

A Case Study of Key Effective Practices in Ohio's Improved School Districts: Report of Results from Phase I

Aaron Kercheval
Research Associate

July 27, 2001

Indiana Center
for
 **Evaluation**

Smith Research, Suite 100
2805 East 10th Street
Bloomington, Indiana 47408-2601
<http://www.indiana.edu/~iuce/>

Kim K. Metcalf, Ph.D.
Director

812-855-4438
800-511-6575
Fax: 812-856-5890
iuce@indiana.edu

Background

In September of 2000 the Ohio Department of Education issued a request for proposals for a program evaluation project entitled *A Case Study of Key Effective Practices in Ohio's Improved School Districts*. The evaluation was to identify those practices used by school districts that led to significant academic improvement and document those practices in sufficient detail such that other districts could make use of those examples. The RFP specified that the study should employ the Delphi process to identify and rank those practices considered most effective in improving student academic performance.

Following the identification and ranking of the five to 15 most effective practices, the request for proposals called for telephone interviews to further describe the details of implementing practices identified as effective. Finally, site visits were to be conducted to create a detailed description of the practices, steps of implementation, resources needed, and conditions under which the practices were likely to be effective. The emphasis of the evaluation was to be the identification and description of key effective practices in sufficient detail that other districts might be able to replicate those practices.

Proposal

In November 2000 the Indiana Center for Evaluation was awarded a contract to conduct *A Case Study of Key Effective Practices in Ohio's Improved School Districts*. The initial proposal followed the three-phase research design outlined in the request for proposals. However, it was suggested that a modified on-line Delphi process be employed in the first phase of the evaluation. On-line data collection would allow participants to contribute ideas at any time, respond only to those ideas that were most relevant to them, and to see and respond to the contributions of other participants while maintaining anonymity. It would also allow the Indiana Center for Evaluation to continuously monitor contributions, pose guiding questions if necessary, and issue instructions and guidance as needed. The three-stage Delphi process would consist of identifying key effective practices, ranking those practices identified, and responding to the rankings as well as the contributions of other participants.

As specified in the request for proposals, the second phase would use telephone interviews to gather more detailed information about the implementation of the various practices identified in the on-line Delphi process. Having identified effective practices and gathered information about implementation, visits to select districts that would provide the richest examples of effective practices and their implementation would be conducted. An additional purpose of the site visits would be to identify potential sites for production of informational videotapes and other multi-media products to be distributed to school districts throughout the state.

Defining Improvement

Immediately upon award of the contract, questions were raised about the definition of district improvement. The request for proposals had stated that improvement would be

defined by a move to a higher accountability rating on the Ohio Local Report Card. Thus any district that improved its standing on the Report Card from 1999 to 2000, would be included in the study. The total number of districts that did so was 189, roughly one third of all Ohio districts.

This definition of improvement was problematic for the purposes of this study. First, Report Card classifications are such that a district can make very significant progress on proficiency test scores without improving its classification. It is also possible for a district to make a relatively small improvement, but enough to push it into the next classification. Thus, accountability classifications do not necessarily reflect the degree of improvement districts make. Using that measure does not yield a sample of the most improved districts.

Further, it was recognized that improvement is easier or more difficult to demonstrate depending on the size of the district. For a small district, for instance, a change in the test scores of but a few students can have an impact on the overall performance of the district. Also, there is a ceiling effect that makes improvement among districts already performing well more difficult to achieve, while districts performing at the lowest levels have the greatest room to demonstrate improvement.

Lastly, it was recognized that what works for one type of school district may not be effective for a different type of district. It is necessary to consider the contexts in which improvement occurs. Thus it was decided that rather than select the most improved districts statewide, the study would use the most improved school districts within each of the eight district “types” used by the Department of Education. These include (1) small rural, low SES, (2) rural, moderate SES, (3) small town, moderate SES, (4) urban/non rural, low SES, (5) urban/non-rural, high SES, (6) major urban, (7) urban, moderate SES, and (8) very high SES.

Logistical concerns also led to changes in the initial research design. The proposal suggested that two representatives from each of the 189 districts that had moved up one or more accountability classifications would be invited to participate. However, 378 respondents would be an unwieldy number for an on-line discussion. It was reasoned that for the on-line discussion to be successful the number of participants should be limited to 100. Thus, it was decided to use the 50 most improved school districts distributed proportionally across the eight district types. It also seemed more reasonable to look at fewer districts having made the most significant improvement, rather than the nearly one third of districts that had raised their accountability ratings.

