Pilot Test of the Instruction and Learning Appraisal Process

Patricia E. Ceperley, Edvantia
Georgia K. Hughes, Edvantia
Kavita Mittapalli, Research Fellow

December 2005

Appalachia Educational Laboratory (AEL)
at

EDVANTIA™
Partners in education. Focused on results.

Pilot Test of the Instruction and Learning Appraisal Process

Patricia E. Ceperley
Georgia K. Hughes
Kavita Mittapalli

December 2005

The Appalachia Educational Laboratory (AEL) at Edvantia
P.O. Box 1348
Charleston, WV 25325-1348

Edvantia was founded in 1966 as the Appalachia Educational Laboratory, Inc. (AEL); on September 1, 2005, AEL became Edvantia, Inc. The Regional Educational Laboratory for the Appalachian region is known as the Appalachia Educational Laboratory at Edvantia.

Edvantia is a nonprofit education research and development corporation, founded in 1966, that partners with practitioners, education agencies, publishers, and service providers to improve learning and advance student success. Edvantia provides clients with a range of services, including research, evaluation, professional development, and consulting.

For information about Edvantia research, products, or services, contact

EDVANTIA™

P.O. Box 1348, Charleston, WV 25325 • 304.347.0400 • 800.624.9120 • fax 304.347.0487
One Vantage Way, Suite D-210, Nashville, TN 37228 • 615.565.0101 • fax 615.565.0112
info@edvantia.org • www.edvantia.org

Edvantia maintains policies and procedures that govern the Institutional Review Board in ensuring the welfare and protecting the rights of human subjects who choose to participate in research and evaluation activities.

This publication is based on work sponsored wholly or in part by the Institute of Education Sciences (IES), U.S. Department of Education, under contract number ED-01-CO-0016. Its contents do not necessarily reflect the positions or policies of IES, the Department, or any other agency of the U.S. government.

Edvantia is an equal employment opportunity/affirmative action employer.

Contents

List Of Tables

Introduction

The purpose of this report is to document the pilot test of the Instruction and Learning Appraisal (ILA) and describe the quality of the ILA process. The ILA process was developed by Edvantia staff who serve as technical assistance providers to schools and districts in the region. Low-performing schools and districts (i.e., those that fail to achieve adequate yearly progress) want help with determining the problems that underlie their inadequate progress and with identifying solutions to those problems.

The appraisal process uses a case study approach to provide the desired assistance. According to Yin (2003),

> A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries and the phenomenon and context are not clearly evident. (p. 13)

The case study approach to helping schools permits technical assistance providers to seek answers to the question: How and why is the school's educational process failing to improve the school's performance on state assessments? The theoretical framework that guides the appraisal process is based on factors found in high-performing schools. The framework has also guided the school improvement work of the Appalachia Educational Laboratory (AEL) at Edvantia and has led Lab staff to develop propositions about the factors that may explain why a school is not successful (see Table 1). The ILA process uses the theoretical framework depicted in Table 1 for collecting information at a school. For example, if shared leadership is found in high-performing schools, researchers may propose that the lack of shared leadership may contribute to low school performance. The appraisal process looks specifically for evidence of shared leadership. If evidence of shared leadership is not found, researchers may conclude that the lack of shared leadership may explain, to some extent, a school's poor performance. Data collected during an appraisal can be used to support such a conclusion.

Table 1: Theoretical Framework Explaining School Performance

Factors Found in High-Performing Schools	**Related Propositions Explaining Low School Performance**
Shared leadership	School principal has poor leadership skills. School principal does not share school leadership with teachers.
Shared goals	School faculty do not have a shared vision of how the school will improve. There is no focus on improvement.
Learning culture	Teachers have no opportunities to learn new skills. Teachers are not working together to explore ways to improve performance.
Aligned curriculum	Teachers do not work together to ensure a well-articulated, aligned curriculum.
Effective teaching	Teachers exhibit poor teaching strategies and classroom management skills.
Purposeful assessment	School staff are not monitoring student progress to ensure mastery.
Family and community involvement	Little evidence of family or community involvement is seen.

A variety of tools is used during the appraisal process, including document analyses, interviews, and observations. The information collected is directly linked to the propositions about school performance. The factors that define high-performing schools serve as the criteria for interpreting the findings at each school.

Because of the need to conduct an intensive site visit at the school(s), the ILA is a labor intensive, time-consuming process and must be scheduled well in advance of the visit. When a request to conduct an ILA is received, a team of trained appraisers is pulled together to conduct the appraisal. The composition of the team may vary depending on availability of trained appraisers. Thus, it is imperative to ensure that the composition of the team will not influence the outcome of the appraisal process. Equally important, all of the elements of the ILA process—appraisers, tools, and procedures—must work effectively to produce a high-quality appraisal. To test the quality of the appraisal process, a pilot test of the ILA was conducted by research staff of the Appalachia Educational Laboratory at Edvantia. Four tests of quality were relevant to the pilot test. They are construct validity, internal validity, external validity, and reliability. Researchers sought evidence of each of these during the pilot test.

