THE CHILD CARE QUALITY RATING SYSTEM (QRS) ASSESSMENT

The Quality Rating and Improvement System (QRIS) Evaluation Toolkit

OPRE Report 2011-31

August 2011

This page has been left blank for double-sided copying.

The Quality Rating and Improvement System (QRIS) Evaluation Toolkit

OPRE Report 2011-31

August 2011

Mathematica Policy Research
Julieta Lugo-Gil
Samina Sattar
Christine Ross
Kimberly Boller
Gretchen Kirby

Child Trends
Kathryn Tout

Prepared for:
Office of Planning, Research and Evaluation
Administration for Children and Families
Department of Health and Human Services
370 L'Enfant Plaza Promenade, SW
7th Floor West, Room 7A011
Washington, D.C. 20447

Project Officers:
Ivelisse Martinez-Beck
Kathleen Dwyer

Prepared by:
Mathematica Policy Research
600 Maryland Ave., S.W., Suite 550
Washington, DC 20024-2512

Project Director:
Gretchen Kirby, Mathematica

Co-Principal Investigators:
Kimberly Boller, Mathematica
Kathryn Tout, Child Trends

This document was prepared under Contract #HHSP233200800394G with the Administration for Children and Families, U.S. Department of Health and Human Services. The views expressed in this publication are those of the authors and do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services. This report and other reports sponsored by the Office of Planning, Research and Evaluation are available at http://www.acf.hhs.gov/programs/opre/index.html.

Suggested citation: Lugo-Gil, J., Sattar, S., Ross, C., Boller, K., Tout, K., & Kirby, G. (2011). *The Quality Rating and Improvement System (QRIS) Evaluation Toolkit*. OPRE Report # 2011-31. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

OPRE
Office of Planning, Research and Evaluation
Administration for Children and Families

MATHEMATICA
Policy Research

Child TRENDS

This page has been left blank for double-sided copying.

ACKNOWLEDGEMENTS

The Quality Rating System (QRS) Assessment team at Mathematica Policy Research and Child Trends produced the Toolkit under contract with the Office of Planning, Research and Evaluation in the Administration for Children and Families, U.S. Department of Health and Human Services. The project officers, Ivelisse Martinez-Beck and Kathleen Dwyer, provided valuable advice and guidance in the planning and production of the Toolkit.

The QRS Assessment team appreciates the participation of the project's Expert Panel and the wisdom and experience they shared in the development of the Toolkit. The team is also grateful for the review and contributions to the Toolkit by Karen Tvedt and Shannon Christian of Christian & Tvedt Consulting, as well as our colleagues at Mathematica Policy Research: Jessica Ziegler for research on external resources, Cindy George and John Kennedy for editing, and Jennifer Baskwell, Linda Heath, Laura Sarnoski, and Jill Miller for production.

Finally, the team would like to thank the members of the INQUIRE group who graciously answered questions and provided us with information from their QRIS evaluations. Their willingness to share their knowledge has helped us create a useful and practical tool for the field, and the team expresses gratitude for their collaboration in this work.

This page has been left blank for double-sided copying.

CONTENTS

TABLES

FIGURES

This page has been left blank for double-sided copying.

I. INTRODUCTION

The use of Quality Rating and Improvement Systems (QRISs) is continually expanding across the country; in 2010, there were 25 states operating or piloting QRISs and now nearly every other state is planning to or has already begun developing a QRIS or another quality improvement (QI) initiative (Tout, Starr, Soli, Moodie, Kirby, and Boller 2010; U.S. Department of Health and Human Services, 2010b). As federal and state funds are invested in efforts to improve the quality of child care and early childhood education through the development and implementation of QRIS initiatives, there is an increasing need for informative and rigorous evaluation of the implementation of QRIS; the weaknesses and strengths in its operation; the validation of the tools used by the QRIS; and the progress made in achieving desired outcomes for child care programs, families, children, and the early childhood education system. Recognizing this need, the Office of Planning, Research and Evaluation (OPRE) in the Administration for Children and Families (ACF), U.S. Department of Health and Human Services (DHHS) is supporting the Child Care Quality Rating System Assessment (QRS Assessment) project.[1] The goal of the QRS Assessment is to provide information, analysis, and resources about QRISs for states and other stakeholders.

The *QRIS Evaluation Toolkit*[2] is a product of the QRS Assessment and is intended to serve as an informational resource for state administrators, child care and early education practitioners, and other stakeholders on how QRISs work; why it is important to conduct evaluation of QRISs; and on how to plan and design an evaluation of QRISs. Several useful and extensive guides for conducting

[1]Mathematica Policy Research is conducting the QRS Assessment in partnership with Child Trends and Christian & Tvedt Consulting.

[2]For simplicity the QRIS Evaluation Toolkit is referred to in this document as "the Toolkit".

evaluation of social and education programs are already available,[3] however, this Toolkit focuses specifically on evaluation of QRISs.

The Toolkit brings resources pertinent to the evaluation of QRIS together in one place and translates general guidelines into specific activities appropriate for QRIS evaluations through examples. The intent of this tool is not to be the *only* resource or to duplicate existing resources, but to provide a one-stop shop for planning a QRIS evaluation. That is, the searching and culling of resources has been done for you. The Toolkit covers the key topics in planning a QRIS evaluation—building a logic model, creating research questions, selecting a design, and selecting measures. When you need additional detail on a particular topic, the resource lists organized by topic and specific measurement tools (both included in appendices) will get you where you need to go.

The next chapter provides information on documenting how QRISs work by presenting a step-by-step guide to constructing logic models for QRIS initiatives that can guide and support evaluation; Chapter III discusses research questions by type of evaluation and stage of QRIS, and then discusses the different evaluation approaches (designs) that can address the research questions of interest; Chapter IV describes measures that can be used in the evaluation of QRISs and discusses data collection best practices; Chapter V presents a discussion on how costs differ for different types of QRISs evaluations and provides examples of funding levels for actual QRISs evaluations; and Chapter VI provides guidelines on developing a Request for Proposals (RFP) and selecting and working with an evaluator. A list of online, external resources on building logic models, evaluation (including evaluations of existing QRISs), measures, and selecting and working with an evaluator, as well as summaries of the contents of the resources and the methods used to find them are presented in the appendices.

[3]For example, see "The Program Manager's Guide to Evaluation" (U.S. Department of Health and Human Services 2010d) and the W. K. Kellogg Foundation Evaluation Handbook (W. K. Kellogg Foundation 2004a).

II. HOW DOES YOUR QUALITY RATING AND IMPROVEMENT SYSTEM (QRIS) WORK?

Before planning and conducting evaluations of Quality Rating and Improvement Systems (QRISs), it is important to understand how the initiatives work. A logic model can help with this because it illustrates the goals, components, sequence of activities, and achievements of a Quality Rating and Improvement System (QRIS) in a systematic way. This chapter presents a step-by-step guide to constructing a logic model, which can help clarify expectations for the goals or outcomes of the QRIS and thus help guide the implementation, evaluation, and continuous improvement of these systems. The information presented here can also be applied to the implementation and evaluation of other quality improvement (QI) interventions. Throughout this chapter, we use the term "intervention" to denote any program or initiative that is expected to change the environment and, thereby, change outcomes. Links to several external resources on the construction of logic models for further reference are presented in Appendix A.

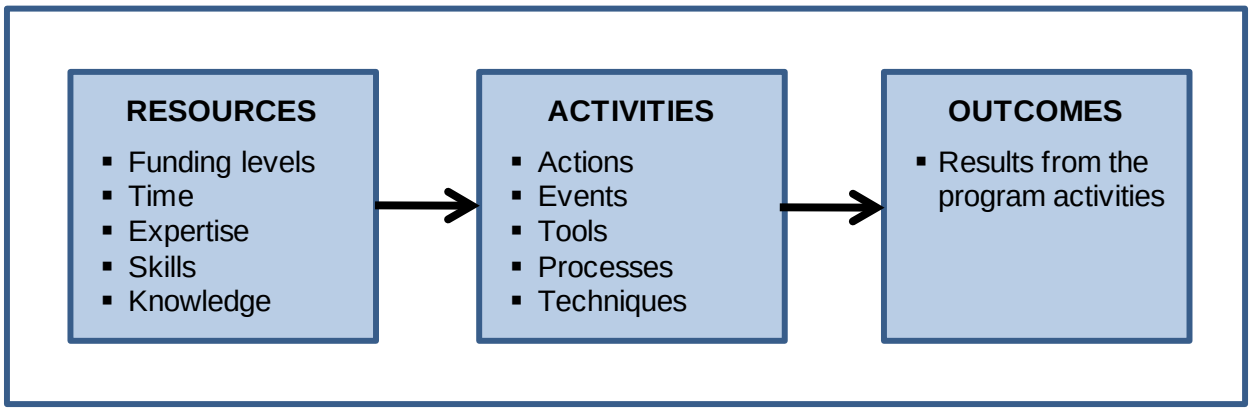
CHAPTER CONTENTS

A. What Is a Logic Model?

A logic model is a systematic and visual way to present expected relationships among (1) the resources available to the program; (2) the activities or policies that are to be put in place; and (3) outcomes, or the changes or results that are expected to follow (W.K. Kellogg Foundation, 2004b). In other words, logic models depict the links between the environment, the intervention, and the changes in the environment that result from the intervention.

Figure II.1 shows an example of a basic logic model, with the program resources in the first box, the activities or policies in the second box, and the outcomes in the third box. The resources include the funding, expertise, time, skills, and knowledge that the program can direct towards doing the work. The activities refer to what the program does with its resources, that is, the actions, events, tools, and techniques that constitute the work of the program. And finally, the outcomes include all the results of the program activities (W. K. Kellogg Foundation, 2004b).

Figure II.1. Basic Logic Model



B. Why Is It Important to Use a Logic Model?

Logic models are crucial for effective program implementation and evaluation because they can help stakeholders clarify their goals and expectations for the QRIS program. QRISs may articulate the broad goal of "quality improvement" differently and, therefore, the systems designs may also vary. For example, a QRIS with a goal of improving the quality of care that children supported by child care subsidies receive may emphasize child care quality standards and financial incentives that encourage achievement of those standards (for example, tiered reimbursement rates that increase with a provider's rating level). Another QRIS might have a broader goal of improving the cognitive and social-emotional development of all preschool children; such a QRIS might aim to help a wide range of child care providers achieve the standards for high rating levels, provide financial incentives to support achievement of high rating levels, and help families gain access to higher-rated child care providers.

Clarifying the program's resources, activities, and policies and identifying realistic, measurable indicators of outcomes will help stakeholders assess whether these components are likely to be adequate to achieve the targeted goals. Documenting the expected outcomes is important to ongoing monitoring of program implementation and performance, and serves as a basis for evaluating the QRIS. A logic model thus enables planners to clearly state the purposes, components, sequence of activities, and accomplishments of the QRIS. The earlier a logic model is applied, the more it will inform the implementation, success, and evaluation of the QRIS.

C. Elements of a QRIS Logic Model

For QRIS-specific logic models, we expand the basic logic model presented in Figure II.1 in three ways. First, we specify the community environment and goals that underlie the QRIS design. Second, we divide "outcomes" into direct program *outputs* (measurable direct results of activities, such as the number of child care providers receiving quality improvement support) and more distant *outcomes* for providers, families, and children (for example, increased quality of care, parents systematically using the QRIS ratings to choose child care providers, and improved child outcomes). Third, we specify external influences (for example, social and political influences) that may affect the QRIS resources and activities.

We present descriptions of the key elements of a logic model and some QRIS-specific examples of them in Table II.1. These examples are meant to provide a quick reference to the elements of a QRIS logic model to help you in building a logic model for your program. More detailed descriptions and examples of these elements are presented in Section II.D.

Table II.1. Key Elements of a QRIS Logic Model

Logic Model Element	Description and Examples
Environment and Goal Statement	• Describes the children, families, child care settings, and community (the environment) that will participate in the QRIS and benefit from the resources provided by it. • Outlines the stakeholders' expectations and goals for the program. For example: - Rating a certain number of child care providers - Supporting child care providers to reach a certain rating level - Helping a certain numbers of families learn about the providers' ratings and use them in making decisions about care for their children • Lists the QRIS stakeholders. For example: - State administrators - Representatives from child care providers - Families - Accreditation agencies - Early education researchers - Evaluation partner • Could explain assumptions regarding how the planned activities of the QRIS will link to expected results. For example: - Technical support provided to child care providers is expected to increase the quality of care provided and improve providers' ratings - Dissemination activities and engagement of parents is expected to result in parents learning about the QRIS ratings and using them in choosing child care providers - Improved quality of child care providers participating in the QRIS is expected to improve cognitive and emotional outcomes for children in the long run
Inputs	• Inputs are the resources invested in the QRIS, including knowledge, expertise, and skills, and financial, political, and community assets. • Inputs support the activities of the QRIS. • The inputs in a QRIS logic model include: - Legislation - Funding levels - Administrative structure: the agency or organization that leads the QRIS, the "administrative entity," provides the support to operate the QRIS and implement its activities. It can be a state government agency, a local government agency, a nonprofit organization, or another agency - Number and qualifications of the QRIS agency staff - Resources to reach specific communities and groups of parents - The standards that constitute the basis of the system: quality standards for professional development, training, the early childhood learning environment, and parent involvement - The organizations partnering with the QRIS to provide support in reaching the QRIS goals, such as state agencies, resource and referral agencies, community colleges, universities, or other nonprofit organizations
Activities	• The actions, processes, techniques, events, tools, and technology that need to be implemented so that the QRIS reaches its ultimate goal of improving the conditions in education and care settings for children • The activities vary in scope and intensity across QRISs. For example, QRIS rating systems include different numbers of quality standards and assess different qualities of child care providers. • Examples of QRIS-specific activities include: - Facilitating child care providers' application to and entry into the QRIS - Ensuring the integrity of the process to assess child care providers - Validating rating levels/components and the overall rating for each provider (that is, determining whether the QRIS levels represent different levels of quality)

Logic Model Element	Description and Examples
	- Providing technical assistance for quality improvement and/or connecting child care providers to other agencies that can provide these services - Disseminating information on the QRIS ratings and accomplishments to parents and the community
Outputs	• Outputs are the results of the QRIS activities, and are usually described in terms of the size and/or scope of the activities undertaken and the products created by the QRIS. They indicate whether the intervention targets were reached at the intended "dose." • Examples of QRIS-specific outputs include: - Number of child care providers/programs rated - Number of programs receiving technical assistance and/or connected with agencies that can provide assistance - Number of families provided with information on child care provider ratings - Number of sessions held to inform families and the community about the QRIS goals, activities, and achievements - Number of families using the QRIS resources such as websites and consultations with staff
Outcomes	• The outcomes are the results of the QRIS activities and can be short-, medium-, or long-term • Examples of short-term outcomes are: - Providers learning about the QRIS and choosing to participate - Parents learning about program ratings from child care resource and referral agencies or state child care websites • Examples of medium-term outcomes are: - An increased number of providers adapting their physical space, staffing structure or qualifications, and/or instructional practices in response to observational assessments or to improve their QRIS-rating level - The use of the QRIS ratings by parents in selecting a child care provider • Examples of long-term outcomes of QRIS activities are: - Providers consistently earning high ratings - Increased supply of high-quality child care - Parents consistently base their choices of providers on the QRIS ratings - Increased demand for high-quality care that leads to closings of low-quality programs - Consistent availability and access to high-quality care - Improved cognitive and emotional child outcomes, including improved school readiness
External Influences	• External influences refer to the influence that the social, physical, political, and institutional settings can have on the inputs, activities, outputs, and outcomes of the QRIS. For example, the following can influence the funding opportunities, timeline of implementation, and ultimate success of a QRIS: - Changes in support provided by the government, funding agencies, and other organizations with important roles in the child care sector - Changes in support from parents and community organizations that cater to families' interests

Logic Model Representations. Logic models can be created in different ways, but all approaches have common elements. Logic models are frequently presented as flowcharts with the elements of the model arranged in boxes connected by arrows. However, logic models can also be

presented in a table (Westmoreland et al., 2009). Figures II.2 and II.3 show examples of QRIS-specific logic models in different formats.

"Reading" a logic model from left to right in the case of Figure II.2, or from bottom to top in the case of Figure II.3, illustrates how the resources invested in the program lead to activities that in turn lead to achieving the program's goals (Zellman et al., 2011). Although logic models are often developed and "read" from left to right or from bottom to top, sometimes they are developed and used starting from the other end of the model (the right side of the model in Figure II.2 or the top of the model in Figure II.3). That is, sometimes goals (desired outcomes) are established first and then the elements that will result in those goals (for example, the activities that are likely to achieve the goals and the resources necessary to implement the activities) are developed (Breitner et al., 2010).

