What are the characteristics of a successful IB teacher? A successful IB classroom?</td>
</tr>
<tr>
<td></td>
<td>What support is offered to teachers to ensure success in the classroom?</td>
</tr>
<tr>
<td></td>
<td>What are important steps/actions taken to successfully implement an IB curriculum in a classroom? In a school? What steps/actions are taken to maintain it?</td>
</tr>
<tr>
<td></td>
<td>How can you tell the IB program at your school is on track? Be specific.</td>
</tr>
<tr>
<td></td>
<td>What are some signs the program is off track? Be specific.</td>
</tr>
<tr>
<td></td>
<td>What are some of the most important/valuable characteristics of a successful IB program?</td>
</tr>
<tr>
<td></td>
<td>If you had to point to specific characteristics that distinguish IB from traditional programs, what would they be and why?</td>
</tr>
<tr>
<td>Non-IB Principals</td>
<td>What are the characteristics of a successful teacher? A successful classroom?</td>
</tr>
<tr>
<td></td>
<td>What support is offered to teachers to ensure success in the classroom?</td>
</tr>
<tr>
<td></td>
<td>Is there continuity in the learning that occurs within grade levels? Between them? If so, why and how is it maintained? If not, why not?</td>
</tr>
<tr>
<td></td>
<td>Are there methods/means for collaboration among your teachers?</td>
</tr>
<tr>
<td></td>
<td>What factors are taken into account in your school to ensure student success? Think about teacher characteristics, school community, professional development, etc.</td>
</tr>
<tr>
<td>Non-IB Students</td>
<td>Questions for non-IB students were the same as for IB students, except for the elimination of the first question regarding IB program participation choice.</td>
</tr>
<tr>
<td>Non-IB Teachers</td>
<td>What are some characteristics of a successful student?</td>
</tr>
<tr>
<td></td>
<td>Tell me about instructional strategies you use in class. Do you differentiate their use? Which ones do you find most successful and why?</td>
</tr>
<tr>
<td>Focus Group</td>
<td>Question</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>Which ones do you find least successful and why?</td>
</tr>
<tr>
<td></td>
<td>Are there other techniques/components you include in your classroom to ensure student success? Describe them.</td>
</tr>
<tr>
<td></td>
<td>How do you organize your instruction overall (themes, units)? What do you take into consideration when developing lessons/units in your classroom?</td>
</tr>
<tr>
<td></td>
<td>Tell me about a successful lesson. What made it so successful?</td>
</tr>
<tr>
<td></td>
<td>What support do you have in the classroom? In the school? Do you work with other teachers at your grade level? In the school?</td>
</tr>
</tbody>
</table>

After the completion of the focus groups, the study administered a survey to a random sample of teachers in the district. Equal numbers of IB and non-IB teachers were included. The survey excepted teachers of kindergarten, first grade, part-time teachers, and teachers with fewer than three years experience. Drawing on information gathered from the interviews and focus groups regarding the effectiveness of certain teaching strategies, it attempted to determine the attributes of programs deemed to be the most effective. Most concepts and components addressed were specific to the IB program. Unfortunately, the study does not list the specific questions posed by the survey.

Classroom observation occurred for grades corresponding to the PYP, MYP, and DP programs. Non-IB classrooms at those levels were likewise observed as a basis for comparison. The survey observed 15 classrooms in all for full 45 minute or 60 minute classes. In order to further examine strategies and concepts identified by the interviews, focus groups, and the survey, observers specifically examined instructional strategies and teaching methods, student engagement and student-teacher/student-interactions in the classroom.

The study analyzed the collected data for recurring themes related to student involvement, student-teacher interaction, and more specific areas emerging from the data itself. The goal was to find characteristics that could “help to explain superior student success across grade levels and throughout the district’s classrooms.”

Themes from interviews, focus groups, surveys, and classroom observations were crosschecked and verified against one another. Themes that emerged from one part of the study but not from the others were discarded or merged into another theme as a subset, while themes that recurred in various aspects of the study were considered strongly supported.

Many of the techniques adopted in ASD 20’s qualitative study - interviews, focus groups, surveys, and classroom observation - would likely prove to be useful in qualitatively assessing whether the implementation of an IB program is achieving the expected goals.

19 Ibid. 4.
Academic Studies of the IB program

**PYP Implementation Study by the University of Georgia**

In February 2009, the Education Policy and Evaluation Center (EPEC) in the College of Education at the University of Georgia published The Primary Years Program Study, an analysis of the implementation of the PYP program in Georgia schools. The IBO commissioned the study, but it was conducted independently by the EPEC.

The EPEC study was comprised of two parts. Its first part examined how schools move from interested candidate to authorized provider status for the PYP program. In the second part, the EPEC focused on the implementation of the program, examining successful strategies, challenges, and support from the IB. Key steps for successful program implementation included:

- Important preparatory steps included networking with other IB schools, sending teachers to IB workshops, and meeting with various stakeholder groups to ensure buy-in from all stakeholders, particularly teachers.

- By spearheading a school vision, developing a plan of action, and communicating it, successful school leaders boosted understanding of the PYP, strengthened teachers’ commitment and collaboration, and fostered the use of new teaching strategies.