Methods

The details of the methods and findings of this pilot test project are divided into three phases: (1) pre-pilot test activities, (2) pilot test observation, and (3) ILA Appraisers' Feedback survey. The following section presents methodological information for each phase of data collection. Researchers sought and received the approval of Edvantia's Institutional Review Board (IRB) to collect survey data from the appraisers.

Participants

For two of the three phases (document review and pilot test observations), three lab researchers were the primary data collectors. A senior research and development (R&D) specialist and a research and evaluation (R&E) specialist worked closely with a minority research fellow to develop data collection and analysis protocols, collect the information, and analyze and interpret findings.

For the ILA Appraisers' Feedback, appraisers who participated in the pilot test process served as the respondents. In all, 16 appraisers (15 trained appraisers and one leader/developer) took part in the pilot test. Two lab researchers were involved as participant observers. All appraisers were either Edvantia staff or were semiretired or retired school- or district-level educators.

Instruments

For the pre-pilot test activities and pilot test phases, researchers recorded observations and notes for the purpose of later analysis. Thus, no formal data collection instruments were used for those phases. For the third phase, researchers developed the ILA Appraisers' Feedback survey to assess appraisers' perceptions of the clarity, comprehensiveness, adequacy, and ease of use for the ILA data collection instruments and processes. In all, the instrument contained 66 items (see Appendix for a copy of the instrument). Appraisers were instructed to rate the extent to which each item was true using a 6-point Likert-type scale ranging from 1 (*not at all true*) to 6 (*completely true*).

Data Collection

In January 2005, researchers observed a two-day training session for ILA appraisers and took notes about how the training process was conducted and the appraisers' involvement. Further, the first day of the pilot test visit included a session to refresh appraisers' understanding of instructions. Researchers also participated in that event.

Beginning in February of 2005, lab researchers began reviewing the Instruction and Learning Appraisal (2004), a manual that contained a description of the process and all data

collection instruments. All sections of the manual were reviewed: (1) district/school administrator interview; (2) identify goals/focus/purpose of appraisal; (3) document review; (4) classroom observations; (5) focused walkthrough; (6) analysis of student academic work; (7) interviews with teachers and administrators; and (8) building consensus, aggregating and analyzing data, and drawing conclusions and recommendations. During the document review, researchers took notes concerning the clarity and completeness of descriptions and instructions. The document review process concluded in early April of 2005.

Researchers participated in the 5-day pilot test site visit. On the first day, appraisers met with Edvantia staff, who were charged with leading the ILA process for the pilot test. The purpose of the meeting was to review and refresh training in the ILA processes and to establish the specific focus of the ILA process in response to the needs of the district. In addition, the pilot test included 3 days for school visits and one day for a debriefing with district staff. During the school visits, the senior R&D specialist and the minority research fellow observed the data collection as well as portions of the analysis and interpretation. Throughout the week, research staff kept careful field notes and records of their observations.

At the close of the pilot test, appraisers were asked to complete the Pilot Test Appraisers' Feedback survey (included in the appendix). Feedback surveys were given to all 16 appraisers at the end of the last day of the pilot test. Appraisers were asked to complete only those sections of the survey that corresponded with the sections of the ILA for which they had collected information. Completion of the questionnaire required approximately 30 minutes.

Analysis

Data collection related to the pilot test was conducted in three phases. The procedures and criteria for data analyses for each phase of the pilot test are described below.

Pre-Pilot Test Activities. Prior to the pilot test, lab researchers participated in appraiser training sessions and conducted a review of the ILA manual. For the training sessions, researchers used Yin's (2003) explanation of case study seminar training (p. 63) to guide the analysis of the training. According to Yin, training should

> take the form of a seminar rather than rote instruction. . . . much time has to be allowed for reading, preparing for the training sessions, and the sessions themselves. . . at least a week's worth of preparation and discussions. The seminar will cover all phases of the the planned case study investigation, including readings on the subject matter, the theoretical issues that led to the case study design, and case study methods and tactics. (p. 63)

Yin also explains, "The goal of the training is to have all participants understand the basic concepts, terminology, and issues relevant to the study" (p. 63). Researchers also used Yin's list of what each investigator needs to know as criteria for evaluating the training sessions. That list includes the following:

- Why the study is being done
- What evidence is being sought
- What variations can be anticipated (and what should be done if such variations occur)
- What would constitute supportive or contrary evidence for any given proposition (p. 63)

For the manual review, the three researchers systematically examined the ILA manual. The criteria used to assess the quality of the manual were drawn from the tactics advocated by Yin (2003). For the data collection phase of case study research, Yin suggests the following tactics: use multiple data sources, establish a chain of evidence, use a case study protocol, and develop a case study database (p. 34).