The step-by-step guide to build a QRIS-specific model presented in the next section is based on the approach to developing and "reading" logic models from left to right (or from bottom to top). That is, our approach to build a QRIS-logic model begins with a description of the QRIS environment and ends with a description of the desired outcomes for the QRIS. It should be noted, however, that there exist other approaches to build a logic model. For example, the field of implementation science, an approach to assessing the fidelity of implementation to evidence-based practices or programs, relies heavily on documenting program logic and assessing progress toward targeted outcomes (Fixsen et al. 2005). One method used by program evaluators and technical assistance providers in that field engages program developers/operators in articulating a "cascading logic model" that depicts the causal chain between each step in what is expected to change and how that affects the next target of change (Blase 2008). For example, if teachers using a specific evidence-based practice is one step in the logic model, the preceding behavioral change is that technical assistance providers have to engage lead teachers in learning about the evidence-based practice. The cascading behaviors can be depicted in pairs and linked to the previous behavior changes that must

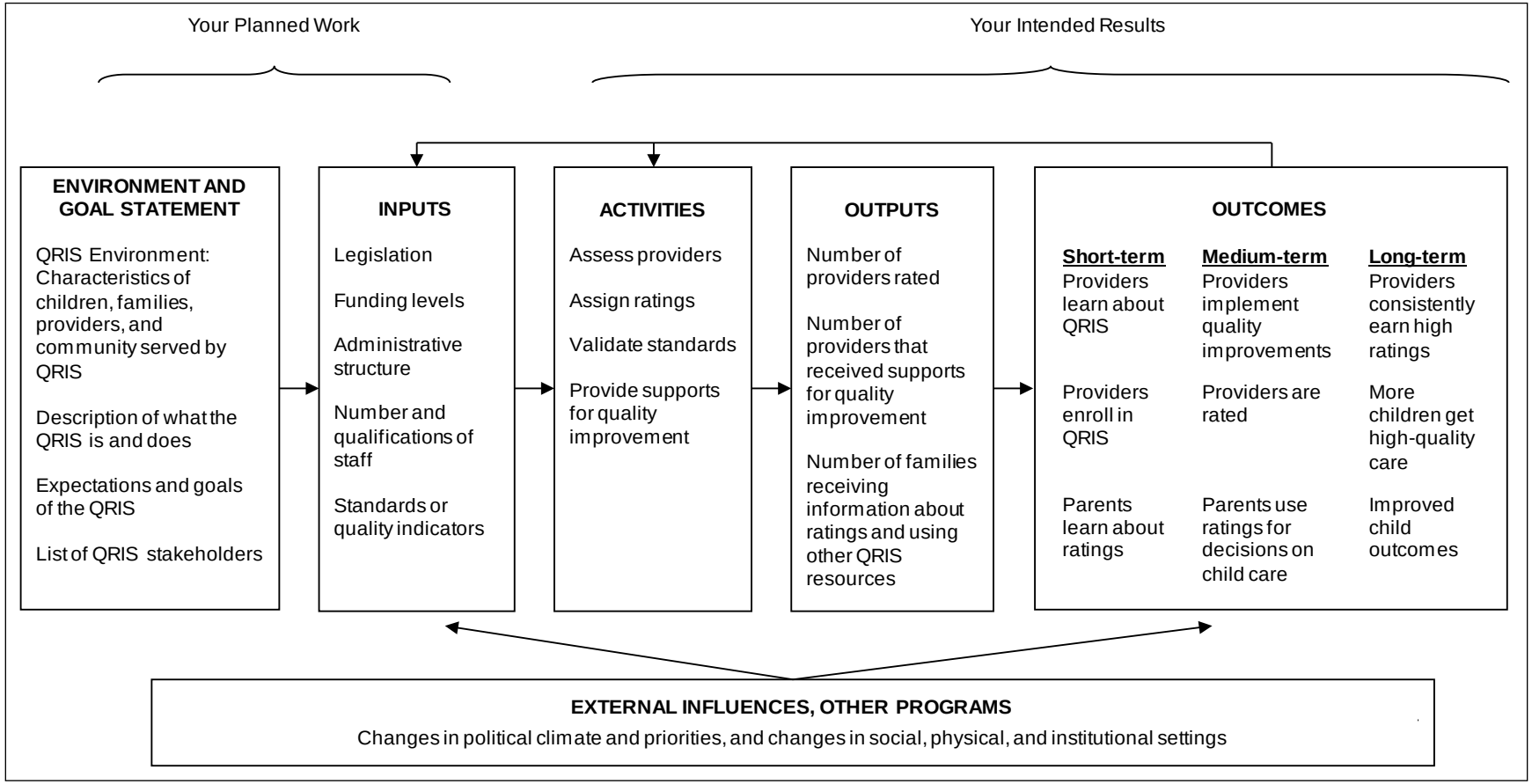
happen at each step between the overall goal of the program and the overall desired outcomes. One benefit to this approach is that by graphically representing these links and how they cascade, program developers/operators and evaluators can quickly identify gaps in the program logic and work to fill them.

D. Developing a QRIS Logic Model

As stated earlier, the purpose of the logic model is to inform the implementation, success, and evaluation of the QRIS by clearly depicting the goals, components, sequence of activities, and achievements of the initiative. However, producing streamlined descriptions of all aspects of a QRIS that can be used to fill in logic model "boxes" like those presented in Figures II.2 and II.3 requires that state administrators, QRIS managers, and other QRIS stakeholders first work together to achieve consensus on establishing (1) the goals of the initiative, (2) the ways in which the resources available to the program will be used, and (3) the desired outcomes of the program activities. Stakeholders need to confer and agree on the information about the QRIS that will be included in the logic model. The types of activities and the amount of time required to complete that process depends on the stage of development of the QRIS. For example, consider the following three QRISs in different stages of development:

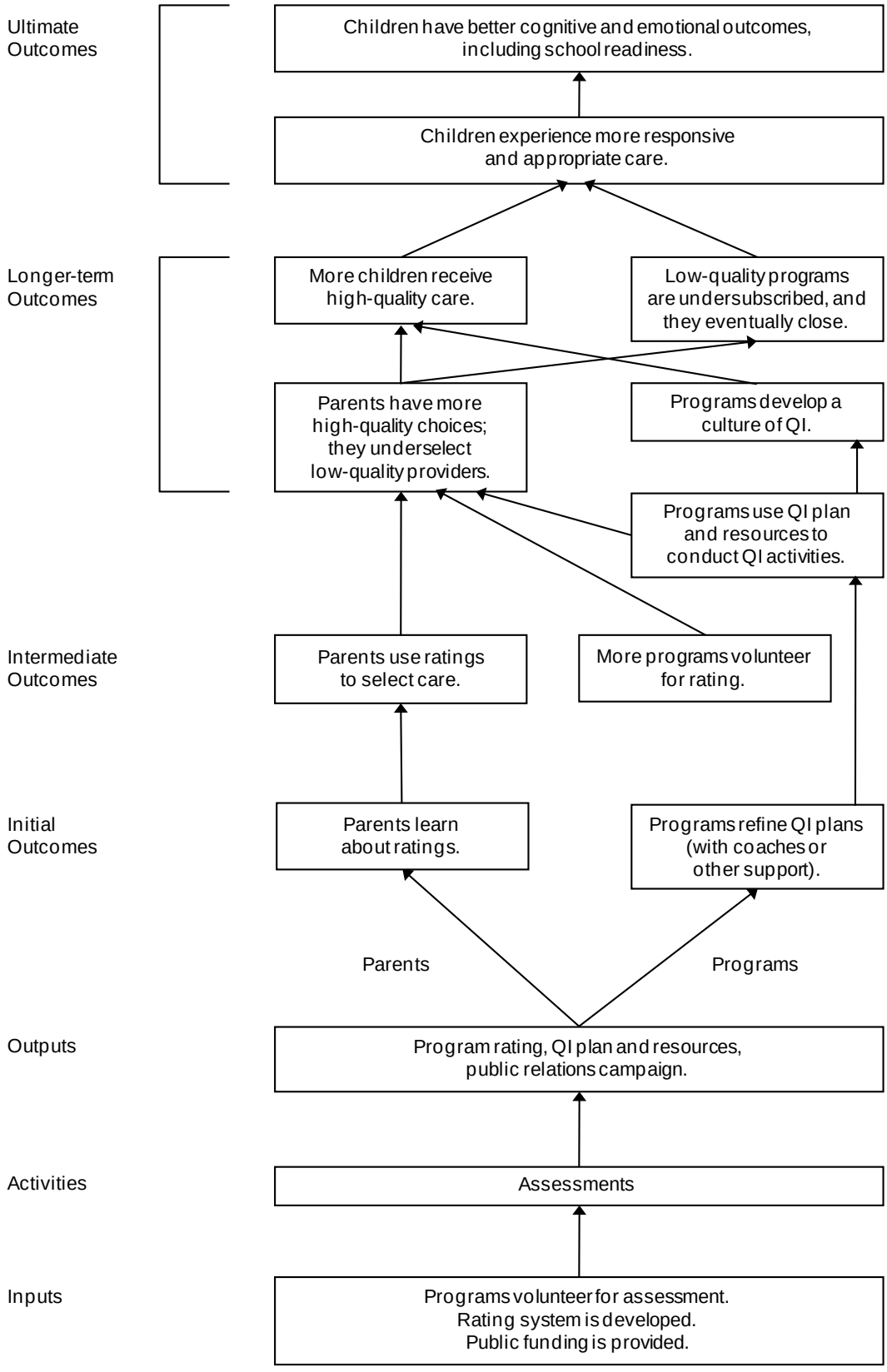
- **A QRIS that is about to be launched or has just started operations**. To build a logic model for such a QRIS, state officials and QRIS managers might first convene other stakeholders to make sure that all agree on the goals, activities, and outcomes of the QRIS. This can be done in one meeting or in a series of meetings. After that, the state officials and QRIS managers might work together with a smaller group of stakeholders to organize the information gathered from the initial meetings; draft detailed descriptions of the agreed-upon QRIS goals, activities, and outcomes; and prepare work plans. A logic model can then be constructed based on those longer descriptions of the QRIS goals, activities, and outcomes. The team could then present the logic model to the larger group of stakeholders and refine it further.

Figure II.2. Example of a Basic QRIS Logic Model—Horizontal Display



Source: Adapted from McCawley, 2001; Westmoreland et al., 2009; W. K. Kellogg Foundation, 2004b.

Figure II.3. Example of a Basic QRIS Logic Model—Vertical Display



Source: Zellman and Perlman, 2008.

- **A QRIS that is already using its resources to conduct activities**. To build a logic model for such a QRIS, state administrators and QRIS managers could summarize information on the inputs used and the activities conducted to date and include it in the "inputs" and "activities" boxes, respectively, of the logic model for the QRIS. Then, they could place the implicit goals of using those activities in the "outputs" and "outcomes" boxes of the logic model. For example, a QRIS that has been in operation for about a year may have already enrolled and rated some child care providers. The "inputs" box in the logic model for this QRIS could show the amount of financial resources and staff time (and qualifications) used thus far, and describe the standards upon which ratings of child care providers are based. The "activities" box could list the actions already conducted to assess the quality of child care providers (such as observations) and to provide quality improvement assistance (such as technical assistance for quality improvement and/or connecting child care providers to other agencies that can provide these services). The "outputs" box of the logic model could state the number of child care providers that have been rated and received technical assistance (for example 25 family care providers and 10 child care centers) and the number of families who have received information on the QRIS ratings. The "outcomes" box could state the results of those activities and outputs (for example, child care providers received technical assistance to improve quality and start implementing quality improvement changes, and parents in the community learned about the QRIS ratings).
- **A QRIS for which a logic model already exists**. A QRIS that has already developed a logic model for its program should revise it periodically (for example, every year) to verify that it is current (that is, that the logic model reflects the resources that are currently available to the QRIS, the activities that have been conducted since the previous revision, and the outputs of those activities) and that the results of the activities described in the logic model are indeed the desired outcomes for the QRIS (to check that the program is on track to achieve its goals). Thus, a current and accurate logic model can serve to support the evaluation of the QRIS (see section II.F for more details on how a logic model can help in planning an evaluation of a QRIS).

The process of building a logic model can involve multiple discussions with stakeholders and the logic model may go through a number of iterations before the stakeholders agree upon a final version. Logic models can become extremely complex when they include details about the relationships between inputs, activities, outputs, and outcomes. Therefore, the logic model should remain a straightforward organization and communication device for planning, monitoring, and evaluation. To maintain the clarity of the logic model, it is best not to include details on the activities' time frames, priorities, and responsible parties. Instead, present them in separate, more detailed documents (such as work plans for specific activities) that are aligned with the logic model.

Step-by-Step Guide to Creating a QRIS Logic Model

This section presents a detailed guide to creating a QRIS logic model. For each box in the logic model, we describe the information that needs to be included and the questions that need to be asked about the particular QRIS in order to complete it. The section includes worksheets with examples of information that could be included in each box and space for specific details about your QRIS.

Step 1: Describe the Environment and Articulate the Goals of the QRIS

An optimal way to begin building a QRIS logic model is to prepare a statement describing the environment in which the QRIS operates and the goals that it expects to achieve. This step is useful because it:

1. Establishes a baseline to be used as a comparison as the program progresses or when it ends (McCawley, 2001)
2. Encourages stakeholders to define the goals of the QRIS and determine whether the goals are realistic given the available resources and the priorities of the QRIS (Westmoreland et al., 2009)
3. Underscores the need for stakeholders to work collaboratively on building the QRIS as a hub for quality improvement (Tout et al., 2009)
4. Emphasizes that the various targets (child care providers, parents, children, parents, and the community) of the QRIS have different needs, which should be considered when identifying the resources and activities to be included in the logic model

This statement should include the following information about the QRIS:

1. A description of the **environment** in which the QRIS operates. Describe the characteristics of the children, families, child care settings, and community that participate in the QRIS and will benefit from the resources provided by it. The description of the environment will help stakeholders understand the context for the QRIS. Each QRIS is implemented in an environment with a unique set of characteristics and thus, it cannot simply be a copy of another QRIS. For example, QRISs in states with limited licensing and enforcement may need to focus more attention on health and safety than those in states with rigorous licensing regulations and strong enforcement. Similarly, the incentive structures in the QRIS might look different in states with adequate subsidy reimbursement rates than in states with low provider reimbursement rates.

2. A **definition** of the QRIS, for example:
   - The QRIS is a method to measure and improve the quality of child care programs

- The QRIS is a program designed to increase awareness about standards in child care quality, recognize child care programs that provide high-quality care, and help parents learn about high-quality child care programs
- A QRIS is a "consumer guide, a benchmark for program improvement, and an accountability measure for funding" (Brian A. Gallagher, President and CEO, United Way of America, "First Things First," included in Mitchell, 2005)

3. The **expectations** and **goals** for the QRIS, for example, rating a certain number of child care providers, supporting child care providers to reach a certain rating level, and/or helping a certain number of families learn about the providers' ratings and use them in making decisions about care for their children. The goals included in the QRIS logic model should be specific, measurable, achievable, realistic, and time-bound (SMART, Doran, 1981), as follows:
   - **Specific:** Goals are well defined and clear to anyone who has a basic knowledge of the QRIS
   - **Measurable:** Goals include a numeric or descriptive measure that makes it possible to know whether the goal is obtainable, how much additional work is needed to achieve the goal, and when the goal has been achieved
   - **Achievable:** Goals are within reach of the QRIS staff and stakeholders and have been agreed upon by all the stakeholders
   - **Realistic:** Goals are within the resources, knowledge, and time constraints of the QRIS
   - **Time-bound:** Goals should be achieved within a specific deadline, and the proposed time frame should be sufficient to achieve the QRIS goals

   The following are examples of SMART QRIS goals:

   - Assign ratings to 40 percent of licensed child care programs/providers within one year
   - Increase the number of child care programs/providers that attain a higher rating by 50 percent during the period 2010-2011
   - Within six months of the launch of the state marketing campaign, receive 500 "hits" per week for the information on QRIS ratings provided in the QRIS website

The assumptions about how the planned activities of the QRIS will link to results can be included here. For example:

- Technical support to child care providers will lead to quality improvements and higher ratings for the providers
- Dissemination activities and engagement of parents will result in parents learning about the QRIS ratings and using them in choosing a child care provider

- Improved quality of child care providers participating in the QRIS will lead to improved cognitive and emotional outcomes for children in the long run

4. A list of the QRIS **stakeholders**. Provide a list of the individuals and agencies who participate in planning, developing, and implementing the QRIS. For example, list the members of the QRIS task force or committee, such as:

    - The department of human or social services of the state
    - State's office of early learning
    - Licensing and accreditation agencies
    - Early education advocates and organizations
    - Child care resource and referral agencies
    - Professional organizations
    - Early education researchers
    - Teacher/educator representatives
    - Representatives from child care centers, family care, Head Start, and school-age care
    - Representatives from health care agencies or organizations
    - Parent/families representatives

The statement describing the environment and the goals of the QRIS can be included in the first box of a logic model like the one in Figure II.2. Alternatively, although not currently in the figure, it could be entered in a box below the "inputs" box in a logic model like the one in Figure II.3. Before entering information about the QRIS environment and goals in the corresponding logic model box, it is helpful to prepare a one- or two-page statement describing the environment, defining the QRIS, stating the initiative's goals, and listing the stakeholders. Once that statement has been reviewed and agreed upon by the program stakeholders, key elements can be abstracted from it to be included in the logic model. Table II.2 presents examples of the statements from four QRISs describing their environment, defining the QRISs, outlining the programs' expectations and goals, and listing their stakeholders. Worksheet 1 summarizes the information to be included in the environment and goals statement and provides a space to enter a description of the environment and the goals for your QRIS.

Table II.2. Statements Describing the Environment and Goals of Four QRISs

QRIS Program	Program Name	Description of the Environment	Definition of the QRIS	QRIS Goals	QRIS Stakeholders
Maine	Quality for ME	Based on meetings with focus groups and on examination of national research on quality indicators, the Early Childhood Division (ECD) of the Department of Health and Human Services (DHHS) created a tiered system to identify programs based on their level of quality and to determine the supports programs need to increase their quality as measured by the Quality for ME system.	Quality for ME is a four-step program designed to increase awareness of the basic standards of early care and education, recognize and support providers who are providing care above and beyond those standards, and to educate the community about the benefits of higher-quality care.	The goals of Quality for ME are: • To recognize child care programs that provide quality care • To encourage providers to increase their level of quality • To provide parents with identifiable standards of quality	• Early Childhood Division (ECD) of the Department of Health and Human Services (DHHS) • Maine Department of Education • Child care providers representatives • Head Start representatives • Parent representatives • Early education advocates and organizations
Maryland	Maryland Child Care Tiered Reimbursement Program	Not included in source.	The Maryland Child Care Tiered Reimbursement Program is a system to recognize child care programs that go beyond the requirements of state licensing and registration regulations.	The goals of the program are: • To promote a well-qualified workforce • To increase the amount of Purchase of Care payments to child care programs • To promote a high level of program quality through the application of standards for program accreditation and environmental rating scales • To encourage parent involvement in child care settings	• Maryland State Department of Education • Child care providers representatives • Parent representatives
Missouri	Missouri Quality Rating System	The Missouri Quality Rating System (MO QRS) initiative was established to provide a reliable, valid, easy-to-understand measure of quality (with ratings from 1 to 5 stars) to state policymakers, early childhood and school-age professionals, and parents.	The Missouri Quality Rating System (MO QRS) is a method to assess and continually improve the quality of early childhood and school-age/after-school programs.	The goals of the program are to: • Improve the quality of early childhood and school-age/after-school programs for Missouri's young children and youth • Raise public understanding about high-quality programs • Allow subsidy-receiving	The Missouri Quality Rating System State Committee is composed of representatives from: • Department of Health and Senior Services • Department of Social Services • Department of Mental Health • Missouri Department of

QRIS Program	Program Name	Description of the Environment	Definition of the QRIS	QRIS Goals	QRIS Stakeholders
				children access to higher-quality programs by linking state child care subsidies to the various QRIS levels	Elementary and Secondary Education • Missouri After School Network • Missouri Child Care Resource and Referral Network • Association for Education of Young Children of Missouri • State Accreditation Agencies • TEACH Missouri Scholarship • Center for Family Policy & Research, Institute for Human Development, Start Up and Expansion, University of Missouri
North Carolina	North Carolina Star Rated Licensing System	• In 1993, the creation of the Division of Child Development reflected the growing importance of child care to North Carolina families. • North Carolina boasts one of the highest rates of working mothers with young children in the nation, making the availability of child care essential for the state's economic development and stability.	Not included in source.	The goals of the program are to: • Implement standards • Increase access for families • Collaborate to promote enhanced service delivery of care and education across the state	• Division of Child Development, North Carolina Department of Health and Human Services • Child care providers representatives • North Carolina Child Care Resource and Referral Network representatives • Early education advocates and organizations

Source: Quality for ME. Maine Department of Health and Human Services, Office of Child and Family Services Early Care & Education, Early Childhood Division, Maine, 2007. Accessed on November 2, 2010 [http://www.maine.gov/dhhs/ocfs/ec/occhs/qualityforme.htm]; Maryland Child Care Tiered Reimbursement Program Website, Maryland State Department of Education. Accessed on November 2, 2010 [http://www.marylandpublicschools.org/MSDE/divisions/child_care/credentials/tiered]; OPEN: Missouri's Professional Development Initiative for Early Childhood and School-Age/After-School Professionals, Curators of the University of Missouri, 2007. Accessed on November 2, 2010 [https://www.openinitiative.org/content.aspx?file=QRIS.txt]; NC Division of Child Development Home Page. North Carolina Department of Health and Human Services. Accessed on November 2, 2010 [http://ncchildcare.dhhs.state.nc.us/general/mb_aboutus.asp].

