



A Collaborative Partnership Evaluation Project: Assessing a Middle School Initiative

*James E. Witte & Maria Martinez Witte
Auburn University*

*Iris Saltiel
Troy University*

*Paul T. Hackett
Columbus State University*

*Kathy Hesler & Mike Johnson
Richards Middle School, Columbus, Georgia*

Abstract

A collaborative partnership provides school administrators and university researchers an opportunity to evaluate a school-based initiative. Richards Middle School (RMS) in Columbus, Georgia identified school improvement goals and involved university researchers in a collaborative evaluation with implications for strengthening leadership, training, and instructional practices in schools. This article reports initial baseline data for International Baccalaureate and Non-International Baccalaureate 6th grade students and their academic performance and achievement results. An overview of the project and recommendations for further research for students and professional development are also provided in this article.

Overview

In the spring of 2004, faculty members from Troy University and Columbus State University proposed using the Richards Middle School (RMS) as a site to conduct a study of professional learning practices. During initial conversations, it became apparent that RMS had ambitious school

improvement goals for the upcoming school year. Implementation of the International Baccalaureate (IB) Programme, incorporation of IB classroom teaching methods throughout the school, continued work on Annual Yearly Progress as required by *No Child Left Behind*, as well as integrating new faculty into the RMS instructional culture were all components of an ambitious program needing assessment and evaluation. Three questions emerged from these initial conversations:

1. How can RMS teachers most effectively be trained regarding IB goals, objectives, and practices?
2. How effective will the IB Programme be in improving student achievement?
3. How can these two questions best be assessed and evaluated?

Given the scope and complexity of the task defined by the school principal and the RMS IB coordinator, selected members from Auburn University were brought in as research team contributors. Each selected university educator offered a different skill set to the project and their interests as researchers played a part in the negotiation and evolution of roles and responsibilities.

As a former school superintendent and principal, one team member had extensive public school experience, while another brought expertise in guiding collegiate collaborative partnerships. Others had expertise in adult learning and program evaluation.

Richards Middle School applied for and has requested authorization for the use of the Middle Years IB Programme (MYP). The International Baccalaureate (IB) Programme is part of the International Baccalaureate Organization (IBO) that uses a curriculum based on interdisciplinary, multi-nationalistic, international advanced curriculum education. The curriculum employs five areas of interaction (community and service; health and social education; environment; *homo faber*; and approaches to learning) and encourages students to be active learners, well-rounded individuals, and engaged world citizens (International Baccalaureate Organization, 2006).

Analysis of the connection between RMS and the IB Programme included an assessment of student and faculty information that answered the following question: In what ways does the IB program impact RMS? The design included three snapshots of student and faculty data over a two-year period. Students and faculty not in the IB program served as a control group. The student evaluations include student-centered assessments (performance

based assessments in music, art, technology, physical education, and Spanish), as well as formal reading and math scores from the Iowa Tests of Basic Skills (ITBS) and the Muscogee Assessment Program (MAP) tests. Other data sources included student essays, along with teacher recommendations.

Demographic variables included race, gender, neighborhood, and elementary school attended, as well as previous and current teachers. Faculty assessment focused on the IB as an innovative change to the overall curriculum. The IB Programme was acknowledged as a new and major staff development initiative within RMS and the school district. Assessments occurred at three primary points (beginning, mid-year, and end of year). The content of teacher planning meetings was also examined at three primary points: in the beginning of the year, mid-year, and end of year.

Collaborative Evaluation Team

Collaborative relationships between university researchers and local school faculty have not always been effective. Goldstein (2000) pointed out dilemmas created for researchers when roles are not clearly defined and when relationship issues for team members impact the work of implementation and assessment. Baldwin and Austin (1995) indicated that “productivity is greatest among collaborative teams mature enough to have well-defined procedures (an infrastructure) in place to operate efficiently but not so old that creative tension has diminished” (p. 67). Baldwin and Austin (1995) also considered that collaborations consisting of members with diverse backgrounds (differing with respect to gender, ethnicity, social status, and so on) would need more “system maintenance” and negotiations about goals, roles, procedures, and responsibilities than homogeneous collaborations.