Pilot test observation. For the pilot test participation-observation phase, researchers participated as appraisers in the ILA process. They recorded observations and notes for the purpose of later analysis. In addition, they debriefed their experiences each day. Criteria for assessing the week-long process included those used for the pre-pilot test activities. In addition, researchers observed the extent to which the appraisers followed a case study protocol.

For each of the tests of quality for case studies Yin (2003) developed a list of tactics to be used to ensure the quality. Table 2 replicates Yin's table that lists the tactics and the phase of research in which the tactic is to be used. Researchers used Yin's list of tactics to guide the analyses of the ILA process. Researchers' notes from the document/manual review and from observations during the pilot test (including documentation of the ILA process) were reviewed by research staff for evidence of the use of case study tactics.

Table 2: Case Study Tactics for Four Tests*

Tests	**Case Study Tactic**	**Phase of Research**
Construct validity	• Use multiple sources of evidence • Establish chain of evidence • Have key informants review draft case study report	data collection data collection composition
Internal validity	• Do pattern-matching • Do explanation-building • Address rival explanations • Use logic models	data analysis data analysis data analysis data analysis
External validity	• Use theory in single-case studies • Use replication logic in multiple-case studies	research design research design
Reliability	• Use case study protocol • Develop case study database	data collection data collection

*Yin, 2003 (p. 34)

Pilot Test Appraisers' Feedback Survey. Responses to the feedback survey were aggregated and organized for reporting. Analysis of data was used to triangulate information with the other data collection techniques used for the pilot test.

Findings

Information about the Instruction and Learning Appraisal (ILA) was collected in three phases: (1) pre-pilot test activities, (2) pilot test observation, and (3) ILA Appraisers' Feedback survey. Results for each phase of data collection and analysis for the pilot test are presented in this section.

Pre-Pilot Test Activities

Two pre-pilot test activities were conducted. The first was participation in the training of ILA appraisers. The second was the review and analysis of the ILA manual (AEL, 2004).

Appraiser training. In all, training for the ILA appraisers totaled 2 ½ days. The format involved some direct lecture, but there was a substantial amount of appraiser involvement and discussion about the terminology, concepts, and procedures. Readings were not part of the training, but theoretical issues were described and discussed. The study design, methods, and tactics were presented. Finally, appraisers learned why the appraisal was being conducted, what evidence was being sought, and what would constitute supportive or contrary evidence for each of the propositions related to the school improvement framework. Researchers did not observe any evidence of what variations might be expected with the data collection process and how to handle such variations. This appeared to be left to the professional judgment of the appraisers.

ILA manual review. Researchers reviewed the ILA manual (AEL, 2004), which contained a brief description of the process and a compilation of all data collection instruments and protocols to be used during the ILA process. The three researchers looked for the case study tactics recommended by Yin (2003, p. 34), i.e., a theoretical framework, the use of multiple data sources, methods for establishing a chain of evidence, the use of a case study protocol, and methods for developing a case study database (see Table 2).

Researchers found that a theoretical framework was used to design the ILA process. Further, the ILA manual, especially interview protocols, contained numerous citations for best-available research about each of the seven components of Edvantia's framework. Thus, the documents and instruments appeared to be well-researched and related to the best available evidence for each of the seven components.

Researchers found that the data sources included document review, focused walkthroughs, interviews, reviews of student work, and classroom observations. This variety of instruments and data sources indicated that multiple sources of evidence were used during data collection.

Researchers found that the ILA manual serves as a case study protocol, guiding the appraisers through the data collection process and providing instruments to use for collecting data. A description of data collection procedures and instructions for using the instruments and recording the data were not found, however. The large number of interview questions suggests

that there may be too many questions and that the questions may be too specific to elicit the kind of information that appraisers are seeking. Broader questions would allow respondents to tell their story, which may increase the amount and richness of information collected.

Researchers also found that the instruments and data sources in the manual were designed to help appraisers establish a chain of evidence. The protocols were linked to the theoretical framework and questions to be answered by the ILA process. Instruments were designed to organize data according to the framework.

Finally, researchers found that procedures for the appraisal process and instructions for using the instruments were not included in the ILA manual. This finding was consistent for each of the data collection methods.

