

Migrant Education  
Program Evaluation Toolkit  
A Tool for State Migrant Directors

Summer 2012

Developed by the U.S. Department of Education Office of Migrant Education through a contract with The SERVE Center at the University of North Carolina-Greensboro.

## Section A: Introduction & Overview

### Letter to State Directors

- A.1 Purpose of the *Program Evaluation Toolkit*
- A.2 General Approach to Developing the *Program Evaluation Toolkit*
- A.3 Organization of the *Program Evaluation Toolkit*
- A.4 How to Use the *Program Evaluation Toolkit*
- A.5 How to Adapt the Process for Small States
- A.6 Resources and Tools in Appendix A
- Appendix A.1 Terms and Definitions Related to the MEP Comprehensive Needs Assessment, Service Delivery Plan, and Program Evaluation
- Appendix A.2 Summary of Findings from a Review of a Sample of State Migrant Education Evaluation Reports

## Section B: Overview Of Statutes, Regulations, And Non-Regulatory Guidance Related To Program Evaluation

- B.1 Elementary And Secondary Education Act (ESEA) Evaluation Terms
- B.2 Office of Migrant Education Evaluation Terms
- B.3 State Requirements for Evaluation
- B.4 Local Requirements for Evaluation
- B.5 State Responsibilities to Local Education Agencies for Evaluation
- B.6 Reflection Questions
- B.7 Resources and Tools in Appendix B
- Appendix B.1 Checklist For State Migrant Education Program (MEP) Evaluation

## Section C: Planning The Evaluation

- C.1 Evaluating Implementation
- C.2 Evaluating Results
- C.3 Generating Evaluation Questions
- C.4 Evaluation Timeline
- C.5 Summary of Key Points
- C.6 Reflection Questions
- C.7 Online Resources
- C.8 Resources and Tools in Appendix C
- Appendix C.1 Setting Measureable Program Outcomes (MPOs)
- Appendix C.2 Evaluation Timeline Template

## Section D: Collecting Evaluation Data

- D.1 Using Existing Data
- D.2 Conducting Interviews
- D.3 Administering Surveys
- D.4 Observing Activities
- D.5 Evaluation Matrix
- D.6 Summary of Key Points
- D.7 More Online Resources

- D.8 Reflection Questions
- D.9 Resources and Tools in Appendix D
- Appendix D.1 Simple Evaluation Matrix Template
- Appendix D.2 Complex Evaluation Matrix Template

### Section E: Analyzing And Interpreting Data

- E.1 Analyzing Quantitative Data
  - E.1.1 Descriptive Statistics for Categorical Variables
  - E.1.2 Descriptive Statistics for Continuous Variables
- E.2 Analyzing Qualitative Data
- E.3 Summary of Key Points
- E.4 Reflection Questions
- E.5 Resources and Tools in Appendix E
- Appendix E.1 Using Excel To Analyze Quantitative Data
- Appendix E.2 Using Access To Analyze Qualitative Data
- Appendix E.3 Using Inferential Statistics

### Section F: Communicating Evaluation Findings

- F.1 Writing Evaluation Reports
- F.2 Organizing the Evaluation Report
- F.3 Displaying Data
- F.4 Summary of Key Points
- F.5 Reflection Questions
- F.6 Other Resources

### Section G: Using Evaluation Findings

- G.1 Continuous Improvement
- G.2 Links to Other Online Resources
- G.3 Resources and Tools in Appendix G
- Appendix G.1 Protocol For Using Evaluation Findings In A Continuous Improvement Cycle

# Section A: Introduction & Overview

---

## Letter to State Directors

*The Program Evaluation Toolkit* is the third in a series of three resource documents developed by the Office of Migrant Education (OME). Previous documents include the *Comprehensive Needs Assessment Toolkit* and the *Service Delivery Plan Toolkit*.

Taken together, these three toolkits provide a guide for a continuous improvement cycle of program planning, implementation, and evaluation. In this cycle, states identify the current needs and priorities of migrant students and families, select appropriate strategies to meet those needs, implement the services that reflect such strategies, and assess the degree to which planned services have been successful at meeting identified needs. States complete the cycle by reflecting on the evidence they have gathered and consider where program recalibrations or improvements need to be made, what new or remaining needs exist, and whether to re-design existing programs or create new ones.

The *Program Evaluation Toolkit* is intended to help you plan and conduct an evaluation that is aligned with such a continuous improvement cycle and framed by the program strategies and measurable outcomes laid out in your Comprehensive Needs Assessment and Service Delivery Plan. It also provides program evaluation templates and resources that will help facilitate effective state Migrant Education Program (MEP) evaluations. Because the Evaluation Plan itself is part of the service delivery planning process, evaluation planning is also covered in the *Service Delivery Plan Toolkit*.

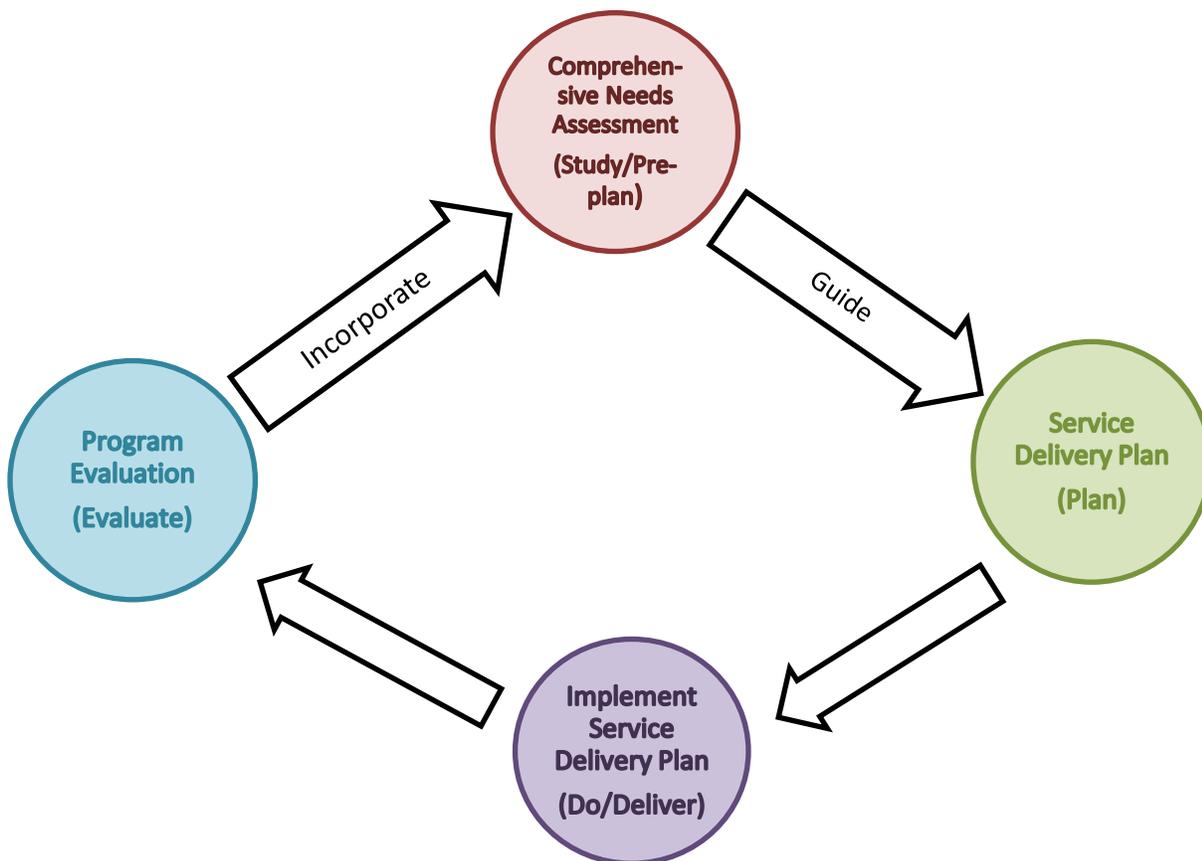
We hope that you find the *Program Evaluation Toolkit* useful, and we invite you to send us your comments and share your administrative and governance best practices. We would like to express our gratitude to those states that shared their policies, procedures, and forms with us and made their usage available as examples.

Thank you,

The Office of Migrant Education

## A.1 Purpose of the *Program Evaluation Toolkit*

The purpose of this *Program Evaluation Toolkit* is to provide state and local directors of Migrant Education Programs (MEPs) with guidance and resources to conduct useful evaluations of the services provided to migrant students and their families. A useful evaluation is one that generates reliable information about the quality of program implementation and the results that have been accomplished through MEP program activities. Using this information, state and local education agencies can revise or re-design program plans as needed to improve results and help more migrant students achieve academic success. Program evaluation is an important step in the Continuous Improvement Cycle (CIC) of a state MEP. It serves to inform and improve both the Comprehensive Needs Assessment (CNA) and the Service Delivery Plan (SDP). (See *Appendix A.1 Terms and Definitions Related to the MEP Comprehensive Needs Assessment, Service Delivery Plan, and Program Evaluation* for a list of terms and definitions used across all three processes: comprehensive needs assessment, service delivery planning, and program evaluation.)



State Migrant Education Program Continuous Improvement Cycle  
(Office of Migrant Education, 2011)

While using data to guide decision making is part of every stage of a CIC, the following resources offered in the *Program Evaluation Toolkit* may be particularly helpful to guide program improvement in each stage of the planning and implementation of the MEP:



When you reach the point of developing an evaluation plan for your SDP, the tools described in *Section C: Planning The Evaluation* will help you organize the plan. The data collection methods described in *Section D Collecting Evaluation Data* will help you decide which methods to use to gather meaningful data.



When you are ready to make sense of the data you have gathered, *Section E Analyzing And Interpreting Data* will explain several ways to analyze and learn from your data. *Section F Communicating Evaluation Findings* will show you how best to report and display your data to communicate what you have learned.



And finally, when you have analyzed and interpreted your data and are ready to reconvene the CNA and SDP committees to review progress toward state MEP performance targets, *Section G Using Evaluation Findings* will guide you through a process for using your evaluation data to decide what to do next.

## A.2 General Approach to Developing the *Program Evaluation Toolkit*

The *Program Evaluation Toolkit* addresses common problems in MEP program evaluation reports that were identified from a sample of nine MEP evaluation reports. (See *Appendix A.2 Summary of Findings from a Review of a Sample of State Migrant Education Evaluation Reports* for a summary of the findings included in a review of the nine evaluation reports.)

The following are areas of improvement that were identified in the review and are directly addressed in the *Program Evaluation Toolkit*:

- Using state performance targets to determine the gap between actual migrant student performance and where they should be performing
- Using consistent data disaggregation, including the disaggregation for priority for service migrant students, all other migrants, and non-migrant students
- Developing strong Measurable Program Outcomes (MPOs) that define concrete expected outcomes for migrant students who received specific educational or educationally related services

- Including contextual information related to specific needs, baseline performance, barriers, and MPOs and related results to increase the significance of reported findings and the implications for reassessing needs and strategies
- Understanding the difference between description and evaluation
- Evaluating systematically the local program implementation aggregated in the state report
- Using expenditures and related outcomes to reconsider or prioritize strategies for helping migrant students achieve academically and personally

With these challenges in mind, the *Program Evaluation Toolkit* was developed to fill in the gaps between the available guidance and OME’s expectations for high-quality evaluation reports. OME drew upon the expertise of its program staff, state MEP directors, and program evaluation experts to develop the information in the *Toolkit*.

### **A.3 Organization of the *Program Evaluation Toolkit***

The *Program Evaluation Toolkit* is divided into seven sections, beginning with this *Introduction and Overview (Section A Introduction & Overview)*.

*Section B Overview Of Statutes, Regulations, And Non-Regulatory Guidance Related To Program Evaluation* summarizes the statutes, regulations, and *Non-Regulatory Guidance for Title I, Part C Education of Migratory Children* governing evaluation of the implementation and results of MEPs and clarifies state and local requirements.

*Section C Planning The Evaluation* describes steps in the evaluation planning process and includes tools to help you organize your thinking and present the final plan.

*Section D Collecting Evaluation Data* describes different ways to collect meaningful information about the quality and effectiveness of MEP activities and services.

*Section E Analyzing And Interpreting Data* explains how to analyze and interpret quantitative and qualitative evaluation data.

*Section F Communicating Evaluation Findings* suggests ways to organize evaluation reports and data displays to communicate effectively with your intended audiences.

*Section G Using Evaluation Findings* describes how to use evaluation results to re-assess needs, re-think strategies, and re-design programs as needed for the benefit of continued progress toward state performance targets.

In each section, there are illustrative examples and links to online resources that address each topic in more detail. In most of the sections, there are sample forms and templates that can be adapted to suit your needs.

#### **A.4 How to Use the *Program Evaluation Toolkit***

Clearly, there is no “one size fits all” or one “best” way to evaluate programs in the complex world of migrant education. However, there are some methods and tools that can help improve the quality and relevance of MEP evaluations. The *Toolkit* provides some tools and resources to increase your understanding of evaluation methods and your capacity for using these methods to customize evaluation plans for your own state. If your state budget is adequate to pay for an external program evaluation, the information in this guide will help prepare you to be an informed contributor to the evaluation plan and a knowledgeable consumer of the evaluation findings.

As a state director, you need to assess what your expertise is, and that of those who may be able to assist you through the MEP evaluation process, as well as determine how the *Toolkit* will best benefit you. Some state directors will want to read the *Toolkit* as a total document, while others will want to utilize only those sections and tools that they feel they need.

Some state directors may feel unprepared to understand the technical information on data analysis and interpretation. While not all state directors have the background needed to conduct a program evaluation, any can benefit from the material presented in the *Toolkit* to develop a greater understanding of the methods and analytic approaches that underlie strong program evaluation. This information will help you define what you need for the MEP program evaluation and oversee the process in an informed way as you solicit the expertise of an experienced program evaluator.

The *Toolkit* is made available on the OME website in two formats. One is a downloadable document that includes all sections and appendices. The other is a list of individual sections and appendices that may be downloaded individually.

Please note that the *Toolkit* provides only suggestions and recommendations for conducting the MEP evaluation. While the state is required to evaluate its migrant program, it is not required to utilize the methods, tools, and resources provided in the *Toolkit*.

#### **A.5 How to Adapt the Process for Small States**

If you are a MEP director in a small state (defined as those with \$1m or less in MEP funding), you will need to develop an evaluation plan that is appropriate for the size of your program and budget. The following are some considerations to help you evaluate the MEP in a quality, yet practical, fashion:

- Prioritize evaluation questions according to those that will provide the most important information to help you improve the MEP, and select a limited number of questions for the evaluation
- Enlist help from data experts in your state agency to guide you in conducting a program evaluation

- Work with a local college or university that may assist with the evaluation for free or provide interns or graduate students who can assist while building their skills
- Hire an external consultant and work with this person on developing an evaluation plan that will enable you to get the most out of what you can budget for the process
- View program evaluation as an ongoing initiative and identify a limited set of questions to explore each year over a period of several years
- Small states have the advantage of having fewer local programs for which to develop and oversee a consistent data collection process that will contribute to the state MEP; work with local project directors to develop a consistent set of tools and processes for data collection and provide technical assistance and opportunities for them to communicate with you (or your consultant) and other project directors during the evaluation process

## **A.6 Resources and Tools in Appendix A**

Appendix A.1 Terms and Definitions Related to the MEP Comprehensive Needs Assessment, Service Delivery Plan, and Program Evaluation

Appendix A.2 Summary of Findings from a Review of a Sample of State Migrant Education Evaluation Reports

## Appendix A.1 Terms and Definitions Related to the MEP Comprehensive Needs Assessment, Service Delivery Plan, and Program Evaluation

**Concern Statements:** Clear and consistent interpretations of the points that the NAC discussed that should be used to guide the work in developing the CNA. Concern Statements identify particular areas that require special attention for migrant students.

**Expert Work Groups:** Technical experts who provide input on research and evidence-based strategies that support solutions which contribute to closing the gaps identified during the needs assessment.

**Implementation Question:** An evaluation question that addresses the extent to which a strategy is implemented.

**Management Team:** A core group of advisors that helps develop the management plan and oversees the process.

**Measurable program outcomes (MPOs):** Outcomes (i.e., objectives) that a State’s migrant education program will produce to meet the identified unique needs of migratory children and help migratory children achieve the State’s performance targets.

**Need:** The difference between “what is” and “what should be.”

**Needs Assessment Committee (NAC):** A broad-based committee of stakeholders that provide input and direction throughout the CNA process.

**Need Indicator:** A measure that can be used to verify that a particular gap/discrepancy exists for migrant children and sets a parameter to specify the severity of that gap.

**OME Seven Areas of Concern:** A broad area based on the root causes of the unique characteristics of the target group. The Office of Migrant Education has identified seven areas of concern which are: Educational Continuity, Instructional Time, School Engagement, English Language Development, Educational Support in the Home, Health, and Access to Services.

**Priority for Services:** Section 1304(d) of the Elementary and Secondary Education Act establishes a Priority for Services (PFS) requirement. In accordance with this requirement, MEPs must give PFS to migrant children who are failing, or are most at risk of failing, to meet the state’s content and performance standards and whose education has been interrupted during the regular school year.

**Results Question:** An evaluation question that addresses the level of improvement resulting from a program or strategy.

**Service Delivery Plan:** A plan for delivering and evaluating Migrant Education Program-funded

services to migratory children. It is based on the results of an up-to-date statewide Comprehensive Needs Assessment and is intended to meet the unique needs of migrant children and their families.

**Solution Strategy:** A strategy that addresses an identified need.

## Appendix A.2 Summary of Findings from a Review of a Sample of State Migrant Education Evaluation Reports

In 2012, the Office of Migrant Education (OME) oversaw a review of a sample of nine state Migrant Education Evaluation reports in order to identify areas of need that could be addressed in the *Program Evaluation Toolkit*. The project began with a review of evaluation requirements specified in the statutes, regulations, and *Non-Regulatory Guidance for Title I, Part C Education of Migratory Children*. In addition, OME posed four questions to provide a full picture of expectations for high-quality evaluation reporting:

1. How does the state document its effectiveness in improving the results of migratory children in meeting state performance targets in reading, mathematics, high school graduation, and (where applicable) school readiness?
2. How does the state measure progress toward performance measures related to other strategies designed to overcome barriers that prevent migratory children from participating effectively in school?
3. How does the state determine that it has effectively implemented other requirements of the program and state-level activities, especially for Priority for Services (PFS) students?
4. What is there that expenditures are aligned with activities and outcomes supported by the expenditures?

