

Analytic Technical Assistance and Development

Forming a Team to Ensure High-Quality Measurement in Education Studies

Ellen Eliason Kisker
Twin Peaks Partners, LLC

Kimberly Boller
Mathematica Policy Research



REL 2014-052

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased, large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

September 2014

This report was prepared for the Institute of Education Sciences (IES) by Decision Information Resources, Inc. under Contract ED-IES-12-C-0057, Analytic Technical Assistance and Development. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Kisker, E. E., & Boller, K. (2014). *Forming a team to ensure high-quality measurement in education studies* (REL 2014-052). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

This report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

Contents

Introduction	1
Forming a measurement team	3
Study leader	4
Subject matter expert	5
Measurement expert	5
Data collection expert	5
Statistician	6
Planning a measurement strategy	6
Specify constructs	7
Identify existing measures	7
Gather additional information	7
Weigh the strengths and weaknesses of existing measures	8
Consider creating new or adapted measures	8
Ensuring the collection of high-quality data	10
Develop and follow a data collection plan	10
Create variables for analysis and assess their quality	10
Summary	11
Case study: The roles and tasks of a measurement team in a large impact evaluation study	11
Additional resources	13
Notes	14
Tables	
Table 1. Overview of measurement team members and responsibilities	2
Table 2. Overview of key steps in developing and carrying out a sound measurement strategy	4

Introduction

Forming a team whose members can make key measurement decisions that lead to high-quality data is an essential step at the beginning of a study. Investing the time and resources to find appropriate experts who can help develop and carry out an effective measurement strategy will set the course for obtaining accurate and reliable answers to important research questions.

This brief is intended to help identify the measurement expertise needed to formulate an appropriate and effective measurement strategy and to collect high-quality data. The primary audience for this brief is research teams – such as those at Regional Educational Laboratories (RELs) – who work in partnership with school districts or states. The brief provides tips for finding staff and consultants with the needed expertise and outlines their main responsibilities. It also outlines typical measurement tasks and discusses how these measurement team members can work together to complete these tasks successfully. This information is summarized in Table 1 and discussed in the following sections of this brief.

Table 1. Overview of measurement team members and responsibilities

Team member	What to look for regarding the expert	What this person will do	How to find an expert
Study leader	<ul style="list-style-type: none"> The project director or a principal investigator or task leader who has responsibility for new measurement in the study 	<ul style="list-style-type: none"> Recruit members of the measurement team Oversee the work of the measurement team and ensure that its work reflects study goals 	Not applicable
Subject matter expert	<ul style="list-style-type: none"> A record of research on the intervention or phenomenon to be studied Knowledge of or experience with relevant constructs and measures Knowledge of the population to be studied 	<ul style="list-style-type: none"> Identify and define the constructs to be measured Help identify existing measures, assess their alignment with the constructs, and examine the value of developing new or adapted measures Review any measure specifications and procedures developed for the study Help address content-related questions about measures that arise during data collection Review and use variables constructed for analysis 	<ul style="list-style-type: none"> Contact authors of previous relevant studies that measured similar constructs Consult colleagues and members of relevant professional networks Examine recent relevant conference programs to identify researchers who are conducting related work currently Perform an Internet search
Measurement expert	<ul style="list-style-type: none"> Training in psychometrics Experience in developing measures Knowledge of or experience with relevant measures 	<ul style="list-style-type: none"> Help identify existing measures Assess the strengths and weaknesses of existing measures for use with the study population Specify and pilot test new measures or adaptations to existing measures Help address questions about implementing measures that arise in the data collection process Construct measures and assess their psychometric properties; identify any variables that should not be included in analyses 	<ul style="list-style-type: none"> Contact authors of previous relevant studies that measured similar constructs Search for and contact developers of measures used in previous research to explore whether they are qualified and available and, if appropriate, seek recommendations Examine the programs of relevant conferences to find measurement experts who may be doing relevant work on measurement methods Contact relevant university centers, such as the Kansas Center for Research Methods and Data Analysis or the Interuniversity Consortium for Political and Social Research
Data collection expert	<ul style="list-style-type: none"> Experience in implementing similar types of measures with similar populations in similar settings If the measures are multi-item tests, scales, or ratings, experience in training data collectors to administer such measures with 	<ul style="list-style-type: none"> Assess the resources needed to implement existing measures Lead pilot testing of adapted and new measures Assess the feasibility of implementing adapted and new measures in the current study and the resources required to do so Develop a data collection plan Oversee training and data collection 	<ul style="list-style-type: none"> Contact data collection experts and leaders of previous relevant studies that measured similar constructs and explore whether these experts are qualified and available. If appropriate, ask for recommendations for other experts Consult colleagues and members of relevant professional networks