ILA Pilot Test Observations

Researchers served as participant observers during the pilot test of the ILA process between April 18 through April 22, 2005. On April 18, Edvantia staff led sessions to refresh appraisers' training in ILA and to establish the focus of the visit. On the afternoon of April 18, ILA leaders met with the district's central office staff and superintendent.

During the review session, the ILA developers/trainers who were involved with both the development and use of ILA reviewed the training manual and data collection instruments with the appraisers who participated in the pilot test. Researchers noted that among the appraisers, some were Edvantia staff and some were external consultants. All appraisers either currently served a school district in some capacity or had retired from working in a school system. Appraisers further had many years of experience as district-level education professionals, and some were experienced with the ILA process. As is standard with ILA methods, no appraisers were employed in any way by the school district participating in the ILA pilot test.

Researchers observed that the ILA developers/trainers referred often to the manual and gave many examples during the review session. The leaders engaged the appraisers by asking many questions and soliciting feedback and suggestions. After the initial review was complete, appraisers were separated into groups of three or four to review each of the seven components of the underlying theoretical framework and to review the ways the concepts were operationalized in the ILA instruments. The small appraiser subgroups then were instructed to select appropriate questions for the school visits, which were to begin the next day. This was part of the process for tailoring the focus of the pilot test district's needs. At the full-group level, attempts were made to build consensus about which indicators would signal technology integration in the curriculum (the focus of the pilot test) as it related to each of the seven framework components. Appraisers were asked for their professional judgment, and differences in professional opinion were resolved by the ILA developers/trainers.

On April 19, 20, and 21, ILA appraisers visited multiple schools. In all, five schools were included (three elementary schools, one middle school, and one high school). Lab

researchers accompanied appraisal teams to different schools to observe the process. Each of the two researchers observed the process at different schools.

On the first day of school visits, researchers noted some confusion among the appraisers about the instructions they had been given and about the nature of their work as ILA appraisers. Appraisers further appeared to be unclear about the definition of certain terms and concepts included in the ILA (e.g., shared leadership, shared goals, and so on). There seemed to be differences across appraisers, however, in the level of confusion and level of confidence in their ability to conduct the ILA. Researchers noted that some appraisers were very experienced and seemed to be less confused about instructions and definitions. Other appraisers who did not have similar long-term experience in school systems or with the ILA appeared to be more confused than their peers. Additionally, appraisers commented that the classroom observation form used was not sufficient to assess all aspects of classroom activities and behavior; nor was the time limit for observations sufficient for appraisers to fully review, comprehend, and assess classroom instruction. At the conclusion of the first day of the school visits, a debriefing session was held, and appraisers took that opportunity to bring up their concerns, ask questions, and get clarification from the ILA developers/trainers managing the site visit.

Researchers noted that the second and third days of school visits were progressively better than the first in terms of appraiser understanding and comfort with the ILA process. Confusion about instructions and definitions was apparently resolved through discussion and debriefing. Further, as the school visits progressed, appraisers seemed to become more familiar with the structure both of the ILA and of the pilot test leadership. Therefore, appraisers were better able to direct questions to appropriate staff and get answers more quickly. Researchers noted that appraisers seemed to be much more comfortable with each other, with the ILA, and with the school environment by the third day of school visits.

Appraisers discussed the evidence they found related to the integration of technology into the learning and teaching processes. Appraisers also discussed the positive aspects of their daily experiences as well as the challenges they faced. They specifically focused on their use of ILA processes and instruments. For instance, appraisers discussed their perceptions that the time limits established for classroom observations and focused walkthroughs were not long enough for them to gain an accurate understanding of classroom or school practices or characteristics. They also discussed their difficulties using various ILA instruments (e.g., classroom observation forms, interview protocols, and so on). The Edvantia leaders discussed possible solutions with the group and reinforced appraisers' successes.

Other topics for discussion during debriefing meetings included appraisers' general observations about the school, their perceptions of school and/or district functioning, and other information or professional judgments they gained or formed as a result of their school visits.

After all data were collected, ILA developers/trainers synthesized and summarized the ILA findings and related all data to technology integration. The summarized information was presented to the district's central office staff and superintendent on April 22, 2005, the last day of the visit. Following the site visit, a final, written report was developed and sent to the district.

Table 3 presents the purpose, length, and number of appraisers involved in each of the various steps and processes that were included in the Manassas ILA pilot test.