The following is a summary of the major challenges identified in the evaluation reports:

- The states in the sample presented data using state performance indicators but did not include the related state performance targets. Instead of noting gaps between actual and expected performance, some states reported gaps between migrant student performance and the performance of all students (including migrant students themselves).
- Across these states and even within individual state reports, the categories used for data disaggregation were inconsistent. Performance data were rarely presented in mutually exclusive categories for PFS migrant students, all other migrants, and non-migrant students.
- Measurable Program Outcomes (MPOs) took different forms in the reports reviewed. Many states used state performance targets—reported for the state as whole—as their MPOs. Other states used subjective perceptions of effectiveness to define results for categories of services (e.g., all instructional services). MPOs rarely defined concrete expected outcomes for migrant students receiving specific educational or educationally related services.
- Contextual information related to specific needs, baseline performance, barriers, and strategies was generally not reported in relation to MPOs and related results. Therefore,

the significance of reported findings and the implications for reassessing needs and strategies could not be determined.

- The state evaluation reports described rather than evaluated the implementation of state-level administrative programs. At the local or service delivery level, the implementation of instructional and support services was evaluated by administering state MEP surveys to small samples of local staff and parents to gather their subjective assessments of the effectiveness of broad categories of services. The state evaluation reports did not include any examples in which the implementation and results of specific services were systematically evaluated at the local or service delivery level and then summarized or aggregated in the state report.
- None of the states reviewed used expenditures and related outcomes to reconsider or prioritize strategies for helping migrant students achieve academically and personally. States will be required to collect this information when the OME's selected measures for compliance with the Government Performance and Results act (GPRA) are approved by the Office of Management and Budget.<sup>1</sup>

The *Program Evaluation Toolkit* was developed with these findings in mind, in order to build state MEP capacity to develop strong evaluations that will inform program improvement.

---

<sup>1</sup> **The Government Performance and Results Act (GPRA) of 1993** was enacted by Congress to provide for the establishment of strategic planning and performance measurement in the Federal Government (made up of an annual performance plan and an annual performance report). Federal programs are required to establish GPRA indicators to measure the progress of their program.

# Section B: Overview of Statutes, Regulations, and Non-Regulatory Guidance Related to Program Evaluation

---

This section outlines the guiding statutes, regulations, and Federal guidance related to the evaluation of Migrant Education Programs (MEPs). Three main documents inform and guide MEPs and the evaluation of their service delivery:

- [Code of Federal Regulations \(C.F.R.\), Title 34, Sections 200.83, 200.84, and 200.85](#), *Responsibilities of state education agencies for evaluating the effectiveness of Migrant Education Programs, and using the results of evaluations to improve services to migratory children.*
- [Elementary and Secondary Education Act, Section 1304 and 1306](#), *Comprehensive needs assessment, service delivery plan, program evaluation; authorized activities*
- [Non-Regulatory Guidance for Title I, Part C Education of Migratory Children](#) (specifically Chapter VIII – Program Evaluation)

While a summary of the requirements is provided in this section, it is important to read the referenced documents fully, especially to clarify questions regarding program requirements.

## **B.1 Elementary and Secondary Education Act (ESEA) Evaluation Terms**

The following terms are defined for the evaluation of ALL students under ESEA:

- *State Performance Goals:* Part I of the Consolidated State Performance Report (CSPR) collects data related to the five ESEA Goals, established in the approved June 2002 Consolidated State Application, information required for the Annual State Report to the Secretary, as described in Section 1111(h)(4) of ESEA, and data required under Homeless Collection (added in FY 05-06). These goals broadly define the results that every state is expected to achieve for ALL students. For example, Performance Goal 1 states that all students will attain proficiency or better in reading/language arts and math by School Year 2013–14.
- *State Performance Indicators:* For each state performance goal, state performance indicators are the specific kinds of data that states are required to use as measures of progress toward the state performance goals. For example, one of the performance indicators for Goal 1 is annual state assessment data in reading/language arts for grades 3–12.
- *State Performance Targets:* Upon identification and consideration of its unique needs, a state education agency establishes specific state performance targets. These are annual

benchmarks for the progress of all students on each state performance indicator. These performance targets are the same as the Annual Measurable Outcomes (AMOs) that states include in their definitions of Adequate Yearly Progress (AYP). Using the example of Performance Goal 1 and the performance indicator for reading/language arts, a state sets a specific level of expected performance for ALL students in each grade. These expected levels of performance are the annual performance targets.

## **B.2 Office of Migrant Education Evaluation Terms**

State and local MEPs are required to assess the performance of migrant students on Performance Goals 1 and 5 using the state performance indicators for each goal, disaggregating the data by migrant status, and comparing it to state performance targets for each grade:

- Performance Goal 1: By 2013–2014, all students will reach high standards, at a minimum attaining proficiency or better in reading/language arts and math.
  - o Performance Indicator 1.1: The percentage of students at or above the proficient level each year on the state assessment in reading/language.
  - o Performance Indicator 1.2: The percentage of students at or above the proficient level each year on the state assessment in math.
- Performance Goal 5: All students will graduate from high school.
  - o Performance Indicator 5.1: The percentage of students who graduate from high school each year with a regular diploma.
  - o Performance Indicator 5.2: The percentage of students who drop out of school each year.

In compliance with the Government Performance and Results Act (GPRA) of 1993, the Office of Migrant Education (OME) has adopted four GPRA measures for monitoring progress and maintaining accountability in the federal MEP. MEPs are now required to report the following data annually to the U.S. Department of Education:

- (1) Percentage of MEP students that scored at or above proficient on their state's annual reading/language arts assessments in grades 3-8 and high school
- (2) Percentage of MEP students that scored at or above proficient on their state's annual Mathematics assessments in grades 3-8 and high school
- (3) Percentage of MEP students who were enrolled in grades 7-12 and graduated or were promoted to the next grade level
- (4) Percentage of MEP students who entered 11th grade and received full credit for Algebra I, or who were enrolled in a non-remedial Math course for which Algebra I was a prerequisite

Through the Comprehensive Needs Assessment (CNA), the state develops a sophisticated understanding of instructional and other factors that affect migrant students' participation and success in school. The CNA committee studies evidence-based solution strategies and selects specific strategies that address the underlying factors inhibiting migrant students' academic progress. The state then develops a Service Delivery Plan (SDP) to communicate the types of instructional and support services that are consistent with selected strategies and establishes Measurable Program Outcomes (MPOs) for these services.

- MPOs are established by the MEP for services to migrant students. State and local MEPs that provide services to migrant students are evaluated by the extent to which actual participation and academic performance of migrant students compare to these expected MPOs.
- The nature of an MPO varies depending on whether it is written for an instructional service or a support service:
  - MPOs related to Direct Instructional Services should be expressed as expected increases in performance on measures of academic achievement, including state assessments.
  - MPOs related to Non-instructional Support Services should be expressed as expected increases in students' active participation in school or other instructional programs.
- MPOs are not the same as either state performance targets or the AMOs that States use in their definitions of AYP.
  - State performance targets and AMOs are used to monitor the accomplishments of all educational initiatives combined. They are macro-outcomes.
  - MPOs are defined for specific educational or educationally related services. They are micro-outcomes.

Among all migrant students, those who are failing or at-risk of failing and who experience interruptions in their schooling are designated as having a Priority for Services (PFS). When funding is insufficient to serve all migrant students, the needs of PFS migrant students must be addressed first. Consistent with this priority:

- Achievement on state performance targets should be disaggregated for PFS and other migrant students to determine the overall effectiveness of state and local MEPs.
- Results related to MPOs should be disaggregated for PFS and other migrant students in order to evaluate the impact of services for migrant students who have the greatest needs.

MEPs are responsible for evaluating the implementation of programs as well as program results.

- An implementation evaluation examines how well a program is carried out to meet the needs of migrant students, especially PFS students.

- Results include the actual performance and participation of migrant students compared to the MPOs established for MEP services.

### **B.3 State Requirements for Evaluation**

If your state receives Title I, Part C funds for migrant education, you must evaluate the effectiveness of your state MEP using the following information:

- The four state performance targets related to Goals 1 and 5 (included in Section B.2 of the *Toolkit*) —disaggregated for PFS and other migrant students;
- MPOs established for specific activities and services disaggregated for PFS and other migrant students at the service delivery level and summarized at the state level.

States that adopt a performance target for school readiness determine their performance indicators, usually adopting some measure of early literacy. If your state adopted a performance target for school readiness or any other state performance targets, you must be prepared to provide services that enable migrant students to meet those targets and to disaggregate performance data for PFS students, other migrant students, and non-migrant students related to those targets.

Other requirements for state MEPs:

- The comprehensive state plan for service delivery must determine the effectiveness of its program through a written evaluation (34 C.F.R. Section 200.83). (See *Section C Planning The Evaluation*)
  - o The MEP should examine program implementation within the first or second year of the program and every two-three years thereafter. (*Guidance*, Chapter VIII, C5).
  - o MEP results based on performance measures, state performance targets, and measurable program outcomes should be examined every year (*Guidance*, Chapter VIII, C5).
  - o The state must focus on migrant children who are PFS students and develop methods for disaggregating state assessment data and measurable outcomes in order to determine the impact of the MEP on PFS students (34 C.F.R. Section 200.84; *Guidance*, Chapter VIII, C8).
- The state MEP must use the results of the evaluation to improve services to migrant children (34 C.F.R. Section 200.84 and 200.85) (See *Section G Using Evaluation Findings*).
- The state MEP must periodically document the evaluation in a written report (34 C.F.R. Section 200.84). (See *Section F Communicating Evaluation Findings*)
  - o OME requests that states submit a written program evaluation report once every two to three years (*Guidance*, Chapter VIII, C5).

- o States should report the *purpose* of the evaluation, *methodology* for what data were collected and how they were collected, *results* of the implementation evaluation, *results* for PFS and other migrant students, and the *implications* for making decisions about the program (*Guidance*, Chapter VIII, D2).

#### **B.4 Local Requirements for Evaluation**

Local operating agencies (LOAs) that receive subgrants from the state MEP for migrant education must also evaluate the effectiveness of their services for meeting the needs of migrant students, especially PFS students. The local project evaluation should measure both the implementation of the project and student performance against the project's MPOs, the state's MPOs, and the state's performance targets. (*Guidance*, Chapter VIII, C3)

- The LOA should develop MPOs that are aligned with the state's MPOs (*Guidance*, Chapter VIII, B6).
- The LOA must use the results of the evaluation to improve services to migrant children (34 C.F.R. Section 200.85).
- LOAs should evaluate progress of migrant children in the project against MPOs, report these outcomes to the state MEP, and use evaluation results to improve services for children in MEP preschool projects (*Guidance*, Chapter VIII, C10).

States should require that:

- The LOA project application include a description of the project and the services it will provide in accordance with the state's Service Delivery Plan (SDP), as well as the MPOs and a plan for evaluating whether the project achieves these outcomes.
- Additional needs not covered in the SDP, but identified by the LOA, may also be addressed in the project application along with appropriate MPOs for related activities and a plan for evaluating whether these additional activities achieve the specified MPOs.
- In addition to examining the implementation or delivery of services, the local MEP measures the performance of PFS and other migrant students against other state performance targets and MPOs.

#### **B.5 State Responsibilities to Local Education Agencies for Evaluation**

State MEPs must provide guidance to local MEPs for evaluating their projects and services and must also ensure that these evaluations are conducted properly. In addition, the state MEP must ensure that the LOA conducts the evaluation properly (*Guidance*, Chapter VIII, C3).

States should:

- Notify local MEPs in advance of specific data they will need for the statewide evaluation and provide guidance for how to collect the necessary data (*Guidance*, Chapter VIII, C3).
- Monitor the progress of local MEPs against state and local MPOs.
- Require appropriate changes to ineffective projects before additional funds can be subgranted to LEAs for these projects.

## **B.6 Reflection Questions**

1. Do I understand the requirements for evaluating the MEP on the state and local levels?
2. Are these requirements reflected in the Evaluation Plan included in the SDP?
3. Are we collecting the appropriate data to determine the impact of implementation and results?
4. What changes are needed to ensure that the evaluation system meets requirements for measuring program effectiveness?
5. What technical assistance and monitoring should the state provide to ensure that LOAs conduct proper evaluation and use results for program improvement?

## **B.7 Resources and Tools in Appendix B**

### Appendix B.1 Checklist for State Migrant Education Program (MEP) Evaluation

## Appendix B.1 Checklist for State Migrant Education Program (MEP) Evaluation

**The following checklist is based on the statutes, regulations, and Federal guidance related to state and local evaluation of MEPs (see *Section B Overview of Statutes, Regulations, and Non-regulatory Guidance Related to Program Evaluation* for links to the full documents). It is important to read and be very familiar with the law and guidance. You can use this checklist as a quick reference guide to help determine if the evaluation plan meets the Federal requirements. The state MEP Evaluation:**

- Includes a written evaluation plan in the statewide Service Delivery Plan (SDP), which specifies how the state will collect data related to the implementation of MEP activities and services as well as the results achieved through these services and activities (34 C.F.R. Section 200.84).
- Collects data on state performance targets related to Performance Goals 1 and 5 for each grade, disaggregated for Priority for Services (PFS), other migrant and non-migrant students (34 C.F.R. Section 200.84, *Guidance*, Chapter VIII, B3, C8).
- Collects data on additional state performance targets for school readiness and other needs, disaggregated for PFS, other migrant and non-migrant students, if applicable (*Guidance*, Chapter VIII, B5)
- Collects data on six Government Performance and Results Act (GPRA) measures and reports it annually to the Office of Migrant Education, to be used in the evaluation of the Federal MEP with an anticipated start date in 2004.
- Collects data on Measurable Program Outcomes (MPOs) established for all MEP activities and services, disaggregated for PFS, and other migrant students (*Guidance*, Chapter VIII, B5).
- Notifies local MEPs in advance of specific data needed for the statewide evaluation and provides guidance for how to collect the necessary data (*Guidance*, Chapter VIII, C3).
- Provides guidance to local MEPs on what to evaluate locally and how to evaluate it (*Guidance*, Chapter VIII, C3).
- Develops a plan for reviewing all evaluation findings and using the results to improve services to migrant children (34 C.F.R. Section 200.84 and Section 200.85).
- Documents the evaluation in a written report, including the purpose of the evaluation, what data were collected and how they were collected, the findings of the implementation evaluation, results for PFS and other migrant students compared to all other students, and the implications for making decisions about MEP activities and services (34 C.F.R. Section 200.84; *Guidance*, Chapter VIII, D2).

# Section C: Planning the Evaluation

---

This section provides additional guidance on planning for the evaluation. It is important to bear in mind that evaluation of Migrant Education Programs (MEPs) takes place at two levels. State performance targets and MEP performance measures are used to monitor the progress and success of the entire program statewide. Measurable Program Outcomes (MPOs) are used to gauge the success of specific instructional and support services at the local level.

- At the state or program level, you will monitor the results of all MEP activities and services combined by disaggregating state performance targets for migrant students who are Priority for Services (PFS) eligible, all other migrant students, and all non-migrant students.
- At the level of specific services, you must craft one or more MPOs that state the results you expect to achieve if services are implemented effectively. MPOs for instructional services should relate to improvements in academic performance, while MPOs for other support services should reflect the changes you expect in migrant students' ability to effectively engage in their schooling.

In this section, the *Program Evaluation Toolkit* picks up where the *Service Delivery Plan Toolkit* left off (See *Section E* of the *Service Delivery Plan Toolkit*) at the point of deciding how to evaluate the implementation and results of specific educational and educationally related services outlined as part of the Service Delivery Plan (SDP).

## C.1 Evaluating Implementation

Implementation evaluation, also known as formative evaluation, focuses on the extent to which programs are delivered as intended. The results of the implementation evaluation improve programs. Throughout the needs assessment and service delivery planning processes, you have already made numerous decisions about what services to provide for migrant students and how to implement them. For example, members of the Comprehensive Needs Assessment (CNA) committee considered the greatest needs of migrant students in your state and evidence-based strategies for meeting those needs. The SDP planning committee then decided how best to deliver these services, balanced with their knowledge of the practicalities of providing these services to a specific group of people. The following questions might have guided these decisions:

- For planning decisions related to instructional services
  - What teaching strategies are most effective with migrant students?
  - What qualifications, training, and experience do teachers need to use these strategies effectively?
  - With what content and skills do migrant students need the most help?

- What instructional and assessment materials are most appropriate to address these particular needs?
- How much time is needed to adequately instruct students in the areas they need most?
- For planning decisions related to support services:
  - Which students need additional help with English language acquisition, transportation, basic nutrition, medical and dental care, and mental health services?
  - Which of these needs are the greatest?
  - What is the best way to communicate information about available services to parents who do not speak English as their first language?
  - What are the best times and locations for providing services to the greatest number of students in need?

These and other decisions related to these questions or these and other similar questions represent the specific intentions of the services you provide. Evaluating program implementation, that is, whether a service was delivered as intended involves determining whether your intentions were actually carried out in practice. Table C.1 shows how planning decisions relate to implementation evaluation.

**Table C.1 Sample Planning Questions and Possible Strategies**

Sample Planning Questions – Instructional Services	Possible Implementation Evaluation Strategies
What teaching strategies are most effective with migrant students?	<b>Observation:</b> Use a list of effective teaching strategies to create an observation protocol or checklist. Include specific teaching behaviors or activities that are consistent with these strategies. Have an internal or external evaluator use the protocol to periodically observe instructional activities (supplemental, after-school, summer sessions) and determine the extent to which the most effective strategies are being used and how effectively they are used.
What qualifications, training, and experience do teachers need to use these strategies effectively?	<b>Interview:</b> Make a list of the ideal qualifications, training, and relevant experience(s) that you feel teachers need in order to be effective in the instructional program. Have an internal or external evaluator interview teachers, ask them to describe their qualifications, training, and experience working with migrant students and using the teaching strategies for this program. The evaluator can then compare teachers’ answers to the list of ideal qualifications, etc. to determine the extent to which they are qualified to implement the instructional program.
With what content and skills do migrant students need the most help? What instructional and assessment materials are most appropriate to address these particular needs?	<b>Document Review:</b> Review needs assessment data, lesson plans, curriculum materials, and assessments. Determine the extent to which the materials being used in the program are aligned with identified needs. Also confirm that the needs of the students being served by the program are the same as the needs the

	program was intended to address.
How much time is needed to adequately instruct students in the areas with which they need the most help?	<b>Attendance Records:</b> Compare the time individual students actually spend in the program with the amount of time that is recommended in order to achieve progress.