Team member	What to look for regarding the expert	What this person will do	How to find an expert
	high inter-rater reliability	for using the measures <ul style="list-style-type: none"> • Develop procedures for administering the measures and checking the quality of the data • Provide information to the measurement expert to assess the psychometric quality of the measures 	
Statistician	<ul style="list-style-type: none"> • Experience in developing sample designs and conducting statistical power analyses 	<ul style="list-style-type: none"> • Help measurement team members design a pilot test of new measures or a substudy to answer research questions with measures that cannot be administered to the entire sample 	<ul style="list-style-type: none"> • Consult the measurement expert, who may have the statistical expertise or be able to recommend colleagues with the needed skills • Contact relevant university centers or departments that may employ statisticians with the required expertise • Contact relevant professional networks

Forming a measurement team

When a study requires the collection and analysis of data, it is essential to create a sound measurement strategy—that is, identify the constructs to be measured, identify and assess existing measures of these constructs, and select measures for use in the study (or create new measures when necessary). Such a strategy is a critical foundation for collecting data that will support answering the research questions. Researchers who carry out the measurement strategy must ensure that the measures are administered appropriately and the variables constructed from the data have adequate psychometric properties. Table 2 outlines the key steps in planning and implementing a measurement strategy.

The remainder of this section identifies the key roles for a study's measurement team and presents ideas for identifying one or more people to fill these roles. The next section examines how members of the measurement team can work together to carry out the key steps in planning and implementing a measurement strategy. Although this brief assumes that different people will fill each measurement team role, in some cases, one person may be able to fulfill more than one of these roles. In other cases, multiple people may be needed to fill one of these roles.

The psychometric properties of a measure include its

- **Variability: the extent to which a measure detects differences between or within persons**
- **Reliability: the consistency and stability of a measure and the extent to which it is free of error**
- **Validity: the degree to which a measure accurately assesses what it is designed to measure for its intended purpose**

Table 2. Overview of key steps in developing and carrying out a sound measurement strategy

Key step	Specific types of tasks	Experts needed
Develop a measurement strategy	<ul style="list-style-type: none">• Identify constructs to be measured• Identify existing measures of constructs• Assess the measures' alignment with constructs• Gather information on psychometric properties and cost of administration• Weigh strengths and weaknesses of existing measures• Consider creating new or adapted measures if needed	<ul style="list-style-type: none">• Subject matter expert• Measurement expert• Data collection expert• Study leader
Collect high-quality data	<ul style="list-style-type: none">• Develop and follow a data collection plan• Hire qualified data collection staff and provide needed training• Monitor data quality and address any problems• Collect information needed for assessing psychometric properties of measures in the study	<ul style="list-style-type: none">• Data collection expert• Subject matter expert• Measurement expert
Create variables for analysis and assess their quality	<ul style="list-style-type: none">• Construct scores and assess psychometric properties of multi-item measures• Advise researchers who use the data about the performance of the measures in the current study	<ul style="list-style-type: none">• Measurement expert• Data collection expert

Study leader

At the outset, a study leader—a project director, principal investigator, or task leader with responsibility for aspects of the study that involve new measurement—must identify the expertise needed to develop and carry out the study's measurement strategy. If the study leader does not have the technical knowledge to identify all of the expertise needed, he or she should seek guidance from a colleague who has led a similar study or should consult technical reports of completed studies with similar measurement requirements to learn about the issues and tasks involved and the individuals who addressed them. Then the study leader should recruit members for the measurement team who have the needed expertise and should oversee their work to be sure that it reflects study goals and remains aligned with other aspects of the study.