Table 3: ILA Process Used at the Pilot Test Site

Process	Purpose	Number of people involved*	Length	When
Structured interviews with school/district leaders	To state the purpose of ILA; To clarify concerns and priorities for the sites chosen	5 people (one team leader, instructional technology specialist, two appraisers and a researcher)	About 2 hours	Day 1- before the school visits
Structured interviews with teachers and students	To document perceptions of current practice	13	About 1 hour	Days 2-4
Classroom observations, focused walkthroughs	To obtain "snapshots" of instructional practices and interventions	12	2-3 hours	Days 2-4
Review of relevant school documents	School policies, curriculum frameworks	5	3-4 hours	Days 2-4
Analysis of student achievement data (e.g., grades, test scores)	To identify the relationships among educational inputs (e.g., instructional interventions, policies)	5	3-4 hours	Days 2-4
Baseline report of findings, including recommendations and resources	To guide schools or districts in improvement planning and assessing continuous improvement	2 (team leaders)	About 2 weeks	Within 15 days of the pilot test visit
Debriefing session with school or district leaders	To inform them of the findings and suggest recommendations, if any	5 (team leader, instructional technology specialist, two appraisers, and a researcher)	About 2 hours	The day after all the school visits

* Data from personal observation and survey

ILA Appraisers' Feedback

On the last day of the school visits, appraisers were asked to complete a survey assessing their perceptions of the ILA process. The survey was composed of eight sections, one for each phase of the ILA process. Appraisers were asked to rate each item on a 6-point, Likert-type scale ranging from 1 = (*not at all true)* to 6 = (*completely true*). Results of the appraisers' survey are presented in Table 4.

Table 4: Tally of Responses to the ILA Appraisers' Survey

Items	N	1	2	3	4	5	6	Percent-age Rating Item 5 or 6
1. Terms and concepts use in this process were clearly defined.	76	0	0	1	3	28	44	99%
2. The training I received was adequate for successful implementation of this process.	78	0	0	2	3	27	46	97%
3. Amount of time scheduled for this process was adequate to achieve the purpose of the process (e.g., comprehensive data collection).	74	0	0	2	12	21	39	97%
4. The forms for recording data were easy to understand and use.	73	0	3	1	5	21	43	95%
5. Information gathered in this process was adequate for making meaningful recommendations.	80	0	1	4	13	28	34	94%
6. Information gathered in this process was adequate for drawing meaningful conclusions.	79	0	1	4	13	25	36	94%
7. Instructions for this process were clear and unambiguous.	79	0	2	4	8	29	36	92%
8. Instructions for this process were comprehensive and complete.	79	0	1	7	9	26	36	90%

The percentage of appraisers rating each item 4-6 ranged from 90% to 99%. Item 6 (Terms and concepts use in this process were clearly defined) was the highest rated item (99%). The lowest rated item was Item 2 (Instructions for this process were comprehensive and complete). The percentage of appraisers rating items 1-3 was small, and no appraiser gave a 1 rating to any item.

As part of the research plan for this study, the research fellow who was a participant observer during the pilot test recorded observations during the administration of the appraisers' feedback survey. These observations are presented here. First, the feedback survey was administered at the end of the last day of the site visit, just prior to the final appraisers' debriefing session. The research fellow perceived that some appraisers seemed frustrated at being asked to complete the survey at that time. She surmised that the frustration may have arisen from the length of the survey or the lack of sufficient advance notice that appraisers would be asked to complete it.

Second, the research fellow noted that the instructions for completing the feedback survey were not clear. In designing the survey, researchers had anticipated that all appraisers would participate in all ILA processes; however, appraisers participated in only some of the processes. Therefore, at the time the instrument was administered, researchers requested that appraisers complete only the sections of the survey that corresponded with those ILA processes in which they had participated. For instance, appraisers who participated only in the document review and analysis of student academic work processes were instructed to complete only Sections III and VI of the survey. However, the research fellow noted that there seemed to be some confusion among the appraisers about which portions of the survey they were to complete. She observed that some appraisers may have responded to sections corresponding to processes in which they had not been involved. Some of these appraisers may have participated in previous appraisals and therefore may have considered themselves qualified to complete those sections.

Finally, the research fellow observed that the 16 appraisers who completed the questionnaire included one of the ILA developers/trainers.

Conclusions and Recommendations

The following section presents conclusions and recommendations based on researchers' findings during the ILA pilot test site visit.

Appraiser training. According to Yin (2003), training is an essential part of ensuring that a case study such as the ILA process is carried out as intended. Nearly all (97%) of the appraisers indicated on the survey that they felt the training was adequate. However, researchers noted some confusion among the appraisers about the procedures to follow on the first day of school visits. Since the one-day training session is substantially less than the week-long training seminar suggested by Yin, the training may need to be lengthened. Training may also need to include participation in the appraisal process as a "trainee" prior to becoming a fully qualified appraiser.