## C.2 Evaluating Results

Services are expected to have a positive impact on migrant students' ability to achieve academic success. Evaluating program results, also known as summative evaluation, involves two key steps:

Step 1: Determine how much change is both necessary and realistic in order to make the related investments of time, money, and other resources worthwhile

Step 2: Determine how much change actually occurred as a result of these investments.

MPOs, based on identified needs and appropriate services, should express the amount and type of change that will make it worthwhile to continue offering a specific service in a particular way helping to complete Step 1. Remember, MPOs are focused, detailed, and quantifiable, and they should clearly define "success" in meeting a particular need of migrant students. An appropriate MPO is one that articulates the difference that participation in the MEP will make for migrant students. Key components of an MPO define:

- Which students participate
- What happens in the program
- What is expected to happen as a result of participation in the migrant program
- The timeframe in which activities occur

Setting specific and quantifiable MPOs in Step 1 will make it easier to achieve Step 2, which involves measuring the changes specified in order to determine how much change actually occurred. For example, review the following need, related strategy, and MPO:

- Need: Less than half of migrant students in the 3rd grade score at or above the proficient level on the state reading assessment. Most of these students (82%) are also Limited English Proficient (LEP).
- Strategy: For migrant LEP students who are reading below grade level, research supports the use of structured phonetics programs that emphasize language development in both native-language and English instruction. A minimum of 108 hours of supplemental instruction is recommended.
- MPO: 85% of migrant LEP students in 3rd grade who read below grade level at the beginning of the year and who complete supplemental instruction will demonstrate progress equivalent to at least one grade level in their use of phonics knowledge and

word parts to decode unfamiliar words as measured by the Test of Word Reading Efficiency (TOWRE).

This sample MPO provides clear direction to evaluate or determine the value of a supplemental reading program for 3<sup>rd</sup> grade English language learner (ELL) students. At the beginning of the school year, ELL migrant students in 3<sup>rd</sup> grade who read below grade level take the TOWRE as a pre-test of their ability to recognize sight words and decode unfamiliar words phonetically. Following completion of a supplemental instructional program, these same students will take the TOWRE as a post-test to determine how much progress they have made. More exercises on developing MPOs can be found in *Appendix C.1 Setting Measureable Program Outcomes (MPOs)*.

The data from pre and post tests can be used in different ways to measure how much change actually occurred as a result of these investments (Step 2). For example:

- Pre-test and post-test assessments of students can be used to measure students' gains in comparison to themselves. In this case, the value of the supplemental reading program can be measured by the number or percentage of students who show gains or by the increase between average pretest and average posttest scores for all students who completed the program.
- The value of the program can also be measured by comparing the percentage of students who gain a full grade level in word recognition and decoding skills to the performance target that is stated in the MPO.

### **C.3 Generating Evaluation Questions**

Having considered the essential elements of the service to be delivered, as well as the type and extent of change that will make the service worthwhile, it is time to pose a set of specific questions related to the implementation and results of the program.

- Evaluation questions about program implementation relate to the presence or absence of the essential features of the program or service that is being delivered.
- Evaluation questions about program results relate to the achievement of MPOs.

A credible evaluation can be based on as few as three to five good evaluation questions. For example, below are evaluation questions about the implementation and results of the supplemental reading program described above.

- Evaluation questions about implementation:
  - What training or experience does the teacher have to provide phonics instruction in English as well as the students' native language?
  - Are students focused and engaged when receiving instruction? What tasks or exercises appear to hold their attention most, or least?

- Are students able to complete the minimum number of hours recommended? If not, what prevents them from receiving the recommended amount of instruction?
- Evaluation questions about results:
  - Do students who complete the program make gains in their ability to recognize sight words or decode unfamiliar words? Are these gains substantively significant?
  - How many, or what percentage of, students gain the equivalent of one full grade level in word recognition and decoding? How does this percentage compare to the expected outcome?

#### **C.4 Evaluation Timeline**

An evaluation timeline can help to organize evaluation questions and identify key dates and deadlines. As you begin to consider methods for collecting and analyzing evaluation data, a timeline can also track when you need to conduct specific evaluation activities in order to generate timely information for decision making and reporting. An evaluation timeline template can be found in *Appendix C.2 Evaluation Timeline Template*.

Keep in mind the following when developing an evaluation timeline:

- Before you plan how to evaluate your program, you already know some important dates to include in your timeline. These include service delivery dates, specific dates on the school calendar (e.g., beginning and end of marking periods, testing dates), and state and local report deadlines.
- Start by entering these dates, as they can either limit your evaluation activities at certain times or provide you with benchmarks for beginning or completing these activities.
- Keeping an eye on your timeline as you develop your evaluation plan will help you to think critically about what you can reasonably accomplish by the program and its evaluation in the time available to you.

Consider the following example: Program staff will offer the parents of migrant students a variety of services to make them more knowledgeable about how they can support their child's schooling and academic achievement. Having considered their data collection options and modest evaluation budget, they believe they can adequately address their evaluation questions using a combination of parent focus group(s), one-on-one interviews with service providers, and state test data.

- In the case of the parent involvement program described above, putting established dates on the timeline allowed program staff to see that preliminary state test data would not be available until after the end-of-year district budget meeting (see Table C.2). In order to report on outcomes at this meeting, they decided to examine results on district benchmark tests instead of state test results.

**Table C.2 Program Evaluation Timeline**

Program and Evaluation Activities	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
<b>Program Dates</b>												
Workshop: How to read with your child		X	X	X	X		X	X	X	X		
Workshop: How your child is learning math		X	X	X	X		X	X	X	X		
Workshop: How to encourage and support homework time		X	X	X	X		X	X	X	X		
<b>Evaluation Dates</b>												
Parent Focus Groups						X				X		
Service Provider Interviews						X				X		
District benchmark tests				X		X		X		X		
Report of benchmark test results										X		
Preliminary state test results												X
District budget meeting											X	

**C.5 Summary of Key Points**

- State performance targets are used to assess the impact of all MEP programs and services combined.
- MPOs define the type and extent of success you expect to achieve through the delivery of specific instructional and support services.
- Implementation evaluation, also known as formative evaluation, focuses on the extent to which programs are delivered as intended and how this information can be used to improve programs.
- Evaluating program results, also known as summative evaluation, involves making meaningful comparisons to assess the value or worth of a program according to the terms stated in the MPOs.
- Summarizing expectations for program implementation and results in a set of specific evaluation questions helps focus the evaluation on what you most want to know about your program.
- A timeline of program activities, evaluation data collection, and reporting or decision making deadlines helps you think critically about what can and should be accomplished in the time available for your program.

**C.6 Reflection Questions**

1. Do our MPOs define the type and extent of success we expect to achieve through the delivery of specific instructional and support services?

2. Do the evaluation questions focus on what we most want to know about our program?
3. Do we have a timeline of program activities that helps us think critically about what can and should be accomplished and evaluated in the time available for our program?

### **C.7 Online Resources**

U.S. Department of Education. (2006). *Reading First Implementation Evaluation: Interim Report*. Executive Summary retrieved August 23, 2012, from <http://www2.ed.gov/rschstat/eval/other/readingfirst-interim/readingfirst.pdf>

World Health Organization. (2000). *Workbook7: Outcome Evaluations*. Retrieved August 23, 2012, from [http://www.emcdda.europa.eu/attachments.cfm/att\\_5869\\_EN\\_7\\_outcome\\_evaluations.pdf](http://www.emcdda.europa.eu/attachments.cfm/att_5869_EN_7_outcome_evaluations.pdf)

W.K. Kellogg Foundation. (2004). Implementation evaluation: understanding how the project was implemented. In *Program Evaluation Handbook* (24-27). Retrieved August 23, 2012, from <http://www.wkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx>

### **C.8 Resources and Tools Appendix C**

Appendix C.1 Setting Measurable Program Outcomes (MPOs)

Appendix C.2 Evaluation Timeline Template

## Appendix C.1 Setting Measurable Program Outcomes

Now that you know more about what useful Measurable Program Outcomes (MPOs) should entail, you can practice setting appropriate MPOs below. After you have worked through both of these exercises, you can find helpful suggestions on the following page.

**Exercise 1:** Let's say that your planning committee developed the following MPO:

**Measurable Program Outcome:** By the end of the 2013–2014 school year, and each year thereafter, the gap between migrant students and all students will be reduced in reading/language arts in grades 3–8.

How would you rework this MPO to be more useful?

Articulate in the box below a more useful MPO.

**Reworked Measurable Program Outcome:**

**Exercise 2:** Now let's say that your planning committee identified the following need:

**Need:** By the end of grade 9, only 40% of migrant students pass Algebra 1, as compared to 70% of non-migrant students.

Propose at least one corresponding strategy and MPO that will address this need?

**Strategy:**

**Measurable Program Outcome:**

### ***Exercise 1: Helpful Hints***

*There are several ways to make this MPO more effective, including:*

- Establishing the assessment or measure of reading/language arts you will measure change
- Establishing by what percentage of migrant students to establish incremental success in closing the “gap”
- Establishing better comparison groups to measure progress – (ideally broken out by 3 mutually exclusive groups: Priority for Services migrant students, other migrant students, and non-migrant students)
- Establishing either explicitly or implicitly through program logic that such incremental successes should occur as a result of what specific services or supports

### ***Exercise 2: Helpful Hints***

*There are a wide range of strategies and MPOs you could have proposed here. See Section E of the Service Delivery Plan Toolkit for examples of working through program logic.*

Appendix C.2 Evaluation Timeline Template

Program and Evaluation Activities	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
<b>Program Activities (Service Delivery Plans)</b>												
<b>Evaluation Activities (Data Collection &amp; Reporting)</b>												

# Section D: Collecting Evaluation Data

---

As you plan and focus your evaluation, you must also think about what kind of information you will need to answer your evaluation questions. Section D provides guidance for locating existing education data and collecting original data through interviews, surveys, and observations. Links to online resources are included to point you to more detailed treatments of the topics we identify.

As you consider your data collection options, keep in mind the following general principles:

- Try to collect data from more than one source to answer each of your evaluation questions. You might analyze assessment data for all of the students in your program and also conduct a focus group with selected students to find out what aspects of the program were most and least helpful. Or, you might survey students and parents to identify the ways in which the program impacted students.
- Collect a mix of quantitative and qualitative data. Quantitative or numerical data are useful for discovering the magnitude of a phenomenon (e.g., outcomes, barriers, facilitators). Qualitative or narrative data can help you to better understand the phenomenon (e.g., who benefits most from a program, why a program works in one setting but not another, what additional support is needed to improve outcomes).
- Only collect data that can actually be used to evaluate your program. Asking for more information than you need taxes the people who are providing you with data in the short-run and taxes your own ability to analyze and use the information in the long-run. Always keep your evaluation questions front and center in order to stay focused on the information you really need.

## D.1 Using Existing Data

Before you spend time collecting new information, become familiar with the kind of information that is already being collected and figure out how to access information that will be useful in the systematic evaluation of your program. In many cases, particularly with respect to program outcomes, you may find the answers to your questions among the wealth of education data that is already routinely collected at the school, district, and state levels. If the available data are not suitable to answer your evaluation questions, you may wish to modify your questions slightly in order to make use of available data. Here are two types of large-scale data collection efforts that may be useful.

- Migrant Student Information Exchange (MSIX) – MSIX is an OME initiative to enable states to share education and health data for migrant students. Consolidated student records include demographics, enrollment, assessment, and course history in every state where a migrant student has attended school. For more information, go to

<https://msix.ed.gov/msix/LocateAnAdmin?submit=d> to find contact information for your state or district user administrators. See also <http://reactsmsix.ed.gov/> for information about technical assistance and other support available to users of MSIX data.

- Statewide Longitudinal Data Systems (SLDS) – The Institute for Education Sciences (IES) in the U.S. Department of Education has awarded grants to 42 states to design, develop, and implement “P-20” longitudinal data collection systems that link preschool, K-12, and postsecondary education using unique state identifiers for every student. To find out whether your state has received an SLDS grant, go to: <http://nces.ed.gov/programs/slds/stateinfo.asp>. To learn more about your state’s progress in developing a P-20 data system, including creation of unique student identifiers, go to: <http://www.dataqualitycampaign.org/>.

Note that the identification numbers in MSIX are different from the identifiers in statewide systems that cover all students. Learning more about both systems and contacting the people in your state who are responsible for them will help you figure out how to use them to compare Priority for Services (PFS) migrant students, other migrant students, and non-migrant students—statewide or in selected local project samples.

Other sources of information about existing data include:

- The school or district data specialist who inputs or uploads attendance, assessment, and other information to the state database
- The district grade-level or subject matter specialist who will know what, if any, benchmark assessments are administered in reading/language arts and mathematics and when they are administered
- The school computer lab manager who will know what, if any, learning software programs have built-in assessments that are routinely used by teachers at the school

## D.2 Conducting Interviews

Interviews can provide in-depth information about how a program is working, the impact on participants, barriers to implementation, and program successes. There are several options for conducting interviews—with individuals (either face-to-face or by phone) and in focus groups.

- Face-to-face interviews are best in the case of sensitive subjects and should be conducted by someone with especially good interpersonal skills who knows how to put people at ease.
- Telephone interviews work well when systematic data collection is desirable, content is not too sensitive, and budgetary constraints prevent the interviewer from travelling to meet the interviewee (or vice versa).
- When you want to test out ideas or assess the extent of group consensus on a topic, a focus group interview can be a cost-efficient alternative to individual interviews. On the

other hand, a focus group is not an appropriate or effective method for collecting sensitive personal information.

Costs related to conducting different types of interviews include:

- The time needed to schedule individual and group interviews, especially in the case of focus group interviews
- In the case of individual interviews, especially telephone interviews, the time needed to re-schedule interviews with “no-shows”
- Time needed for documenting responses (e.g., transcribing notes or recordings), coding them according to themes or prominent ideas, and analyzing responses within thematic groups
- In the case of focus group interviews, the costs of renting space in a neutral location or familiar community place, and providing child care and refreshments as an incentive for participants to take part during their free time
- Time needed to actively recruit focus group participants and follow up to remind them of the date, time, and place where it will be held
- The cost of a second party to take notes and observe the dynamic in a focus group, so the focus group facilitator can concentrate on leading the conversation

For more information about conducting different types of interviews, review the following online resources:

- A good overview of face-to-face, telephone, and focus group interviews: [The Community Tool Box, Chapter 3 Assessing Community Needs and Resources, Section 12. Conducting Interviews](#). A service of the Work Group for Community Health and Development at the University of Kansas.
- Step-by-step instructions for setting up and conducting focus groups: [The Community Tool Box, Chapter 3. Assessing Community Needs and Resources, Section 6. Conducting Focus Groups](#).

### D.3 Administering Surveys

Surveys are generally a good way of getting specific information from a large group of people in a relatively short period of time. There are two types of questions used on surveys: close-ended and open-ended.

- Close-ended questions include a specific set of answers, or response options, from which survey respondents select the best one(s). Response items may include the points on a five-point scale (1 = strongly disagree...5 = strongly agree), a list of services offered (“check all services you have received”), and other types of answers that respondents simply check or circle. To be sure all possible answers are captured, close-ended questions may include a response for “Other (please specify): \_\_\_\_\_.”

- Open-ended survey questions require survey respondents to write answers in their own words. These types of questions are used when the possible response options are unknown, or when we want to know what respondents think without prompting them. For example, the question “How has your child benefited from the Summer Program?” could include a list of all the benefits we hope they received, but ideally we would leave the question open-ended in order to get the most accurate, unprompted answer.

All surveys require time to develop clear and concise questions that respondents will be motivated and able to answer.

- Use plain and simple language so your questions can be understood by the broadest range of people. If English is not the first language of your target population, translate the survey into their language.
- Open-ended survey questions should be used sparingly, only when absolutely necessary. Use interviews and other qualitative data collection methods to identify the range of answers to a question. Then use these answers as close-ended response options to your survey question.
- Stay focused on one idea at a time, avoiding double-barreled questions, such as, “Do you feel your child benefited from instruction and other support services provided by this MEP?” Because instruction and support services are different in nature, respondents may have different answers for each category of services. So, ask two questions instead: “Do you feel your children benefited from the instructional services provided by the MEP?” and “Do you feel your child benefited from MEP support services?”

Writing good survey questions is more involved than it looks, but the payoff in terms of data quality is worth taking the time to learn more. Here are a few resources that provide useful guidance on how to prepare good survey questions:

- For guidance on wording, formatting, and pretesting surveys, see: Taylor-Powell, E. (1998.) [Questionnaire Design: Asking questions with a purpose](#). University of Wisconsin-Extension.
- For examples of “good and not so good” survey questions, go to: University of Wisconsin, Cooperative Extension, Program Development and Evaluation website on [“Local Program Evaluation in Tobacco Control.”](#)

Surveys can be administered in person, over the phone, by mail, or online. Your selection should be based on the method that is most likely to reach your target population without systematically excluding members of that population, and the costs associated with the method. For example:

- The cost of telephone surveys varies depending on the number of people you want to survey, whether you have staff who are trained and available to make the calls and whether you need to contract with a telephone survey center to make the calls. Similarly, the cost of mailed surveys varies depending on the number of people you want to survey, the cost of the initial mailing plus self-addressed and stamped envelopes, incentives to increase response rates, and mailed reminders. Because phone numbers and mailing addresses for migrant families change frequently, these methods may not be the best for reaching this target population.
- A wide variety of online survey tools makes it easier than ever to assemble and disseminate online surveys. See the idealware.org list of [“A Few Good Online Survey Tools.”](#) The cost of these services may vary depending on the number of surveys you wish to collect, the number of questions in your survey, and the length of time you need to keep your survey open for responses. However, online surveys are only useful if you have accurate email addresses and the people you want to reach either own or regularly access the internet.
- The costs associated with administering surveys in person depend on whether respondents come to you (e.g., students on the last day of the program, parents attending a culminating event) or you go to them (e.g., stationed at a popular location, going door-to-door in the community). Given the limitations of other survey methods for reaching the target population, administering surveys in person may be the most effective way to collect survey data from migrant students and their families.