Identifying and Contacting Participants

District improvement was calculated using the Standard Unit of Improvement (SUI) formula developed by the Ohio Department of Education. The formula takes into account both the size of districts and their current levels of performance, and determines appropriate expectations of improvement. Districts were ranked according to the degree to which they exceeded these expectations. The highest performing districts in each of the eight district types were selected totaling 50 districts.

The Indiana Center for Evaluation contacted the superintendent of each district by telephone. He or she was informed of the district's selection as one of the most improved in the state and asked to recommend two individuals who would be able and willing to participate in the on-line study of effective practices. Superintendents were asked to identify someone at the district level (e.g. superintendent, assistant superintendent, curriculum director) and someone at the building level (e.g. principal, assistant principal, teacher) who would be in the best position to suggest reasons for their district's improved academic performance. Several superintendents, especially in smaller districts, volunteered to participate themselves. A total of 107 participants were identified. (Some superintendents asked that more than two individuals from their district be allowed to participate.)

The Indiana Center for Evaluation mailed letters to each of the 107 individuals identified by superintendents. The letter notified participants of the improvement their districts had shown, explained the nature of the study, and acknowledged that they had been recommended for participation by their superintendent. Follow-up telephone calls were made to confirm each person's willingness to participate and to verify e-mail addresses.

Next, an e-mail message was sent to all participants initiating the on-line Delphi portion of the study. The e-mail directed participants to the website where their contributions would be collected. Each person was assigned a username and individual password allowing them access to the site and the ability to contribute their ideas anonymously.

The following instructions appeared on the website:

The purpose of the first stage of this process is to identify as comprehensive a list as possible of practices that have contributed to the improvement made in your district. You may visit this site as often as you like during the next week and contribute as many ideas as you like. When you hit **submit** your ideas will be posted to a bulletin board. At the end of the week you will be able to read what participants from other districts have identified as effective practices for them.

Further, participants were asked to make separate entries for each practice that they believed contributed to their district's improvement. For example, a participant might enter "On-Line Proficiency Tests" in a title box, then describe how teachers made use of previous proficiency tests to familiarize students with the format of the tests, and then "submit" that as an entry. They would then continue to identify, describe and submit factors affecting improvement in the same manner.

Following the first round of submissions, the Indiana Center for Evaluation categorized the responses and asked participants to rank categories of practice in order of importance for improving student academic performance. The following instructions explained the purpose of the second round of the Delphi process:

The purpose of this stage is to prioritize, evaluate, reflect upon, and discuss the practices identified in the first phase. What appears below is a synthesis of those practices. Please prioritize those practices by ranking them according to how important you believe them to be in achieving improved school performance.

After submitting rankings, participants were able to view all entries from the first round, as well as the rankings. Following this ranking process, participants received the following instructions:

Now, we would like to find out more about *how* you have implemented these practices. We ask that you provide specific examples of *how* your district has aligned the curriculum with proficiencies, for instance. What specific steps has your district taken to track student performance and provide remediation? What specific professional development activities contributed to the success in your district?

Also posted on the website at this time were links to research literature relevant to those practices. Participants were encouraged to view the research provided and identify other research that informed their practice.

The third stage of the Delphi process was initiated with the following instructions:

Before moving on to the final stage of this Delphi process, we ask that you once again rank the items identified as contributing to the success of your district as measured by proficiency test scores. Having seen what your colleagues have said, some literature about the strategies identified, and how they were ranked at the end of stage I, you may decide to change your ranking or keep it the same.

Having ranked these items again, we have one final and crucial question to which we need your responses. We need to know how the strategies identified thus far have caused an increase in student test scores. While mapping and aligning the curriculum to state standards, for instance, may be a critical first step, alone it cannot improve test scores. Curriculum alignment must ultimately affect what goes on in classrooms. Please be as specific as you can about how the various improvement strategies identified in this study have affected student learning.

No ranking took place at the end of this final round of submissions. E-mails thanking all participants for their contributions signaled the end of the on-line Delphi process.

Results of Delphi

The first round of the Delphi process yielded 300 entries from 62 discreet contributors. Several factors may have contributed to the lower than ideal response rate. The Indiana Center for Evaluation received several calls from participants who were having difficulty logging on to the site. Some had deleted their user names or passwords. Others were

simply having difficulty navigating the site. Others called to say they wanted to participate, but had not had time to do so. Because of a low response rate by the end of the period designated for the first stage and because of technical confusion, participants were notified by e-mail that the time to contribute would be extended. It was at the end of this extended period that 62 (53%) of the 107 identified participants had posted to the site.

Key effective practices identified in this stage were grouped according to the subject headings that participants had entered for them.