The study leader may need to work closely with the measurement team in making key decisions. For example, if alternative potential measures have different strengths and weaknesses, the study leader will need to play an important role in making tradeoffs and choosing among the measures. The study leader may decide to investigate the feasibility of devoting study resources to adapting a measure or creating a new measure and to seek additional funding to do so. If study resources are not sufficient to support administration of measures of all constructs needed, the study leader may need to manage decisions about scaling back the research questions.

Subject matter expert

The subject matter expert will play a key role throughout the study. As indicated in the specific responsibilities listed in Table 1, the subject matter expert will identify the constructs that must be measured to answer the study's research questions and ensure that valid measures are used to operationalize these constructs. Later, the subject matter expert will address content-related questions that may arise in planning and carrying out the data collection and will review the analyses to ensure that the data are used appropriately in analyses to address the research questions.

Constructs are the characteristics and concepts that must be measured to answer a research question.

If the study leader is not a subject matter expert, he or she may need to find an expert on the phenomena to be studied. If the subject matter expertise available within the organization conducting the study is not sufficient to meet study needs, the study leader should seek outside expertise and may find it useful to try one or more of the strategies listed in Table 1 for finding a subject matter expert.

Measurement expert

The measurement expert will play a central role throughout the study. While the subject matter expert will specify the constructs to be measured, the measurement expert will lead the identification and selection of measures of these constructs. Once the measures have been selected, the measurement expert will help monitor data collection and use of the measures. Later, the measurement expert will be responsible for constructing variables from the data, assessing the psychometric properties of the variables created from multi-item scales, and advising the study team regarding the use of these variables in the study.

In seeking a measurement expert, the study leader should look for someone with:

- Training in psychometrics and a record of research on measurement methods
- Knowledge of relevant constructs and measures
- Knowledge of the population to be studied (especially if the study focuses on a particular group, such as English learners or students with disabilities).

If the measurement expertise available within the organization conducting the study is not sufficient to meet study needs, the study leader may find it useful to use one or more of the strategies listed in Table 1 for finding a measurement expert.

Data collection expert

It is not only essential that the measures selected for the study align with the constructs to be measured and have the appropriate psychometric qualities, but it must be feasible to administer the measures. A data collection expert should be involved in planning the measurement strategy to ensure that it is feasible to implement with the study population within the available resources. Later, the data collection expert will develop detailed plans, oversee data collection using the selected measures, and provide information needed by the measurement expert to assess the psychometric properties of the measures.

To fill this role on the measurement team, the study leader should seek a data collection expert with:

- Experience implementing similar types of measures with similar populations in similar settings;
- If appropriate for the study, experience in training data collectors to administer measures that are multi-item tests, scales, or ratings and experience in establishing and maintaining high inter-rater reliability; and
- Experience in managing the costs and logistical requirements of data collection with similar types of measures.

If the study leader's organization does not have the capacity to conduct the needed data collection, it may be desirable for the data collection expert to come from an organization that can carry out the study's data collection. To find an outside data collection expert if needed, the study leader may find it useful to try one or more of the strategies listed in Table 1 for finding a data collection expert.

Statistician

A statistician—someone with experience in developing sample designs and conducting statistical power analyses—may need to create a sampling plan for the study to ensure that the study will have adequate statistical power for answering the research questions using the selected measures. The statistician can also help other team members design a pilot test of new measures or design a smaller substudy to address a research question that uses measures that cannot be implemented with the full study sample. Later, if needed, the statistician can construct sample weights for use in data analyses. If the organization conducting the study does not employ a statistician with the needed expertise, the study leader may find it useful to use one or more of the strategies listed in Table 1 for finding a statistician.