For more information about survey methods, planning and implementing surveys, plus sample survey materials, review:

Taylor-Powell, E & Hermann, C. (2000). [Collecting Evaluation Data: Surveys](#). University of Wisconsin-Extension Cooperative Extension.

#### **D.4 Observing Activities**

An often overlooked way to collect evaluation information is to directly observe a program, activity, service, or meeting to decide for yourself whether things are going as planned or whether there is evidence of the desired outcomes. Costs associated with collecting observational data include training observers, traveling to observation sites, and time needed to document observations and analyze them.

Some important principles of conducting observations include:

- Scheduling observations at a time that is convenient and least intrusive for the people and activities you are observing.
- Developing a checklist or protocol to guide your observations and ensure that you look for key elements associated with high-quality implementation or desired outcomes.
- Training observers to recognize and document the same key elements of the program.

For more information about how to conduct observations, review:

Taylor-Powell, E & Steele, S. (2000). [Collecting Evaluation Data: Direct Observation](#).  
University of Wisconsin-Extension Cooperative Extension.

## D.5 Evaluation Matrix

An evaluation matrix is a useful tool for keeping track of the methods you are considering for collecting data. In its simplest form, a matrix displays evaluation questions in alignment with the data collection strategies that will be used to answer them. The example below (Table D.1) shows a simple evaluation matrix for the parent involvement program we used in the last section to illustrate the evaluation timeline.

**Table D.1 Sample of a Simple Evaluation Matrix**

<b>Strategy: Provide parents of migrant students with a menu of support services to help them support their child’s academic accomplishments.</b>			
<b>Evaluation Questions</b>	<b>Parent focus group</b>	<b>Service Provider interviews</b>	<b>State test scores</b>
<b>Implementation</b>			
1. Are program staff trained to provide these services?		X	
2. Are program staff qualified to provide these services?		X	
3. What support services are most requested by parents?	X	X	
4. What support services are least requested by parents?	X		
5. Do services help parents support children academically?	X		
<b>Results</b>			
6. Do children whose parents participate in services perform better on the state reading test compared to a sample of migrant children whose parents did not participate?			X
7. Do children whose parents participate in services perform better on the state math test compared to a sample of migrant children whose parents did not participate?			X

In this example, staff members use a program strategy to frame the implementation and results evaluation questions, and then match these questions with data collection methods.

- Program staff members will offer the parents of migrant students a variety of services to make them more knowledgeable about how they can support their child’s schooling and

academic achievement. Having considered their data collection options and modest evaluation budget, they believe they can adequately address their evaluation questions using a combination of parent focus group(s), one-on-one interviews with service providers, and state test data.

- The matrix shows program staff members that they have at least one strategy for answering each of their questions. What remains is for them to decide who will conduct the focus groups and one-on-one interviews, what specific questions should be asked in those interview settings, and who at the school or district level can help them extract state test data for the children of program participants and a sample of other similar children whose parents did not participate.

The basic structure of the evaluation matrix can be as simple or complex as needed to help focus your Evaluation Plan. A more complex matrix might organize the evaluation questions by state performance targets, strategies, and related Measurable Program Outcomes. For the sake of simplicity you might set up a separate matrix for each performance target (see Table D.2).

**Table D.2 Sample of a Complex Evaluation Matrix**

State performance target 1: All students will reach high standards, at a minimum attaining proficiency or better in reading/language arts and math.			
Strategy: Provide parents of migrant students with a menu of support services to help them support their child’s academic accomplishments.			
MPO 1a: Children whose parents attended the reading and homework support workshops will perform on average 10 points higher on the state reading test than children of non-participants. MPO 1b: Children whose parents attended the math and homework support workshops will perform on average 10 points higher on the state math test than children of non-participants.			
Evaluation Questions	Parent focus group	Service provider interviews	State test scores
<b>Implementation</b>			
1. What support services are most requested by parents?	X	X	
2. What support services are least requested by parents?	X	X	
3. Are program staff trained to provide these services?		X	
4. Are program staff qualified to provide these services?		X	
5. Do services help parents support children academically?	X		
<b>Results</b>			
6. How much better do their children perform on the state reading test compared to a sample of children whose parents did not participate in the reading and homework support workshops?			X
7. How much better do their children perform on the state math test compared to a sample of children whose parents did not participate in the math and homework support workshops?			X

In *Appendix D.1*, you will find a simple evaluation matrix template and a complex evaluation matrix template, based on the examples above, that may be customized for your needs.

## D.6 Summary of Key Points

- Try to collect data from more than one source to answer each evaluation question.
- Collect both quantitative and qualitative data.
- Before collect data, find out what data are already being collected and how these data might be used to answer your evaluation questions.
- When it is necessary to collect original data, interviews, surveys, and observations offer a wide variety of options for reaching your target population.
- Learn more about these and other data collection methods by accessing the online resources identified in this section of the *Program Evaluation Toolkit*.

## D.7 More Online Resources

[W. K. Kellogg Foundation Evaluation Handbook](#). (2004)

For a more detailed overview of data collection methods, see Step 5: Determining Data-Collection Methods (pp. 70–84) and Step 6: Collecting Data (pp. 84–87).

Administration for Children and Families. Office of Planning, Research and Evaluation. [The Program Manager’s Guide to Evaluation, 2<sup>nd</sup> Edition](#)

For step-by-step considerations of data collection decisions and a sample data collection plan, see Chapter 7: How Do You Get the Information You Need for Your Evaluation? (pp. 63–76).

## D.8 Reflection Questions

1. Have we accessed available sources to determine if the data we need is currently available?
2. Have we determined the best method for collecting existing and original data?
3. For interviews and surveys, have we created questions that will help us assess our program in an objective way?
4. Have we assessed the costs involved and have we planned for them?

## D.9 Resources and Tools in Appendix D

Appendix D.1 Simple Evaluation Matrix Template

Appendix D.2 Complex Evaluation Matrix Template

Appendix D.1 Simple Evaluation Matrix Template

Evaluation Questions	Possible Data Collections Strategies						
Implementation							
1.							
2.							
3.							
4.							
5.							
6.							
Results							
7.							
8.							
9.							

Appendix D.2 Complex Evaluation Matrix Template

<b>State Performance Target:</b>							
<b>Strategy:</b>							
<b>Measurable Program Outcome(s):</b>							
<b>Evaluation Questions</b>	<b>Possible Data Collection Strategies</b>						
Implementation							
1.							
2.							
3.							
4.							
5.							
6.							
Results							
7.							
8.							
9.							

# Section E: Analyzing and Interpreting Data

---

Program evaluation requires analyzing and interpreting data in order to make decisions about programs such as what program activities to fund, which activities to improve, and how to improve them. Up to this point in the evaluation process:

- You focused your evaluation with a select set of questions about program implementation and results in order to make appropriate decisions about your program.
- You also selected practical and affordable methods for gathering reliable and credible data.

The next step is to make sense of the data you collected so that you can answer specific evaluation questions about the quality of your program, what it accomplished, and the extent to which different categories of participants benefited.

How you analyze, organize, or summarize your data depends on the type of data you collected:

- Quantitative data, or numerical information, is analyzed or summarized using descriptive and inferential statistics.
- Qualitative data, or information gathered in narrative form, is analyzed or organized into categories or themes, which can in turn be summarized by percentages.

In this section of the *Program Evaluation Toolkit*, there is a description of how to analyze and interpret quantitative data using a selection of descriptive statistics, and how to analyze and interpret qualitative data by organizing it into meaningful categories. Data storage options that facilitate subsequent analyses are also discussed.

## E.1 Analyzing Quantitative Data

Numerical data are used to summarize useful and important information about large groups of people. In order to get the most out of this information, it is helpful to know something about the different types of data elements, or variables, which can be analyzed using quantitative methods.

Categorical variables describe people and things in discrete categories. For example, gender is a categorical variable that has two categories – male and female.

- Migrant status is a categorical variable that can have two categories (migrant and non-migrant) or three categories (Priority for Services [PFS], other migrant, and non-migrant).

Continuous variables are expressed as numbers on a continuous scale. For example, age is a continuous variable that can be measured with different levels of precision (e.g., years, months, days, hours).

- A scale score is a continuous variable representing student achievement, but proficiency level is a categorical variable that is also used to describe student achievement.

Even though categorical variables are represented by names or labels and continuous variables are represented by numbers, both types of data can be analyzed using quantitative methods.

In the broadest sense, analyzing data means looking for patterns that are relevant to what you are studying. In the case of evaluation, which relies on making comparisons to assess value, patterns are observed in order to compare two or more groups. For example:

- Those who participated in a program and those who did not
- PFS and other migrant and non-migrant students
- Students who achieved success as defined by measurable program outcomes and those who did not

The following descriptive statistics can help reveal relevant patterns when comparing for two or more groups.

### **E.1.1 Descriptive Statistics for Categorical Variables**

Frequencies are counts calculated for categorical variables to describe the composition of a group such as a group of people, things, or survey responses. These counts express amounts and are usually translated into percentages for easy monitoring of changes over time. For example, changes in the demographic composition of migrant students in your state may explain an increase in demand for certain types of services or a decline in participation in longstanding services.

The two branches of statistics are descriptive statistics and inferential statistics.

Descriptive statistics describe a specific sample of people or things.

Inferential statistics, including the use of statistical significance tests, are used to understand or draw conclusions about the population from which your specific sample was obtained.

For example, you could use descriptive statistics to identify changes in academic performance among migrant students who received a specific instructional service (e.g., one-on-one tutoring) through the local migrant education program, you could then use inferential statistics to determine whether these findings could be generalized to the population of migrant students in your school district, as long as they also received the same instructional service.

A rate is a percentage that expresses the extent of an action or condition. You and your colleagues routinely compare academic success rates for migrant and non-migrant students when you calculate the percentage of students in each group who have achieved proficiency, and you look for changes in the pattern of differences between the two groups to see whether instructional services for migrant students are helping to close the gap with non-migrant students.

A cross-tabulation, or crosstab, is the calculation of a set of frequencies using two or more categorical variables to describe a group. The result of this calculation is a contingency table that shows how the members of your group are sorted into subgroups depending on the categories they inhabit for each variable. Figure E.1 is an example of a simple 2x2 contingency table, in which frequencies are calculated for two variables at the same time and both variables have only two categories:

**Figure E.1 Example of 2x2 Contingency Table**

English as a Second Language (ESL) Program Participation Status	Migrant Status of Limited English Proficient (LEP) Students		Row Totals
	# Migrant LEP Students	# Non-Migrant LEP Students	
# Who participated in the after-school ESL program	50	275	325
# Who did not participate in the after-school ESL program	100	75	175
<b>Column Totals</b>	150	350	500

In this example, a total of 500 Limited English Proficient (LEP) students, some of whom were migrant, some of whom were non-migrant, were identified as being able to benefit from a district-wide after-school English as a Second Language (ESL) program. Space was budgeted for all 500 eligible students. The row totals show that 325 of the eligible students participated. The remaining 175 eligible students did not participate. For eligible LEP students as a whole, the participation rate was 65% ( $325 \div 500$ ). This result is respectable for a new program.

To identify ways to improve participation, consider reasons why 35% of eligible students did not participate. Knowing that the reasons for non-participation may be different for migrant and non-migrant students, examine the cells in each column and note that the participation rate for migrant students ( $50 \div 150$ , or 33%) was much lower than the participation rate for non-migrant students ( $275 \div 350$ , or 79%).

Based on this finding, make it your first priority to identify the reasons why migrant LEP students chose not to participate, or were not able to participate, in the ESL program.

### **E.1.2 Descriptive Statistics for Continuous Variables**

The mean for the distribution of a continuous variable is the same as the average for a set of continuous, numerical values. A mean is one of several “measures of central tendency” used to calculate a “typical” number to describe a whole group of numbers or the people they represent and compare it to the typical value for another group, or for the same group at another point in time.

Consider the case of a state assessment system that defines cut points that divide continuous scale scores into four levels of proficiency: (1) well-below proficiency, (2) below proficiency, (3) proficient, and (4) advanced. Your state performance target is to increase the percentage of PFS students and other migrant students who are proficient in reading by 10% this year in each grade.

- In spite of all your targeted and evidence-based efforts, you are discouraged to find that only small percentages of students in each grade have gained enough to move into the next highest proficiency level. In order to get a more precise view of what has occurred in the lower two proficiency levels, you calculate mean scale scores for migrant students in each of these groups and find positive change in the mean scale scores for each group in every grade. So, even though students did not meet their state performance targets, you can see that they did achieve growth.
- Some of the differences in the grade-by-grade mean scale scores are bigger than others. You compare strategies used in the most successful grades with those used in less successful grades to identify programs that may have made a difference.

A median is another measure of central tendency and is often used to describe a set of numbers that includes some extreme values, or outliers. When the distribution includes outliers, the mean can be skewed and either overestimate or underestimate the middle or “typical” value. While the mean is the computed middle, the median is literally the middle-most number in a set of numbers when you arrange them in order.

For more information about analyzing quantitative data, go to:

[Quantitative Data Analysis: An Introduction, U.S. General Accounting Office, May 1992](#)

[Taylor-Powell, E. \(1996\). Analyzing Quantitative Data. University of Wisconsin-Extension.](#)

[Leahy, J. \(2004\). Using Excel for Analyzing Survey Questionnaires. University of Wisconsin-Extension, 2004.](#)

## E.2 Analyzing Qualitative Data

While quantitative data answer questions related to magnitude (how many, how much, to what extent), qualitative data can help us understand more complex issues related to “how” and “why.” For this reason, many evaluators and program directors prefer working with qualitative data.

Just like quantitative data, qualitative data are analyzed by systematically looking for patterns. If instead of studying numbers, researchers carefully comb through interview transcripts or open-ended survey responses to find the themes and patterns that emerge in people’s answers to our questions. Categories and subcategories are then defined to describe different kinds of answers. In order to convey the relative strength of the themes that are uncovered, the qualitative analysis findings are then summarized by calculating the percentages of respondents who provided answers in the different categories and displaying them in a table.

Step 1: Organizing individual responses into broad categories. Qualitative researchers create categories, or codes, of increasing specificity as they review data. Level 1 codes are the broadest categories that can be used to describe your data.

- In the case of a focus group or interview that includes multiple questions, you can use your interview questions as your Level 1 categories.
- In the case of an open-ended survey question that was answered by people in different roles, you might use the roles of the respondents (student, parent, teacher, Migrant Education Program [MEP] staff) as Level 1 categories.
- You can begin with a set of Level 1 categories or themes that you expect to find in the data based on your involvement in the data collection, past experience, or knowledge of the subject area being discussed.

Consider the example of an open-ended survey question that was answered by a group of migrant students in 10<sup>th</sup> grade and by the parents of these students. Seventy migrant students and 85 parents attended a four-hour Saturday morning information session to learn about preparing for and succeeding in college. Students and parents attended the same sessions and received the same information. Therefore, you might begin with two Level 1 categories: one category for answers given by the students [coded “STU”] and a second Level 1 category for answers given by parents [coded “PAR”].

You have invited students and parents to a brief follow-up session to provide additional information they requested. Forty-nine students and 62 parents attend the follow-up session. Before they leave, you give each of them a half-page questionnaire with two questions. The first question asks them to use a five-point scale (from 1 = “not at all useful” 5 = “very useful”) to rate the usefulness of the information they have received for planning their next steps toward college. The second question reads, “What specific pieces of information, if any, were useful to you?” There is plenty of blank space on the questionnaire to write their answers.

Step 2: Within each broad Level 1 category, organize individual responses into related subcategories. The process of sorting responses in Level 1 categories into narrower and more informative sub-categories is known as Level 2 coding.

- Going back to our example, you have collected surveys from all of the students and parents and are ready to begin sorting their answers into more meaningful categories. You begin reading through each group’s individual answers to identify Level 2 sub-categories.
- The information that was most useful to students fell about equally into three sub-categories – academic, extracurricular, and financial. You code their answers using the abbreviations “ACAD,” EXTRA,” and “FINAN,” respectively. Parents’ answers also fell into the same sub-categories, but most of the information they found useful fell into the “FINAN” sub-category.

Step 3: Summarize results in a table showing percentages of responses in each Level 2 sub-category and for each Level 1 category.

Level 2 Categories	Level 1 Categories	
	Students (n=49)	Parents (n=62)
Categories of useful information		
Academic	39%	37%
Extra-curricular	42%	12%
Financial	29%	77%
Non-specific, positive (e.g., “It was all useful”)	4%	6%
Non-specific, negative (e.g., “Nothing was useful”)	2%	---

\*Percentages add to more than 100% for each group because answers fall into more than one sub-category.

- The findings summarized in the table reinforce the generally high ratings given for usefulness by students and parents who attended the follow-up session. The vast majority of participants in both categories could name at least one piece of useful information they had received in the original four-hour information session.
- Given the different kinds of information that students and parents found useful, you and your colleagues decide to offer the four-hour information session again next year, but you plan to separate students and parents and provide information that will be tailored to what each group finds more useful for college planning.

### **E.3 Summary of Key Points**

- Analyzing and interpreting the data help identify themes or patterns that emerge from the information collected. These themes can be used to answer specific evaluation questions about the quality and impact of programs.
- Quantitative data, or numerical information, is analyzed or summarized using descriptive and inferential statistics. Quantitative data can be used to summarize useful information about large groups of people.
- Qualitative data, or information gathered in narrative form, is analyzed or organized into categories or themes, which can in turn be summarized by percentages. Qualitative data can help us understand more complex program issues related to “how” and “why.”
- Both kinds of data – quantitative and qualitative – can be summarized in tabular form to reveal patterns, show comparisons or demonstrate impact.

### **E.4 Reflection Questions**

1. Have we considered the appropriate variables in the interpretation of the data?
2. Have we identified all of the major themes and subthemes?
3. Have the data been analyzed in a way that answers specific evaluation questions?
4. Are there other ways of interpreting the data that we have overlooked?

### **E.5 Resources and Tools in Appendix E**

Appendix E.1 Using Excel to Analyze Quantitative Data

Appendix E.2 Using Access to Analyze Qualitative Data

Appendix E.3 Using Inferential Statistics

## Appendix E.1 Using Excel to Analyze Quantitative Data

Microsoft Excel, which is widely available, is able to perform many calculations that are useful in analyzing data. The following instructions are based in large part on a document called *Using Excel for Analyzing Survey Questionnaires* published by the University of Wisconsin Cooperative Extension in 2004.