Curriculum Mapping and Alignment

The most often mentioned category of practice that participants believed was responsible for improved academic performance in their district was curriculum mapping and alignment. Typical of the comments made were the following:

We completed a curriculum audit to see how well the performance objectives our district identified aligned with the state proficiencies.

We have had our teachers spend much time and effort to match testing competencies/outcomes with subject content and activities for each class.

We have worked with our staff via curriculum mapping to make sure our curriculum lines up with the testing required.

Like these three examples, the statements made about curriculum mapping and alignment tended to be brief and general.

Professional Development

Second in frequency was the identification of professional development activities. In this category, contributors made general statements about staff dedication and increased teacher planning time, but they also identified specific programs to which they attributed district improvement. For instance, the Schmoker process was mentioned several times as was *Target Teach* and Kathy Gardner's *Elements of Effective Instruction*. Teacher training in the use of several reading programs was identified as instrumental in improving the academic performance of districts. Specifically, contributors identified Ohio Reads grants, Reading Recovery, participation in the Summer Institute for Reading Intervention, and use of the Four Blocks approach to literacy.

Emphasis on Literacy

While this emphasis on literacy showed up under the subject heading of professional development, it was also identified as a subject heading of its own. Under the heading of literacy, participants identified many more programs that had been implemented in their districts and which they felt contributed to academic improvement. In addition to the specific programs mentioned above, Success for All, Cooperative Integrated Reading and Composition, Literacy Collaborative, Power Writing, Reading Excellence Act grants, and

Richard Allington's *Schools that Work, Classrooms That Work*, mentioned, along with the use of methods such as guided reading and the scheduling of family reading nights.

The following statement illustrates the ways districts emphasized literacy:

Our district has worked extensively with Beverly Eisele to implement the 8 components of a Balanced Literacy program as our instructional practice. Leveled reading librarians have been developed for grades K-4 in each of our elementaries. In grades K-2 a focus on reading, writing and math has been established to improve our students' foundation for learning. Science and social studies are taught through reading for grades K-2.

Other general comments included statements like, "elementary teachers have focused intensively on the eight elements of literacy," and "Reading Recovery is used in every elementary building in addition to chapter and district funded remedial reading programs." One person stated that, "during the past two summers we have had approximately 80% of our teachers of reading enrolled in the Summer Institute for Reading Intervention. In another district, involvement in professional development was mandatory for all teachers.

Data Analysis and Tracking

Also key among the categories of effective practices identified in the Delphi process was the use of data analysis and tracking of student performance. Many participants stated that the district had begun to use test scores to track individual students' strength and weaknesses. They reported using subscale analysis of individual student proficiency test scores to guide remediation efforts. One district went so far as to engage students in the process.

Students individually analyze their performance on our state proficiency tests. Using technology, each child maps their scores in each academic area, for each year. Students then compare and determine their personal strengths and weaknesses. By reviewing their bar graphs students readily see how close they are to being proficient in EACH area, every individual strand.

Districts report using a variety of technologies to manage data and track student performance. One district referred to use of School Monitoring System software while another claimed to have developed a "CD used for item analysis that allowed teachers and administrators to effectively find areas that were problems." Another district created a Director of Technology position specifically for the purpose of tracking student performance. The increasing role of data analysis is made clear by the following comment:

Having all staff review the data had been an eye opener. We started this process about 3 years ago. We published our progress to staff, parents, and community members. All critical district decisions are made based on an analysis of appropriate data.

Intervention/Remediation

Related to the tracking of student academic performance data, was the development of intervention and remediation programs. Contributors to the study emphasized the identification of at-risk students and the use of summer school, proficiency test review sessions, homework assistance, and a program called Buckle Down Ohio to help students who did not pass the proficiency tests the first time.

Most of the intervention strategies cited were focused on helping students pass the proficiency tests. For instance:

We provided summer proficiency intervention in the areas of science, math, and citizenship. ..The approach used was emphasis on test taking strategies and not so much on teaching specific content. This proved to be very successful.

and

[An] intervention period was established as part of the regular day. Teachers are responsible for identifying students in need of proficiency intervention. This changed a study hall context into a specific opportunity for students to receive additional instruction.

Thus the subject heading of Intervention/Remediation used by participants might well be combined with the subject heading of Test Preparation.

Test Preparation

Emphasizing test preparation was also among the top categories of practice that participants identified as having improved performance.

First thing every morning the children are given a 10-minute practice test to get them used to responding to the directions and format of the test.

And from another district:

Teachers at the 4th and 6th grade levels developed intervention packets for all students to take prior to taking the 4th and 6th grade tests. The material modeled the format of the 4th and 6th grade tests.