Nevertheless, collaborative academic and school-based partnerships can be instrumental in gaining access to resources that may not have been available otherwise. The East Central Alabama/West Central Georgia area has several post-secondary educational institutions that can develop these kinds of partnerships directed toward the development of promising practices for the improvement of student learning (Clark et al., 1996).

Project Description

The evaluation project implemented at Richards Middle (RMS) in Columbus, Georgia involved a collaborative partnership between several universities and RMS. The desired outcomes of the school-based evaluation initiative involved strengthening leadership, training, and instructional practices in schools. The primary goal was to evaluate the effectiveness of the school's International Baccalaureate (IB) Programme in its first year of implementation in the sixth grade. A second goal of the investigation was to evaluate the effectiveness of the staff training and development process employed during the initial year in terms of effective professional learning practices. The third goal was to investigate the effectiveness of the collaborative process involving the universities and school representatives.

Initial Findings

Throughout the first year evaluation team members have kept, and shared, extensive field notes on the implementation of the IB Programme at RMS, efforts involving school and faculty professional learning, and on the development of the assessment model for the study. A review of these notes yielded two major observations. The first observation was that the roles of collaborative team members were clearly defined. Issues relative to the inappropriate clarification of team roles as cited by Goldstein (2000) have been few. The second observation was that initial resistance to the implementation of the IB model at RMS (teachers initially reported being overwhelmed by IB and *No Child Left Behind*) has yielded to the quick rollout of methods and practices as reported by the IB Coordinator.

Initial evaluation of assessment instruments used to identify candidates for the IB program have shown reliability in terms of discriminating quality of student performance and will assist as a means of evaluating the program. Survey instruments prepared by team members were used to ascertain perceptions of both the IB Programme and the training components.

Evaluation of Initial Baseline Data

Data concerning IB Programme assessment at the High School level are fairly accessible; however, data pertaining to the Middle School IB Programme are fairly rare. The team of researchers elected to establish baseline data for Richards Middle School to allow for the development of future assessment information. The sample included both IB participants and Non-IB participants, as shown in Table 1.

Table 1: *Sample Population*

| Participants | Grade Level | # of Students |
|---------------------|-----------------------|---------------|
| IB Participants | 6 th Grade | 76 |
| Non-IB Participants | 6 th Grade | 244 |

$N = 320$.

Data were analyzed using a discriminate function statistical technique and the following relationships were revealed. A variable called Academic Performance was created as an umbrella for the tested subtests of English, social science, science, reading and mathematics. Tests in these subject areas yielded the five standardized sub scores used for student proficiency determination. Pooled within-groups correlations between discriminating variables and the standardized canonical discriminant functions are shown in Table 2. Cohen and Cohen (1983) recommended a correlation of .30 or better as representing an acceptable level of statistical significance.

Table 2: *Sub-Test Correlations*

| Subject Area | R |
|----------------|-----|
| English | .89 |
| Social Science | .87 |
| Science | .87 |
| Reading | .81 |
| Mathematics | .76 |

When using the subtest scores as a predictor for group membership, Non-IB Programme membership could be identified with 84.5% accuracy and IB Programme membership with 94.5% accuracy. These data indicate clear delineation among the IB/Non-IB student groupings. This was an expected outcome due to the stringent pre-selection procedures for admittance to the IB Programme. However, having established a score supported baseline, future data analysis viability is increased.

Student Achievement

The collaborative evaluation team reviewed data for 5th and 6th grade administrations of the Georgia Criterion Reference Test (CRCT) in reading and math. Data were examined using an Analysis of Variance (ANOVA) Statistical Technique. The sample included 109 males, 123 females; 71 students were enrolled in the International Baccalaureate (IB) Programme and 161 were Non-Baccalaureate (Non-IB) Programme participants. CRCT scores for participants were examined by gender, race (White / Non-White) and program (IB/Non-IB). The purpose of this examination was to identify strengths or weaknesses in IB student performance compared to the Non-IB student.