Worksheet 1. Environment and Goals of the QRIS

What Is This?
A statement describing the following information about the QRIS: 1. A description of the **environment** in which the QRIS operates 2. A **definition** of the QRIS 3. The **expectations** and **goals** for the QRIS 4. The **assumptions** about how the planned activities of the QRIS will link to results 5. A list of the QRIS **stakeholders**
QRIS-Specific Examples
Environment (from Minnesota's Parent Aware) • In 2003, only 11 percent of full-day child care centers and 2 percent of family care providers in the state were recognized by national accreditation, and nearly 50 percent of five-year-olds were not fully prepared for kindergarten • Minnesota lags behind all other states in the K-12 achievement gap between white and black children **Definition** (from Missouri Quality Rating System) • The QRIS is a method to assess and continually improve the quality of early childhood and school-age/after-school programs **Goals** (from North Carolina Star Rated Licensing System) • Implement standards • Increase access for families • Collaborate to promote enhanced service delivery of care and education across the state **Assumptions** (from Minnesota's Parent Aware) • Children who attend quality early care and education and school-age programs can "catch up" to their peers, have more chances to graduate from high school, and go on to lead stable, contributing lives **Stakeholders** (from Maryland Child Care Tiered Reimbursement Program) • State Department of Education • Child care providers representatives • Parent representatives
Your QRIS
Environment _____ _____ _____ _____ **Definition** _____ _____ _____ _____ **Goals** _____ _____ _____ _____ **Assumptions** _____ _____ _____ _____ **Stakeholders** _____ _____

Step 2: Identify the QRIS Inputs

The inputs included in the logic model reflect what is needed for the QRIS to reach the goals stated in Step 1. As is the case for most QI interventions, QRIS inputs include existing legislation and political support, funding, planning, administrative structure, staffing, collaboration across agencies, and use of research evidence to inform the QRIS actions. Below, we describe various types of QRIS inputs; Table II.3 shows examples of inputs reported by existing QRISs.

Types of Inputs

- **Administrative structure**: The agency or organization that leads the QRIS, usually referred to as the "administrative entity." This can be a state or local government agency, a nonprofit organization, or another agency.
- **Rating structure and standards**: This refers to the ways in which QRISs assign ratings and to the quality standards for professional development, training, learning environment, and parent involvement, for example, that provide the basis for the rating system. In addition, other standards in the system such as early childhood learning guidelines, core competencies for early childhood care and education practitioners, and standards from other early childhood programs can be included as inputs in the logic model. The standards or indicators serve as the foundation for the QRIS, they provide the definition of quality, and they send a signal to programs and parents about what is important for programs to do in their daily practices with children.
- **Monitoring/accountability**: Processes to ensure that child care providers/programs participating in the QRIS conduct efforts to improve and obtain high ratings. This can also include other existing processes such as licensing, accreditation, and other early childhood education professional development infrastructure such as registries and career lattices.
- **Quality improvement**: Characteristics of the QRIS staff such as the number of staff, their background, and their qualifications for their specific roles at the QRIS. In addition, QI inputs include facilities, equipment, and technology to help child care providers/programs improve quality. Finally, the professional development system for early childhood education practitioners is also an input related to quality improvement.
- **Collaboration with other agencies**: Partnerships with other organizations such as state agencies, resource and referral agencies, community colleges, universities, or other nonprofit organizations.
- **Dissemination**: Tools in place for communicating information about early care and education programs to parents, such as through Child Care Resource and Referral (CCR&R) agencies.

Table II.3. Examples of Inputs from Existing QRISs

Type of Input	Examples
Funding	QRISs reported one or more of the following agencies as **overall funding sources**: • Child Care Development Fund • State • Specific state agency (such as the State Department of Human Services) • Temporary Assistance for Needy Families (TANF) funds • Local or county funds • Foundations • United Way Funds • Tobacco Settlement Fund • State-wide Tax on Tobacco • Other QRISs also reported as an input the **overall funding amount** for the most recent fiscal year.
Administrative Structure	QRISs reported that the **lead agency** was one of the following: • **State agency**, such as the Department of Human Services or Department of Education • **Local or county agency**, such as the Department of Children's Services • **Independent, nonprofit, or quasi-governmental agency**
Rating Structure and Number of Levels	QRISs reported using one of three **rating structures**: • **Building blocks**: In this structure, all of the standards in one level must be met before moving on to the next higher level • **Points**: In a points system, points are earned for each standard and are then added together • A **combination** of building blocks and points The QRISs **rating levels**: • Represent ranges of possible total scores • Provide the steps for programs/providers to achieve in the QRIS
Standards	The categories of **quality standards** or **indicators** vary, but they include the following: • Licensing compliance • Environment • Staff qualifications • Family partnership • Administration and management • Accreditation • Curriculum • Ratio and group size • Child assessment • Health and safety • Cultural and linguistic diversity • Provisions for children with special needs • Community involvement
Collaboration with Other Agencies	QRISs reported **partnering** with: • Universities, to conduct observations and provide technical assistance or quality improvement services • A community college that provided technical assistance or quality improvement services, managed communication and information dissemination, distributed financial incentives, collected/validated information to assign the rating, provided support in navigating the QRIS system, and conducted observations • Resource and referral agencies, that mostly managed communication and information dissemination, and provided technical assistance or quality improvement services, support in navigating the QRIS system, and financial incentives • Another state agency • A nonprofit organization

Source: Compendium of QRS and Evaluations (Tout, Starr, Soli, Moodie, Kirby, and Boller 2010) and QRIS Profiles produced as part of the QRS Assessment project (U.S. Department of Health and Human Services 2010a).

Note: Information on inputs related to funding was obtained from the QRIS profiles in http://www.acf.hhs.gov/programs/opre/cc/childcare_quality/. Information on the other inputs was obtained from the *Compendium of Quality Rating Systems and Evaluations* as follows: (1) administrative structure, p. 24; (2) rating structure and number of levels, pp. 28-29; (3) standards, pp.45-45; and (4) collaboration with other agencies, pp. 26-27.

Worksheet 2 presents a list of examples of QRIS inputs and provides a space to list the inputs of your QRIS. Keep in mind that the inputs box in the logic model for your QRIS does not need to include a detailed description of all the elements included in Table II.3 and Worksheet 2. The inputs description in the logic model should be parsimonious and clear, and detailed descriptions of the inputs can be included in separate documentation (such as work plans and other supporting documents for your QRIS).

Step 3: Indicate the activities that will be implemented for the QRIS

The activities refer to *what* the QRIS will do to improve the quality of child care programs/providers. That includes the actions, processes, techniques, events, tools, technology, and services that will be implemented or provided as part of the QRIS. The activities specified in the logic model should include the primary actions that will be undertaken in coordination with each of the key stakeholders.

Worksheet 2. Inputs of the QRIS

What Is This?

- The inputs included in the QRIS logic model reflect what is needed for the QRIS to reach its goals.
- Inputs include:
  - Existing legislation and political support
  - Funding
  - Planning, administrative structure, and staffing
  - Collaboration across agencies
  - Use of research evidence to inform the QRIS actions.

QRIS-Specific Examples

Funding

- Child Care Development Fund
- State funds
- Local or county funds
- Foundations
- Other

Administrative Structure

- Lead Agency:
  - State agency: Department of Human Services or Department of Education
  - Local or county agency: Department of Children's Services
  - Independent, nonprofit, or quasi-governmental agency

Rating Structure and Levels

- Rating Structures
  - Building blocks
  - Points
  - A combination of building blocks and points
- Rating Levels
  - Represent ranges of possible total scores and provide steps for providers to achieve in the QRIS

Standards

- Licensing compliance, environment, staff qualifications, family partnership, administration and management, accreditation, curriculum, ratio and group size, child assessment, health and safety, cultural and linguistic diversity, provisions for children with special needs, community involvement

Collaboration with Other Agencies

- Partnerships with universities, community colleges, resource and referral agencies, other state agencies, and nonprofit organizations

Your QRIS

Funding

_____
_____
_____

Administrative Structure

_____
_____
_____

Rating Structure and Levels

_____
_____
_____

Standards

_____
_____
_____

Collaborators

_____
_____

Table II.4 shows examples of activities reported by existing QRISs and Worksheet 3 presents examples of QRIS activities and provides a space to list the activities of your QRIS. The QRIS logic model would include a short statement about what each planned activity entails. Below, we describe various types of QRIS activities:

Type of Activity

- **Application and entry**: Includes any QRIS actions that facilitate application to and entry into the QRIS for child care providers, such as providing orientation sessions to providers before enrollment and offering guidance on preparing for receiving ratings.
- **Monitoring/accountability**: Conducting quality assessments of child care programs/providers that want to participate or are already participating in the QRIS. This activity includes (1) assessments based on observations, reviewing documents, and making reports; (2) the frequency of the assessments; (3) the methods used to ensure the integrity of the assessment process; and (4) the validation of each rating component and of the overall rating.
- **Quality improvement**: Staff and/or financial resources provided by the QRIS to assist participating programs with quality improvement efforts, for example, making self-assessment tools available, providing training and on-site assistance, or providing funds to make specific quality improvements. The QRIS can also offer links to other organizations that provide services to assist with quality improvement.
- **Financial incentives to improve quality of child care providers**: Examples include (1) bonuses or awards that are given based on the rating achieved, the number of subsidized children served, and other factors; (2) tiered reimbursement that provides higher maximum reimbursement rates for each subsidized child in the program; and (3) staff-focused incentives such as scholarships, wage enhancements, or retention bonuses (Tout et al., 2009; Tout, Starr, Soli, Moodie, Kirby, and Boller 2010).
- **Methods of disseminating information to providers and parents**: Creating a QRIS website; publishing ratings in community/local newspapers; conducting meetings with parents and child care providers.

Step 4: Indicate the outputs of the QRIS activities

The outputs are the direct results of the QRIS activities and are measured in terms of the amount of work accomplished. Therefore, the activities detailed in Step 3 should be carefully examined to identify the measurable results (outputs) that are expected to be accomplished through each activity. Thus, outputs can be used to track and monitor implementation and they

Table II.4. Examples of Activities from Existing QRISs

Type of Activity	Examples
Application and Entry	QRISs use a variety of **strategies** to facilitate the **application process**, such as offering/requiring: • A preparatory process for providers before they receive a rating • A period for programs to receive a time-limited "pre" rating or a commitment to enter the QRIS at a later time • An orientation session for the QRIS without making it a requirement for enrollment in the QRIS
Monitoring/Accountability: Ratings and Assessments	QRISs conduct **assessments** of the quality of child care programs/providers using observational measures. For example, QRISs could use: • The Environment Rating Scales (ERS; Harms et al., 2005) to assess features of the learning environment such as the materials, activities, routines, provisions for health and safety, and interactions that influence children's experiences in the setting; or • The Classroom Assessment and Scoring System (CLASS, Pianta et al., 2008) to assess features of the learning environment and the quality of teacher-child interactions.
Standards	QRISs periodically (for example, every two years) review and update standards to reflect the current research evidence regarding how the quality of child care programs is related to outcomes for children, parents, providers, and the community
Quality Improvement	**Trainings** linked to or aligned with QRISs **that are designed to improve program quality.** The frequency, length and duration of assistance vary depending on the needs of the program seeking QI services. Trainings or coaching for staff in child care programs/providers can be on the following topics: • The child care environment • Language and literacy • Specific curriculum • Business practices • Safety • Social and emotional development • Infant/toddler care • Adult-child relationships • Developmental screenings • Inclusion • Specific topics related to the Program Administration Assessment and the state's Early Learning Guidelines QRISs have **on-site assistance** available to child care providers for quality improvement**,** including: • Supporting programs with navigation of the QRIS (that is, assisting with filling out paperwork, explaining the rating process) • Implementation of a developmental screening tool • Training on early learning guidelines • Infant/toddler information • Staff training, and technical assistance on classroom layout • Coaching or mentoring to directors and/or classroom staff on how to meet goals included in the quality improvement plan
Financial Incentives	QRISs use one or more of the following **financial incentives**: • Improvement awards: The award amount can be provided (1) in a matrix, with amounts differing by quality level, type or size of program, and the density of at-risk children served, or (2) as a standard amount or an upper threshold for grant amounts. The grant amount can be left unspecified but should indicate that it will align with items in the program's quality improvement plan. • Tiered reimbursement: QRISs can provide a rate matrix showing the rates programs are eligible to receive at different star levels for serving

Type of Activity	Examples
	subsidized children. The tiered bonus may be offered only to accredited programs, or a flat rate increase per subsidized child based on the star level may be offered to all programs. The density of subsidized children in the program may be included as a factor in the rates (with those serving more subsidized children eligible for higher rates) • Quality awards/bonuses: QRISs can offer a one-time merit or achievement payment upon receipt of the rating. These awards are generally modest in size (between $250 and $2,500), depending on the type of program and its star level. QRISs can also offer similar-size awards for achievement or maintenance of quality on an annual (or biennial) basis. • Awards based on density of at-risk children served: Awards offered to programs that serve higher densities of vulnerable or at-risk children. These awards provide a base rate for being at a particular quality level. In addition, the QRIS may offer a dollar amount per subsidized child served that is factored into an annual payment or an amount based on the density of vulnerable or at-risk children served (either 5-25 percent density or 26 percent and above). • Other: These can include scholarships, wage enhancements, and retention bonuses. Such incentives are directed toward individual staff members, to help them increase their educational attainment or provide incentives for staying in their workplace. They typically are available to all practitioners in the state/municipality, not just those in programs that participate in the QRIS. QRISs offer access to scholarships and, in some cases, wage enhancements or retention bonuses.
Dissemination of Information	QRISs use various **methods of outreach** to providers, the public, and parents including: • A website, the most common method of outreach to parents • Dissemination of written materials by the QRIS contractors/partners • Mailings of QRIS-related information to parents • Posting information in doctors' offices or other public venues • QRISs may provide information in languages other than English or provide assistance to non-English speaking parents

Source: Compendium of QRS and Evaluations (Tout, Starr, Soli, Moodie, Kirby, and Boller 2010) and QRIS Profiles produced as part of the QRS Assessment project (U.S. Department of Health and Human Services 2010a).

Note: The information on QRIS activities was obtained from the *Compendium of Quality Rating Systems and Evaluations* as follows: (1) application and entry, pp. 30-41; (2) ratings, p. 159; (3) assessments, pp. 161-164; (4) quality improvement, pp. 171-176; (5) financial incentives, pp. 177-186; and (6) dissemination of information, pp. 187.

Worksheet 3. Activities of the QRIS

What Is This?

- The activities refer to what the QRIS will do to improve the quality of child care providers. They include the actions, processes, techniques, events, tools, technology, and services that will be implemented or provided as part of the QRIS.
- The activities specified in the logic model should also include the actions that will be undertaken in coordination with each of the key stakeholders.

QRIS-Specific Examples

Application and Entry

- Strategies to facilitate the application process

Monitoring/Accountability: Ratings and Assessments

- Using an observational measure to conduct assessments of providers. For instance, using:
  - The Environment Rating Scales (ERS; Harms et al., 2005)
  - The Classroom Assessment and Scoring System (CLASS, Pianta et al., 2008)
  - Other measures of the quality of the learning environment

Standards

- Reviewing and updating standards to reflect current research evidence on how the quality of child care programs is related to outcomes for children, parents, providers, and the community

Quality Improvement

- Trainings/coaching designed to improve quality and that are provided to staff in child care programs/providers
- On-site assistance to child care providers

Financial Incentives

- Improvement awards
- Tiered reimbursement
- Quality awards/bonuses
- Awards based on density of at-risk children served
- Other

Dissemination of Information

- Create and operate QRIS website
- Disseminate written materials
- Mail QRIS-related information to parents
- Post QRIS information in public venues

Your QRIS

Application and Entry

_____
_____

Monitoring/Accountability: Ratings and Assessments

_____
_____
_____

Standards

_____
_____

Quality Improvement

_____
_____
_____

Financial Incentives

_____
_____

Dissemination of Information

_____
_____
_____

indicate whether the intended targets of the QRIS intervention were reached at the intended dose. Comparisons of expected outputs with actual performance over time can inform monitoring and continuous program improvement efforts. If measured outputs are not meeting targets, either the target can be adjusted or the program design can be modified (Tout et al., 2009).

Table II.5 shows the outputs reported by one QRIS (Delaware Stars for Early Success, Phase 1) and Worksheet 4 presents examples of QRIS activities and outputs and provides a space to list the activities and outputs of your QRIS. Below, we describe examples of QRIS-specific outputs that can be entered in the logic model:

Examples of QRIS Outputs

- **Monitoring/accountability**: Outputs of QRIS monitoring/accountability activities might include the number of child care program dimensions assessed and the number of child care providers that received ratings.
- **Standards**: A process to review and update the QRIS standards is completed every two years to ensure that the standards reflect the current research evidence of how quality in child care programs is related to outcomes for children, families, providers, and the community.
- **Quality improvement**: Examples include (1) the number of QRIS staff hours spent providing assistance with quality improvement efforts, (2) the number of child care providers that receive assistance, and (3) the degree (frequency, length of sessions, and duration) of assistance received (overall and for different types of assistance).
- **Financial incentives to improve quality of child care providers**: Examples include (1) the amount of money given for each type of bonus, award, grant, or other incentive that the QRIS offers; (2) the number of child care providers receiving each type of incentive; and (3) the percentage of the total number of child care programs assessed that earned a reward for improving quality or meeting a standard.
- **Methods of disseminating information to providers and parents**: Examples include (1) the number of families accessing the QRIS website, (2) the number of inquiries submitted to the QRIS (in person or electronically), (3) the number of inquiries submitted to local Child Care Resource and Referral (CCR&R) organizations, and (4) the number of parents and child care providers attending meetings organized by the QRIS.

Table II.5. Outputs from Delaware Stars for Early Success, Pilot Phase 1

Monitoring/Accountability	Quality Improvement	Financial Incentives	Dissemination of Information
Four quality standards or indicators were assessed in the Pilot Phase 1: 1. Qualifications & Professional Development 2. Learning Environment & Curriculum 3. Family & Community Partnerships 4. Management & Administration **15 child care programs,** serving 1,083 children, **participated** in Phase 1 of this QRIS pilot and **were rated**: • 8 small centers (6 or fewer classrooms) • 4 large centers • 2 family child care homes • 1 large family child care home	During 2007, Delaware contracted with **10 technical assistants** to **provide support** to participating child care programs. A total of **678.25 hours of technical assistance** were provided During Phase 1, the technical assistants provided Delaware Stars-required **training** to all participating child care programs	Delaware Stars, Pilot Phase 1 offered four types of awards and grants to participating programs in support of their QI efforts: 1. **Participation Award**: **Fourteen programs** received this award. 2. **Quality Improvement Grant**: **Thirteen child care programs** received QI grants and used these funds for education and training needed to meet qualifications, to purchase books and materials to support child assessment and curriculum planning standards, and to purchase materials for record-keeping. 3. **Professional Development Support Grant**: **Six programs** received these grants, which supported 28 staff. 4. **Merit Award**: **Seven programs** received these awards	Delaware Stars disseminates information to parents, providers/programs, and the public using the following methods: • **Parents:** Through the Delaware Stars **website** and a **brochure**. • **Providers/programs**: Through the Delaware Stars **website** and information sessions. QRIS contractors and/or partners disseminate written materials. • **Public**: Through the Delaware Stars **website**. Information on the number of parents, providers/programs, and others reached by these methods was not available.