The keys to effective test preparation, according to the respondents in this study, appear to be the teaching of test-taking strategies, the use of parallel, or mirror, proficiency tests to provide students with familiarity with the instructions and format of the test, ample opportunity to practice test-taking, and controlling testing conditions. One district held proficiency pep rallies, inviting the high school cheerleaders to provide some inspiration and excitement. "This sets a positive tone for the importance of a good effort from all students," said the respondent from this district.

Tutoring and volunteers

Mentioned somewhat less frequently was the use of tutors and volunteers. Again these references were mostly related to literacy and several established programs and sets of materials were noted including, Blast-Off, Show what You Know, and L.A.M.P. (Literacy and Mentoring Partner). Districts said they offered tutoring before during and after school, using teachers, trained volunteers, and other students.

Other categories of effective practices

Under the subject heading of Instructional Practice, participants in the study identified hands-on learning, developmentally appropriate curricula and differentiated instruction to accommodate personal learning styles as having contributed to academic improvement. Under the subject heading of Instructional Organization a few participants believed that block scheduling, interdisciplinary and thematic units of instruction, multi-age classrooms and looping were responsible for improvement.

Another issue was raised several times – that of “buy-in.” A number of people suggested that everyone from the school board, superintendent, principals, and teachers down to students and parents had to accept improved proficiency test scores as the defining goal of their efforts. “The Stakeholders (students, parents, business personnel, local government personnel, senior citizens, and senior citizens) must “buy-in” to the importance of meeting the standards of the State’s 27 criteria areas, thus making a commitment of supporting, promoting, assisting in the design of programs and procedures designed to meet the established criteria.” As another person said:

Clear Vision and Mission are integral for success in improvement. Staff must know the mission and see relevance to them and to their students. [It’s] easy to keep all on task with a clear mission.

A variety of other isolated ideas were suggested. Some felt that their district had improved because of a commitment to arts programs, which challenge students to excel in other areas. Another district used art, music and physical education as a way of optimizing instruction in proficiency skills. For instance, batting averages were calculated in physical education classes. Improved school safety through peer mediation program, student identification badges, lock-downs, and anti-ridicule policies were credited for having contributed to success. Various forms of celebration and incentives were also mentioned as being important. Students’ performance on proficiency tests were rewarded with days off, t-shirts, scholarship money, waiver of parking fees, and elimination of class final exams.

Ranking Effective Practices

At the end of this stage of the Delphi process, participants were asked to rank these categories of effective practices according to how responsible they believed them to be for their district’s improved academic performance as indicated by proficiency test results. The most highly ranked practiced was mapping and alignment of curriculum with

state proficiency standards. According to participants in this study, mapping and aligning the curriculum was the single most important factor in improving academic performance as measure by proficiency test scores.

Rounding out the top five practices deemed to have the greatest effect were, intervention and remediation programs, data analysis and tracking of performance, professional development, and test preparation. It should be noted, however, that these categories are not mutually exclusive. Entries were made under the heading of literacy, but an emphasis on literacy is also seen in practices identified under the headings of professional development, tutoring and volunteering, and intervention and remediation programs. Efforts to improve literacy as a means of improving over all performance on proficiency tests were apparent across the categories created to facilitate ranking in the Delphi process.

Delphi – Second Stage

In the second stage of the Delphi process, participants were asked to elaborate on *how* they implemented the practices they identified in the first stage. Participation dropped off significantly with only 72 responses. Two people who had not contributed to the first stage did so in the second stage. Most notable among the elaborations provided in this stage is the emphasis on data driven change. Districts report having used proficiency test results to determine changes in curriculum, design intervention and remediation, and to determine the focus of professional development activities. The following statement is evidence of the extent to which successful districts have focused on proficiency test results:

We have made some tough decisions to cut out things that are nice, but not necessary. When an interesting in-service comes along, we give it the litmus test to see if it is nice or necessary to help us hit our targets. We do the same with the experiences we give students. Is there is evidence that shows this activity helps students understand a concept or, or is it just fun to do?

Another participant offered that each teacher in the district codes his or her lesson plans according to the proficiency the lesson is intended to develop. Many of the teacher development activities identified in this stage were related to helping teachers use data to design class lessons, remedial activities, and assessments. As one person said, “Our professional development is data driven.”

When asked to rank once again the practices that had been identified, and to some extent elaborated upon in the second stage, participants’ rankings did not vary notably. Given the decrease in participation in the ranking from the first round to the second, it would be difficult to give much weight to changes if they occurred. However, some participants did indicate that they had looked at what others had posted to the site and the literature provided by the Indiana Center for Evaluation. A few identified websites with literature that guided their district’s improvement efforts.