While this publication assumes the use of Microsoft Excel 2002, we have updated the instructions for those using Microsoft Excel 2010, and we have included examples that are relevant to education in general or to Migrant Education Programs (MEPs) specifically.

The instructions below assume that the user is proficient in the use of databases and familiar with Microsoft Excel. *Using Excel for Analyzing Survey Questionnaires* is an excellent resource for users who need more detailed instructions on how to use Excel. The document is available on the web at <https://learningstore.uwex.edu/assets/pdfs/G3658-14.pdf>.

As we walk through some instructions for using Excel to analyze data, we will use a very simple survey as an example. There are two questions in our survey:

### Q1. Does the student have a Priority for Services (PFS)?

- “Yes” responses are coded “1,” and
- “No” responses are coded “2.”

### Q2. Did the student participate in MEP-funded instructional programs this year?

- “Yes” responses are coded “1,” and
- “No” responses are coded “2.”

Assume that surveys for 250 students were completed and returned to us. Responses were as follows:

1. Does the student have a Priority for Services?	Yes		No		Total	
		200		50		250
2. Did the student participate in MEP-funded instructional programs this year?	Yes	No	Yes	No	Yes	No
	190	10	15	35	205	45

Figure 1 depicts how this simple survey database might be set up in Excel.

**Figure 1.1 Example Survey Database in Excel**

ID#	Q1	Q2
1	1	1
2	1	1
3	1	1
4	1	1
5	2	1
6	2	1
7	2	1
8	2	1
9	1	1
10	1	1
11	1	1
12	1	1
13	1	1
14	1	1
15	1	1
16	1	1
17	1	1
18	1	1
19	2	1
20	1	1
21	1	1
22	1	1
23	1	1
24	1	1
25	1	1
26	2	2

### E.1.1 Calculating Frequencies

Frequencies are percentages used to describe the composition of a group, and there are a couple of different ways to calculate frequencies in Excel.

One method of calculating frequencies in Excel is to use the COUNTIF function to determine how many times a response occurs, and then use a formula to determine what percentage that number represents of the whole survey pool. In a blank cell, type the formula:

`=COUNTIF(range, criteria)`

Where:

- “range” includes the cells that hold your data, with the first cell and the last cell separated by a colon (:), and
- “criteria” is the specific response code for which you are calculating the frequency.

For the sample survey data above, the range for Question 1 is B5:B254. The criteria will be 1 (Yes) or 2 (No). We will create a separate COUNTIF calculation for each of the criteria.

- First, we’ll calculate how many students had a Priority for Services (PFS). The formula is:  
$$=COUNTIF(B5:B254, 1)$$

This formula tells Excel to look at the survey responses in cells B5 through B254 and count how many people responded “Yes” to Question 1.
- To calculate how many students did not have a PFS, the formula is  
$$=COUNTIF(B5:B254, 2)$$

This formula tells Excel to look at the survey responses in cells B5 through B254 and tell us how many people responded “No” to Question 1.

To convert these absolute numbers into percentages, we’ll use simple formulas to divide the number of “Yes” responses by the total number of survey responses and the number of “No” responses by the total number of survey responses.

Figure 2 displays the results of our frequency calculations in Excel.

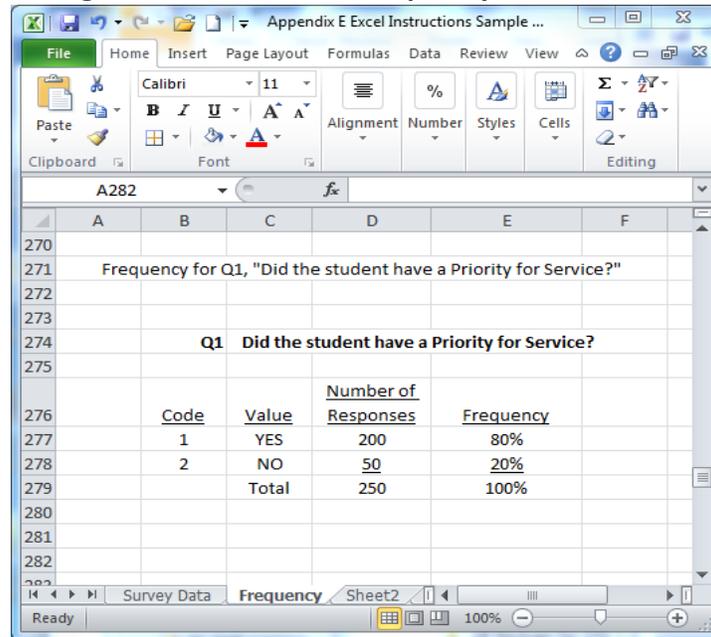
For “Yes” responses, the formula is:

$$=D277/D279$$
$$200/250 = 80\%$$

For “No” responses, the formula is:

$$=D278/D279$$
$$50/250 = 20\%$$

**Figure E.1.2 Results of Frequency Calculations**



### E.1.2 Calculating Rates or Percentages

While frequencies represent the *amount* of something, rates convey the *extent* of something.

To calculate rates in Excel, use the COUNTIF function to determine how many survey respondents indicated a certain answer for each response available on the survey, and then simply divide the number for each response by the total number of responses to convert the answers into percentages.

For our sample survey, we might want to know whether the rate of participation in Migrant Education Program (MEP)-funded instructional programs differs between students who have a PFS and those who do not.

Recalling the responses on our surveys, we know from the 250 surveys returned to us:

- 200 students have a Priority for Services (PFS)
  - 190 of these students participated in MEP-funded instructional programs;
  - 190 divided by 200 = 95%; therefore,
  - 95% of students who have a PFS participated in MEP-funded instructional programs.
- 50 students do not have a PFS
  - 15 of these students participated in MEP-funded instructional programs;
  - 15 divided by 50 = 30%; therefore,
  - 30% of non-PFS students participated in MEP-funded instructional programs.

In this very simple example, calculating the percentages manually is easy. However, for a question that has more possible combinations of responses, this could be a tedious and time-consuming task. Instead, we can use the COUNTIFS function in Excel to do these calculations for us. [Notice the “S” in the formula when we are calculating a frequency based on more than one criterion. This is not a typo!]

In essence, the COUNTIFS function calculates frequencies for each possible combination of responses. We can then convert these absolute numbers into percentages in order to compare results across groups.

Using Excel for our calculations, the formula would be:

```
=COUNTIFS (criteria_range1, criteria1, criteria_range2, criteria2, ...)
```

Where:

- “Criteria\_range1” includes the cells that hold data for the first criterion you wish to count;
- “Criteria1” is the first response code for which you are calculating the frequency;
- “Criteria\_range2” includes the cells that hold data for the second criterion;
- “Criteria2” is the second response code to count, etc.

Let us return to our simple survey example to illustrate the use of the COUNTIFS function. Our example involves two questions, each with its own corresponding range of data and two criteria (or response codes) for each question. However, you can use the COUNTIFS function to calculate rates for more than two questions and more than two criteria per question (e.g., Yes, Maybe, No).

As a reminder, we want to know if the rate of participation in MEP-funded instructional programs differs between students who have a PFS and those who do not. First let us look at students who do have a PFS and who did participate in MEP-funded instructional programs:

- The range for Question 1 responses is B5:B254, and the code for “Yes” is 1 (criteria1 = 1);
- The range for Question 2 responses is C5:C254, and the code for “Yes” is again 1 (criteria2 = 1).

The formula, therefore, will be:

```
=COUNTIFS(B5:B254, 1, C5:C254, 1)
```

This formula tells Excel to give us the number of students who answered “Yes” to Question 1 AND “Yes” to Question 2. There are 190 students in our survey who answered “Yes” to both Question 1 and Question 2.

We can do the same to students who do not have a PFS but who did participate in MEP-funded instructional programs:

- The range for Question 1 responses is B5:B254, and the code for “No” is 2 (criteria1 = 2);
- The range for Question 2 responses is C5:C254, and the code for “Yes” is 1 (criteria2 = 1).

The formula, therefore, will be:

=COUNTIFS(B5:B254, 2, C5:C254, 1)

Figure 3 shows what this exercise might look like on our Excel spreadsheet.

To convert the absolute numbers into percentages, we’ll use simple formulas to divide the number of responses that meet our criteria by the total number of students in that group:

- For students who do have a PFS and did participate in instructional programs,  
=D291/D290 = 190 divided by 200 = 95%.
- For students who do not have a PFS and did participate in instructional programs,  
=H291/H290 = 35 divided by 50 = 30%.

In answer to our question, we can say that the rate of participation in MEP-funded instructional programs in our survey is higher for students who do have a PFS (95%) than for students who do not have a PFS (30%).

**Figure E.1.3 Example Exercise in Excel**

	Code	Value	Number of Responses	Code	Value	Number of Responses
Q1. Did the student have a Priority for Service?	1	YES	200	2	NO	50
Q2. Did the student participate in MEP-funded instructional programs?	1	YES	190	1	YES	15
<b>Rate of Participation in Programs</b>			<b>95%</b>			<b>30%</b>

### E.1.3 Creating Cross-tabulations

Another approach to comparing data across groups involves cross-tabulations, or crosstabs. Excel uses the Pivot Table Wizard to create crosstabs.

We will work with the same question we used earlier: Does the rate of participation in Migrant Education Program (MEP)-funded instructional programs differ between students who have a Priority for Services (PFS) and students who do not? We will also use the database we have already set up (see Figure E.1.1).

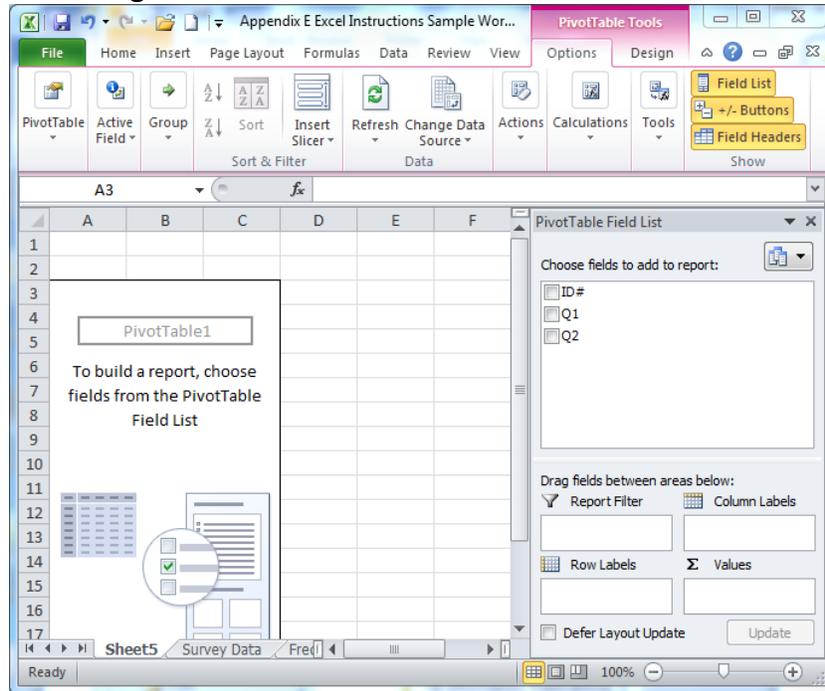
To start creating a crosstab, go to the Insert tab in Microsoft Excel 2010, click on PivotTable, and choose PivotTable.

The first dialogue box will ask you to select a data range. Our data range is \$A\$4:\$C\$254. You can type that in or simply highlight the range on your screen. Notice:

- We use “\$” to anchor the columns (as in, \$A and \$C) and rows (as in, \$4 and \$254), and
- In the range, we also now include the cells with the column headers, which is necessary in order to identify the variables we are using.

Excel also wants to know where to put the pivot table, so choose New Worksheet, and click OK. A PivotTable Layout window will appear in the new worksheet (see Figure 4).

**Figure E.1.4 New Worksheet for the Pivot Table**

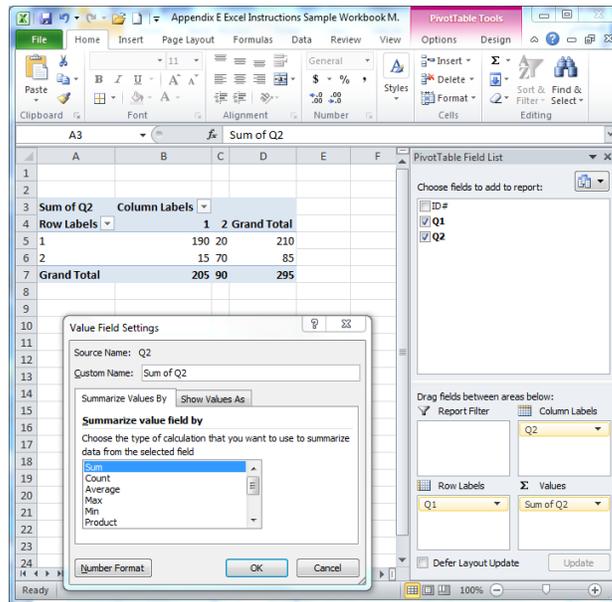


For the Pivot Table, we need to designate an independent variable and a dependent variable. For this example, the independent variable will be Question 1 (*Does the student have a Priority for Services?*), and the dependent variable will be Question 2 (*Did the student participate in instructional programs this year?*). Since independent variables are used to explain something about dependent variables, we are in effect asking whether participation rates differ by PFS status. Therefore, we are using PFS status (the independent variable) to explain participation rates (the dependent variable).

On the PivotTable worksheet, click on Q1 (PFS status, the independent variable) in the Pivot Table Field List, and drag it down to the Row Labels section. Click on Q2 (participation, the dependent variable) and drag it down to the Column Labels section. Click on Q2 again (in the Pivot Table Field List), and drag it down to the Values section.

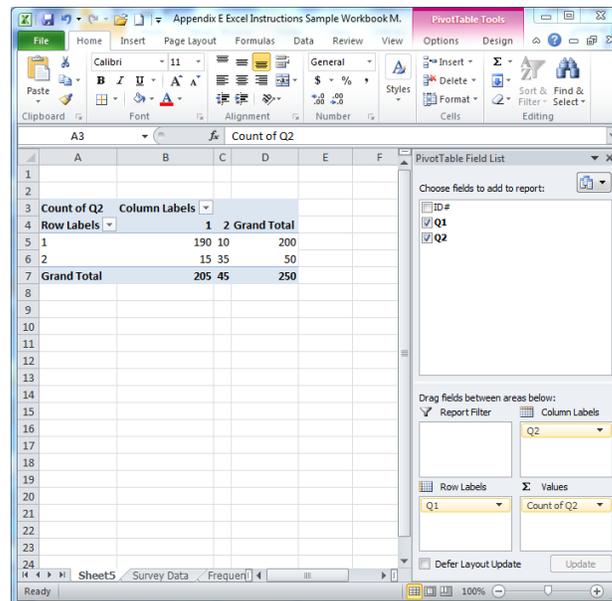
In the Values section, you can choose what you want Excel to do with the data. In our example, we want Excel to count the number of responses for each value. However, the default in Values is to calculate the sum of the responses (see Figure 5.)

Figure E.1.5 Sum of Values (Default)



To change the default operation in Values, click on the drop down menu in Values, choose Value Field Settings, and choose Count. You will notice that the label in the Values box changes from “Sum of Q2” to “Count of Q2” (see Figure 6).

Figure E.1.6 Count of Values



While some of the numbers in the PivotTable might look familiar (190, 200, 250), we do not know what we are looking at without our variable (or field) names. We will have to type them in ourselves:

- In the PivotTable itself, click on “Row Labels” (which we know represents Q1) and type “Does student have PFS?” Next, let’s define the values in the rows: Change “1” to “Yes” and change “2” to “No.”
- Now click on “Column Labels” (which we know represents Q2) and type “Participated in MEP-funded instructional programs?” Next, define the values in the columns: Change “1” to “Yes” and change “2” to “No.”

Now we can now see that the number of students who have a PFS and participate in instructional programs is 190. We can also see that the number of students who do not have a PFS but do participate in instructional programs is 15. But, again, percentages are more helpful in order to compare participation rates for the two groups.

To display percentages in the pivot table, go back to the Values Section and click on the drop down box next to “Count of Q2” (the column variable), choose Value Field Settings, and click on the tab called “Show Values As.” Then, click on the drop down menu next to “No Calculation,” and select “% of Row Total.” Now you should see the familiar result that 95% of PFS students participated in instructional programs this year compared to 30% of non-PFS students.

In this case, we select “% of Row Total” because the independent variable Q1 is the row variable, and we are interested in how Q1 explains the dependent variable Q2, participation. The results in the Pivot Table show us that participation rates are very different depending on whether the migrant student does or does not have a PFS. In this case, PFS status does indicate something important about participation!

**Figure E.1.7 Table with Percentages**

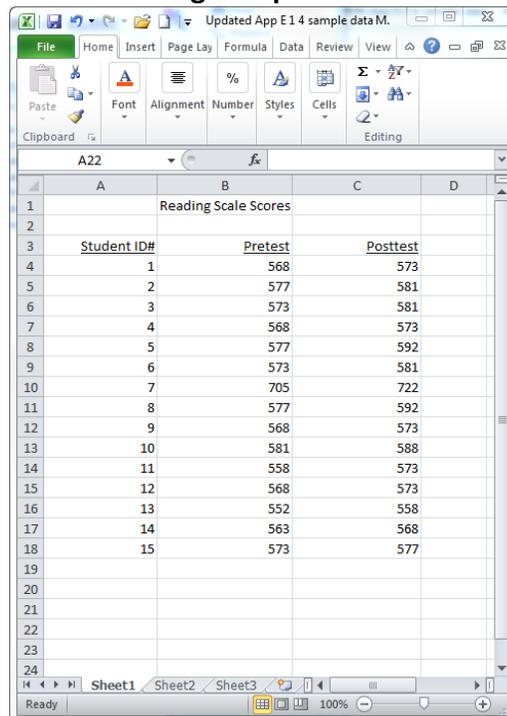
Count of Q2	Participates in MEP-funded instructional programs?		Grand Total
Does student have PFS?	No	Yes	
Yes	5.00%	95.00%	100.00%
No	70.00%	30.00%	100.00%
<b>Grand Total</b>	<b>18.00%</b>	<b>82.00%</b>	<b>100.00%</b>

### E.1.4 Calculating Means and Medians

As mentioned earlier in this section of the *Toolkit*, means and medians can be useful when individuals in a group are described by numbers on a continuous scale. Excel contains functions for calculating both means and medians.