Source: Delaware Department of Education. *Delaware Stars for Early Success*. Phase 1 Report, 2007. Accessed on October 26, 2010. [http://www.doe.k12.de.us/infosuites/students_family/earlychildhood/files/DE%20Stars%20for%20Early%20Success%202007%20report.pdf]; U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. *Delaware Stars for Early Success: QRIS Profile*. Washington, DC: Child Trends, Mathematica Policy Research, ACF, OPRE, 2010a. Profiles accessed on November 2, 2010 [http://www.acf.hhs.gov/programs/opre/cc/childcare_quality/].

Note: Delaware Stars is a five-level system (Star Level 1 to Star Level 5). The lowest level (Star Level 1) requires programs to meet licensing rules. Participation awards refer to one-time awards given upon completion and approval of a Quality Improvement Plan. Quality improvement grants refer to awards based on program type and star level. Professional development support grants were used to pay for Training for Early Care and Education (TECE) 1 and TECE 2. These 60-hour classes are pre-service requirements for teachers and/or assistant teachers in the new Delaware center licensing rules. Merit awards are one-time awards given to programs when they complete all requirements for the next higher Star level. The award amount is based on the type of program and the Star level achieved. Programs entering the Delaware Stars system at a Star Level 5 through national accreditation also receive this one-time award to recognize their high level of quality.

ITERS-R = Infant/Toddler Environment Rating Scale-Revised.

ECERS-R = Early Childhood Environment Rating Scale-Revised.

Worksheet 4. Activities and Outputs of the QRIS

What Are the "Activities"?	What Are the "Outputs"?
• The activities refer to what the QRIS will do to improve quality of child care providers. • The activities in the logic model should also include the actions that will be undertaken in coordination with each of the key stakeholders	• The outputs are the direct results of the QRIS activities and are measured in terms of the amount of work accomplished • Outputs can be used to track and monitor implementation and they indicate whether the intended targets of the QRIS intervention were reached at the intended dose.
QRIS-Specific Examples of Activities	**QRIS-Specific Examples of Outputs**
Monitoring/Accountability: Ratings and Assessments • Using an observational measure to conduct assessments of providers	**Monitoring and Accountability: Ratings and Assessments** • Number of quality dimensions assessed • Number of child care providers that received ratings • Number of child care providers (or percentage of all providers participating in the QRIS) at each rating level
Standards • Standards are reviewed and updated every two years to reflect current research evidence on how the quality of child care programs is related to outcomes for children, parents, providers, and the community	**Standards** • Process to review and update standards is completed every two years
Quality Improvement • Trainings and/or coaching for staff in child care programs • On-site assistance to child care programs participating in the QRIS	**Quality Improvement (QI)** • Number of QRIS staff hours spent providing assistance with QI efforts • Number of child care providers receiving QI assistance • Degree (frequency, length of sessions, and duration) of assistance received (overall and for different types of assistance)
Financial Incentives • Provide the required financial incentives (improvement awards, tiered incentives, etc.) to child care programs that obtain a rating above a certain level (for example, a level of 3 stars) • Award a certain amount of money (for example, $2,000 per child care program) to support staff training in quality improvement practices	**Financial Incentives** • Percentage of child care programs participating in the QRIS who met the standard and were awarded financial incentives • Number of child care providers who received awards to support staff training and amount of money awarded
Dissemination of Information • Create and operate QRIS website • Disseminate written materials • Mail QRIS-related information to parents • Post QRIS information in public venues	**Dissemination of Information** • Number of families accessing the QRIS website • Number of inquiries submitted to the QRIS • Number of parents and child care providers attending meetings organized by the QRIS
Activities of Your QRIS	**Outputs of Your QRIS**
Activity 1: _____ Activity 2: _____ Activity 3: _____ Activity 4: _____ Activity 5:	Output(s) of Activity 1: _____ Output(s) of Activity 2: _____ Output(s) of Activity 3: _____ Output(s) of Activity 4: _____ Output(s) of Activity 5:

Step 5. Define Desired Outcomes of the QRIS

The outcomes are the *results* that are expected from investing in and implementing the QRIS. Outcomes include the changes in behaviors, knowledge, skills, level of functioning, or attitudes that are expected to result from the QRIS activities and outputs.

In defining the outcomes of the QRIS, consider the following two issues:

1. **Setting Expectations that Align Activities and Outcomes.** Desired outcomes are more likely to be achieved when the activities and the outcomes are aligned so that there is a logical flow between the efforts of the QRIS and the changes that are expected to result from them. For instance, if one of the desired outcomes of the QRIS is that the practices of child care providers improve in ways that lead to better cognitive outcomes for children, then that goal would more likely be achieved if the QRIS implemented as many activities as possible to support improvement of providers' practices. Such activities could include combinations of the following: professional development, reduction of class sizes, providing additional classroom materials, and/or classroom management training. A QRIS that focuses on just one of these activities instead of a combination of them may make little progress in changing behaviors and, in turn, outcomes for children. For example, providing classroom materials might change the resources available and possibly the physical environment, but it may be unrealistic to expect changes in provider behavior or practices without additional activities such as participation of providers in professional development. In addition, if class sizes are not reduced and providers have large workloads, then it may be difficult for them to manage their classrooms/care settings effectively.
2. **Setting Expectations About the Timing of Change.** Outcomes are not immediate and some of them take place in the long term. In addition, some outcomes materialize only after others have occurred. For example, knowledge of practices related to quality improvement can be readily acquired in the short term, but changes in behavior and skills based on the new knowledge can take longer. Another example is that parents can learn about the ratings in the short term, but it may take time for them to incorporate this knowledge into their decisions about child care providers.

In a logic model, the outcomes are entered after the outputs. The logic model can also specify what will change at each level targeted by the QRIS (for example, the early care and education system, the market and community, child care providers, parents, and children). Examples of the outcomes of QRIS initiatives are presented in Table II.6 and Worksheet 5. A space to list outcomes of your QRIS is also provided in Worksheet 5.

Table II.6. Examples of QRIS Outcomes

Short-Term Outcomes	Medium-Term Outcomes	Long-Term Outcomes
• Parents increase their knowledge about rating levels and quality of child care programs/providers • Programs/providers participating in the QRIS increase their knowledge about the standards for quality • Programs/providers focus increased effort on quality improvement (reflecting changes in the level of functioning and attitudes about quality care) • Data obtained through the QRIS are routinely reviewed and used by stakeholders at all levels in the early care and education system	• Parents use the ratings and guidance provided by the QRIS and CCR&Rs to inform their search for child care providers • Increased involvement and communication by parents with their child care program/provider • Child care programs/providers continue making changes focused on quality improvement and receive incentives for having reached higher rating levels • Programs/providers increase commitment to their work • Provider-child interaction quality improves	• Quality of participating child care programs/providers improves overall • Decreased staff turnover in child care programs/providers • The overall supply and access to quality early care and education programs increase • Parents select high-quality child care programs/providers • Children receive higher-quality care • Access to high-quality care results in improved cognitive and social-emotional outcomes for children • Reduced parenting stress • Cooperation across agencies in the system is codified

CCR&Rs = Child Care Resource and Referral organizations.

Worksheet 5. QRIS Outcomes

What Is This?

- The outcomes are the results that are expected from investing in and implementing the QRIS, and include the changes in behaviors, knowledge, skills, level of functioning, or attitudes that are expected to result from the QRIS activities

QRIS-Specific Examples

What outcomes are expected for child care providers/programs?

In the short term, child care providers focus increased effort on quality improvement, reflecting changes in the level of functioning and attitudes about quality care

In the medium term, child care programs/providers continue making changes focused on quality improvement and they receive incentives for having reached higher rating levels

In the long term, quality of participating child care programs/providers improves

What outcomes are expected for parents?

- In the short term, parents learn about using ratings to select higher-quality care
- In the medium term, parents use the ratings and guidance provided by the QRIS and resource and referral agencies to inform their search for child care providers
- In the long term, parents identify providers offering high-quality care and choose these providers over others perceived as providing lower-quality care

What outcomes are expected for children?

- In the long term, children's cognitive and social-emotional outcomes improve

What outcomes are expected for the community?

- In the short term, data obtained through the QRIS are routinely reviewed and used by stakeholders at all levels in the early care and education system
- In the long term, the overall supply and access to quality early care and education programs in the community increase

Your QRIS

Outcomes expected for child care providers/programs

_____
_____
_____
_____
_____

Outcomes expected for parents

_____
_____
_____
_____
_____

Outcomes expected for children

_____
_____
_____
_____
_____

Outcomes expected for the community

_____
_____
_____
_____

Step 6: Identify External Influences

Funding opportunities and the QRIS activities, outputs, and outcomes are influenced by supports from the government, funding agencies, and parents (Zellman and Perlman, 2008). For instance, changes in the political and economic environment can affect support for and funding of QRISs. Changes in state or federal policies or reporting requirements might require shifting resources among inputs and/or activities (for example, lower caps for monetary quality improvement incentives might lead QRISs to provide incentives in other forms, such as providing classroom equipment or facilities improvement assistance). In addition, a change in licensing standards might necessitate a change in QRIS standards or rating levels and the movement of children into and out of care settings affects the amount of their exposure to certain levels of quality care. Although it is not possible to document in the logic model all possible external influences on the QRIS inputs, activities, outputs, and outcomes, keep in mind during the QRIS planning process that the social, physical, political, and institutional environments can influence the components of the QRIS. Doing that can help stakeholders realize how the QRIS component and their expectations of its outcomes might need to adjust to a changing environment.

E. Updating the Logic Model

Completing the logic model is an important milestone, and it is the first step in the process of monitoring the performance of the QRIS and continually updating and improving its activities and outputs (Westmoreland et al., 2009). Logic models can facilitate communication among key stakeholders to help identify the need for any changes or improvements to make the QRIS a more effective quality improvement tool.

Logic models can be adapted to reflect any changes in the QRIS' context, priorities, and focus. Key stakeholders should monitor changes to the logic model to ensure that it remains relevant to the QRIS' goals.

F. Using the Logic Model as a Basis for Evaluating the QRIS

Logic models can be useful tools for developing an evaluation strategy because they provide a formal, written articulation of specific assumptions, activities, and expected outcomes of an intervention. For evaluations, they can inform stakeholders' decisions regarding the following features of the QRIS:

- **Specification of the research questions.** The logic model outlines the resources (inputs), activities, and outcomes of the QRIS, and documents its goals and the measurable indicators of progress toward achieving those goals. The evaluation may investigate how resources were used, activities were implemented, and whether the desired outcomes were achieved.
- **Methods to be used in the evaluation.** Since the logic model provides information on the stage of development of the QRIS and the expectations for it at that stage, it can inform the decision on the type of evaluation needed. Implementation evaluations and outcome evaluations can be done for QRISs.
- **Measurement strategies.** The logic model helps focus on (1) the information that the evaluators need to know, (2) when to collect the information, (3) best sources of the information, and (4) the most efficient way to collect the information.

In addition, since a QRIS logic model can show the pathways to outcomes for communities, markets, child care programs/providers, parents, and children, it can be a helpful tool for guiding the evaluation of the QRIS at each of those levels, and it can be used to establish feedback links between the evaluation findings and decisions about the QRIS inputs and activities (Tout et al., 2009). Finally, keep in mind that changing program design and other circumstances (for example, an economic crisis that threatens state-level investment in the QRIS) may require updating evaluation strategies and design. Evaluators must revisit the logic model over time to update the evaluation focus and measures as needed to align with program and contextual changes (Zellman et al., 2011).

QRIS Logic Model and the Type of Evaluation

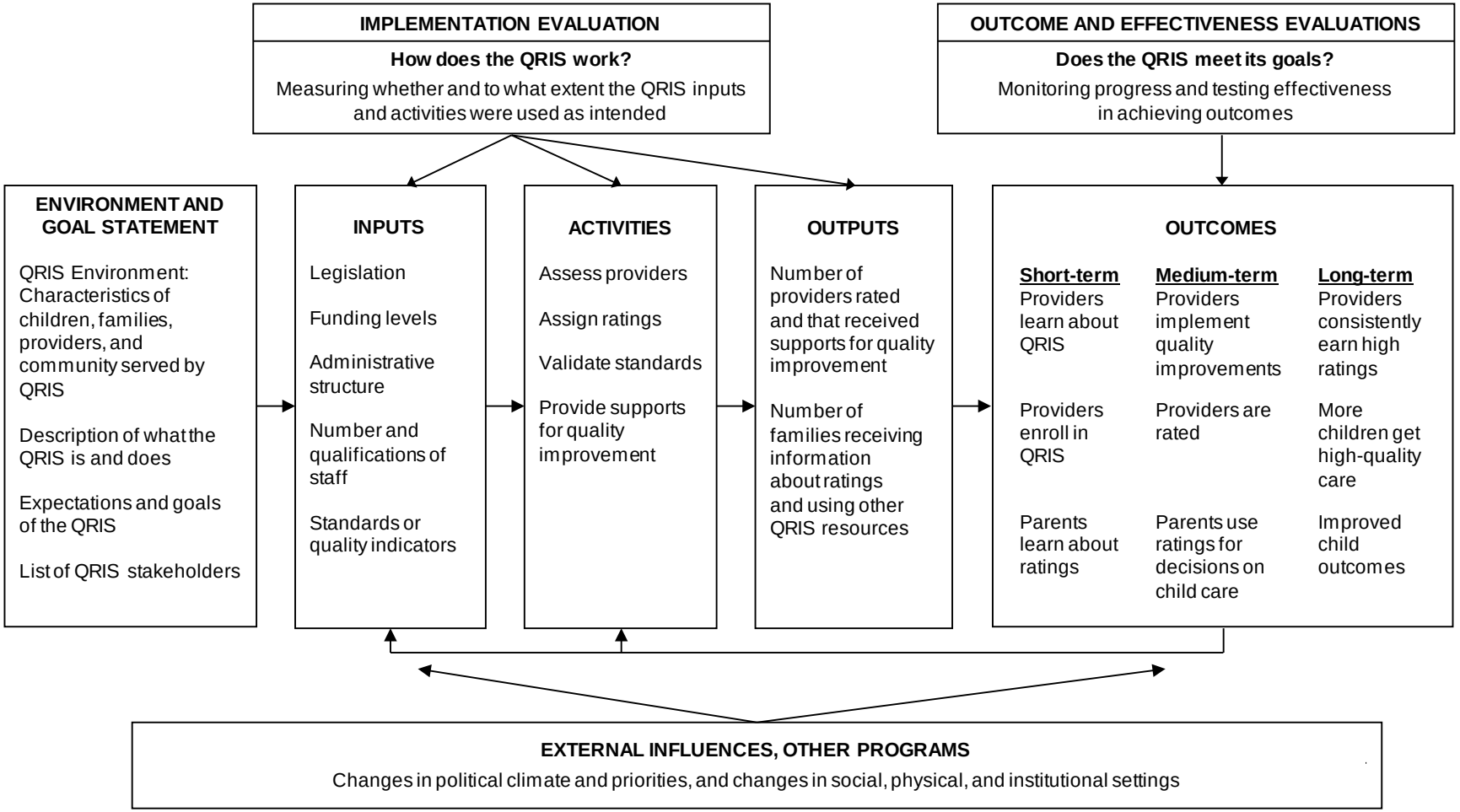
The elements of the QRIS logic model can be linked to different types of evaluations. Figure II.4 shows the relationships between the purpose and type of evaluation and the elements of the logic model. In this figure, implementation evaluations[4] point to the QRIS inputs and activities, as these evaluations focus on *how* the QRIS works, and outcome evaluations point to the QRIS outcomes, because these evaluations focus on the goals of the QRIS. Chapter III provides a detailed discussion about how to develop research questions and select the appropriate evaluation design to answer them. This section explains the relationship between the logic model and the type of evaluation most suited to assessing progress in implementation and achieving outcomes.

Implementation evaluations assess whether and to what extent the QRIS' inputs and activities (the elements in the left side [Figures II.2 and II.4] or bottom [Figure II.3] of a logic model) were used and implemented as intended. That is, this type of evaluation determines the fidelity of implementation. The results from implementation evaluations are essential for continuous program improvement as they can inform decisions about whether and how to change or adapt resources (inputs) and activities in ways that could improve the QRIS functions. Implementation evaluations might examine whether:

- The QRIS inputs were used as planned, in terms of amount, quality, and timing
- The QRIS activities were conducted as planned, in terms of timing, content, format, and quality
- The rate of enrollment of child care programs/providers in the QRIS met program goals
- Child care program/provider participation occurred as intended
- Quality levels were validated
- The QRIS participants were satisfied with the activities and information (including technical assistance and supports to parents) offered by the QRIS

[4]Process evaluation is another term used to describe an evaluation of how an intervention/quality improvement activity works. Throughout we use the term implementation evaluation.

Figure II.4. Using the Logic Model as a Basis for Evaluation of the QRIS



Source: Adapted from McCawley, 2001; Tout et al., 2009; Westmoreland et al., 2009; W. K. Kellogg Foundation, 2004b.

Outcome evaluations measure progress toward meeting the QRIS' desired outcomes (the right side [Figures II.2 and II.4] or top [Figure II.3] of the logic model). The level of rigor in the design of these evaluations determines whether they monitor progress (referred to throughout as an outcome evaluation) or assess effectiveness (referred to throughout as an effectiveness evaluation). An evaluation may focus on monitoring progress in achieving outcomes (for example, the benchmarks set by the state regarding the proportion of children assessed by teachers as ready for kindergarten) in order to assess whether the expected outcomes are reasonable and headed in the right direction given the resources and activities that have been invested and implemented in the QRIS. However, an outcome evaluation cannot necessarily attribute changes to the effects of QRIS. An outcome evaluation might measure whether:[5]

- Child care programs/providers receiving more quality improvement supports (for example, trainings, coaching, and financial incentives) from the QRIS achieve higher ratings than those that receiving fewer supports
- Higher ratings are associated with better outcomes for children
- Program ratings improve over time
- Over time, parents eligible for child care subsidies are more likely to choose care that is rated as higher in quality.
- QRIS supports are associated with less staff turnover and higher observed quality

The ultimate question for evaluation is whether the QRIS or QI initiative is effective in achieving the expected outcomes. Assessing the effectiveness of the QRIS as a whole or of specific activities that are part of it requires an evaluation that can determine causal effects through an experimental design. Effectiveness evaluations can help in answering questions such as:

- What is the impact of the QRIS on improvements in the level of quality provided by child care programs?