Final Stage

In the final stage of the on line portion of the study, participants were asked to explain *how* the practices they had identified as effective had been translated into classroom practice. All participants were administrators either at the district or building level. All had been to some degree the architects of their districts improvement. However, as was noted in the instructions at the beginning of stage three, changes must manifest themselves at the classroom level if they are to effect student performance.

Having ranked these items again, we have one final and crucial question to which we need your responses. We need to know how the strategies identified thus far have caused an increase in student test scores. While mapping and aligning the curriculum to state standards, for instance, may be a critical first step, alone it cannot improve test scores. Curriculum alignment must ultimately affect what goes on in classrooms. Please be as specific as you can about how the various improvement strategies identified in this study have affected student learning.

Only 32 entries were made in the final stage. It should be noted that the end of the on-line Delphi process corresponded with the end of the school year and that timing may have been a factor in declining participation through the stages. However, other factors may also have affected participation. As one person said when contacted by phone, “ I’m damned if I do and damned if I don’t. If my school does well they want to study me. If my school does poorly they want to study me. I give up.” Clearly educators feel that participating in research initiated from outside is an added burden to an otherwise overflowing schedule. At the beginning of the third stage of the Delphi process, the Indiana Center for Evaluation received an e-mail from some who had not contributed to the first two stages, stating tersely that he had never agreed to participate in the first place. Another factor that may have affected participation might be individual’s comfort level in using computers and accessing websites. Although superintendents were asked to recommend participants who were willing and able to participate on-line, the Indiana Center for Evaluation did receive numerous calls from participants requesting technical assistance.

Finally, there is one last factor to consider. It is possible that some participants may not know what is responsible for the improvement made by their district. For example, one participant simply noted that, “our students are bright, respectful, and responsible,” and that, “our student population has little turnover – people move here and they stay.”

Focus of future phases

Several categories of effective practice emerged from the first phase of this study. Representatives of the districts in Ohio that have shown the most significant improvement in test scores believe that aligning what is taught with what is tested is the single greatest determinant of improvement. They also believe that professional development is key. Districts that are most improved also appear to rely heavily on data, using data analysis and tracking of performance to inform everything from curriculum

changes to remediation activities to profession development. Also emphasized across the categories was the importance of building literacy.

What remains to be determined in successive phases of this study is how these general principals have been translated into classroom practice. As noted previously, participants in the Delphi process were administrators at both the district and building levels. While there was significant reference to the work of teachers and “buy-in” from everyone involved, it will be important to describe the ways in which these key effective practices have manifested themselves at the classroom level. This may require that telephone interviews in the next phase of the study not be limited to those who participated in the first phase but also include teachers. Teachers are in the best position to explain classroom changes. It is also possible that what teachers report may not be perfectly aligned with what administrators reported, and that other clues to improvement may be revealed.

It will also be important to follow up by telephone with those who were identified by superintendents as potential participants but who did not contribute to the on-line Delphi process. Contacting these individuals may confirm some of the reasons for non-participation suggested earlier, but, more importantly, it may provide different information about how improvement has occurred in those districts.

Another critical task for subsequent phases of this study is to describe in detail the steps districts have taken to implement the practices they identified as most effective. Who was involved in the process? How long did it take? What had to happen in order for teachers to successfully align the content of their instruction with proficiency outcomes? How were resources for professional development obtained? How was it decided that certain commercial curricular packages or computer software programs would be purchased and used? How were new practices communicated to teachers, students and parents? It is also important to document those actions taken by districts that proved ineffective and why they did not succeed. What might these most improved districts suggest were missteps or unproductive attempts to reach their goals? Also, what are the necessary conditions and resources for successful implementation of the practices that have proven to be so successful at improving student performance on proficiency tests.

While focusing on the specific steps different districts have taken, it is important not to simply create a set of directions for others to follow, but rather to glean from the details of implementation the principles that guide successful implementation. Other districts cannot simply copy what successful districts have done, but must understand the mechanisms that lead to success and decide how best to apply them in their own contexts.

Regarding the importance of context, the data collected in this first phase of the study awaits further analysis to determine if there is a relationship between the kinds of practices identified as effective and the types of districts in which they were implemented. Districts need to know what has been effective for other districts like them. Data driven practice may not look the same in a large district than in a small one. A large

district may employ a Technology Coordinator while a small district, which cannot afford added personnel, may rely on the initiative of individual teachers.

A fairly clear picture of what Ohio's most improved school districts are doing has emerged from the first phase of this study. However, the picture needs greater detail, if the processes of implementing practices successfully are to be understood and adapted to different districts around the state.