ILA manual. The review of ILA documents and instruments revealed that all included multiple citations of best-available evidence. The development of instruments and individual items founded on a sound literature base indicates that developers took into consideration contemporary conceptualizations and multiple facets of each component. Additionally, multiple approaches and perspectives for assessing each of the seven components of the underlying theoretical framework are evident in the ILA materials, especially the interview protocols, which contain questions designed for many different facets of each component. Linkages of the data collection instruments to the theoretical framework enhance the opportunity for appraisers to establish a chain of evidence. Using multiple questions and approaches for assessing each component (also called triangulation) ensures that most or all aspects of a construct will be assessed, thereby reducing the mono-operation bias (Cook & Campbell, 1979) and increasing the construct validity of the process (Yin, 2003).

Finally, the ILA manual serves as a case study protocol, which increases the reliability of the information gathering process. Although the ratings related to instructions were lower that other responses, most appraisers (92%) thought the instructions for the process were clear and unambiguous. In addition, most (90%) also thought that the instructions were complete and comprehensive. Nearly all (99%) agreed that the terms and concepts used in the process were clearly defined, and 95% reported that the forms for recording data were easy to understand and use. The high level of agreement may reflect the extensive experience of the appraisers recruited for the ILA process. It is unclear whether less-experienced appraisers would report the same level of agreement with these statements. To ensure the high level of reliability of appraisals is maintained, developers may want to consider the addition of procedures and instructions for collecting and recording data.

ILA Pilot Test Observations

With the exception of some confusion observed during the first day, the ILA process appeared to go smoothly. Appraisers carried out the data collection at the schools without notable problems. At the end of the day, appraisers returned to debrief and create a database of

information that would later be used to write the report. Although the appraisers appeared to understand the processes and instructions for aggregating and analyzing data at the end of each day, as indicated previously, researchers did not find any written instructions for this process.

Ensuring that all appraisers have a precise, thorough, and shared understanding of the definitions of terms and concepts used in the ILA is an important step in ensuring reliable data collection and analysis. Nearly all (99%) ILA appraisers felt that the terms and concepts for each data collection process were clearly defined.

Researchers also observed the building of the database at the end of each day's school visit. Appraisers were able to reach consensus on the findings at each school using the various data collection instruments. Further, most (94%) agreed that the information they gathered was sufficient for drawing conclusions and for making recommendations. Researchers did not observe the report-writing process, so it is not clear whether key informants had the opportunity to review the draft report to the district.

This initial pilot test of Edvantia's ILA was conducted to determine the extent to which ILA processes and instruments are reliable and valid for the intended use. The findings from the ILA pilot test were compared to Yin's tactics to determine the extent to which the ILA met the tests of quality (see Table 5).

Table 5: Determining the Quality of ILA

Tests	**Case Study Tactic**	**Phase of Research**	**ILA Findings**
Construct validity	• Use multiple sources of evidence • Establish chain of evidence • Have key informants review draft case study report	data collection data collection composition	• ILA used several data sources including interviews, observations, and document review • ILA data collection tools guide the establishment of a chain of evidence • Researchers collected no data related to this tactic
Internal validity	• Do pattern-matching • Do explanation-building • Address rival explanations • Use logic models	data analysis data analysis data analysis data analysis	• Researchers did not find any evidence of pattern-matching, explanation building, addressing rival explanations, or use of logic models • Tactics may have been used in the report-writing phase
External validity	• Uses theory in single-case studies • Uses replication logic in multiple-case studies	research design research design	• Researchers found extensive evidence related to the theoretical framework for school performance • Replication logic was not relevant for this single case study.
Reliability	• Use case study protocol • Develop case study database	data collection data collection	• Researchers found that the ILA manual served as the case study protocol. • Appraisers developed a case study database each day as they debriefed.

As Table 5 indicates, substantial evidence was found to support the conclusion that the process is reliable and has construct and external validity. No evidence was found related to internal validity. However, they may have been used in summarizing and synthesizing the data during the report writing process, which was not part of the pilot test.

Appraisers typically feel that the information collected in each section is sufficient for making meaningful conclusions and recommendations to influence school improvement. Much of the success of the appraisal process may be attributed to the high quality and level of experience of the appraisers. Suggestions have been made that would increase the reliability and validity as new appraisers are recruited and trained. In summary, researchers make the following recommendations:

- Lengthen the training process and/or include trainees in appraisal processes to increase their familiarization with the terminology, process, and tools for data collection.
- Include more information in the ILA manual rather than relying on the training sessions or the professional knowledge and experience of the appraiser to impart that information.
- Develop written procedures for the school visits to minimize any confusion during the appraisal process.
- Develop written procedures for developing the database of findings at each site to increase the consistency of the process and decrease the reliance on the skills of the appraisers.
- Develop tools such as rubrics or Innovation Configuration maps to guide appraisers in the analysis and interpretation processes to increase the internal validity of the data analysis.