[5]The relevance of these outcome evaluation topics depends on the specific logic model elements and research questions policy makers and program operators plan to address through the evaluation.

- Are the ratings achieved by providers participating in the QRIS higher than they would have been in the absence of QRIS?
- Do parents with access to information about the QRIS ratings make decisions different than those they would have made without access?
- Do children enrolled in child care programs participating in the QRIS demonstrate better school readiness than they would have if enrolled in programs not participating in the QRIS?

The time frame needed to produce evaluation findings increases as the focus moves from left to right in the logic model (Figure II.4). Process evaluations can be relatively short term, depending on the purposes for which the information will be used. Outcome and effectiveness evaluations need a longer time frame, depending upon the theory about how long it may take to produce changes in outcomes. Issues pertaining to timing and evaluation design are discussed further in Chapter III.

III. DESIGNING THE EVALUATION OF YOUR QUALITY RATING AND IMPROVEMENT SYSTEM (QRIS)

This chapter presents guidelines for designing evaluations of Quality Rating and Improvement Systems (QRISs). Links to external resources for designing evaluations of QRISs and of other quality improvement (QI) interventions that are not QRIS specific, and links to published evaluation reports of existing QRISs, are presented in Appendix A.

CHAPTER CONTENTS

A. Where to Start?

The first step in designing an evaluation of a QRIS is to understand the objectives of the QRIS and how it operates. Understanding how the QRIS operates involves knowing the stakeholders involved in the QRIS and their expectations; the resources invested in the QRIS; the QRIS elements (for example, processes, standards, QI efforts, incentives, and dissemination efforts); and the expected outcomes for participants (child care providers, children, and families) that are the targets of the QRIS efforts. Developing a well-articulated logic model for the QRIS is a recommended strategy because the logic model provides a description of the initiative's goals, stakeholders, resources (inputs), activities, and the expected outputs and outcomes of those activities. (Chapter II provides detailed information about how to develop a QRIS logic model.) QRIS policymakers and developers can then use the QRIS logic model to guide the design and implementation of the evaluation (Zellman et al., 2011).

The next step in designing the evaluation is to determine what information is needed. This means identifying the research questions (Section B provides more information on identifying research questions) for the evaluation. After identifying the key research questions, subsequent planning steps can be addressed, including the following:

1. Determining whether it is possible to address the research questions given the available funding for the evaluation and the time frame of information needs
2. Identifying the types of analyses and the level of statistical precision required to answer the research questions
3. Selecting the overall research design that best answers the research questions and can be implemented in the selected settings
4. Selecting measures that align with the research questions
5. Developing a data collection plan
6. Conducting a credible and adequate evaluation
7. Developing a plan to communicate the findings of the evaluation

More information on each of these steps is provided in the rest of this chapter; Table III.1 includes a list of the sections in the Toolkit that can help you address the specific needs of your QRIS evaluation.

Finally, another important step is to consider the ethical issues involved in the QRIS evaluation. Evaluators of QRIS must take into account their responsibilities toward participants in the evaluation by making sure that the evaluation is respectful of and relevant to the participants' circumstances, and that the instruments and data collection methods used in the evaluation are culturally sensitive and appropriate for the participants (U.S. Department of Health and Human Services, 2010c). In addition, as in any research initiative involving human subjects, participants in the QRIS evaluation must be informed that they are part of the evaluation and the evaluators will undertake any necessary steps to safeguard the confidentiality of the information that participants provide to the evaluation.

Table III.1. Sections of the Toolkit That Can Help You Design Your QRIS Evaluation

Questions on the Evaluation of Your QRIS	Section of The Toolkit Containing Information That Can Help Address These Questions
• What are the goals of your QRIS? • Who are the stakeholders involved in the initiative and what are their expectations for the QRIS? • What are the characteristics of the environment in which the QRIS operates? • What resources have been invested in the QRIS? • What are the components of the program (monitoring/accountability, standards, QI efforts, financial incentives, and dissemination efforts)? • What outcomes (for child care programs/providers, children, families, and the community in general) are expected to be influenced by the efforts of your QRIS?	Chapter II: How Does Your QRIS Work?
• What are the key research questions to be answered by the evaluation? • Are the research questions appropriate given the stage of development of your QRIS?	Section III.B: Identifying Research Questions
• Can the research questions be answered with the available funding for the evaluation? • What is the timing of your information needs? • How much time will be needed/is available to answer the evaluation research questions? • What headline would you want for an article based on the results of the evaluation? • Do you want or need to make causal statements (for example, QRIS caused changes in quality and child outcomes)? • What level of rigor is needed to answer the research questions? • What types of analyses are needed to answer the research questions? That is, what evaluation design is appropriate to answer the research questions? • What methods (quantitative, qualitative, a combination) will be used in the evaluation?	Section III.C: Selecting the Evaluation Design
• What sample design and size are required/are feasible? • What data collection methods should be used? • Who should be interviewed? • How will data be managed?	Section III.D: Selecting Data Collection Methods and Developing a Data Collection Plan Chapter IV: Choosing Measures and Data Collection Best Practices
• How can the evaluation analyses be planned and conducted effectively to answer the research questions?	Section III.E: Considerations for an Adequate and Credible QRIS Evaluation
• Who are the audiences for the products of the evaluation? • How will the findings be communicated to each audience?	Section III.F: Communicating the Findings of the Evaluation
• What are the funding sources and costs of the QRIS evaluation? • Are the sources of funding for the QRIS and the research stable over the expected life of the evaluation?	Chapter V: Evaluation Costs and Funding Sources
• Who should conduct the evaluation? • What level of independence from the QRIS designers and operators is needed to ensure credibility of the results?	Chapter VI: Guidelines for Selecting and Working With an Evaluator

B. Identifying Research Questions

Research questions are the foundation of any evaluation because they identify the goals of the evaluation and narrow the scope to the most important topics to be addressed. The research questions are building blocks for the evaluation and therefore play a key role in determining how the evaluation will be designed and what needs to be measured to achieve the objectives of the evaluation.

The decision about which research questions to pursue in an evaluation should come early in the planning stages of an evaluation and are driven by what the team designing the evaluation wants and needs to be able to report about the results when the evaluation is over. For example, if as described in Chapter II the goal of the evaluation is to guide, assess, or improve program implementation, the types of statements that can be made will be descriptive. If the goal is to make statements about causality—the QRIS changed the skill level of the workforce, the quality of care, or children's outcomes—an evaluation of the effectiveness of the initiative is required. The importance of refining the questions should not be overlooked, as they will guide the selection of the design and the selection and development of measures and other data collection instruments further along in the evaluation.

A QRIS evaluation might ask a number of different questions about the QRIS as a whole or about each component of the QRIS (monitoring/accountability, standards, QI efforts, financial incentives, and dissemination efforts). However, it is often not feasible to address every question in one evaluation or at the same time. It then becomes important to choose the questions that will provide the best answers for the purpose the evaluation aims to serve. In deciding which questions the QRIS evaluation should answer, it is important to seek a balance between the value of the knowledge that can be gained by answering specific research questions and the resources and capabilities available to collect and analyze the data needed to answer such questions. There is no science to finding that balance, but it helps to have realistic expectations about the availability of

resources for the evaluation such as funding and time, the accessibility of data sources, and the receptiveness of subjects (child care providers, families, and children) to participate in the evaluation. Mapping the research questions to the QRIS logic model (see Chapter II) can also help in the process of refining the list of research questions that the evaluation will pursue because the logic model provides a complete picture of the components of the QRIS and of how they relate to one another.

This section provides a list of examples of research questions that might be asked in evaluations of the QRIS as a whole or of a particular component. Although this list of research questions examples is comprehensive, it does not include every possible question that could be asked in a QRIS evaluation. The examples are organized by (1) the elements of the logic model (inputs, activities, outputs, and outcomes) to which the questions can be mapped and (2) the focus of the questions (Sections B.1 and B.2). As described in Chapter II, the focus of the research questions can be on the *implementation* of the QRIS—that is, on *how* the QRIS works—or on the QRIS *outcomes*—that is, on *whether* the QRIS meets its goals. Answers to questions focused on *how* the QRIS works are important to inform and improve the development and operations of the QRIS, and answers to questions focused on *whether* the QRIS meets its goals provide information on the extent to which the resources invested in the QRIS have resulted in the expected outcomes for the community, child care programs/providers, parents, and children. Those two types of research questions are not mutually exclusive, and both types of questions are necessary in gathering critical information for an evaluation. Examples of research questions that are appropriate for evaluations of QRISs at different stages of development are also presented (Section B.3).

1. Research Questions Focused on How the QRIS Works (Questions Focused on Implementation)

Questions about *how* the QRIS works address the need to know information about current conditions or conditions in a previous time period that can be observable or quantifiable; that is,

they provide information about what happened with the initiative and why. For example, research questions that address the implementation of the QRIS can focus on identifying the strengths in developing the QRIS and the barriers in implementing QRIS activities, assessing whether the available resources can sustain QRIS activities, determining the nature of the interactions between QRIS staff and families, measuring families' perceptions of the QRIS, and monitoring other stakeholders' experiences with the QRIS.

An evaluation focused on *implementation* of the QRIS might examine whether the following occurred:

- QRIS inputs were used as planned, in terms of amount, quality, and timing.
- QRIS activities were conducted as planned, in terms of timing, content, format, and quality.
- Enrollment of child-care providers for ratings and provider participation in the QRIS occurred as intended.
- Quality levels were validated.
- QRIS participants were satisfied with the activities and information (including technical assistance and supports to parents) offered by the QRIS.

Table III.2 presents a list of examples of research questions asked in evaluations focused on how the QRIS works, organized by the QRIS logic model element to which they can be mapped.

2. Research Questions Focused on Whether the QRIS Meets Its Goals (Questions Focused on QRIS Outcomes)

An evaluation focused on whether the QRIS meets its goals might center on *monitoring progress in achieving outcomes* to assess whether they are reasonable and moving in the right direction given the resources that have been invested in the QRIS. Alternatively, the evaluation might focus on *determining the effectiveness* of the QRIS as a whole or of specific activities that are part of the QRIS. An evaluation of the effectiveness of the QRIS can provide evidence on whether the investments in the QRIS are paying off and can help answer questions about what works for whom and in what circumstances and about how to improve program delivery and services.

Table III.2. Examples of Research Questions Focused on How the QRIS Works (Questions Focused on Implementation)

Logic Model Element	Examples of Research Questions
Inputs	• How were QRIS inputs used? • What "dimensions" of quality (licensing compliance, teacher-student ratio, group size, professional development, training, learning environment, curriculum, parent involvement, program management and operations, cultural and linguistic diversity) were included in the standards or indicators? • Do the dimensions of quality included in the standards or indicators provide a comprehensive picture of the child care programs' characteristics? • Are the standards evidence-based? • How many child care programs participated in existing licensing, accreditation, registries, and career lattice processes as part of the QRIS rating requirements? • How many staff in child care programs participated in early childhood education professional development as part of the QRIS program? • What were the qualifications of the QRIS staff? Are these qualifications appropriate for their roles in the QRIS? • What were the facilities, equipment, and technology used to provide QI assistance to child care programs? • Were resources (facilities, equipment, and technology) adequate in addressing the needs of QRIS participants? • What tools were used to communicate with parents? How often were these used? • Are state agencies coordinating with the QRIS lead agency/organization to implement, monitor, and share data about QRIS and other related information? • How many Child Care Resource and Referral (CCR&Rs) or other agencies worked with the QRIS to reach providers and parents?
Activities	• What types of assessments of child care programs were conducted (observations, reviewing documents or reports)? • What types of observational measures were used to assess providers? How often were the assessments conducted? • What processes were used to ensure the integrity of the assessment process (for example, training and certification of observers, conducting quality control (QC) observations)? • What resources were provided to assist child care programs with QI efforts (for example, availability of self-assessment tools, training, on-site assistance, funding for specific quality improvements)?
Outputs	• How many child care providers enrolled/volunteered for ratings? • How many enrolled providers were rated? • How do the number of rated child care providers compare with the number of providers expected to participate in the QRIS? • What was the rate of participation of child care providers in the trainings and on-site assistance provided by the QRIS? • What was the focus of the trainings and on-site assistance provided by the QRIS? • Were quality levels validated? • Are the QRIS rating components/levels related to measures of classroom quality, for example, the Environmental Rating Scales (ERSs)? • Are the QRIS rating components related to teacher qualifications (experience, education, and wages)? • Are the QRIS rating components related to measures of the quality of teacher-child interactions (for example, the Caregiver Interaction Scale, the Pre-Kindergarten Snapshot)? • What is the level of satisfaction with QRIS services reported by participants in training and on-site assistance offered by the QRIS? • To what extent do parents find the QRS website to be informative and user friendly?

The choice between the two alternatives depends on the overall goals of the evaluation and, as described earlier, the types of statements and conclusions the program funders and stakeholders want to be able to make about the findings. Selection of the level of rigor and statistical precision of the evaluation design is discussed in more detail in Section C.

Examples of research questions that might be asked in evaluations focused on whether the QRIS meets its goals are presented in Table III.3.

3. Research Questions Aligned With the Stage of Development of the QRIS

The QRIS evaluation questions should also focus on factors that are likely to experience change during the time frame of the evaluation. For example, an evaluation could look at whether parents are using ratings to make decisions about where to send their children soon after a QRIS begins publishing ratings, but it might be too early to look at whether the providers have improved their service quality as a result of participating in the QRIS.

An evaluation of a QRIS can be conducted at any stage of development of the QRIS, and in fact, the operation and activities of the QRIS can benefit from the findings of an evaluation at all of its stages of development. However, to conduct an informative and useful evaluation, the focus of the research questions should be consistent with the stage of implementation of the QRIS.

For the purposes of conducting an evaluation, a QRIS can be thought of as having the following three stages of implementation (Zellman et al., 2011):

1. The pilot and scale-up stage
2. The early operation stage (the first two to five years of operation of the QRIS after the pilot phase has been completed)
3. The mature operation stage (QRISs with more than five years of operation after the pilot phase has been completed)

Table III.3. Examples of Research Questions Focused on Whether the QRIS Meets Its Goals (Questions Focused on Outcomes)

Outcomes for	Examples of Questions Focused on Monitoring Progress Toward Meeting QRIS Goals
Child Care Providers	• Are different types of child care providers (center- and home-based) in different geographic areas being reached for participation in the QRIS? • Do child care programs/providers receiving QRIS QI achieve high ratings? • What are the characteristics of providers who increased their quality ratings? • Are high quality ratings for providers sustained over time? • Are the QRIS ratings reliable indictors of quality of child care settings? • Are child care centers or family child providers with lower scores closing? Are businesses with higher scores experiencing more demand?
Parents	• Is the number of parents receiving information from the QRIS on child care providers' quality ratings increasing? • What are the characteristics of the parents/families who receive information on provider's quality ratings? • Do parents who receive information on the QRIS quality ratings use it in selecting child care providers?
Children	• What are the characteristics of the children enrolled in child care programs in the areas/communities covered by the QRIS? • Is the number of children receiving care from providers that achieved high QRIS ratings increasing? • How do the characteristics of children enrolled in child care programs that achieved high QRIS ratings compare with those of children enrolled in child care programs that achieved lower QRIS quality ratings? How do they compare with the characteristics of children enrolled in child care programs not participating in the QRIS?
Outcomes for	**Examples of Questions Focused on Assessing the Effectiveness of the QRIS to Influence Outcomes**
Child Care Providers	• Does participation in the QRIS result in improvements in the level of quality provided by child care programs? • How does the quality of care provided by child care programs participating in the QRIS differ from the quality provided by programs not participating in the QRIS? • Do child care providers receiving QI assistance achieve higher ratings than those not receiving this assistance? • Does participation in the QRIS contribute to increasing the level of involvement in professional development activities for staff of child care programs/providers?
Parents	• Do parents feel the QRIS provides adequate access to information about the quality of child care providers in their community? • Are parents satisfied with the quality of care their children receive from providers participating in the QRIS? • Do parents choose higher quality settings than they would have in the absence of QRIS?
Children	• Does the QRIS create access to high quality care for more children? For children of parents receiving subsidized child care? • Has attending child care programs that participate in the QRIS (and at higher levels) resulted in improved learning and progress toward school readiness for all children? • Has attending child care programs that participate in the QRIS resulted in improved developmental outcomes for at-risk children?

Pilot and scale-up stage. States and municipalities commonly conduct evaluations of their efforts to implement the QRIS and to move from a relatively small-scale pilot to state-wide implementation because the pilot and scale-up stage is an adequate development stage at which to test the QRIS approach and also to assess the results of the activities of the QRIS (Zellman et al., 2011).

An implementation evaluation during this stage would be guided by the inputs, activities, and outputs included in the QRIS logic model so that it can help answer important questions about the operations and activities of the QRIS. Examples of research questions, by logic model component, are presented in Table III.4.

It is important to keep in mind that the pilot or scale-up stage might be too early to draw conclusions from the evaluation about the progress of the QRIS in meeting its goals and on whether the QRIS influences outcomes for child care providers, parents, and children. However, an evaluation conducted during the pilot or scale-up stage could provide important information about how the implementation of the QRIS can be improved so that the program can reach the level of intensity and quality that is more likely to lead to improved outcomes for children, families, child care providers, and the community (Zellman et al., 2011).

Early operation stage. The first years of operation of the QRIS after completion of the pilot stage provide an opportunity to conduct longitudinal evaluations of a larger scale than what can be accomplished during the pilot or scale-up phase. Evaluation conducted during the early operation stage might focus on (1) aspects of the QRIS that require changes in order to maintain effective operation of the program (implementation evaluation), (2) the process of assessing QI of child care providers/programs, and (3) the costs of operating the QRIS. During this stage, the evaluation can also focus on gathering information on child outcomes that could be used for comparison in later years—that is, used as baseline measures of child development

Table III.4. Examples of Research Questions for an Implementation Evaluation Conducted During the QRIS Pilot and Scale-Up Stage

Logic Model Element	Examples of Research Questions
Inputs	• Are the quality standards (indicators) of the QRIS based on current research evidence? • Do the QRIS rating levels actually reflect different levels of quality in child care programs? • How many child care providers/programs enrolled/volunteered for rating in the QRIS during the pilot and scale-up stage? • What is the program type and location/area serviced by the providers/programs that enrolled in the QRIS during the pilot and scale-up stage? • What are the characteristics of the communities served by the QRIS (family income level, racial/ethnic composition, home language, availability of early childhood education services)? • Who are the partners of the QRIS and what are their roles? • What do key stakeholders view as the major implementation challenges and successes during the pilot and scale-up stage? • How were the resources of the QRIS allocated (in administration, QI efforts [including staff], monitoring/assessing, incentives, and dissemination) during the pilot and scale-up stage?
Activities/Outputs	• What strategies were used to facilitate the application process of child care providers/programs to the QRIS? • What are the perceptions of the providers participating in the QRIS about the application and enrollment process? • How many child care providers/programs were rated during the pilot and scale-up stage? What ratings did they achieve? • What strategies (observation tools, analysis of administrative data, interviews with stakeholders) were used for assessing providers during the pilot and scale-up stage? • Were the QRIS ratings validated during the pilot and scale-up stage? • What were the supports offered to child care providers/programs to help them navigate the QRIS processes and requirements? • What were the supports offered to child care providers/programs for QI? • What are the perceptions of child care providers/programs about the supports they received during the pilot and scale-up stage? • Do the incentives offered in the pilot and scale-up stage cover the costs of QI efforts that are being implemented in this stage? • Do parents have access to the Internet? How many families are accessing the QRIS website? What do parents think about the QRIS website? • What are parents' perceptions about the support they receive from the QRIS about understanding and using the information provided by the ratings?

that can be compared with child outcomes measured in later years with the purpose of assessing a trend in child outcomes or gauging the effect of a specific aspect/component of the QRIS on child outcomes. Examples of research questions that could be included in evaluations conducted during the early operation phase are presented in Table III.5. The questions included in Table III.5 are examples of the types of questions that are most relevant to QRISs at the early operation stage. However, the examples of research questions for evaluations conducted during the pilot/scale-up stage (Table III.4) can be asked in an evaluation conducted during the early operation stage as well.