Planning a measurement strategy

All members of the measurement team need to work together effectively to develop the optimal feasible measurement strategy. Planning a sound measurement strategy consists of:

- Developing a list of constructs to be measured
- Identifying existing measures and assessing their alignment with the study constructs
- Gathering additional information on the quality and cost of each potential measure
- Considering the adaptation or creation of measures if no appropriate measures exist
- Considering an appropriate sampling plan for collecting data using the measures
- Assessing the feasibility and logistics of collecting data using the measures

The following sections describe these tasks and how measurement team members work together to accomplish them.

Specify constructs

If the study's research questions are stated precisely, focused on clearly defined phenomena at a specific time or age, and based on a logic model or theory of change, then creating a list of constructs may be straightforward. For example, in the highlighted research question to the right, the words in italics identify the constructs that must be measured to answer this question. Otherwise, the subject matter expert, along with the study leader, may need to articulate the underlying theory of change and refine the research questions.

Research Question: What impact did Curriculum X have on student *engagement in classroom learning*? Specifically, what impact did Curriculum X have on (a) *class preparedness*, (b) *attention*, and (c) *persistence*?

Even if the research questions describe clearly the constructs that they address, some constructs may have multiple dimensions that must be further specified into narrower constructs as the measurement plan is formulated. For example, a research question may ask about the impact of a basic second-grade curriculum on student achievement. In order to develop a plan for measuring student achievement, the measurement team, including the study leader, should specify the relevant dimensions of student achievement, such as achievement in math and reading.

Identify existing measures

An early and important challenge for the measurement expert, with the help of the subject matter expert, is to compile a list of existing measures of the constructs to be studied. Developing new measures is often costly and time intensive, so the consideration of existing measures for each construct is essential, taking into account the study's schedule and resources for data collection. Using existing measures also facilitates comparisons with other studies and takes advantage of the established psychometric properties of the measures.

In seeking existing measures, the measurement team should consider the following questions:

- Have prior studies used survey questions to measure the same constructs? Can these surveys be obtained, and can the psychometric properties of the relevant questions or test items be examined?
- Are instruments—such as tests, scales, or observation protocols—that are designed to measure the constructs either published or available from commercial vendors or instrument developers?

The study's subject matter expert can use his or her knowledge of the literature and other resources, such as measures compendia, to identify existing measures and assess how well the measures align with the constructs needed to address the study's research questions. Those that align well with the needed constructs should then be examined further by the measurement expert.

Gather additional information

To make an informed selection of existing measures for the study, the team should gather information and examine each potential measure further. Together, the measurement and data collection experts should address the following questions:

- What evidence exists for the psychometric quality of the measure? Is there evidence from past studies that data collected for studies that used the measure can meet standards such as the What Works Clearinghouse evidence standards?¹
- How similar is the measure's development sample to the population in the current study in terms of age, cultural and linguistic diversity, location, timeframe, and accommodations?
- Is there evidence that the measure is sensitive to the intervention being studied and the outcomes of interest?
- What is the cost of administering the measure? Does the developer charge a fee for its use? How much time and what special materials are required to administer the measure? If necessary, can the measure be shortened?
- What qualifications must staff who administer the measure have, and what training must they receive to administer the measure? Who can provide the training? Is certification required to administer the measure, and what is the process for receiving certification?
- What ongoing monitoring of measure administration is required to ensure high-quality data?
- What scoring procedures or other data processing will be required to produce data for analysis?

The measurement expert and data collection expert may need to contact other experts to gather the needed information and seek advice on the use of the measures in the study.