Consider the example of a group of migrant students who are rising 9<sup>th</sup> graders. They participated in an eight-week reading and math enrichment program at the high school they will be attending in the fall. On the first day of the program, every student was tested in Reading Comprehension and Mathematics Problem Solving using the SAT-10 Advanced 2 test booklet. The teacher did not go over the correct answers with the students, but he did emphasize these skills throughout the enrichment program. On the next-to-last day of the program, students took the same two sections of the SAT-10 again. Their raw scores were converted to scale scores. Results for reading comprehension are displayed in Figure 8.

**Figure E.1.8 Reading Comprehension Scale Scores**



Student ID#	Pretest	Posttest
1	568	573
2	577	581
3	573	581
4	568	573
5	577	592
6	573	581
7	705	722
8	577	592
9	568	573
10	581	588
11	558	573
12	568	573
13	552	558
14	563	568
15	573	577

We want to know whether the average performance for the group at the end of the summer was better than their average performance at the beginning of the summer. This can be determined this by comparing the means of the two groups of scores.

To calculate the mean in Excel, use the formula  
`=AVERAGE(range)`

Where:

- “Range” indicates the first and last cells that define the range of data for which you want to calculate the mean.

Figure 9 shows the means of the pretest and posttest scores, which we calculated using the following formulas:

=AVERAGE(B4:B18) for the pretest, and

=AVERAGE(C4:C18) for the posttest

As you can see, the mean of the posttest scores is 587 compared to the pretest mean of 579, an increase of 8 points.

**Figure E.1.9 Pretest and Posttest Means**

Student ID#	Pretest	Posttest
1	568	573
2	577	581
3	573	581
4	568	573
5	577	592
6	573	581
7	705	722
8	577	592
9	568	573
10	581	588
11	558	573
12	568	573
13	552	558
14	563	568
15	573	577
<b>MEAN</b>	<b>579</b>	<b>587</b>

Reviewing the data, you notice that both the pre- and post-scale scores include an outlier (highlighted in Figure E.1.10). You can see that the outliers inflate the average pretest and posttest scores. You could just remove that student’s scores from the data and analyze it again. Instead, you decide to compare the median pretest and posttest scores in order to get a more realistic idea of the middle or “typical” value in each set of scores.

To calculate the median in Excel, use the formula

=MEDIAN(range)

Where:

- “range” indicates the first and last cells that define the range of data for which you want to calculate the median.

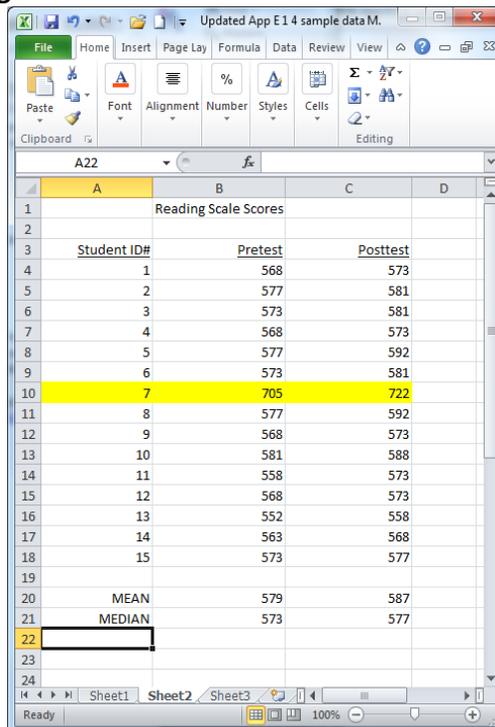
Figure 10 shows the medians of the pretest and posttest scores, which we calculated using the following formulas:

=MEDIAN (B4:B18) for the pretest, and

=MEDIAN (C4:C18) for the posttest

The median pretest score is 573 (compared to a pretest mean of 579). The median posttest score is 577 (compared to a posttest mean of 587). And the pre-post difference between the medians is a more modest 4 points (compared to 8 points for the pre-post means). In addition to illustrating how to calculate means and medians, this example also demonstrates how outliers can influence your analysis and conclusions.

**Figure E.1.10 Pretest and Posttest Medians**



For more information about using Excel to analyze quantitative data, go to: *Using Excel for Analyzing Survey Questionnaires* at <https://learningstore.uwex.edu/assets/pdfs/G3658-14.pdf>.

## Appendix E.2 Using Access to Analyze Qualitative Data

In *Section E.2 Analyzing Qualitative Data* of the *Program Evaluation Toolkit*, we explained how to analyze qualitative data by organizing focus group responses to open-ended survey questions into categories and assigning codes to each category. Level 1 codes are given to responses that fall into the broadest categories you see in your data. Level 2 codes are used to organize data in Level 1 categories into narrower and more descriptive categories.

In the past, we might have taken one of the following approaches to analyzing qualitative data:

- Cutting responses out of the paper on which they were printed and physically re-organizing them into groups that fit together in some relevant or interesting way, then counting the number of scraps of paper we included in each category grouping; or
- Reading through the responses, generating a list of the categories we found, creating a code for each category, then penciling codes in next to each response, and counting the number of times we assigned each code.

Between the paper cuts and eraser shavings, these approaches could be very messy indeed!

### E.2.1 Microsoft Access as Qualitative Analysis Software

Today you can purchase software (or download freeware) to code, sort, and summarize qualitative data electronically (e.g., NVivo, Atlas.ti, QDA Miner, etc.). These packages have a lot to offer the serious qualitative researcher, but this may be more firepower than you actually need. Before you purchase or download something new, consider a tool that you may already have in your Microsoft Office Professional arsenal, the database software, Access.

NOTE: For the purpose of this discussion, we assume that you have some familiarity with Access. If you are not familiar with it, and you have it on your computer, we encourage you to learn more about this flexible data management tool. One way to do this is by taking an online short course through a university computing center. For example, the Computer Training Unit in the McKimmon Center at North Carolina State University offers self-paced online courses in Access for users at three levels of experience. Learn more by visiting:

<https://onece.ncsu.edu/mckimmon/divisionUnits/ctu/index.jsp>

We routinely create databases in Access to store both quantitative and qualitative data. We export the quantitative data into Excel or a statistical software package (e.g., SPSS, SAS, R) for further data analysis. However, we keep our qualitative data in Access and create forms, queries, and reports to help us code and summarize the data.

All data are entered into tables in Access. It can either be entered directly into the table in DataSheet View, or in a data entry “form.” Forms are often more visually appealing than

datasheets. In Form View, you can create data entry “windows” that are large enough for you to see your entire text block (e.g., an excerpt from a focus group interview or a single open-ended survey response).

### E.2.2 Qualitative Data Example

Before we show you how we use Access forms, queries, and reports to analyze qualitative data, consider the following example:

- Your Migrant Education Program (MEP) office serves as a “broker” for all non-instructional, support services to migrant students, families, and out-of-school youth (OSY) in your area. The recruiters in your office do their best to make sure the people they identify are given information about all available services.
- Once a year, you go out into the migrant community to find out what your target population thinks about the services you offer and what services they still need. To do this, you set up focus groups at several local churches that have sizable numbers of migrant parishioners. You will use the information gathered in the focus groups to improve communication about the services you offer, improve the services themselves, and find out what other services are needed in the community.
- With the help of the pastor or priest and church staff, you schedule two types of focus groups: one for parents of students who are enrolled in school, and another for older teens and young adults who are out of school and living on their own while they work in your community (e.g., OSY, ages 17–21). Each focus group has 8–12 participants.
- At the beginning of each focus group, you give participants a one-page survey that lists all of the services provided through your office. They are asked to check all of the services they have used during the past year. In addition to asking for some basic demographic information (e.g., age, gender, OSY or parent; if parent, number and ages of children; etc.), the survey also includes two open-ended questions:
  - Of all the services we offer, which ones have been the most helpful to you? How have they been helpful to you?
  - Which services have been the least helpful? Why do you think they were not helpful?
- After everyone has completed the survey, you collect them (to analyze later) and begin the focus group. Over the course of 45 minutes to an hour, you ask the following questions:
  - After seeing our list of services, were there any that you did not know about? If so, what services are you just hearing about? Are these services you think you can use?
  - What other services do you think are needed in this community? Why are they needed?
- Each focus group has a facilitator who asks the questions, makes sure everyone gets a chance to contribute to the conversation, and generally keeps the discussion on track to end on time. Each group also has an observer who takes detailed notes and records the

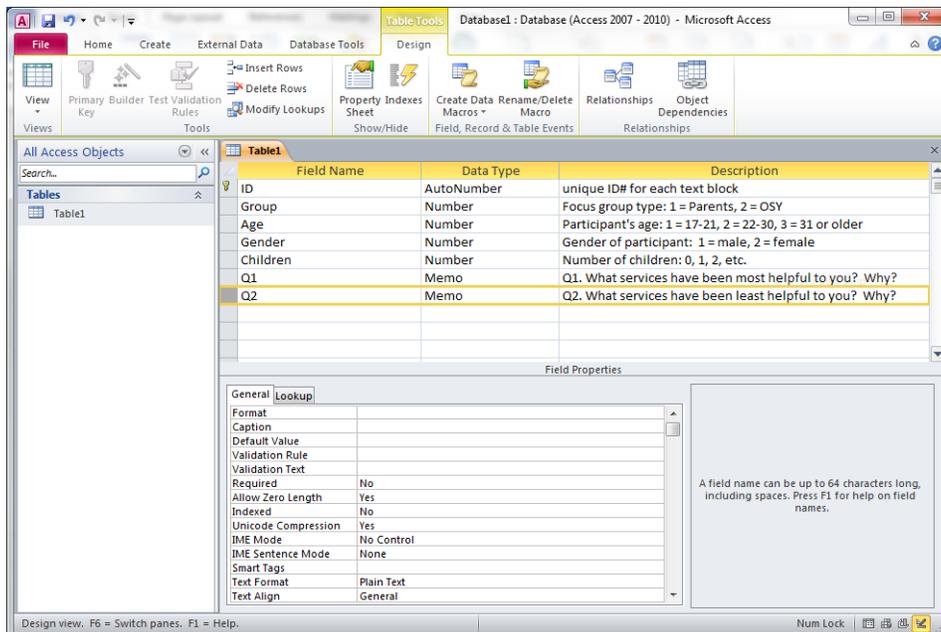
discussion (using a digital or tape recorder) in case parts of it need to be reviewed later for more details or longer quotes.

- When all of the focus groups are done, you will have a focus group dataset that includes responses from all or most participants to your two discussion questions. For coding purposes, you will need to separate each discussion into distinct chunks or text blocks. If we have a transcript of the entire focus group, we try to excerpt an individual’s entire answer to a question to create one text block. If we are working from notes, we summarize a single individual’s response from our notes to create a text block.
- In addition to the focus group dataset, you will also have a survey dataset with information about all of the participants, including open-ended responses with their thoughts about the most and least helpful services they have received.

Creating and coding text blocks from focus group data is complicated by the fact that participants may repeat themselves, or interrupt each other, or reconsider their answers. This does not make for nice, neat text blocks! We encourage you to try it and see what we mean. But for now, we will continue with the simpler example of creating text blocks from written responses to open-ended survey questions.

### E.2.3 Creating Qualitative Data Analysis Forms in Access

When you are ready to analyze the surveys you collected before the focus groups, you can start by entering all of the data into an Access table. First, create a table in Design View.



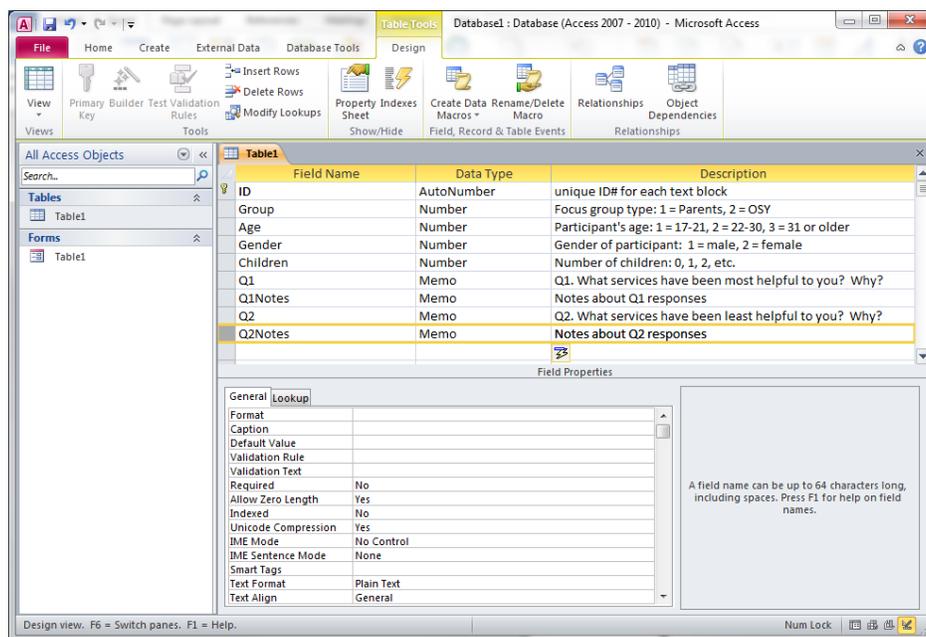
Next, enter your survey data. You can enter data directly into the table in DataSheet View, or you can create a form for your table (e.g., using the Form Wizard) and enter your data into the form.

No matter how we enter the survey data in the first place, we like to create separate forms just for analyzing answers to open-ended questions. We start with a very simple form and add to it as we identify different themes or categories in the survey responses.

Our qualitative data analysis forms include the following features:

- A unique survey number, so we can refer back to the original written response as needed
- A data entry “window” for the memo field that contains each text block in our dataset
- Another data entry window for the memo field where we will write notes about possible categories for each answer, the main idea, and quotable bits (e.g., specialized or colorful vocabulary, descriptive phrases, and sentences or passages that are typical for a particular category)
- Multiple “Yes/No” check boxes, one for each code (Level 1, Level 2, etc.) that we assign to the open-ended responses

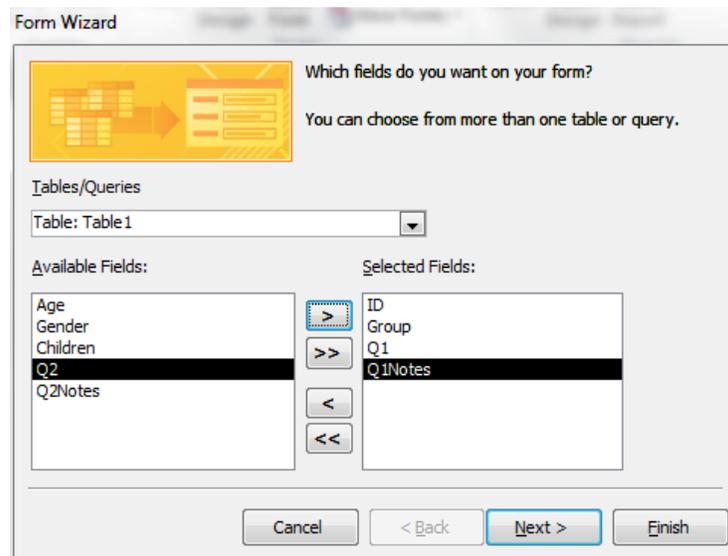
We always begin our analysis by reading through survey answers and taking notes as we go. So, before we start creating a form for our first question, add fields in the table for notes about the answers to each question. Now, save and close the table (Table 1).



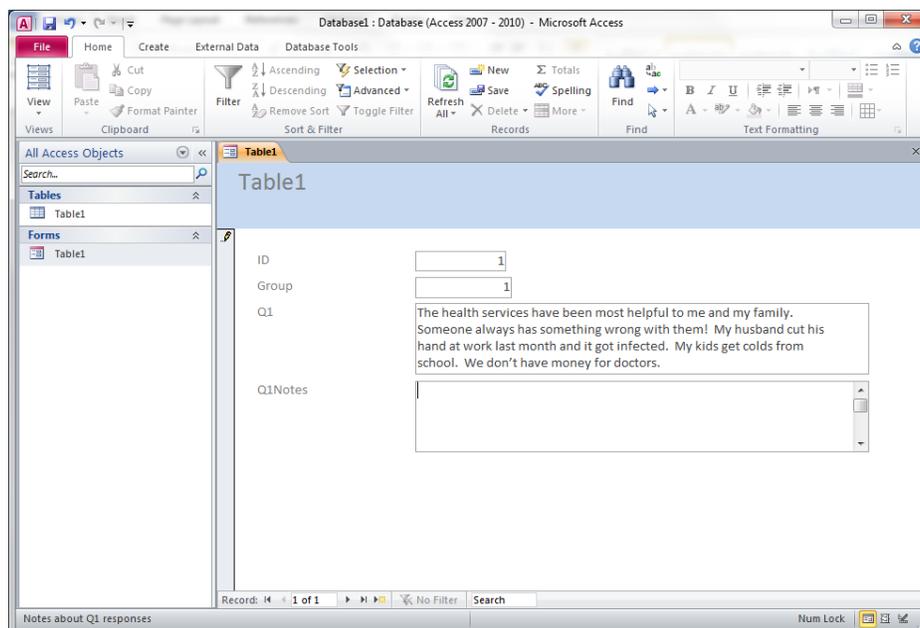
To begin creating a form for Question 1:

- Select Create > Form Wizard.
- In the Tables/Queries drop-down menu, select the table with your survey data (Table 1).
- Next, click on each variable you want to include in your analysis, including Q1.

- Click on the Field Name in the Available Fields window to select, then click the single “>” to move the variables you want into the Selected Fields window.