Table III.5. Examples of Research Questions for an Evaluation Conducted During the Early Operation Stage

Focus of Evaluation	Examples of Research Questions
Quality and Description of Program Operation	• Do child care providers/programs believe that the incentives provided by the QRIS compensate them for their efforts to reach the next rating level? • Do parents use the QRIS ratings to choose child care providers/programs? • Do child care providers/programs participating in the QRIS serve the children who are in the greatest need of high quality care?
Provider Assessment Process	• How were participants in the QRIS assessed (observation tools, analysis of curricula, interviews with stakeholders) during the early operation stage? • What strategies were used to ensure the integrity of the rating process? • What efforts were conducted to validate the standards/indicators during the early operation stage?
Costs of Operating the QRIS	• What are the financial and in-kind costs of operating the QRIS?
Trends in QI	• How has the quality of care provided by the programs participating in the QRIS changed during the early operation stage? • How long did child care providers participating in the QRIS take to move from one rating level to the next higher level? • What are the changes in the population of children and families served by the child care providers that move from lower to higher rating levels? How do these changes relate to the type of financial incentive (for example, improvement awards or tiered reimbursement) received by the providers?
Measurement of Initial Characteristics of Children	• What are the physical, cognitive, language, and/or social-emotional development outcomes of children receiving care by providers/programs participating in the QRIS?

Mature operation stage. Evaluations conducted during this stage can assess trends in QI and outcomes and can also gather data that are useful for improvement of the QRIS (Zellman et al., 2011). After a number of years of operation (for example, five or more years of operation) it is possible that several years of data on child care providers/program quality and child outcomes are available so that outcomes assessed before the implementation of the full-scale QRIS can be

compared with the same outcomes assessed after several years of QRIS implementation. These types of evaluations, however, require that data are collected at relevant points in time (for example, every year starting on the first year of QRIS implementation after the pilot phase has been completed); that the quality of the data is adequate; and that the evaluators have access to existing state data sets and other data collected as part of the QRIS QI assessments (Zellman et al., 2011).

Evaluations of mature QRISs can provide information on whether additional training and technical assistance resources should be provided to all or some of the child care providers/programs participating in the QRIS. For example, some providers might not show increases in quality over time or might stay at the same rating level for a number of years. Data collected over time about those providers can be helpful to identify why quality has not improved and to target additional supports most effectively. Additionally, at this stage of implementation alternative QRIS services and activities (for example, alternative financial incentives or alternative QI support services, such as providing guidelines/recommendations to child care providers regarding the physical environment of the child care setting [facilities, classroom materials], professional development activities/programs, or instructional strategies such as alternative curricula) could be tested strategically. Evaluations conducted in the mature operation stage can also focus on reassessment of the QRIS quality standards and approach to measuring providers' quality to ensure that they reflect up-to-date research findings and promote quality changes most linked with child outcomes (Zellman et al., 2011). Examples of research questions that could be included in evaluations conducted during the mature operation stage of the QRIS are presented in Table III.6. The questions included in Table III.6 are examples of the types of questions that are particularly relevant to QRISs at the mature operation stage. However, the examples of research questions for evaluations conducted at earlier stages (Tables III.4 and III.5) are also relevant to an evaluation conducted during the mature operation stage of the QRIS.

Table III.6. Examples of Research Questions for an Evaluation Conducted During the Mature Operation Stage

Focus of Evaluation	Examples of Research Questions
Trends in QI and Outcomes	• How have ratings of child care providers participating in the QRIS changed from the pilot or scale–up stage (or from the first time they were rated) to the mature operation stage (fifth year of operation, for example) of the QRIS? • What are the characteristics of the child care providers that show quality increases over time? • What are the characteristics of the child care providers that do not show quality increases over time? • Have the financial incentives been adequate to cover costs of QI incurred by child care providers/programs in the QRIS?? • How has enrollment changed over time in the child care programs participating in the QRIS? Are changes in enrollment in child care programs associated with changes in the quality ratings of the programs? • How has the quality of care provided for the most at-risk children changed over time? • Have centers and family child care providers with high ratings helped provide mentoring or give supports to those with lower ratings? • How have kindergarten readiness assessment results for children who received care from providers participating in the QRIS changed over time? Do children show better school readiness outcomes at the mature operation stage of the QRIS? • How many child care providers have enrolled in the QRIS since it started operating? What percentage of the total number of licensed child care providers in the state does that represent? • What is the percentage of families in the community who indicate they are choosing child care providers based on information they learned about the QRIS ratings?
Reassessment of Quality Standards	• At the mature QRIS operation stage (for example, after five years of full operation and beyond), do the quality standards on which QRIS ratings are based reflect recent research findings? • Do the QRIS ratings foster the types of changes in the quality of care provided that are most linked with improved child outcomes?
Testing Alternative Services and Activities	• How do quality ratings for child care providers receiving improvement awards compare with the ratings of providers receiving awards based on density of at-risk children served? • How do quality ratings for child care providers with staff using professional development program "X" compare with the ratings of providers with staff using professional development program "Y"? • How do outcomes of children in child care programs receiving training on early childhood learning curriculum "A" compare with the outcomes of children in programs receiving training on curriculum "B"?

Examples of QRIS evaluations of systems that have been in operation for different lengths of time. Of the 26 existing QRISs, 18 have conducted or are currently conducting an evaluation (Tout, Starr, Soli, Moodie, Kirby, and Boller 2010). Table III.7 lists the research questions used in the evaluations of three QRISs (Minnesota, Missouri, and Pennsylvania). These QRISs have been in operation for varying numbers of years. The focus of Minnesota's evaluation is both the QRIS implementation and outcomes, whereas the focus of the Missouri and Pennsylvania evaluations is the QRIS outcomes. Appendix A presents summaries of all reports the QRIS assessment team located as of January 2011 on the evaluations of existing QRISs.

Table III.7. Research Questions Used in the Evaluations of Three QRISs

QRIS Program	Research Questions	Year Launched
Minnesota Parent Aware	• The evaluation of the Parent Aware program includes an implementation study and an outcomes study, and it is focused on the following five research questions: 1. Does the supply of high quality early care and education programs increase in Parent Aware communities during the pilot? 2. Do parents in Parent Aware communities pay more attention to quality and change their early care and education choices during the pilot? 3. Are the high quality programs identified by Parent Aware promoting improvement in the school readiness of low-income children? 4. What changes, if any, should be made to the Parent Aware quality indicators, levels and rating process if the program is implemented statewide? 5. What components of program quality are the most critical supports for children's school readiness? What resources are needed to facilitate program improvements in these critical areas?	• 2007[a]
Missouri Quality Rating System (QRS)	• The focus of the Missouri Quality Rating System evaluation is the following question: Do preschool children who attend higher quality early childhood programs—as measured by Missouri QRS—show greater gains in school readiness than their peers who attend lower quality programs?	• 2003[b]
Pennsylvania Keystone STARS	• The evaluation of the Pennsylvania Keystone STARS includes an outcomes study. • The focus of the evaluation is the following question: Is Keystone Stars improving the quality of child care programs?	• Began as a pilot in 2002 • Launched statewide in 2004

Sources: Information on research questions: Minnesota: Tout, K., R. Starr, and J. Cleveland. *Evaluation of Parent Aware: Minnesota's Quality Rating System Pilot: Year 1 Evaluation Report*. Child Trends, December 2008. Accessed on October 26, 2010 [http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7B61078E31-3393-49B1-B301-D5CBC9A0DAC3%7D.PDF]; Tout, K., R. Starr, T. Isner, J. Cleveland, M. Soli, and K. Quinn. *Evaluation of Parent Aware: Minnesota's Quality Rating and Improvement System Plot: Year 2 Evaluation Report*. Child Trends, March 2010. Accessed on December 17, 2010 [http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BE0D4D742-A334-4305-BAF1-8432F580856D%7D.PDF]; Tout, K., R. Starr, T. Isner, J. Cleveland, M. Soli, and K. Quinn. *Evaluation of Parent Aware: Minnesota's Quality Rating and Improvement System Plot: Year 3 Evaluation Report*. Child Trends, November 2010. Accessed on December 17, 2010 [http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BB5ADD2AE-D080-4290-A698-A972B0A93B6A%7D.PDF]. Missouri: Thornburg, K. R., W. A. Mayfield, J. S. Hawks, and K. L. Fuger. "The Missouri Quality Rating System School Readiness Study." Columbia, MO: University of Missouri Center for Family Policy and Research, October 2009. Accessed on October 26, 2010 [http://mucenter.missouri.edu/MOQRSreport.pdf]. Pennsylvania: Barnard, W., W. E. Smith, R. Fiene, and K. Swanson. *Evaluation of Pennsylvania's Keystone STARS Quality Rating System in Child Care Settings*. University of Pittsburgh Office of Child Development, Pennsylvania State University Prevention Research Center, December 2006. Accessed on October 26, 2010 [http://www.pakeys.org/docs/Keystone%20STARS%20Evaluation.pdf]. Information on launch dates: Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G. & Boller, K. (2010, April). *The Child Care Quality Rating System (QRS) Assessment: Compendium of Quality Rating Systems and Evaluations*, OPRE Report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Accessed on July 30, 2010 [http://www.childcareresearch.org/childcare/resources/18554].

[a]This QRIS was launched as a pilot in 2007 and will continue until 2011. The program has not been launched statewide as of January 2011.

[b]As of October 2009, Missouri is not actively operating the QRIS pilot due to lack of funding.

C. Selecting the Evaluation Design

This section presents key factors to consider when selecting an evaluation design (that is, a method or approach to conduct the evaluation) and describes design options for evaluation of QRISs.

1. Key Factors in Selecting a QRIS Evaluation Design

Three key factors influence the selection of a design for the evaluation of the QRIS that is appropriate to answer the research questions of interest:

1. The available funding
2. The time frame in which the evaluation will be conducted and reports completed
3. The level of rigor and statistical precision that is required to answer the research questions

Available Funding. Establishing realistic expectations for what can be accomplished with a QRIS evaluation by selecting an evaluation design that is affordable given the amount of available funding is a necessary first step. The costs of different evaluation designs vary considerably depending on the sample size, the frequency of data collection, the number of data sources, and the types of analysis that the designs require.

Some evaluation designs might be more (or less) affordable than others depending on the level and type of data already being collected as part of the QRIS activities (Zellman et al., 2011) and whether the data can be managed and analyzed internally or externally at a reasonable cost. For example, if the activities of the QRIS include conducting observations of a large number of classrooms on a regular basis, surveying child care providers/programs in the QRIS regularly (for example, every year), and/or collecting data on child outcomes every year, then the evaluation design can potentially capitalize on these data to reduce the costs of data collection and thus reduce the costs of the QRIS evaluation. Similarly, if QRIS staff or other people already involved in QRIS activities (for example, a research consortium) possess the qualifications to conduct the analysis

required for evaluation, then the evaluation can capitalize on these resources to reduce the costs of data management and analysis.

Combining funds from different sources could enable states to afford a rigorous evaluation design that might otherwise be out of reach (Zellman et al., 2011). Funding for the evaluation can be secured from several sources, such as states and the federal government offering grants and contracts to fund QRIS evaluation, public-private collaborations established by states supporting QRIS activities (including evaluations), and foundations also providing funding for QRIS evaluations.

Time Frame to Conduct the QRS Evaluation. The timing of information needs must be considered when selecting the evaluation design, because the time required to complete an evaluation depends on how long it would take the evaluation team to obtain institutional review board (IRB) approval (if applicable),[6] the types of data that will be collected, and the analyses that will be conducted. Although state stakeholders and other policymakers might need information and data in the near term to guide decision making regarding the operations of the QRIS, many of the most rigorous evaluation designs require that data on QRIS impacts be collected longitudinally, which could take months or even years to complete (Zellman et al., 2011). QRIS developers and advocates, managers, state leaders, and other stakeholders should work together in establishing an evaluation time frame that is adequate to address both the most important research questions and necessary schedules for collecting and analyzing data. For example, if the most important questions to be answered with the evaluation are about QRIS effects on child care quality and children's school readiness, then the evaluation time frame should span a period long enough so that changes in providers' quality have time to materialize and child outcomes in preschool can be measured.

[6]Research studies on human subjects (even when based only on secondary or administrative data) usually must obtain IRB approval before the start of the investigation.

Level of Rigor and Statistical Precision in the QRIS Evaluation. The level of rigor and statistical precision needed to answer the research questions is a key factor in selecting the evaluation design because it determines the credibility of the findings and the conclusions that can be drawn from these findings.

The level of rigor in an evaluation refers to the accuracy and degree of bias of the conclusions that can be drawn from the findings of the evaluation. For example, implementation evaluations (which are focused on *how* the QRIS, or some if its components, works) that are the most thorough and systematic provide information about whether *all* the components of a QRIS are operating as intended; information that is crucial to identify problems in the way the program works should be corrected before the effectiveness of the QRIS is tested.

The most rigorous outcome evaluations (that is, focused on *whether* the QRIS meets its goals) allow for causal statements such as, "Child care programs' participation in the QRIS *significantly improved* the quality of care provided to children" and "Children attending child care programs that participated in the QRIS show *significantly improved* language and emotional development outcomes," to be drawn from the findings. Less rigorous evaluation designs enable statements such as, "Provider participation in the QRIS *is associated* with improvements in the quality of care provided to children" and "Attending child care programs that participated in the QRIS *is associated* with improvements in children's language and emotional development outcomes," from which a causal inference cannot be derived. The differences in those two types of designs might appear to be subtle; however, each provides a different level of scientific rigor to what can be said about the way the QRIS influences outcomes for children, families, child care providers, and the community. For the first, the effects of the QRIS (or a component of the QRIS) can be isolated and the changes can be attributed to the presence of the QRIS (or that particular component of the QRIS). For the second, the influence of the QRIS cannot be fully disentangled from other influences; an association is present, but a direct cause cannot be fully determined.

The level of statistical precision in an evaluation refers to the extent to which measurements could be repeated with little variation due to sampling. For example, if it were possible to include all the members of the population of interest (for example, if the universe is a state, then the population of interest could be all children eligible to receive child care services in the state or it could be all child care programs in the state) in the evaluation, then evaluators could be completely confident that the evaluation measurements reflect the true characteristics of the participants because all eligible participants (programs, families, children) would have been included in the evaluation. Usually it is not possible to base evaluations on entire populations of subjects. This means that evaluations must be conducted using a sample drawn from the population of interest. A larger sample size increases confidence that the evaluation findings truly represent what is happening in the full population. Therefore, an evaluation's findings are more precise when sample sizes are larger (the required sample size to detect effects is derived based on prior experience in studies of the targeted outcomes). Regardless of the level of statistical precision, there may be other sources of variation in measurement, such as differences in raters' judgments, reporting inconsistencies by survey respondents, or variations in children's performance on assessments due to fatigue or other immediate circumstances.

2. Approaches to Conduct Evaluation of QRISs.

In this section, we describe approaches (or designs) to conduct evaluation of QRISs. For each approach, we (1) discuss what can (and cannot) be learned from it, (2) provide examples of the research questions it can address, and (3) present key factors to consider when employing it.

Implementation Evaluation

Implementation evaluations can include one or a combination of the following types of studies: (a) a process study of the QRIS as a whole; (b) an in-depth study of an essential element in one QRIS, or across QRISs (also known as a comparative study); (c) a cost study; and (d) a validation study. Next, we describe each of these types of implementation studies.

a. Process Study of a QRIS as a Whole

What is this approach? This type of implementation evaluation assesses whether and to what extent the QRIS's inputs and activities were used and implemented, and whether this is in line with the original intent. That is, this type of evaluation examines how the QRIS works.

What can be learned with this approach? This approach to evaluate QRISs provides information on the plans to operate a QRIS and how these plans are actually carried out, the decision processes in which QRIS managers and staff take part as they conduct their work for the QRIS, and the characteristics of partners and other agencies that contribute to the activities of the QRIS. Therefore, conducting a process study of the QRIS can inform decisions about whether and how to change or adapt resources (inputs) and activities in ways that could improve the QRIS functions. This approach does not provide information about whether (or how) the QRIS changes outcomes for children, families, providers, or the community. That is, a process study can help explain why and how changes in outcomes might have occurred, but it cannot assert whether the QRIS has produced the outcome changes.

What research questions can be answered with this approach? The following are examples of questions that can be addressed with a process study (also, refer to Chapter III, Table III.2 for examples of implementation questions):

- Were QRIS inputs (resources) used as planned, in terms of amount, quality, and timing?
- Were QRIS activities conducted during the planned time frame? Were QRIS activities conducted using the planned content, format, and quality?
- What is the rate of enrollment of child care programs/providers in the QRIS? Was this in line with expectations about participation overall and by provider type and geographic area?
- How were the standards selected and the rating levels defined?
- Was the rating process systematic? What types of procedures and tools are used to determine whether and how each provider meets the specified elements for a particular level?

- What were the experiences of child care programs participating in the activities offered by and receiving information from (including technical assistance and supports to parents) the QRIS?
- How did child care programs use the financial incentives provided by the QRIS?
- What types of dissemination activities were pursued to build awareness among parents with children or the community at large about child care quality and/or the QRIS?

Many existing QRISs have conducted process studies. Most recently, for example, the evaluation of *Parent Aware*, Minnesota's QRIS (see Tout et al., 2008,[7] 2010 [March],[8] 2010 [November][9]), examined the program's activities to recruit and enroll providers and described the processes to rate programs and to provide financial incentives. Parents were also interviewed in this process study about their experiences in participating in the QRIS. The following are examples of the research questions asked in the Minnesota's QRIS process study:

- How many child care programs enrolled in *Parent Aware* during Year 1? What was the enrollment by type of program in Year 1?
- What factors facilitated the enrollment process in Year 1?
- What are the challenges to recruiting new programs after programs with initial interest have enrolled?
- What type of promotion and marketing activities did *Parent Aware* conduct in Year 1? What was the focus of these activities? What was the extent to which these activities reached their intended audiences?
- What is the process programs follow in order to receive ratings? What type of activities did *Parent Aware* conduct to assign ratings? What were the challenges in conducting these activities?
- What activities were conducted to provide quality improvement supports for programs?
- What are the successes and challenges experienced in the implementation of *Parent Aware*?