- School strategies to establish the legitimacy of the PYP among families and the local community included offering IB information sessions, involving parents in the preparation of IB reports and events, and partnering with institutions such as local churches, hospitals and non-profit organizations.

- Six particularly successful implementation strategies were identified:
  - Whole school immersion: surrounding students and teachers with the PYP by making it visible in the school and promoting the IB philosophy throughout the school day, as well as ensuring that all special areas teachers also integrate the IB units in their instructional plans.
  - Collaborative planning, through both grade-level meetings and meetings across grade levels, together with sufficient planning time.

---

21 Ibid. 8-10.
22 Ibid. 11-13.
Continuous training, ranging from IB workshops and visits to other schools to onsite weekly professional development meetings.

- The availability of resources to purchase materials to support each IB unit, and the support of a media center.
- Community involvement efforts beyond newsletters and websites: inviting all stakeholders to School Council and PTA meetings, and making special efforts to communicate with parents in their native languages when necessary.

- Strong support for IB by the school leadership: The PYP coordinator was identified as the most important person for keeping teachers on task and maintaining the overall focus on IB goals.

- The report identified four areas of particular challenge:
  - Limited resources, particularly time to support the IB units
  - Integration of Georgia Performance Standards with the PYP
  - The trans-disciplinary nature of the program
  - District and state expectations

As Hanover’s report focuses on the metrics and tools used to assess implementation, please follow the above footnotes to the full report to view its conclusions in detail.

EPEC’s study used two methods to reach its conclusions. It administered an online survey to administrators and teachers at sixteen Georgia schools authorized to administer the PYP. Secondly, it initiated three detailed case studies of Fair Street Elementary in Gainesville, the International Community School in Atlanta, and Clubview Elementary in Columbus. The schools were chosen based on their varied sizes and demographic compositions. At each school, EPEC conducted 13 interviews with district and school-level administrators, held nine focus groups with teachers and three focus groups with parents, and conducted 16 classroom observations and various document reviews. The study analyzed collected data for analysis by categories of program standards and practice: philosophy, organization, curriculum, students and their families, and overall implementation.

Although the EPEC study focused on implementation of the PYP rather than measuring PYP efficacy, many of the tools it used are potentially translatable. Selected techniques follow that are likely to be useful not just for implementation evaluation, but also to assess program outcomes.

---

23 Ibid. 13-15.
Document Reviews

EPEC reported that it collected a number of documents for analysis including: 24

- Documentation of feasibility study (work products, meeting agendas and minutes, participants)
- Sample communications with parents and community
- Professional development plan, including records of attendance
- Written goals, strategies, timelines
- Student/parent handbooks
- Student work samples, projects
- Records of student’s learning (report card)
- Products of professional development (strategic plans, lesson plans, rubrics)
- Meeting minutes, memos, emails, journals, logs
- Newsletters, announcements, memos, communications with parents
- Community support documentation (communications, services, publicity)
- Web-site, newspaper articles

Some of these documents are clearly more applicable to assessing program implementation than measuring a program’s success in achieving the IB’s objectives. However, a detailed analysis of these documents could be a tool to assess not only student achievement gains, but also student progression along the ten dimensions of the IB learner profile. For example, meeting minutes, memos, emails, journals, and logs may provide an informal picture of student progress along the profile objectives of becoming open-minded and caring. The above documents may also prove useful as baseline records that will serve as a benchmark for progress for the school’s implementation of the program and for student outcomes.

Surveys

EPEC collected a total of 561 surveys responses from administrators and teachers at sixteen IB-authorized Georgia schools. The survey was divided into two parts: one addressed the authorization process, the other implementation of the program. It used a four point scale, allowing for responses of strongly disagree, disagree, agree, and strongly agree. EPEC’s survey organized questions about attitudes toward implementation of the IB program into the categories of philosophy, organization, curriculum, students and their families, and overall implementation. While this structure may not directly apply to evaluating program outcomes, surveys may prove to be one of the more powerful tools to measure program efficacy.

As in ASD 20, administrators and IB coordinators, grade-level teachers, other school personnel, parents, and students could be surveyed to gauge progress toward the IB

---

24 Quoted verbatim from Ibid. 7.
program goals. A survey could include questions focusing on goals of the program such as the IB learner profile, global thinking, or school spirit and togetherness.

_interviews_

Interviews formed an important component of EPEC’s review. EPEC conducted interviews with school and district-level administrators, IB coordinators, and other school personnel (non-teachers). These interviews provided valuable insight into the administration’s view of the implementation process. While the study’s interviews focused on implementation, in-depth interviews with as many stakeholders as time and resources allow would likely be an effective tool to evaluate a program’s efficacy from multiple viewpoints.