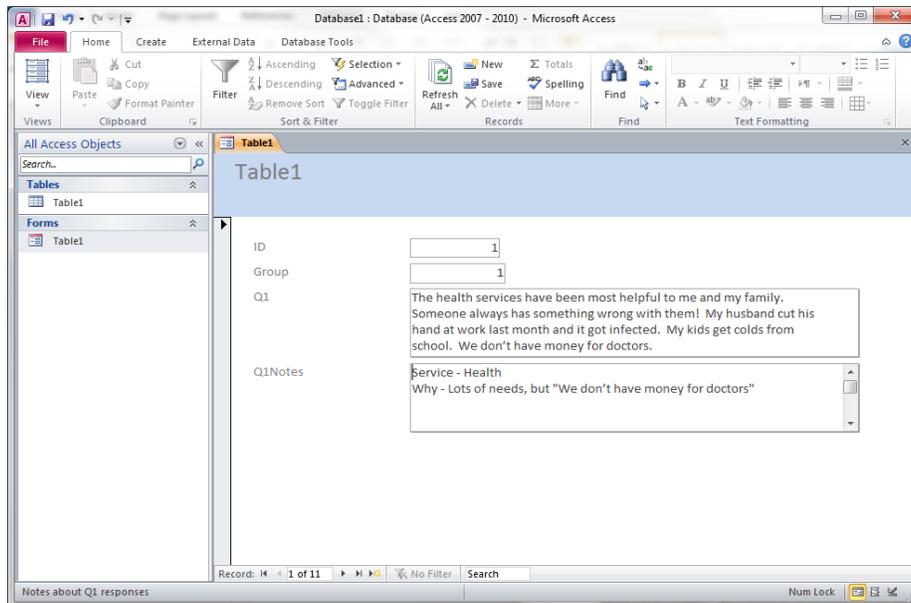
- Click Next > to select a pre-designed layout for your form. In this example, we will use the Columnar layout.
- Click Next > and rename the form, if you wish.
- Then click Finish to see what your form looks like so far.



## E.2.4 Selecting Level 1 Coding Categories

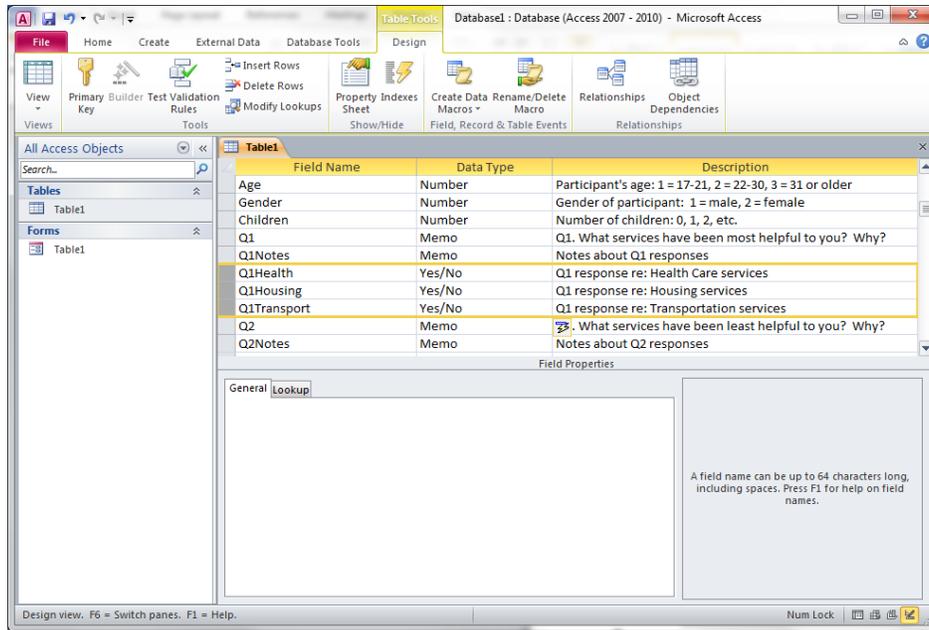
Notice that the answers to Question 1, which were already entered in Table 1, are now visible in the form for Table 1. Now you can read each answer and add some notes, while you get familiar with what participants wrote and begin to identify possible coding categories.

For example:



Eventually, you will read through and write notes for all of the answers to Question 1, or you may read many of the responses and find that you are not seeing any new categories. Now you can decide which broad categories you will use to organize the answers to Question 1.

Categories of services are the obvious choice for this question. All the answers you have read have had to do with health, housing, and transportation services. So, save and close the form, then go back to Table 1 to add fields for these categories.

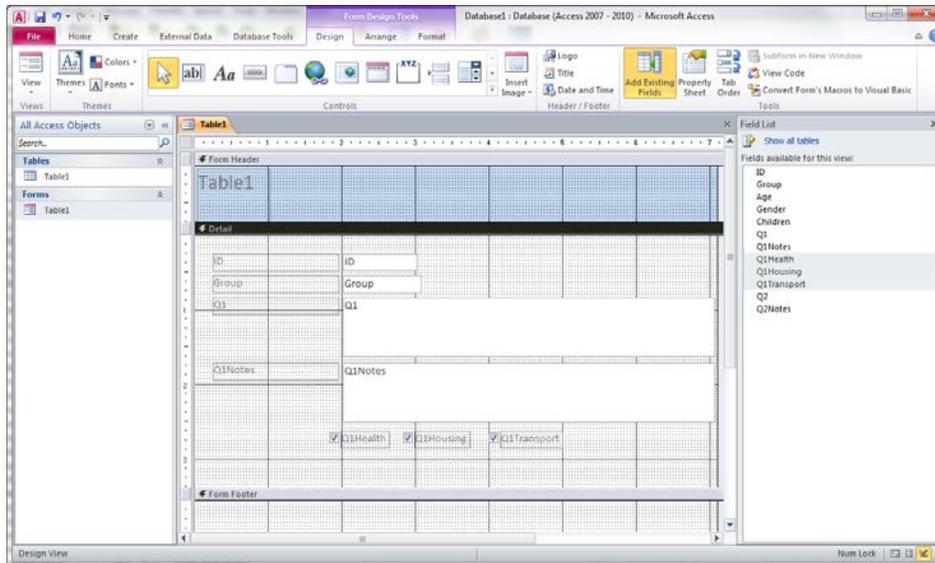


We just created a set of Level 1 codes! We can add more Level 1 codes for categories of services later if we run across some that we missed in our first pass through the data. All we have to do is save and close the data entry form, then open Table 1 in Design View to insert rows and add fields.

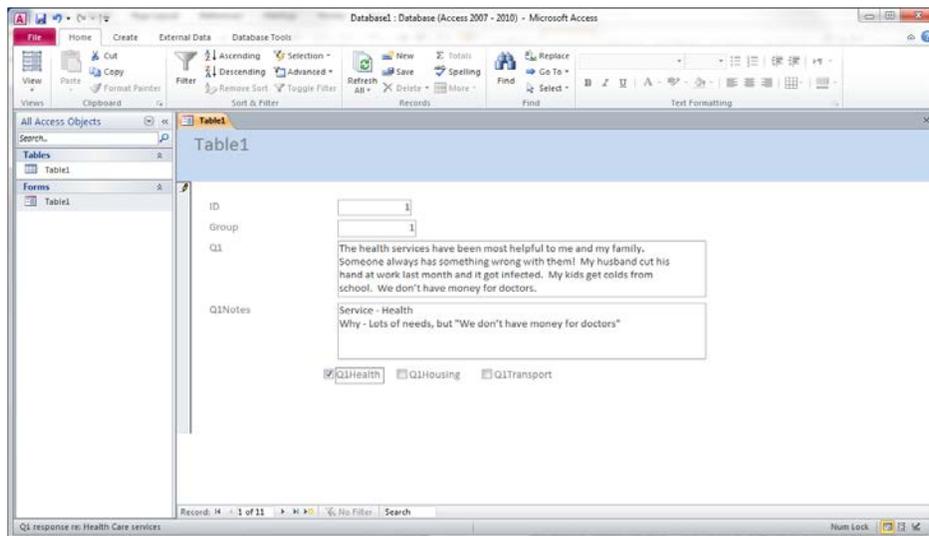
We use the “Yes/No” data type for all of our coding fields. Basically, they are simple check boxes. They are quick and easy to use as we read through all responses and select the category that is the best fit.

To add our Level 1 coding check boxes to the form, save and close Table 1. Then:

- Open the form for Table 1 and switch to Design View.
- Click on “Add Existing Fields” in the Toolbar at the top of the screen.
- Click each coding variable and drag it down to the form, which now appears in grid format.
- Place each coding variable check box wherever you like, as long as it does not overlap with another field on the form. Try different arrangements to find the layout that works best for you.



Save the new layout of your form and switch back to Form View. Once again, you will see the responses that have been entered into Table 1 as well as the notes you have already entered. Now you can just click on the box next to the appropriate coding category, and voilà – you have just coded an entry!

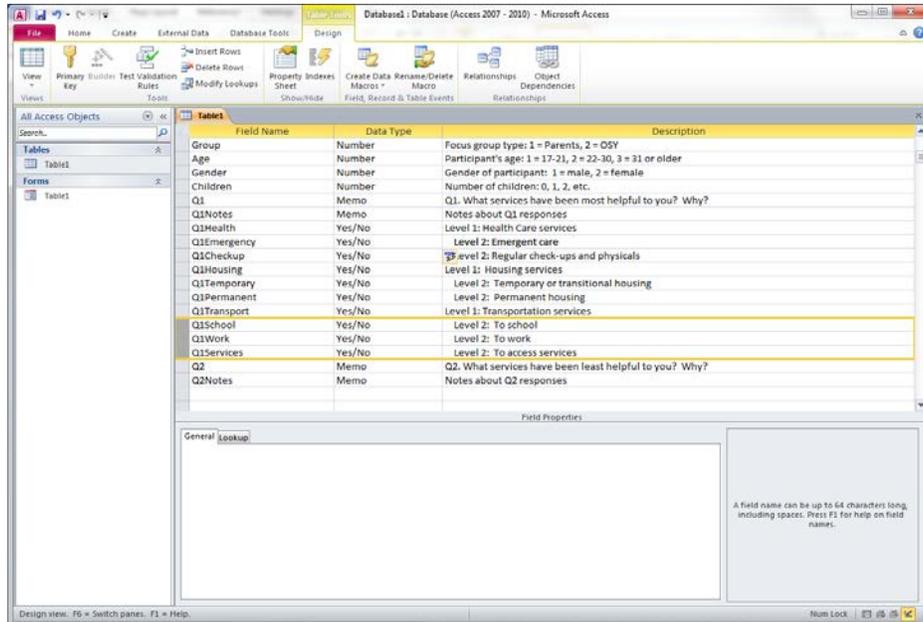


### E.2.5 Selecting Level 2 Coding Categories

At some point, you will begin to see more narrow categories within your broad Level 1 categories. As you decide what these next categories are, you can add them to your table and data analysis form the same way you added the Level 1 coding categories and check boxes:

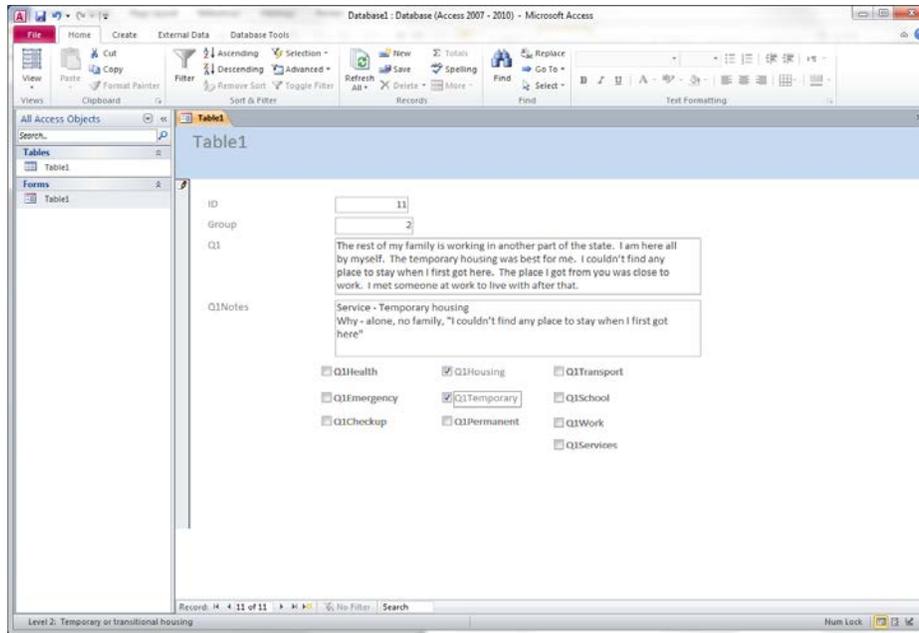
- Save and close the form.
- Open Table 1 in Design View.

- Insert rows, as needed.
- Add Field Names, select “Yes/No” data type, and write a brief description for each new coding category.
- Revise Level 1 descriptions as needed.



Then:

- Open the form for Table 1 and switch to Design View.
- Click on “Add Existing Fields” in the Toolbar at the top of the screen.
- Click each new Level 2 coding variable and drag it down to the form.
- Arrange Level 2 coding boxes in close proximity to their corresponding Level 1 coding boxes. Re-arrange Level 1 check boxes as needed.



## E.2.6 Organizing Data in Categories

By using Access to create and apply codes to your data, you will have the flexibility to look at your data in a variety of ways. Once you have coded all of your data, you can sort it into coding categories for further analysis. You can do this by:

- Setting up queries for each coding category (Create > Query Wizard > Simply Query Wizard > OK, etc.); and
- Using queries to create reports with all of the responses that were coded into that category (Create > Report Wizard, etc.)

You may see interesting patterns or differences in why or how different services in the same category were especially helpful (or not helpful) to participants. From there you could decide to create Level 3 coding categories in order to include these patterns in your analysis.

You can also export the data in your check boxes to Excel and then run frequencies on these fields to create summary tables. We encourage experienced Access users to give this coding method a try, and we encourage new Access users to learn more!

## Appendix E.3 Using Inferential Statistics

In *Section E.1 Analyzing Quantitative Data* of the *Program Evaluation Toolkit*, we explained how to use descriptive statistics to summarize information about migrant students, their participation in programs, their attitudes, their achievements, etc. Descriptive statistics such as frequencies, rates, means, and medians are relatively easy to calculate and display. Using descriptive statistics to compare two or more groups may reveal important differences or instructive patterns.

We often make programmatic decisions based on available descriptive information. However, the results obtained from descriptive statistics are at best suggestive and cannot be used to draw valid conclusions about programs. In order to make more definitive and convincing statements about program impacts, we analyze data using inferential statistics that test whether meaningful or significant differences have been achieved and whether these findings can be generalized beyond the specific group of people we studied.

NOTE: The information provided in this Appendix is no substitute for the expertise of a qualified statistician who understands the complexities of inferential statistics. Nevertheless, we have included information about this topic for MEP directors and staff who are curious about how to go beyond merely describing the patterns in their data to testing for more definitive conclusions about these patterns.

The intention of this Appendix is to explain some basic concepts and provide examples for using specific inferential statistics to evaluate Migrant Education Program services and activities. Because we cannot adequately cover the complexity of this topic here, we have included links to several online resources (searchable textbooks, tutorials, and applets) at the end of this Appendix in case you want to learn more or refresh your memory.

### E.3.1 Basic Concepts in Inferential Statistics

NOTE: This section of the Appendix includes a number of technical terms. If you are not interested in the underlying principles of inferential statistics, you can skip this section. However, if you would like to learn more, we encourage you to use the terms that appear in italics to search the references provided at the end of this Appendix.

Inferential statistics are used to draw conclusions about a *population* of people based on information from a *sample* of that population.

For the purposes of statistical analysis, a population is the entire group of people about whom we are interested. In order to draw the strongest possible conclusions about the population of interest, we would collect and analyze relevant or interesting data from every person in that

population (e.g., data about race/ethnicity, gender, migrant status, education level, attitudes, behaviors, etc.)

When it is not possible to collect data from every subject in the population of interest, we collect data from a subset or sample of the population instead. For example, we might select a sample of migrant students from the high schools in our school district and use the information we collect from them to make general statements about the population of all migrant high school students in the district.

The appropriate use of inferential statistics involves collecting a *representative sample* of adequate size to detect outcomes if, in fact, they have been achieved. Traditional inferential statistics are based on the measurement of numerical characteristics of a sample, which in turn are assumed to be the same as the characteristics of the population from which the sample was drawn.

- Examples of numerical characteristics include measures of *central tendency* (mean, median, mode) and *spread* or *dispersion* of the data (variance, standard deviation).
- These numerical characteristics are also known as *parameters* and are calculated for continuous variables. *Parametric statistics* is the more traditional and purist branch of inferential statistics.
- *Non-parametric statistics*, which are not based on the numerical characteristics of a sample, can be used to analyze categorical and continuous variables. Instead of using measures of central tendency and dispersion, non-parametric statistics are based on the rank order of data values.

In order for parametric statistics to produce valid and unbiased estimates of population parameters, certain conditions must be met. A representative sample, specifically a *random sample* of the population, is a fundamental condition for correctly using parametric statistics.

- If the sample does not represent the population, then the statistics you calculate for the sample will not represent the true parameters of the population. They will only represent the sample.
- A *simple random sample* is a single sample drawn at random from a population (e.g., all migrant students in your district or in your state). Every subject in the population has an equal chance of being selected. This type of sample is representative of the population as a whole.
- A *stratified random sample* is comprised of simple random samples drawn from specific subgroups (or strata) in the population (e.g., Limited English Proficient (LEP) migrant students and non-LEP migrant students). This type of sample is representative of specific subgroups in the population.

The size and “shape” of a sample must also be considered when deciding whether you can correctly use parametric statistics.

- Random selection alone does not guarantee that a sample will be representative of the population. The size of the sample is also important. A sample of 10 books in a library of 3,000 is very unlikely to represent the depth and breadth of fiction vs. non-fiction or the range of genres and topics included in the entire collection. The larger the sample, the more representative it becomes.
- We use *power analysis* to determine how large a sample we need in order to be very confident (e.g., 95-99% confident) that any estimates we calculate from the sample will have small *margins of error* (e.g. 1-5%).
- Another condition for using parametric statistics is that the population from which you draw your sample must have a *normal distribution*. In a normal distribution, most of the observations in the sample are clustered fairly close to the mean or average value in the sample. The number of increasingly larger and smaller values tapers away on either side of the mean, which results in a bell-shaped curve when you graph the distribution.
- Fortunately, according to the *Central Limit Theorem*, a sample of at least 30 observations approximates a normal distribution enough to allow the use of parametric statistics. Nevertheless, we still strive for larger samples in order to increase confidence that the margins of error will be small.