Weigh the strengths and weaknesses of existing measures

Bringing together multiple perspectives to assess the strengths and weaknesses of the existing measures for the study will enhance the selection of measures, as follows:

- The subject matter expert can help ensure that selected measures align well with the constructs that must be measured.
- The measurement expert can interpret evidence for the psychometric quality of the measures and ensure that selected measures meet relevant thresholds for measure quality.
- The data collection expert can help ensure that selected measures can be implemented with study resources.

Together with study leaders, the measurement team can consider the tradeoffs among the existing measures under consideration. Ideally, study leaders and the measurement team can reach agreement on the optimal set of measures to select for the study, but even if the team cannot reach consensus, study leaders will be able to consider the multiple perspectives in making final measure selections.

Consider creating new or adapted measures

If the measurement team is unable to identify existing measures of a construct or if the assessment of the strengths and weaknesses of existing measures reveals that the existing measures are not adequate, the team may need to consider adapting an existing measure or creating a new one. If a construct can be measured

with a straightforward survey question, then developing a new measure may not be difficult or take a long time. Constructs requiring more complex measures, however, may require significant time and resources.

For constructs that require more complex measures, such as tests of knowledge or multidimensional scales based on ratings of different aspects of the classroom environment, the development of new measures may involve substantial effort. Permission from the original developer may be needed in order to make adaptations to an existing measure. The process of making adaptations to existing items or developing new test or scale items may entail many rounds of review and revision before agreement on a final set of items can be reached. Staff who develop a new measure may need to consult with several experts and conduct analyses to establish the validity of a new measure.

Pilot testing is important for establishing the psychometric properties of new measures. While it may be tempting to proceed directly to using a newly developed or adapted measure in the study without pilot testing, doing so is highly risky. If the data collected by using the new measure show that it was not valid or reliable, the study will not be able to provide credible answers to the research questions that the collected data were intended to address. If a measure does not perform well in pilot testing, the testing may reveal insights into why the measure is not performing well and point to improvements that can be tested.

Collaboration among measurement team members will strengthen the development of new or adapted measures and the pilot testing of these measures, as follows:

- The subject matter expert can assess the value of developing new measures or adapting existing ones. He or she can also review new measure specifications, address content-related questions that arise, and help make any revisions to the measures that may be needed.
- The measurement expert can help plan and oversee the pilot test to ensure that it provides the data needed to assess the psychometric quality of the measures. He or she can also conduct psychometric analyses and assess the strengths and weaknesses of the new or adapted measures.
- The data collection expert can lead staff members in developing data collection materials (questionnaires, assessment materials, programming, debriefing protocols, and so forth), identifying and training data collectors, recruiting subjects to participate in the pilot test, overseeing data collection for the pilot test, and preparing the pilot test data for analysis.
- The statistician can design a pilot test of new or adapted measures with the statistical power needed to inform final decisions about use of the new or adapted measures in the study.

A pilot test entails administering the new measures to individuals similar to those who will be participating in the study. The pilot test makes it possible to assess the feasibility of administering the measures and to establish evidence for the psychometric quality of the measures.

Together with study leaders, the experts can assess the strength of the new or adapted measures as well as the costs and logistical requirements for using them in the study.

Not all studies will have the resources to adapt existing measures or create new ones, so scaling back the research questions and constructs to be measured may be inevitable if existing measures are not available. If

resources are available, however, the study may be able to address all of the original research questions and contribute to the research field by expanding the available measures.

Ensuring the collection of high-quality data

Once the measurement team has developed a measurement strategy and determined what measures will be implemented in the study, the measurement team should turn its focus to ensuring the collection of high-quality data. The team's next step should be to develop a plan for administering the selected measures as specified by the measure developers and according to best practices in research data collection. Once that plan has been developed, the team will be responsible for following the data collection plan, creating variables for analysis, and assessing the quality of the variables.

Develop and follow a data collection plan

Just as important as the measures themselves is their administration in the study. If the staff members who are collecting data cannot administer the measures consistently or do not follow study procedures for administration, the data may not lead to accurate and reliable answers to the research questions.