[7]Available at http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7B61078E31-3393-49B1-B301-D5CBC9A0DAC3%7D.PDF. Accessed on February 1, 2011.

[8]Available at http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BE0D4D742-A334-4305-BAF1-8432F580856D%7D.PDF. Accessed on February 1, 2011.

[9]Available at http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BB5ADD2AE-D080-4290-A698-A972B0A93B6A%7D.PDF. Accessed on February 1, 2011.

- What were the experiences of programs participating in *Parent Aware*?
- What were the experiences of parents participating in programs enrolled in *Parent Aware*?

What are the key factors to consider when using this approach? Process studies mostly involve qualitative data collection methods (such as semistructured interviews, focus groups, observations of activities or specific staff/program interactions, and review of program documents), but they can include quantitative data collection methods (for example, surveys and analyses of administrative data). An important consideration for conducting a process study is that this approach requires adequate investments of time and resources to conduct field work, for example, the services of an evaluator skilled in conducting interviews, moderating focus groups, abstracting data from program documents, and analyzing qualitative data will be needed.

Because process studies provide valuable information about how a program operates and about the experiences of the participants in the program, they often accompany other studies/evaluations that are focused on the outcomes of that particular program. However, a process study can serve as an independent evaluation effort.

b. In-Depth Study of an Essential Element in One QRIS, or Across QRISs (Comparative Study)

What is this approach? This approach is used to closely examine a particular element in one QRIS or across multiple QRISs to achieve a level of depth that might not be possible in a broad process study. This type of evaluation uses a systematic approach to examine in-depth the inputs, activities, and outputs of a specific QRIS element such as quality measurement or QI. The approach typically relies heavily on qualitative data collection methods, but can also be supported with quantitative data from surveys or administrative data.

What can be learned with this approach? An in-depth study can provide a rich, "soup to nuts" description and examination of a core issue. A comparative examination of the same topic across QRISs provides the multiple perspectives and contexts in which the topic is based to help

explain the differences and commonalities that exist and possible reasons why. Comparative, in-depth studies are particularly useful when a program or initiative—such as a QRIS—is taking root in many states and localities but there is no common consensus on how to build and structure a QRIS and rigorous evidence is lacking. Comparative studies can help illuminate practices that have potential to lead to positive outcomes, as well as call attention to issues that need further refinement and direction.

What research questions can be answered with this approach? Research questions for this type of evaluation are often fine-tuned subsets of those that would guide a broader process evaluation. Examples of research questions for an in-depth study could include the following:

- What are the distinctions in quality rating components or different (or new) measurement strategies that QRISs have implemented and what are the reasons for these distinctions?
- What strategies and supports do QRISs use to promote progress up the quality levels among participating providers and what goals guide the improvement process and use of resources?
- What strategies have QRISs implemented to extend participation by specific providers or programs (such as home-based providers) or to serve a specific population (such as low-income families, infants and toddlers, or minority populations)?
- How do QRISs serve culturally diverse populations of parents and providers? How do the systems address their preferences or needs directly within ratings, through provider outreach, or through parent outreach?
- How do QRISs engage parents in the design and implementation processes? What planning and attention is given to outreach and dissemination strategies?

A number of comparative, in-depth studies currently underway are initiated and/or funded either by the federal government or by foundations. Two focus on QI activities (one led by the National Child Care Information Center (NCCIC) and one by Child Trends). The QRS Assessment project also includes two in-depth, comparative studies. One focused on quality measurement and the other on the role of QRIS in integration of the early care and education system. The research questions for each study include the following:

Quality Measurement in Five Select QRISs

- What is the variation in how select QRISs define and measure quality?
- What processes are used to measure components and determine an overall rating?
- What is the availability (and use) of consistent and reliable data on quality measurement?

The QRIS's Role in Early Childhood Education System Integration in Two QRISs

- How and to what extent do QRISs contribute to early childhood education (ECE) system development and integration?
- How could QRISs assess the extent of their contribution to ECE system development?

The in-depth study of quality measurement also includes a secondary data analysis. This analysis is being conducted with a subset of three QRISs included in the quality measurement in-depth study to assess the relationships between quality rating components, observed quality, and rating levels. The analysis will build on the 13 quality component categories described in the QRIS Compendium (Tout, Starr, Soli, Moodie, Kirby, and Boller 2010). The data sources include QRIS databases, professional development registries, and ERS data and will examine the following topics:

- Variation in the presence and prevalence of quality components across QRISs
- Differences in the prevalence of quality components between quality rating levels and for certain types of centers (Head Start and accredited programs)
- Unique effect of each quality component on observed environment quality
- Profiles of quality based on the number and combination of components that demonstrate unique effects on observed quality

What are the key factors to consider when using this approach? The following are factors to consider when planning and conducting an in-depth and/or comparative evaluation of a QRIS initiative:

- An in-depth study of an element within one QRIS includes the same considerations as those noted earlier for a process study, such as having a skilled qualitative evaluator. The additional consideration is that substantive expertise might be beneficial given the "one-issue" focus of this type of effort.
- To initiate a comparative approach across QRISs requires teaming up with other sites or states to identify funding sources and evaluators. Sites and states could also be

approached by evaluators to participate in this type of research effort when initiated by a national organization or funder.

- Most comparative studies are not based on a random sample of the full study population (all QRISs, all child care providers, or all families) and the findings are not generalizable beyond the targets of the evaluation. Nonetheless, the findings can provide a good sense of the patterns and themes that are likely to be encountered in other sites, beyond those included in the study. For a comparative study, the process for selecting sites should be made clear. It might be that a sample of convenience (for example, states in the Mid Atlantic region) is necessary and appropriate given the scope and resources of the evaluation. In such cases, the limitations should also be acknowledged. Other methods for purposefully selecting sites should be based on select criteria that are tied to the research questions.
- Systematic methods of qualitative data collection are necessary in cross-site comparative studies to ensure consistency across sites in the depth and quality of the information that is gathered. Consideration should be given to what tools and methods will achieve these objectives—such as whether telephone interviews will suffice or on-site visits are necessary to include multiple activities, as well as observations or demonstrations of procedures.
- Cross-site, comparative studies can include collection or analyses of quantitative data. When using administrative data across multiple sites, be prepared for the challenges that can arise from different methods of collecting data, the layering of files by different unit of analyses (such as classrooms or child care facilities), and defining and creating common analysis variables.

c. Economic Analysis: Cost Analysis, Cost Effectiveness Study, Cost-Benefit Study

A QRIS economic analysis is conducted in different ways depending on whether the goal is to assess (1) the pre-implementation or start-up costs required to build the required QRIS infrastructure, (2) the implementation or steady-state costs that best reflect the level of resources needed to operate the QRIS after start-up is completed, and (3) the cost-effectiveness or cost-benefit of the QRIS for children and for society.

What is this approach? Economic analysis values the costs of an intervention and guides decision makers as they allocate resources. *Programmatic cost analysis* is the approach used to assign an economic value to the costs QRIS program developers and operators incur as they develop the supports needed to prepare for implementation as well as the costs needed to implement the QRIS when the supports and required infrastructure are in place. Programmatic cost analysis considers every relevant cost category and usually encompasses direct, indirect, and in-kind costs. Program

cost analysis is one component of *cost-effectiveness* or *cost-benefit analyses*, approaches to valuing the outcomes of an intervention and determining whether the benefits outweigh the costs to society. For QRIS evaluations, program cost analysis in the pre-implementation period might include valuing the cost of staff time in attending planning meetings, training raters, and developing a QRIS management information system. In the implementation phase, programmatic cost analysis might include (1) the direct costs of staff members who conduct ratings, develop quality improvement plans, provide the incentives and quality improvement supports, and enter data and maintain the QRIS database; (2) the indirect costs associated with driving to the child care settings, paying for staff to attend courses, and making the financial awards; and (3) in-kind costs associated with such things as donated space for conducting child care provider training events. The valuation of QRIS effectiveness or benefits might include cost savings associated with reductions in kindergarten retention or the benefits associated with children's higher test scores and reading achievement in early elementary school that are a direct result of children's exposure to QRIS-rated settings.

What can be learned with this approach? An economic analysis can provide policymakers and the public with information about what a QRIS costs to develop and maintain, and whether the costs to society are worth the benefits. A programmatic cost analysis provides policymakers with detailed information about what the QRIS costs and, depending on the stage of QRIS implementation, different audiences can use the information to meet specific needs. For example, one state's analysis of the costs of developing a QRIS and getting all of the infrastructure in place to support it can be used by policymakers and program developers in another state as they consider how best to develop and launch their QRIS. A programmatic cost analysis of implementation in a steady-state post-start-up can inform both the state conducting the analysis as well as other states about the expected and unexpected costs associated with sustaining and potentially improving the QRIS. Many states use a programmatic cost analysis as part of developing the cost per child or cost per provider. This is one way to gauge or benchmark the costs across different types of

interventions. It is important to note that the start-up costs for a QRIS can be quite large; therefore, it is important to be clear about the goals of a cost analysis at that stage and to ensure that the start-up costs are presented along with the implementation costs.

A cost-effectiveness or cost-benefit study uses the data from a programmatic cost analysis and brings it together with the valuation of specific, measureable benefits in the context of an effectiveness evaluation. Only in the context of an effectiveness evaluation of the targeted child, family, or provider outcomes that uses an experimental or quasi-experimental research design can a true estimate of the size of the benefits be made based on the differences between outcomes of those who had access to the intervention and those who did not.

Economic analyses and the development of tools that allow for simulating costs for different components can also be used to identify ways to reduce costs and increase efficiency. NCCIC maintains a cost estimation model states can use to generate cost estimates and consider alternatives to existing QRIS components or to help guide decision making as states are developing a new QRIS [see http://qriscostmodel.nccic.acf.hhs.gov, accessed on March 8, 2011]. Although these estimates are updated regularly, the estimation model makes a number of assumptions (for example, about the size of an average child care center) and thus cannot take the place of an economic analysis conducted within a state using data on existing child care businesses.

What research questions can be answered with this approach? These economic evaluation approaches can provide information about QRIS costs and, if conducted in the context of an effectiveness evaluation of QRIS outcomes, a cost-benefit analysis provides estimates of the savings or costs to society of implementing the QRIS. The following are types of research questions that can be addressed for different types of economic analyses.

Programmatic cost analysis questions can address the following questions:

- What are the component costs (both financial and in-kind) of building the infrastructure required to implement a QRIS?

- What are the component costs of implementing a QRIS?
- How do changes in investments in a component of the QRIS affect overall costs of the system?

Cost-effectiveness and cost-benefit analysis questions can address the following question:

- Are the outcomes realized by the QRIS worth the costs of implementation?

Although different types of financial studies of child care have been conducted, rigorous economic analyses of QRIS have not. The QRIS Compendium (Tout, Starr, Soli, Moodie, Kirby, and Boller 2010) did not identify any overall economic analyses among the 26 QRISs profiled. A few studies have estimated the costs of moving from one rating level to another or using observational measures of environment quality. Clearly these types of analyses are needed to help guide decision making and resource allocation.

What are the key factors to consider when using this approach? The following are key factors to consider in planning and conducting economic analyses that include a programmatic cost analysis or cost-benefit analyses:

- Programmatic cost analysis requires that cost categories are clearly defined, documented, and valued consistently across all participating settings and types of child care programs.
- Any statements about the cost-effectiveness or cost-benefit of a QRIS cannot be made unless the analysis is conducted in the context of rigorous, well-implemented experimental or quasi-experimental evaluation of QRIS impacts on targeted family, child, or provider outcomes.
- There are challenges in valuing the benefits of early childhood programs such as QRISs because the economic benefits to society are generally not viewed as robust until data from impact evaluations demonstrate that the intervention improved long-term impacts, such as children graduating from college, entering the labor market, or otherwise making contributions to society. Extrapolations of projected economic impacts from child outcomes measured at kindergarten entry or even in early elementary school might not reliably estimate benefits or costs.

d. Validation Study

A validation study assesses the degree to which the quality standards component of the QRIS reflects meaningful quality levels that are linked to desired outcomes. Validity can be established at multiple levels of the QRIS and across multiple stakeholders.

What is this approach? A validation study asks critical questions about the tools used in a QRIS and how they are functioning. Zellman and colleagues (2008) note that a thorough validation of QRIS is a complex process that requires multiple sources of evidence. Therefore, it may be useful to create a long-term plan for validation that outlines different research questions and a strategy for addressing the questions in a systematic way. A comprehensive validation study would address the following broad components: (1) quality standards, (2) construction of the quality levels including weighting and thresholds, (3) feasibility of detecting improvement on the quality levels in response to quality improvement strategies, and (4) linkages between quality levels and desired outcomes. Each of these is described in more detail below.

Validating the QRIS quality standards. Across the 26 QRIS described in the Compendium (Tout, Starr, Soli, Moodie, Kirby, and Boller 2010), certain domains of quality standards are included in the majority of systems including licensing, staff qualifications, family partnerships, environment, administration and management, and accreditation . Domains related to curriculum and assessment are included in fewer than half of the QRIS. Within each of the quality domains, QRIS vary considerably in the indicators that are included to represent each domain. Thus, it is useful to document that the specific quality standards and indicators used in the QRIS resonate with experts in the fields and the stakeholders the QRIS intends to target. Validation of quality standards ideally would occur prior to the launch of the QRIS to document agreement from research experts and practitioners that an appropriate set of standards and corresponding indicators is included (see Elicker et al., 2007 for an example). Validation would also include demonstration of programs' perceptions of the standards as a meaningful reflection of their work and priorities, and of parents'

perceptions of the standards as important to their process for reviewing and selecting child care for their children. For example, in the pilot of Minnesota's QRIS, the standards and indicators were tested in focus groups with parents prior to finalizing the rating tool. While major changes were not made to the standards as a result of the focus group findings, new names for the quality categories were developed to make them more meaningful to parents (Swenson-Klatt & Tout, 2011).

Validating the construction of quality levels. While the components included in most QRIS standards and ratings are based on research, the research base does not provide precise guidance on the processes that should be used to assign points to different indicators, combine indicators into quality levels, assign weights to certain indicators, and set thresholds for different star levels (Tout et al., 2009; Zellman et al., 2008). Therefore, validation is needed to demonstrate that the decisions made using expert opinion or best judgment about construction of quality levels have indeed resulted in quality categories and levels that are distinct. Analytic strategies to address this issue could include: (1) review of a correlation matrix of the indicators to identify any indicators that overlap substantially with other indicators (see Zellman et al., 2008 for an example); (2) review of the average scores and the variance on indicators to identify any that are achieved by all participants in the QRIS (and therefore, do not distinguish adequately among quality levels) or that are achieved by no participants (and therefore, may be unrealistic or unattainable) (see Tout et al., 2010 for an example), (3) examine the correlations between the rating levels and an alternative measure of quality to document how providers rated at different levels score on the alternative measure (see Zellman, et al., 2008 for analyses that examine how providers at different levels scored on the Caregiver Interaction Scale and the Pre-Kindergarten Snapshot). Additionally, analyses could be conducted to model different versions of the rating scale and document how different combinations and/or weighting of the indicators might result in different patterns of programs receiving different rating levels.

Detecting improvement in quality levels. An important issue for programs participating in the QRIS is the degree to which the supports provided in the QRIS can facilitate movement up the quality levels. This type of validation analysis is best conducted once a sufficient number of providers have received at least a second rating so that change in overall rating level and change on individual quality standards/indicators can be tracked. Ideally, a QRIS should be able to document some positive movement on the quality levels (Zellman et al., 2008). However, if all programs move quickly up the levels, or very few programs make positive movement on the quality levels, it may indicate a need to review the quality standards/indicators, the quality improvement supports offered to programs or both.

Linkages between quality levels and desired outcomes. Validation of the QRIS levels is perhaps most salient to stakeholders if it demonstrates that QRIS levels are related to meaningful differences in key targets of QRIS such as children's functioning and parent satisfaction. To date, the majority of QRIS evaluations that involve validation components have not employed an experimental design. As such, findings from those studies examining QRIS levels and measures of children's development do not purport to demonstrate that the QRIS has *caused* particular outcomes but that the practices and other quality features identified at each level of the QRIS are related in predictable ways to children's functioning (thus validating that the levels are meaningful distinctions of quality). Findings to date using this type of validation strategy are mixed with one study noting clear distinctions between quality levels and children's functioning (Thornburg et al., 2009) and others noting no clear patterns of relationships (Tout et al., 2010; Zellman et al., 2008). Across all of the studies, a number of technical issues (sample size, attrition, adequate variation of programs across rating levels) limited the conclusions that could be drawn from the analyses. Thus, this type of validation study is best approached with clear plans and criteria for the type of evidence needed to warrant particular conclusions.

What can be learned with this approach? Depending on its focus, a validation study can provide information to QRIS developers and policymakers about how the QRIS is perceived among key stakeholders, the degree of buy-in to the rating process and levels among programs and parents, the integrity of the rating structure, and the linkages between QRIS levels and outcomes such as children's development and parent perceptions. Perhaps the most compelling way to define the importance of a validation study is that it can quantify if the quality ratings actually mean something important to programs, parents and children.

What research questions can be answered with this approach? As noted, validation studies address a range of research questions that relate to how well the quality standards and the rating process are working to produce distinct levels of quality. These questions include:

- Do the quality standards and indicators reflect the current research base and expert opinion on quality in early care and education programs?
- Do the quality standards represent distinct areas that do not overlap with other standards in the QRIS? Can some standards/indicators be removed?
- Do scoring patterns on the rating tool indicate adequate variation given the population of programs participating in the QRIS?

Other validity questions relate to the functioning of the quality ratings for programs, families and children:

- Are the quality levels related in a linear way to alternative measures of quality?
- Can improvement be detected on the quality levels?
- Are quality levels related to children's functioning?
- Do improvements in quality levels relate to improved outcomes for children?

What are the key factors to consider when using this approach? A validation study (or series of studies) can happen at all phases of QRIS development. Validation questions can be asked before the QRIS is launched and each time significant changes are made to quality standards/indicators (including decision rules for meeting the indicator or scoring changes) or construction of levels (including changes in weighting or cut-points for different indicators).

Therefore, a key consideration is timing of a validation study and how findings can inform decision points in QRIS design and implementation. It is expected that validation analyses will be conducted even when the QRIS is mature as it is very likely that changes will be made to the system to reflect new knowledge.

A second (and related) consideration is the degree to which the QRIS includes a representative sample of providers from the communities included in the QRIS. Validation analyses that are conducted using a select group of program participants (for example, the first group of providers to enroll in the QRIS) may be biased if the sample represents providers at only one end of the quality spectrum or who are homogeneous on other key variables (education level of staff or geographic location). It is important to acknowledge the limitations of validation analyses under these conditions and to consider alternative strategies for improving the diversity of programs included in the QRIS (such as recruitment and incentives).