### E.3.2 Using Inferential Statistics in Program Evaluation

In the context of program evaluation, there are several common challenges to using parametric inferential statistics:

- It is rarely possible to randomly select people to participate in programs. There are ethical issues involved in deliberately withholding a potential benefit from people whose lives may improve with it. There are also practical barriers to identifying and recruiting every person who could benefit just so you can draw a random sample.
- Suppose you are able to randomly select participants. The simplest form of experimental design involves collecting baseline data, exposing participants to the program, and then collecting follow-up data to see whether participants achieved the desired outcome. Unfortunately, between the baseline and follow-up data collections (or the pre-test and the post-test), some participants will inevitably drop out of the program. So the random sample you had at the time of the pre-test is no longer random at the time of the post-test.
- Because of budget constraints and/or eligibility requirements for targeted interventions, activities and services may be available or appropriate for only a small group of participants. Small samples sizes (e.g., fewer than 30 subjects) are not uncommon in program evaluation.

The bottom line is that program evaluation data rarely meet the conditions needed to correctly use and draw conclusions from parametric statistics. Fortunately, these limitations are not fatal to the cause of program evaluation.

- The most common question asked in program evaluation is, “Did the program have the desired effect on the participants?” Therefore, the population of interest is all program participants. In a small program, you can collect data from every participant. If you need to collect a sample of participants in a large program, all participants (e.g., all members of the population) are either known to you or at least identifiable and therefore may be easier to sample at random.
- When you lose participants to attrition from the program or have a low response rate from your data collection activities, look for ways to compare the posttest or low-response sample to what you know about the pretest sample or program participants as a group. The point is to determine how similar the former is to the latter. If the groups continue to be similar on characteristics that matter in your program, then you have some basis for arguing that the posttest or low response sample is still representative of the original sample.
- Non-parametric statistics do not require the same conditions as parametric statistics in order to draw valid conclusions. They can be used to analyze small samples of categorical and continuous variables with non-normal or skewed distributions.

The following table provides some examples of evaluation questions and the appropriate tests for analyzing related data.

**Table E.3.1 Examples of Evaluation Questions and Related Tests**

Type of Analysis	Sample Evaluation Question	Parametric Statistic	Non-parametric Statistic
1. Compare means between two distinct (i.e., independent) groups.	Is the mean reading scale score for LEP migrant students in a language enrichment class significantly different from mean reading scale scores for LEP migrant students who did not take part in the class?	Two-sample t-test	Wilcoxon rank-sum test
2. Compare two measurements of a continuous variable taken from the same individual.	For a sample of migrant students in a summer enrichment program, was there a statistically significant difference between mean math pretest and posttest scores?	Paired t-test	Wilcoxon signed-ranks test
3. Compare means between three or more independent groups.	Comparing LEP migrant students in three supplementary language instruction groups (pull-out group instruction, one-on-one tutoring, summer language immersion program), are there statistically significant differences in their mean pretest English Language Proficiency scores? Are there significant differences in their mean	Analysis of variance (ANOVA)	Kruskal-Wallis test

Type of Analysis	Sample Evaluation Question	Parametric Statistic	Non-parametric Statistic
	posttest ELP scores?		
4. Estimate the degree of association between two continuous variables.	Is there a statistically significant association between the amount of time spent in supplementary language instruction and scores on an English Language Proficiency test?	Pearson coefficient of correlation	Spearman's rank correlation
5. Estimate the degree of association between two categorical variables.	Is there a statistically significant association between participation in specific non-instructional support programs (e.g., transportation, nutrition, medical care, counseling) and proficiency levels (I, II, III, IV; or I-II and III-IV) on the state reading and mathematics assessments?	[Not appropriate for categorical data]	Chi-squared test

NOTE: Table adapted from [Hoskin, T. \*Parametric and Nonparametric: Demystifying the Terms\*](#)

### E.3.4 Online Resources about Inferential Statistics

In this Appendix, we tried to provide enough basic information about inferential statistics to convey the underlying importance of representativeness, sample size, and a normal distribution. Another important topic that we have not discussed is *hypothesis testing*. For more information about all of these topics, we encourage you to check out the following resources.

Online Statistics Education: A Multimedia Course of Study [<http://onlinestatbook.com/>]

HyperStat Online Statistics Textbook [<http://davidmlane.com/hyperstat/>]

StatSoft Electronic Statistics Textbook [<http://www.statsoft.com/textbook/>]

Sample size calculator [<http://www.raosoft.com/samplesize.html>]

Random number generator [<http://stattrek.com/statistics/random-number-generator.aspx>]

# Section F: Communicating Evaluation Findings

---

Conducting an evaluation is necessary but not sufficient for accountability. Being accountable also requires that the findings of the evaluation are communicated to program stakeholders. To be effective, communication of evaluation findings should follow the same basic principles for all good communication:

- Keep the needs and interests of your audience in mind.
- Be clear about your purpose for communicating with a given audience and state that purpose clearly.
- Avoid including irrelevant information that will distract your audience from getting the information they want and the information you want them to have.
- Provide adequate contextual information so audience members can interpret evaluation findings for themselves.
- Organize information so it is easy to find.

## F.1 Writing Evaluation Reports

State Migrant Education Programs (MEPs) are required by statute to develop a written evaluation of the MEP and utilize the evaluation to guide program improvement. The following suggestions will help you develop a written evaluation report that will be useful for those involved with reviewing and improving the state and local MEPs.

- The *Guidance* suggests that states produce a written program evaluation report once every two to three years to document the evaluation of program implementation and results.
- Audiences for the state evaluation report include state and local MEP staff. All of these audiences have an interest in knowing what was accomplished, how it was accomplished, and how to accomplish more in the future.
- Parents and other advocacy groups may be most interested in how well migrant students are doing in school and what strategies work best to help them be successful in school. You might consider preparing targeted data briefs or “report cards” for communicating this kind of information with these audiences.
- The *Guidance* states that written reports should include the purpose of the evaluation, what data were collected and how they were collected, the findings of the implementation evaluation, results for Priority for Services (PFS) and other migrant students, and the implications for making decisions about the program.

- It is recommended that MEP evaluation reports also describe the identified needs of migrant students in your state, the general strategies recommended to address those needs, and the specific services that were provided to implement those strategies. This contextual information will help the reader understand and evaluate the significance of your Measurable Program Outcomes (MPOs).

## F.2 Organizing the Evaluation Report

While there is no fixed format for state evaluation reports, it is helpful to begin with the purpose of the report and some contextual information. This still leaves you some flexibility to organize the results of the evaluation in a logical sequence that works for you.

Rather than rely on the outline for a traditional research report (e.g., program description, evaluation methodology, research results, conclusions), professional evaluators are beginning to experiment with more reader-friendly ways of organizing their reports, and so should you. Put yourself in the place of the reader and devise an organizational structure that would make it easy to find the information that is of greatest interest to you.

For example, you might try the following outline:

1. **Purpose of the evaluation report** – In the case of your state evaluation of the MEP, your purpose is to document the evaluation of program results and implementation in order to improve services for migrant children. For other audiences, the purpose of your evaluation communication might be to recognize the accomplishments of students and staff, invite feedback from interested stakeholders including local MEP staff, or educate funders and migrant student advocates about the challenges of serving migrant populations.
2. **MEP target population** – Details about the migrant population you serve provide useful context for understanding the challenges faced by migrant students in your state. This section would be a good place to include your state’s criteria for identifying Priority for Services students who are failing or at risk of failing because of interruptions in their schooling. You might also present your state’s Migrant Student Profile in a table and provide some narrative explanation of the unique characteristics of the population that influence how and when they are served (e.g., the relative proportions of migrant children in different grade levels including out-of-school youth, or whether migrant children predominantly reside in your state during the school year or the summer).
3. **Evaluation results** – Evaluation results are best understood in the context of the needs of the target population, the specific services that were delivered to address those needs, and the outcomes you expected to achieve. For this reason, you should experiment with ways to present contextual information in conjunction with evaluation questions, data collection methods, and findings. For example, you might decide to organize your results section by Areas of Need and create a table like the one in Figure F.1 to provide context. These tables could be used to introduce the evaluation questions

related to MPOs and service delivery in each Area of Need. Each table might then be followed by the related evaluation questions, a description of how data were collected, and the findings.

**Table F.1 Example Table of Organizing Results Around Needs**

<p><b>Need:</b> In order to complete school, migrant out-of-school youth (OSYs) require access to schooling during non-traditional hours and in settings that are easily accessible from the places where they live and work.</p>	<p><b>Strategies:</b> Collaborate with employers to offer classes at or near places where large numbers of migrant OSYs work. Schedule classes during the late afternoon, early evening, or on weekends, depending on typical work schedules.</p>
<p><b>MPO1.</b> In work places where classes are offered, at least 60% of identified migrant OSYs will be enrolled in these classes.</p> <p><b>MPO2.</b> 90% of migrant OSYs enrolled in these classes will successfully complete their classes with a passing grade.</p>	<p><b>Service Delivery Plans:</b></p> <ol style="list-style-type: none"> <li>1. Identify employers</li> <li>2. Recruit eligible OSY employees to attend classes</li> <li>3. Negotiate space for classes</li> <li>4. Negotiate instructional time that accommodates work schedules</li> <li>5. Recruit teachers to work non-traditional hours</li> <li>6. Staff an academic help desk after class hours to help OSY students with homework as needed.</li> </ol>

**Evaluation Questions (Results):**

1. At all participating work places combined, what percentage of eligible migrant OSYs are enrolled in a class?
2. Of the OSYs who completed a class, what percentage successfully completed it with a passing grade?

**Evaluation Questions (Implementation):**

3. How many work sites were recruited? How many eligible migrant OSYs are employed at these sites? When are classes offered at these (or nearby) sites?
4. What percentage of eligible OSYs at each work site are enrolled in a class?
5. If classes are under-enrolled, what accounts for low enrollment (e.g., lack of interest, inconvenient class schedules, wrong classes offered)?
6. Were an adequate number of qualified teachers recruited to teach the classes?
7. Are enrolled OSYs using the academic help desk? If so, what kinds of help are they requesting? Is help desk staff qualified to provide the requested help?

4. **Implications for improving services for migrant students** – For the sake of consistency and reader-friendliness, try to organize this section the same way you organized your results. In the case of the example above, results were organized by areas of need; for the sake of consistency, you would also discuss implications as they relate to improving services to meet specific needs. If the results had been organized by MPOs, you might then discuss the implications of improving services to achieve better results with respect to outcomes. In either case, how you approach implications for improving services will depend on how successful you were at achieving your desired outcomes.
- If you were successful, consider what aspects of the services you delivered were most critical to your success. What are the implications for providing these services to more migrant children in the future?
  - If you did not achieve the expected outcomes, consider whether the implementation (or under-implementation) of specific services impacted these results. What are the implications for being able to fully implement these services in the future?
  - Are there other ways to implement these services in order to achieve success?

### F.3 Displaying Data

Again, the basic principles for good communication apply when you are compiling data in tables and other visual display formats (e.g., bar charts, pie charts, line graphs):

- Keep your audience in mind. Provide them with the information they want or need most.
- Be clear about your purpose. Limit your display to communicate a few key points.
- Avoid distractions. Keep the design of your data display uncluttered and free of irrelevant information.
- Provide adequate context. Include a concise but descriptive title, unambiguous headings or data labels, clearly defined units of measure, and the size of the sample represented in the display.
- Organize the data so it is easy to find. For example, provide sample size information in the title or column headings.
- Be consistent in your labeling. For example, you can display Student Groups in columns or in rows, but do not switch back and forth. Similarly, use percentages throughout the table as opposed to using numbers for some variables and percentages for others.

For more information about using graphics to report evaluation results, review:

[Minter, E. & Michaud, M. \(2003\). \*Using Graphics to Report Evaluation Results\*. Program Development & Evaluation. University of Wisconsin-Extension.](#)

#### F.4 Summary of Key Points

- The basic principles for good communication apply whether you are writing a formal evaluation report or presenting data in visual display formats: Know your audience, be clear and concise in your purpose, provide adequate contextual information, and organize the information so that is accessible and understandable.
- The presentation, format, and emphasis of evaluation findings may vary depending on the audience and purpose of the document. For example, documents will look considerably different for policymakers as opposed to parents.
- Evaluation results are best understood in the context of the needs of the target population, the specific services that were delivered to address those needs, and the outcomes you expected to achieve. Look for ways to present contextual information in conjunction with evaluation questions, data collection methods, and findings.

#### F.5 Reflection Questions

1. Are the evaluation findings clearly linked to specific needs addressed and evaluation questions asked?
2. Is the document reader-friendly, that is, simple in language, well-organized, and consistent in styles and labeling?
3. Have we included too much information so that key points are lost on the reader? Have we included enough contextual information so that the reader can better understand the results?
4. Do the stated implications for the evaluation findings make sense based on the data presented?

#### F.6 Other Resources

Lester, W. B. & Braverman, M. T. (2004). Communicating Results to Different Audiences in Braverman, M.T., Constantine, N.A. & Slater, J.K. (Ed), *Foundations and Evaluation for Effective Philanthropy* (pp. 281-304). San Francisco, CA: Jossey-Bass.

[W.K. Kellogg Foundation Evaluation Handbook](#) (Communicating Findings and Insights, pp. 96-98).

# Section G: Using Evaluation Findings

---

Utility is a hallmark of high-quality evaluation and one of the standards that guides and motivates the work of professional evaluators:

*Evaluations should be planned, conducted, and reported in ways that encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased<sup>1</sup>.*

There are a number of recommended practices used by evaluators to increase the utility of their work, including:

- Focusing the evaluation on a specific and limited set of questions about the quality of program implementation and the achievement of expected outcomes
- Collecting reliable and credible evidence of progress and success from more than one source
- Using data analysis techniques that are appropriate for the available data and which maximize the learning that is possible from the data
- Communicating evaluation findings to multiple audiences with diverse information needs

## G.1 Continuous Improvement

In the first section of the *Migrant Program Evaluation Toolkit*, evaluation was situated in a Continuous Improvement Cycle for Migrant Education Programs (MEPs):

1. The cycle begins with the statewide Comprehensive Needs Assessment (CNA).
2. The CNA guides the development of the state's Service Delivery Plan (SDP).
3. The SDP is in turn implemented and evaluated,
4. The findings of the Program Evaluation are incorporated into the next CAN,
5. The Continuous Improvement Cycle is repeated.

When you have summarized the evaluation findings in a written format that can be shared with state Migrant Education Program (MEP) stakeholders, you will be ready to reconvene the CNA and SDP committees to look at evaluation results related to the Measurable Program Outcomes (MPOs) for specific MEP services.

To facilitate the process of incorporating the evaluation findings throughout the Continuous Improvement Cycle, a *Protocol for Using Evaluation Findings in a Continuous Improvement Cycle*

---

<sup>1</sup> [American Evaluation Association Guiding Principles for Evaluators. \(2004\).](#)

is referenced in *Appendix G.1*. You can use the discussion questions in conjunction with the results of your program evaluation as a starting point to:

- Reassess stakeholders' shared understanding of needs, barriers, and assets in the migrant community
- Refine the SDP and MPOs as needed to align with more challenging or more realistic strategies and expectations
- Revise evaluation questions and data collection methods as needed to align with changing strategies and expectations

In addition to reviewing current evaluation findings, committee members should prepare ahead of time for these discussions by reviewing longitudinal performance of Priority for Services and other migrant students on state performance targets, as well as changes in the state's Migrant Student Profile and Government Performance and Results Act (GPRA) measures over time.

## **G.2 Links to Other Online Resources**

[W.K. Kellogg Foundation Evaluation Handbook. \(2004\). Step 9: Utilizing the Process and Results of Evaluation \(pp. 99-104\).](#)

[University of Wisconsin Extension Cooperative Extension. \(1998\). Community Group Member Survey: Using the Results.](#)

## **G.3 Resources and Tools in Appendix G**

Appendix G.1 Protocol for Using Evaluation Findings in a Continuous Improvement Cycle

## Appendix G.1 Protocol for Using Evaluation Findings in a Continuous Improvement Cycle

1. **Reassess needs in the migrant community** – In collaboration with Comprehensive Needs Assessment (CNA) committee members, compare evaluation findings related to program implementation and Measurable Program Outcomes (MPOs) to the committee’s original need statements.
  - a. Are the original needs identified by the CNA committee still the highest priority needs for migrant students?
  - b. What, if any, higher priority needs emerged during the previous program year?
  - c. What, if any, additional barriers to migrant students’ success came to light during the evaluation process?
  - d. What, if any, assets in the migrant community or lifestyle came to light that could be leveraged to support student success?
  - e. Based on the evaluation findings and your current assessment of needs, would you recommend any different or additional strategies moving forward?
  
2. **Refine selected strategies and expected outcomes** – In collaboration with Service Delivery Plan (SDP) committee members, compare evaluation findings related to program implementation and MPOs to the original SDP.
  - a. To what extent have instructional services provided to migrant students produced the desired outcomes in academic achievement?
  - b. To what extent have non-instructional support services provided to migrant students and their families led to increased participation and engagement in their schooling?
  - c. Do findings from the implementation evaluation suggest that changes in service delivery may improve the desired outcomes?
    1. If so, what changes are necessary to improve the delivery of existing services?
    2. If not, what other services may be more effective for achieving the desired outcomes?
  - d. What, if any, adjustments should be made to specific MPOs to reflect realistic but still challenging expectations?

- e. If the CNA committee has recommended new service delivery strategies for the MEP, what are the appropriate MPOs, evaluation questions, and data collection plans?
3. **Revise the Evaluation Plan** – Also in collaboration with the SDP planning committee, review the original plan to evaluate program implementation and MPOs as stated in the original SDP.
- a. To what extent were selected outcome measures appropriate for evaluating specific MPOs?
  - b. What, if any, other measures may be better aligned with specific MPOs?
  - c. To what extent did the questions and methods used to evaluate the implementation of specific MEP services produce useful information?
  - d. What, if any, other questions would you ask about the implementation of specific services?
  - e. What, if any, other methods could you use to get useful information about service implementation?