For reading, no statistically significant main effects differences were found, nor were statistically significant interaction effects found in regard to reading gain by gender, race or program. In discussions of these results, the collaborative evaluation team and school leadership members considered the fact that while students in the regular school program receive significant instruction in the area of reading, students in the IB Programme are required to complete a foreign language course instead.

In regard to math gain and gender several noteworthy findings were revealed. In terms of overall math gain mean scores, females ($M = 7.55$, $SD = 22.70$) scored higher than males ($M = 4.28$, $SD = 26.61$). Statistically there was an interaction effect of gender by program (.02); however, the effect size (.03) was considered modest and of little interest.

No statistical differences were found by race by program among the participants; however, differences by race within program were found. Participants self-reported 134 Black and 80 White, with 7 Hispanic and 11 other. Due to the low representation, the Hispanic and other categories were

combined with the Black participants making a non-white category with 152 total participants.

The results of the three-way ANOVA for Math Gain indicated two statistically significant interaction effects and two statistically significant main effects. The statistically significant interaction effects were, Program X Race, $F(1, 232) = 5.52, p = .02, \eta^2 = .02$, and Program X Gender, $F(1, 232) = 5.36, p = .02, \eta^2 = .02$. The statistically significant main effects were Race $F(1, 232) = 7.10, p = .008, \eta^2 = .03$ and Program, $F(1, 232) = 36.55, p < .001, \eta^2 = .14$. Although statistical significance was found, the effect size (.03) for the relationship of Race and Program is sufficiently low to limit any practical significance. Only the main effect of Program demonstrated a meaningful effect size (.14), with students in the IB Programme obtaining greater math gains ($M = 19.04, SD = 25.89$) than students in the non-IB Program.

Recommendations for Further Research

The team recommended several steps for the remainder of the second year: the continued implementation of the professional learning model developed by the IB Coordinator and the research team; the continued evaluation of professional learning practices at RMS; the continued refinement of assessment methods for evaluating those practices and implementation of those methods; and the continued evaluation of student performance and analysis of teacher, parent, and student perceptions through the remainder of the school year. New programs, such as the International Baccalaureate (IB) Programme, require that teachers connect the new curriculum requirements and knowledge to their teaching responsibilities. O'Shea (2005) stipulated that:

Clearly, teachers are going to be attracted to the ideas of people who can help them perform their jobs in more effective ways. New ideas, however, are not enough. Teachers need opportunities to practice new skills and methods in a sheltered or coached environment that is similar to, or actually includes, their classrooms. (p. 135)

It is the view of the collaborative evaluation team that the RMS IB Project will require a long-term commitment from the team in order to track students and assess student performance in future years. Additionally, the team perceives that there is great promise for the IB Programme as it unfolds in the

coming years in terms of identification of sound instructional practices, the development of productive professional learning practices, and the evolution of multi-institutional collaborative associations aimed at improving student learning.

The project is nearing completion of the second year. Last year's sixth graders are now seventh graders and will be tested at the end of the spring term. The current sixth graders will also be evaluated and a continuing interview schedule with the teachers and survey of the parents will be undertaken. By following the established data gathering procedures into the third year of the program, the study of Richards Middle School will represent the most in-depth examination of an IP Middle School Programme in the United States.

References

- Baldwin, R. G., & Austin, A. E. (1995). Toward greater understanding of faculty research collaboration. *The Review of Higher Education*, 19(2), 45-70.
- Clark, C., Moss, P. A., Goering, S. Herter, R. J., Larmar, B., Leonard, D., et al. (1996). Collaboration as dialogue: Teachers and researchers engaged in conversation and professional development. *American Educational Research Journal*, 33(1), 193-231.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates: Hillsdale, New Jersey.
- Goldstein, L. S. (2000). Ethical dilemmas in designing collaborative research: Lessons learned the hard way. *Qualitative Studies in Education*, 13(5), 517-530.
- International Baccalaureate Organization. (2002). *What are the three programmes?* Retrieved March 7, 2006, from <http://www.ibo.org/programmes/index.cfm>
- O'Shea, M. R. (2005). *From standards to success: A guide for school leaders*. Alexandria, VA: Association for Supervision and Curriculum Development.