References

AEL. (2002). *AEL's framework for research-based school improvement.* Charleston, WV: Author.

AEL. (2004). *Instruction and learning appraisal (ILA)*. Charleston, WV: Author.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.

Yin, R. K. (2003). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Appendix

AEL Instruction and Learning Appraisal (AEL ILA)
Pilot Test Appraisers' Feedback

Recently, you participated on an AEL ILA team conducting a school or district appraisal. We want your feedback about the AEL ILA process, particularly data collection, analysis, and consensus building. Your feedback will help us refine and improve the AEL ILA and how we conduct each appraisal. Each section of this questionnaire addresses a different phase of the AEL ILA process. Please think about your experiences with the stated phase of the appraisal, and answer each question honestly. Please use the 6-point scale to rate the extent to which each item is true of your experience with the AEL ILA, with "1" representing "Not at all true" and "6" representing "Completely true." Circle/highlight the number that indicates how true each item was in your experience with the AEL ILA.

Not at all true		**Somewhat true**		**Completely true**	
1	2	3	4	5	6

AEL ILA Data Collection Procedure:
I. District/School Administrator Interview

Instructions for data collection through the *superintendent interview* were clear and unambiguous.	1	2	3	4	5	6
Instructions for data collection through the *superintendent interview* were comprehensive and complete.	1	2	3	4	5	6
Information gathered through the *superintendent interview* was adequate for determining the focus of the ILA.	1	2	3	4	5	6
Information gathered through the *superintendent interview* was adequate for determining indicators to look for during focused walkthroughs and classroom observations.	1	2	3	4	5	6
The amount of time scheduled for the *superintendent interview* was adequate for determining questions to ask during teacher and administrator interviews.	1	2	3	4	5	6
Terms and concepts used in the *superintendent interview* section of the AEL ILA instructions and indicators were clearly defined.	1	2	3	4	5	6
The training I received was adequate for collecting and recording relevant and useful information during the *superintendent interview*.	1	2	3	4	5	6
The forms for recording data during the *superintendent interview* were easy to understand and use.	1	2	3	4	5	6

AEL ILA Data Collection Procedure:
II. Identify Goals/Focus/Purpose of Appraisal

Instructions for the *goal identification* process were clear and unambiguous.	1	2	3	4	5	6
Instructions for the *goal identification* process were comprehensive and complete.	1	2	3	4	5	6
The *goal identification* process was adequate for reaching a consensus among appraisal team members about instructional indicators to look for during Focused Walkthroughs and Classroom Observations.	1	2	3	4	5	6
Information gathered through the *goal identification* process was adequate for selecting questions to ask during the teacher and administrator interviews.	1	2	3	4	5	6
Information gathered through the *goal identification* process was adequate	1	2	3	4	5	6

for deciding what indicators to look for in student records and student work.						
The amount of time scheduled for the *goal identification* process was adequate for comprehensive data collection.	1	2	3	4	5	6
Terms and concepts used in the *goal identification* process of the AEL ILA instructions and indicators were clearly defined.	1	2	3	4	5	6
The training I received was adequate for successful participation in the *goal identification* process.	1	2	3	4	5	6
The forms for recording data were easy to understand and use.	1	2	3	4	5	6

AEL ILA Data Collection Procedure:
III. Document Review

Instructions for collecting data from *documents* were clear and unambiguous.	1	2	3	4	5	6
Instructions for collecting data from *documents* were comprehensive and complete.	1	2	3	4	5	6
Information gathered through the *document review* was adequate for drawing meaningful conclusions.	1	2	3	4	5	6
Information gathered through the *document review* was adequate for making meaningful recommendations.	1	2	3	4	5	6
The amount of time scheduled for the *document review* was adequate for comprehensive data collection.	1	2	3	4	5	6
Terms and concepts used in the *document review* process of the AEL ILA instructions and indicators were clearly defined.	1	2	3	4	5	6
The training I received was adequate for successful implementation of the *document review* process of the AEL ILA.	1	2	3	4	5	6
The forms for recording data were easy to understand and use.	1	2	3	4	5	6

AEL ILA Data Collection Procedure:
IV. Classroom Observations

Instructions for data collection through *classroom observations* were clear and unambiguous.	1	2	3	4	5	6
Instructions for data collection through *classroom observations* were comprehensive and complete.	1	2	3	4	5	6
Information gathered through *classroom observations* was adequate for drawing meaningful conclusions.	1	2	3	4	5	6
Information gathered through *classroom observations* was adequate for making meaningful recommendations.	1	2	3	4	5	6
The amount of time scheduled for *classroom observations* was adequate for comprehensive data collection.	1	2	3	4	5	6
Terms and concepts used in the *classroom observations* for the AEL ILA instructions and indicators were clearly defined.	1	2	3	4	5	6
The training I received was adequate for successful implementation of the *classroom observations* for the AEL ILA.	1	2	3	4	5	6
The forms for recording data were easy to understand and use.	1	2	3	4	5	6