To avoid these problems, the study leader can require the measurement team to develop a data collection plan that includes procedures for ensuring that the data collected with the selected and developed measures are of high quality. Among other things, the data collection plan should specify the qualifications and training required for data collectors, set standards for establishing inter-rater reliability and for monitoring inter-rater reliability throughout data collection (if appropriate for the measures used), and establish other procedures for checking and maintaining data quality.

The data collection expert should ensure not only that the data collection plan is sound but also that it is followed carefully. He or she should ensure that procedures are in place for protecting the confidentiality of participants, check data quality regularly, identify any problems with data quality early, and work with other members of the measurement team to address any problems. The subject matter expert can help address any questions about the content of the measures that may arise, and the measurement expert can address questions about the administration of the measures.

Create variables for analysis and assess their quality

Once the data have been collected, the measurement expert should construct scores for any measures that are multi-item tests or scales and should assess the psychometric properties of the constructed scores. This assessment may entail calculating internal consistency reliability, comparing the scores constructed from different measures to confirm that the scores are related in expected ways, and obtaining measures of inter-rater reliability from the data collection expert as appropriate. After conducting this assessment, the measurement expert can advise the researchers conducting the study's analyses, including the subject matter expert, regarding the performance of the measures in the current study so that the researchers conducting analyses can make well-informed decisions about whether and how to use the data derived from the measures. By helping the researchers understand how much confidence can be placed in particular findings, the measurement expert can help the study team develop the study conclusions.

Summary

This brief provides an overview of the key types of members needed on a measurement team and their main tasks and responsibilities. The size and composition of a measurement team and the specific tasks of its members, however, will vary from study to study and reflect the diversity of education studies undertaken by the Regional Educational Laboratories and other research organizations. For example, the case study that follows describes the composition of the measurement team for a large evaluation study and highlights some of the key activities that the team undertook to make sure that the data collected for the study would support conclusions about the effectiveness of the program being evaluated. A list of resources that may be helpful in forming a measurement team and identifying the responsibilities of its members for a specific study is included following the case study.

Forming a measurement team that can develop and carry out a sound measurement strategy may be essential to a study's success. Evidence reviews, such as the What Works Clearinghouse, do not include effects based on outcome measures that do not meet thresholds for reliability and validity.² Reviewers or consumers of research may discount or reject study findings if the analyses are based on measures compromised by inconsistent or unreliable data collection or with poor psychometric properties. Investment in a strong measurement team provides a strong foundation for a study that rigorously measures the outcomes or skills of interest.

Case study: The roles and tasks of a measurement team in a large impact evaluation study

In a large evaluation of a comprehensive child development program for infants and toddlers, the measurement team included numerous experts. The program aimed to improve outcomes in a wide range of child development and family development domains; therefore, the selection of measures for the evaluation required that the team consider a large number of potential measures and select many measures for the evaluation. The measurement team for this study included:

- Subject matter experts from the organizations conducting the study and from the agency funding it who had expertise about infant and toddler development in different developmental domains and who had knowledge of child development programs for young children
- Measurement experts from the collaborating organizations who had expertise about measures in specific developmental domains and had experience in using the measures with similar populations
- Data collection leaders from the organizations collaborating on the evaluation study who had experience in collecting child assessment data and in overseeing large data collection efforts

To make the measure selection process systematic and thorough, the measurement team established the following guiding principles for selecting child and family measures for the evaluation:

- Relevance to intervention goals and key hypotheses
- Appropriateness to children's age and developmental level
- Appropriateness for the target population

- Adequate psychometric properties
- Prior use in large-scale surveys and intervention evaluations
- Low cost and respondent burden

The subject matter experts worked closely with the measurement experts to identify potential measures, gather information to assess the quality of the measures, and assess the existing measures based on the guiding principles. They began by searching the literature and contacting measure developers and others who had used the measures previously. When existing measures were not adequate, members of the team modified the existing measures or created new ones (with the assistance of measure developers).