_focus groups_

EPEC conducted hour-long focus groups with teachers and 30 minute focus groups with parents. Focus groups have the advantage over interviews of being less time and resource intensive. A group forum may also provide for discussion and elucidation of points that might not be addressed in a formal interview setting. Some of the questions EPEC asked the focus groups may in fact be directly applicable to measuring program efficacy. For example, one parent said that it was difficult to explain the IB program to others but added, “You see it.”25 The response illustrated that the school had some success in making the PYP visible and promoting its philosophy throughout the day, but needed to better communicate its principles to parents. A teacher stated that the strategy of recording examples of IB learner qualities and recognizing both the student that displays one and his or her teacher had “united us a school.”26 Such feedback is from focus groups would be valuable in assessing PYP efficacy.

_classroom observation_

EPEC observed classrooms representing a cross-section of grade levels and subject areas for approximately 15-20 minutes each. The EPEC’s observation goals were to:

- Document a descriptive record of the structure, content, and activities of the classes.

- Capture the character of the learners’ participation in the classes, paying particular attention to: (a) the level and character of the learners’ engagement; (b) interactions with the teacher, each other, and materials; and (c) any evidence of IB curriculum standards that are enacted in the classroom.

25 Ibid. 11.
26 Ibid.
In particular, EPEC observers looked for:

- Learners’ engagement with instructor and materials (raising hands, collaborative group work, individual work, students’ enthusiasm, as well as boredom, disinterest, frustration).
- Students taking responsibility for their own learning (self-initiated, self-directed).
- Students demonstrating components of the IB learner profile (evidence of students as inquires, knowledgeable, thinkers, communicators, principled, open-minded, caring, risk-takers, balanced, reflective).
- Teacher’s engagement with learners (encouragement and formative-type feedback, etc.).
- Instructional strategies to meet the needs of different learning styles (special equipment, technology, building on students’ previous knowledge, grouping strategies, etc.).
- Transdisciplinary teaching (how different subjects are tied together in a meaningful way).
- Multi-cultural experiences (displays, themes, student actions).
- Individual, local, national and global awareness.
- Show of respect and value of diversity.
- Adequate resources to promote learning for diverse learners.
- Multi-language usage.
- Time available for inquiry learning.\(^{27}\)

Similar criteria would likely be an effective measure of the progress of a program, particularly if observations were conducted regularly over time.

*Data Analysis*

EPEC used a triangulation technique to analyze and synthesize data similar to that employed by the study for ASD 20. Case study data from focus groups, interviews, and classroom observations was examined according to pre-identified themes and by themes emerging from individual case and cross-case analysis. Recurrent themes were highlighted, and isolated themes were marginalized. In order to develop

\(^{27}\) Quoted verbatim from Ibid. 38.
narrative descriptions for the study’s findings, this technique was employed multiple times rearranging themes and data sources.

**Other Academic Studies**

A dissertation for the University of Nebraska analyzed MYP outcomes for participating 7th grade student’s academic achievement, behavior, extra-curricular involvement, and perceptions of life skills. It appears that the study compared MYP participating students to a control group of students following a standard academic program. It compared the two groups across several factors over a period of two years. To analyze student achievement, the study compared local and national testing results from the two groups for reading, math, and language tests. It compared absence statistics, self-perceptions of life skills, and extra-curricular involvement between the two groups as measurements of non-academic achievement factors.

While a control group will not be readily available, a district could analyze student data after MYP implementation against districts with a similar demographic and historical profile. Alternatively, student achievement, perceptions, extra-curricular involvement, and behavioral statistics such as attendance or disciplinary referrals after MYP implementation could be referenced against historical data.

---

28 Wilson, J. 2007. The effect of a founding International Baccalaureate Middle Years Programme on participating seventh grade students' achievement, behavior, extra-curricular involvement, and perceptions of life skills.
Conclusion

The quantitative studies performed by ASD 20 provide a mathematical model for analyzing academic achievement in a school that has implemented the IB program against statewide standardized tests. A school-wide implementation of the IB program would change the analysis to an extent, as a comparison of IB students and non-IB students in the same graduating class would not be possible. Student achievement after a certain length of time spent in the program would be a simple measure of program efficacy. In the program’s early years, some students will have spent differing amounts of time within the program. Such a scenario will allow for the comparison of historical data against new classes, creating the potential to examine the effect of spending not only a certain period of time within the program, but also to isolate the effect of the program on children at a given age. Demographic factors identified by the ASD 20 study such as language background, gender, minority, poverty status (defined by eligibility for free or reduced cost lunches), and other demographics would still need to be controlled for as entering classes may be of a different composition than those represented by historical data.

The ASD 20 and EPEC qualitative studies showed significant similarity. Both employed surveys, interviews, focus groups, and classroom observation to evaluate the efficacy of the IB program. By synthesizing the data gained from those tools, these studies were able to identify themes, which in turn allowed the studies to arrive at their conclusions. In order to evaluate the facets of the IB program that are not easily quantifiable, these techniques might be used as guidelines.
Project Evaluation Form

Hanover Research is committed to providing a work product that meets or exceeds member expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.


Note

This brief was written to fulfill the specific request of an individual member of Hanover Research. As such, it may not satisfy the needs of all members. We encourage any and all members who have additional questions about this topic – or any other – to contact us.

Caveat

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties which extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every member. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Members requiring such services are advised to consult an appropriate professional.