AEL ILA Data Collection Procedure:
V. Focused Walkthrough

Instructions for data collection through the *focused walkthrough* were clear and unambiguous.	1	2	3	4	5	6

Instructions for data collection through the *focused walkthrough* were comprehensive and complete.	1	2	3	4	5	6
Information gathered through the *focused walkthrough* was adequate for drawing meaningful conclusions.	1	2	3	4	5	6
Information gathered through the *focused walkthrough* was adequate for making meaningful recommendations.	1	2	3	4	5	6
The amount of time scheduled for the *focused walkthrough* was adequate for comprehensive data collection.	1	2	3	4	5	6
Terms and concepts used in the *focused walkthrough* of the AEL ILA instructions and indicators were clearly defined.	1	2	3	4	5	6
The training I received was adequate for successful implementation of the *focused walkthrough* of the AEL ILA.	1	2	3	4	5	6
The forms for recording data were easy to understand and use.	1	2	3	4	5	6

AEL ILA Data Collection Procedure:
VI. Analysis of Student Academic Work

Instructions for data collection through the *analysis of academic work* were clear and unambiguous.	1	2	3	4	5	6
Instructions for data collection through the *analysis of academic work* were comprehensive and complete.	1	2	3	4	5	6
Information gathered through the *analysis of academic work* was adequate for drawing meaningful conclusions.	1	2	3	4	5	6
Information gathered through the *analysis of academic work* was adequate for making meaningful recommendations.	1	2	3	4	5	6
The amount of time scheduled for the *analysis of academic work* was adequate for comprehensive data collection.	1	2	3	4	5	6
Terms and concepts used in the *analysis of academic work* of the AEL ILA instructions and indicators were clearly defined.	1	2	3	4	5	6
The training I received was adequate for successful implementation of the *analysis of academic work* of the AEL ILA.	1	2	3	4	5	6
The forms for recording data were easy to understand and use.	1	2	3	4	5	6

AEL ILA Data Collection Procedure:
VII. Interviews with Teachers and Administrators

Instructions for data collection through the *teacher and administrator interviews* were clear and unambiguous.	1	2	3	4	5	6
Instructions for data collection through the *teacher and administrator interviews* were comprehensive and complete.	1	2	3	4	5	6
Information gathered through the *teacher and administrator interviews* was adequate for drawing meaningful conclusions.	1	2	3	4	5	6
Information gathered through the *teacher and administrator interviews* was adequate for making meaningful recommendations.	1	2	3	4	5	6
The amount of time scheduled for the *teacher and administrator interviews* was adequate for comprehensive data collection.	1	2	3	4	5	6
Terms and concepts used in the *teacher and administrator interviews* of the AEL ILA instructions and indicators were clearly defined.	1	2	3	4	5	6
The training I received was adequate for successful implementation of the *teacher and administrator interviews* of the AEL ILA.	1	2	3	4	5	6
The forms for recording data were easy to understand and use.	1	2	3	4	5	6

AEL ILA Data Analysis Process:
VIII. Building Consensus, Aggregating & Analyzing Data, Drawing Conclusions & Recommendations

Instructions for reaching a team consensus on the goals of the AEL ILA in this school/district were clear and unambiguous.	1	2	3	4	5	6
Instructions for reaching a team consensus on the goals of the AEL ILA in this school/district were comprehensive and complete.	1	2	3	4	5	6
The process for aggregating data from appraisal team members was clear and unambiguous.	1	2	3	4	5	6
The process for aggregating data from appraisal team members was comprehensive and complete.	1	2	3	4	5	6
The process for aggregating data from appraisal team members was fair, giving all team members the same opportunity to have their perspectives considered.	1	2	3	4	5	6
The instruction and forms provided for aggregating the data made the process efficient.	1	2	3	4	5	6
Multiple data sources supported the team members' conclusions.	1	2	3	4	5	6
There was a high degree of consensus among team members about the conclusions.	1	2	3	4	5	6
Recommendations were consistent with findings and conclusions.	1	2	3	4	5	6

If you have any additional comments or suggestions, please feel free to include them on a separate sheet attached to the end of this questionnaire.

Thank you for your time and feedback!

If you have any questions or concerns about your rights as a participant in this research, please contact Dr. Merrill Meehan, AEL IRB Chair (800-624-9120, ext. 5432, or meehanm@ael.org). Other questions and concerns may be addressed to Patricia Ceperley (800-624-9120, ext. 5423 or ceperlep@ael.org).