Once a complete set of measures was assembled, the measurement experts worked with the data collection leaders on the team to create interview and assessment protocols for administering the measures. Data collection staff pilot tested the interviews and assessments with children and families similar to those in the study. The pilot tests revealed that administration of the interviews and assessments took longer than planned, and the data collection experts worked with the measurement experts, subject matter experts, and study leaders to delete less important measures from the measurement plan. Pilot testing ensured that collecting data with the remaining measures was feasible and could be completed with available resources. The pilot test experience also provided valuable information to inform future training of data collection staff.

After the data collection protocols were finalized, data collection leaders identified the qualifications and training required of data collectors, hired data collectors, and provided intensive training in the administration of the measures. Other members of the measurement team attended the training and addressed questions related to their expertise. Part of the training consisted of practice in administering the measures and required data collectors to demonstrate inter-rater reliability (that is, to demonstrate that the data they collected using each measure were consistent with the data collected at the same time by a “gold standard” data collector using the same measures).

When data collection for the study began, the measurement team, including the measurement experts, data collection leaders, and study leaders, monitored data collection progress and data quality. Data collection leaders measured inter-rater reliability throughout data collection and retrained or replaced data collectors who could not demonstrate continued adequate reliability in their administration of the measures.

After each round of data collection, a subset of the measurement experts on the team prepared the data for analysis. They scored questionnaire scales and child assessments according to directions provided by the measure developers or by using an approach consistent with the current literature. For new measures or measures that required additional data reduction, the experts conducted factor analyses to create variables for analysis.

The measurement experts examined the psychometric properties of the variables constructed from the measures. The team established the following criteria for including variables in the study’s impact analyses:

- Sufficient data at the item level (the constructed variable should be set to missing if 25% or more of the items were missing)
- Adequate distribution of scores (the constructed variable should have a normal distribution with a distribution similar to that found in other studies using the same measure)
- Adequate internal consistency reliability (.65 or above)
- Consistent reliability across major racial and ethnic groups

In an appendix to the report, the study team provided details of the psychometric properties of the constructed variables used in the analyses.

As the impact analyses were conducted and researchers were preparing the evaluation report, the measurement team discussed the interpretation of the findings with the researchers who conducted the analyses. The measurement team provided important information and insights into the strengths and weaknesses of the variables to help the researchers understand how much confidence could be placed in particular findings and decide which findings to emphasize in arriving at study conclusions.

Because the study formed a strong team to develop a sound measurement strategy and ensure the high quality of the data collected, the study provided needed evidence for the effectiveness of the program. Researchers have continued using the data in secondary analyses to address a wide range of questions about child development, parenting, and family development.

Additional resources

Measurement

The following report provides a useful discussion of criteria for measures selection:

Boller, K., Atkins-Burnett, S., Malone, L. M., Baxter, G. P., and West, J. (2010). "Compendium of student, teacher, and classroom measures used in NCEE evaluations of educational interventions." Volume I, *Measures selection approaches and compendium development methods* (NCEE 2010-4012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

The following IES Analytic Technical Assistance and Development brief describes the information about measures that researchers need to report along with study results. It offers checklists that also may be useful for understanding the range of tasks measurement experts may be called on to perform.

Boller, K., and Kisker, E. (Forthcoming). *Reporting what readers need to know about education research measures: A guide*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

The National Council on Measurement in Education website identifies resources that may be useful on its website:

<http://ncme.org/resource-center/library/>

Data collection

The American Association for Public Opinion Research offers data collection tools and guides on its website:

http://www.aapor.org/For_Researchers/5850.htm

Notes

1. *What Works Clearinghouse Procedures and Standards Handbook*, Version 3.0. Retrieved on 6-18-14 at <http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19>).
2. *What Works Clearinghouse Procedures and Standards Handbook*, Version 3.0. Retrieved on 6-18-14 at <http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19>).