3. Outcome Evaluations

Outcome evaluations of QRISs can be conducted using one of the following approaches: (a) an experimental design, (b) a quasi-experimental design, (c) a time series approach, (d) a post-test approach, and (e) a descriptive approach. Next, we describe these approaches to conducting outcome evaluations of QRISs, and in Table III.8 we describe the elements (subjects, sample type, points of data collection, use of a comparison group, and outcomes assessed) of outcome evaluations that have been conducted for existing QRISs.

a. Experimental Design

What is this approach? In experimental or *random assignment* designs, subjects are randomly assigned to an intervention (treatment) group that is eligible to receive or participate in a specific intervention or service (that is, the *treatment*) or a comparison (control) group that does not receive

Table III.8. Characteristics of the Evaluations Conducted for Existing QRISs

States with QRS Evaluations[a]	Evaluation Approach	Evaluation Subjects[b]	Sample Type	Data Collection Periodicity	Comparison Group	Environment Rating Scales[c]	Other Quality Measures[d]	Child Outcome Measures
Implementation Evaluations								
Colorado	Validation Study	• Centers and FCC • Teachers • Children and families	Cross-sectional	1 wave	None	X	X	X
Florida, Palm Beach	Process Study	• Centers and FCC • Peer coaches • QRIS staff and partners	Longitudinal	2 waves, 18-month interval	None	X		
Indiana*	Validation Study	• Centers and FCC	Longitudinal	2 waves, 6-month interval	None	X	X	
Minnesota*	Process Study (Waves 1 and 2) and Validation Study (Wave 3)	• Centers and FCC • Teachers and program staff • Children and families	Longitudinal	3 waves, 1 year interval	None	X	X	X
Ohio*	Cost Study	• Centers and FCC • Children and families	Longitudinal	Not available	None	X	X	X
Outcome Evaluations								
Colorado	Time Series Approach (Waves 1, 2, and 3)	• Centers and FCC • Teachers • Children and families	Longitudinal	3 waves, 1-year interval	None	X	X	X
Indiana*	Descriptive Approach	• Centers and FCC • Children and families	Cross-sectional	1 wave	None	X	X	X
Kentucky	Descriptive Approach	• Centers and FCC • Teachers and program staff • Children and families	Cross-sectional	6 years	Nonparticipating providers	X	X	X
Missouri	Pre-Post Approach	• Centers and FCC • Children and families	Longitudinal	2 waves, 1-year interval	None	X	X	X
Minnesota*	Pre-Post Approach (Wave 3)	• Centers and FCC • Teachers and program staff • Children and families	Longitudinal	3 waves, 1 year interval	None	X	X	X

States with QRS Evaluations[a]	Evaluation Approach	Evaluation Subjects[b]	Sample Type	Data Collection Periodicity	Comparison Group	Environment Rating Scales[c]	Other Quality Measures[d]	Child Outcome Measures
North Carolina	Descriptive Approach	• Centers and FCC • Teachers and program staff • Children and families	Longitudinal	3-4 waves, 2- to 3-year interval	None	X		X[e]
Oklahoma	Descriptive Approach	• Centers and FCC • Teachers and program staff	Cross-sectional	1 wave	None	X	X	
Pennsylvania	Descriptive Approach	• Centers and FCC	Cross-sectional	1-3 years	Nonparticipating providers	X		

Sources: Evaluation reports when available; presentation materials or QRIS website overviews. Full sources are provided in the reference list.

*QRS evaluations that are planned and/or in progress but not yet completed.

[a]Other evaluations completed or underway are not included due to lack of information on measures given the early evaluation stage (Washington) or sparse data collection (Delaware).

[b]For purposes of this table, FCC refers to Family Child Care and covers both small and/or group home-based care.

[c]Current or former versions of one or more of the following: Early Childhood Environment Rating Scale-Revised (ECERS-R), Family Child Care Environment Rating Scale-Revised (FCCRS-R), Infant/Toddler Environment Rating Scale-Revised (ITERS-R), and/or the School-Age Care Environment Rating Scale (SACERS).

[d]Refers to process quality measures such as the Classroom Assessment Scoring System (CLASS), Caregiver Interaction Scale (CIS), or the Early Language and Literacy Classroom Observation Tool (ELLCO).

[e]Child outcomes were examined in only the last time point.

the intervention or service. For evaluations of QRISs, the subjects of the evaluation can be children, families, child care programs, or the communities served by the QRIS; the intervention or treatment refers to a particular component of the QRIS or a specific service provided by the QRIS.

What can be learned with this approach? Assuming that the assignment of subjects to intervention and comparison groups is done correctly (that is, at random and without other factors influencing this assignment) then the intervention and comparison groups are essentially identical on average (except that the intervention group receives the intervention/service and the comparison group does not) at the start of the intervention and any differences in outcomes observed at the end of the evaluation can be said with confidence to be *caused* by the QRIS component, intervention, or service. In other words, experimental designs allow for "one difference only" between the intervention and comparison groups, and that difference is the QRIS component, intervention, or service that is being examined.

An experimental design allows for learning about the *effectiveness* of a specific QRIS component, intervention, or service. However, assessing the effectiveness of the QRIS as a whole (that is, the effectiveness of the entire system) based on an experimental design might not be feasible. That is because the entire population of child care programs/providers in the state (and also the entire population of children and families eligible for child care services) is by definition the target of the intervention or treatment (which, in this case, would be the QRIS) and the QRIS resources are diffused throughout the entire population of child care programs and of the population of children and families eligible for child care services.[10]

What research questions can be answered with this approach? This outcome evaluation approach can provide information about the effectiveness of a clearly defined QRIS component,

[10]In cases in which it is possible to randomly assign entire communities (which have similar characteristics before implementation of the program to be evaluated) to a group that participates in the QRIS and a group that does not, it might be possible to evaluate the system using an experimental approach.

service, or support, such as training in a specific ECE curriculum or a financial support for specific QIs in the classroom/home environment in which care is provided. The following are examples of research questions that can be answered using an experimental design:

- Does the quality of child care programs increase as a result of participating in the QRIS?
- Do services provided by the QRIS (for example, financial incentives and trainings) improve the qualifications of the staff working in the programs enrolled in the QRIS?
- Do children attending child care programs that received high QRIS ratings show more improved language and emotional development outcomes than they would if participating in programs with lower QRIS ratings?

Existing early learning initiatives have been evaluated using an experimental design. For example, *The Seeds to Success Modified Field Test* (Boller et al., 2010) examined whether the coaching model and financial incentives provided by the *Seeds to Success* (Seeds) program influenced the quality of care provided by family- and center-based programs in two communities in Washington State. Specifically, this evaluation investigated the following three research questions (Boller et al., 2010):

1. Did Seeds improve the quality of child care available in participating child care businesses?
2. Did Seeds increase the amount of education, training, and technical assistance services accessed by participating child care businesses?
3. Did Seeds improve the level of education and experience for the workforce employed in participating child care businesses?

To answer those questions, the Seeds evaluation randomly assigned 66 child care programs (52 family care providers and 14 centers) across the two communities into treatment and control groups. The providers in the treatment group received coaching, quality improvement grants, professional development opportunities, and access to funds. The providers in the control group received only professional development opportunities and access to funds. The effect or impact of Seeds was defined as the differences between the treatment and control groups on key outcomes after the seven-month study period. The outcomes examined included service receipt, education and

professional development, and child care quality. The findings from this evaluation include the following:

- Seeds significantly improved observed quality in centers and family care homes, but the program did not improve observed group size or child-adult ratio in either type of care provider.
- More treatment group center teachers than control group teachers reported enrollment in education and training services, but no differences were found for family child care providers.
- The completion rate of three or more credits increased for treatment group center teachers.

What are the key factors to consider when using this approach? The following are key factors to consider in planning and implementing an outcome evaluation based on an experimental design:

- This approach requires that QRIS administrators and staff, child care providers, parents, and other stakeholders buy in to random assignment. To be able to implement this approach, it is fundamental that QRIS stakeholders (1) understand that having a comparison group that does not receive the services/intervention being evaluated is ethical; and (2) conduct the necessary efforts to preserve the integrity of the design (that is, efforts to prevent the comparison group from receiving the services/intervention while the evaluation is being conducted). Offering access to the services/intervention to the members of the comparison group *after* the evaluation has been completed might be helpful to obtain buy-in from the QRIS stakeholders.
- A large sample is often required for this type of design to detect meaningful differences in observed outcomes.[11] That is, this design requires that the treatment and comparison groups are composed of a number of subjects large enough to allow for finding differences in outcomes between the groups even when these differences are small.
- To maximize the rate at which the treatment group actually receives the intervention, the start of service receipt should be close to the timing of random assignment. For example, if the intervention being evaluated consists of a specific early learning curriculum, then implementation of the curriculum should start soon after random assignment is

[11]The sample size that is relevant for detecting effects is the sample of units or subjects that were randomly assigned to a treatment or a control group. In some cases, it is possible to conduct random assignment of individuals (for example, random assignment of children to an intervention or control group, or random assignment of child care program staff to a professional development program or a control group), but in other cases, random assignment has to be conducted for entire programs or communities. The latter are known as cluster designs (for example, children or teachers clustered within child care programs). Obtaining a large sample of clusters (that is, child care programs or communities) to participate in an evaluation can sometimes be challenging and costly.

conducted so that providers in the treatment group do not use other (similar) curricula in the meantime and the providers in the control group do not have the opportunity to implement the intervention curriculum on their own initiative.

- Strategies to minimize drop-outs (that is, attrition from the study) should be implemented to preserve the integrity of the treatment and comparison groups. If a considerable number of subjects from the treatment or control groups (or from both groups) leaves the evaluation, then the composition of each group changes and the subjects that remain in each group might not be comparable in the same way that the complete groups were when they were first randomly assigned (that is, the equivalence between the groups might no longer exist).

b. Quasi-Experimental Design

What is this approach? In quasi-experimental designs, each subject (for example, child care providers, children and their families, and communities) that is offered the QRIS intervention (the treatment group) is matched with a comparison subject at the start of the evaluation based on selected characteristics, such as the number of children served and the proportion of children receiving subsidies if the evaluation subjects are child care providers, or demographic characteristics if the evaluation subjects are children and their families. This type of evaluation design is also known as a comparison group design.[12]

What can be learned with this approach? Quasi-experimental designs can meet standards in the field for high levels of evidence[13] and allow for making defensible conclusions regarding the effectiveness of QRIS component or intervention if the treatment group has very similar characteristics to the comparison group at the start of the evaluation. In other words, causal inferences can be drawn from the findings of an evaluation based on a quasi-experimental design to the extent that *baseline equivalence* between the treatment and control groups can be shown. If baseline

[12]When the intervention and matching subjects are children (and their families), classrooms, teachers, or child care providers, the quasi-experimental designs are generally called *matched case* designs. These designs include those that use propensity score matching (Rusenbaum & Rubin, 1983), a statistical technique in which subjects are matched using a composite score generated by an algorithm that minimizes variance across any one matching characteristic.

[13]For example, see the What Works Clearinghouse Evidence Standards (http://ies.ed.gov/ncee/wwc/pdf/study_standards_final.pdf) and the study rating process established by the Home Visiting Evidence of Effectiveness (HomVEE) initiative (http://homvee.acf.hhs.gov/document.aspx?rid=4&sid=19&mid=5).

equivalence exists, then the likelihood that factors other than the QRIS component or intervention being evaluated explain any differences in outcomes observed at the end of the intervention is low.

What research questions can be answered with this approach? Evaluations based on quasi-experimental designs, as those based on experimental designs, seek answers to research questions about the effects of a specific QRIS component, service, or support. Examples of research questions that can be answered using a quasi-experimental evaluation design are the following:

- Do child care programs participating in the QRIS provide higher quality care than comparable programs not participating in the QRIS?
- Do staff in child care programs participating in the QRIS have more access to professional development opportunities than staff in comparable programs not participating in the QRIS?
- Do children receiving care in programs participating in the QRIS show better language and emotional development outcomes than children in comparable programs not participating in the QRIS?

What are the key factors to consider when using this approach? The following are factors to consider when planning and conducting an evaluation of a QRIS based on a quasi-experimental design:

- In order to be able to implement this approach, it is necessary to have access to rich information sources (for example, administrative data or secondary databases from national surveys/studies) that provide the data elements needed to build the comparison group.
- In addition to access to the right data, this approach often requires a large sample of subjects for which matches for the treatment group can be found and, when the evaluation sample has been determined, it can be possible to find meaningful differences in the outcomes of interest between the treatment and comparison groups.
- Because the treatment and comparison groups in a quasi-experimental design are not created through random assignment, it is not possible to be fully confident that there are no differences (observed or unobserved) between the groups. For example, in an evaluation of a specific type of QI support (for example, coaching, trainings, or financial incentives), the providers in the treatment group could be more motivated to increase the quality of the services they provide or to increase the qualifications of their staff than the providers in the comparison group. In this case, any observed differences in the outcomes of interest between the treatment and comparison group might appear larger than what they really are (that is, the outcomes will be *biased*) because the members of the

treatment group were not really similar (that is, they were not equivalent) to the members of the control group

- Finding a suitable comparison group could be challenging because the entire pool of subjects from which a comparison group could be drawn might have already been exposed to initiatives similar to those that are being evaluated for the QRIS. For example, all child care providers in the state might already be receiving QI supports from sources different from the QRIS (such as foundations or academic institutions).
- Even when a suitable comparison group can be found, sometimes information (data) on outcomes for the comparison group might not be available in a form that will permit meaningful comparisons, or sometimes it simply becomes difficult to recruit these subjects to participate in the comparison group because they might not be officially connected to a program or particular service delivery mechanism.

c. Time Series Approach

What is this approach? In this type of evaluation design, time itself serves as the basis for comparison and changes in the outcomes of interest are tracked longitudinally. That is, baseline data are compared at regular intervals (for example, every six months or every year) with data collected during and after intervention implementation for the *same study group* (whether it is child care providers, families, or children). In time series designs, longitudinal data are typically tracked in cohorts; that is, the outcomes of the same study group or cohort are measured at different points in time. For example, the study group or cohort can be all children who enter care in programs participating in the QRIS in a given year and the outcomes measured can be language development outcomes in the spring of year X, year Y, and year Z of QRIS operations. In another example, the cohort can be all child care providers that enroll in the QRIS in a given year and the outcomes measured can be the QRIS ratings they obtain in year X, year Y, and year Z of QRIS operations.

The longitudinal characteristics of a time series design include consistency over time in measuring the same study group. Other longitudinal studies may track outcomes over time, but not in the same way as a time series design. For example, a cross-sectional design can examine outcomes at multiple points, but not with the same study group (for example, taking a snapshot of providers participating in the QRS in year X and year Y, but not following the same providers).

What can be learned with this approach? This evaluation approach provides information about the trends or changes over time in QI and other outcomes of interest for the particular study group. For example, it can examine changes over time in the quality of care provided by the programs participating in the QRIS and in the characteristics of the population of children and families they serve. However, because time series designs do not include a comparison group to account for the possibility that factors other than the QRIS might also explain changes in observed outcomes, causal inferences cannot be derived from the findings of evaluations based on this type of design. For example, decreases in average ratings across providers over time might be a result of state cut-backs that allowed for fewer hours of professional development and coaching than the hours allowed before the cut-backs.

What research questions can be answered with this approach? The following are examples of research questions that can be answered using a time series approach:

- How has the quality of care provided by the programs participating in the QRIS changed over time?
- How long (on average) do child care providers participating in the QRIS take to move from one rating level to the next higher level?
- What are the characteristics of the child care providers that show quality increases over time?
- How has the population of children and families served by the child care providers in the QRIS changed over time?
- How has the quality of care provided for the most at-risk children changed over time?

What are the key factors to consider when using this approach? Time series designs represent an alternative when an appropriate comparison group cannot be identified. The following are key factors to consider when conducting an evaluation based on a time series approach:

- The QRIS component, intervention, or service being evaluated should be implemented over an extended period to allow data collection to happen at multiple relevant points in time.

- Availability of baseline data is essential to provide an apt reference or comparison for the changes that occur over time in QI and other outcomes.
- This approach requires that efforts are implemented to maintain as much of the full sample as possible across time. That is, as many subjects as possible from the study sample assessed at baseline (the first point of time in the design) should be assessed at the subsequent points in time that are relevant to the investigation. Also, rules should be established regarding the amount of attrition that could be tolerated to preserve the integrity of the study sample.

d. Pre-Post Test Approach

What is this approach? In this type of evaluation approach, data on evaluation participants are collected at the start of the intervention (at baseline, or at pre-test) and then one time after that, usually at the end of the intervention (at post-test). Data collection does not usually continue after that point; however, sometimes data are collected one more time after implementation of the QRIS component, intervention, or service has ended (at follow-up, which could occur, for example, a year after completion of the QRIS intervention or service). In this type of approach, pre-test outcome measures are compared with the post-test outcomes measures of the same study group (for example, a group of children or providers). However, this approach can include one study group (for example, a group of child care providers that implement an early learning curriculum and that are assessed before and after implementation of this curriculum) or several study groups (for example, child care providers that obtained high QRIS ratings and providers that obtained low QRIS ratings. The outcomes of each of those types of providers are then assessed at pre- and post-test). An important nuance is that although there might be multiple study groups assessed, there are no true comparison groups in the same sense as a treatment and control group in an experimental design because participants in the evaluation are not randomly assigned to either a group receiving the intervention or a comparison group not receiving the intervention and are, therefore, different in some way from the start.

What can be learned with this approach? A pre-post test design allows evaluators to compare outcomes of subjects before they participate in a program or receive services with their

outcomes after they have participated in the program or have received the services. For example, a pre-post test approach provides information on whether the number of child care programs that measure aspects of quality required to maintain a healthy and safe environment increases after child care programs receive guidelines from the QRIS on health and safety procedures or stays as it was before the QRIS provided such guidelines. A pre-post test design can also provide information on whether the number of child care programs using a curriculum that meets the QRIS standards increases after providing training for child care programs' staff on specific curricula or stays the same as it was before the training. However, since pre-post test designs do not use a control or comparison group, the findings from pre-post test designs cannot be interpreted as evidence of the effectiveness of the QRIS as a whole or of one component of the QRIS.

Figure III. 1 illustrates why causal inferences cannot be drawn from an examination based on a pre-post design. The first panel of that figure shows the average emotional regulation scores from the Bayley Behavior Rating Scales (Bayley, 1993) at 24 months (the pre-test) and at 36 months (the post-test) for the children in the treatment group of the Early Head Start impact study (Love et al., 2002). Since the average emotional regulation score at 36 months (3.9) is higher than the average score at 24 months (3.6) for children receiving Early Head Start services, one might think that the increase in the average score reflects an improvement in children's emotional regulation skills caused by participating in the Early Head Start program. However, observing a difference in pre- and post-test scores does not imply that the increase in emotional regulation outcomes of children in this study can actually be attributed to participation in the Early Head Start program. The increase in children's social-emotional skills could be a result of children being more mature at 36 months than at 24 months, or could be because parents have had one additional year of experience in parenting and helping their children regulate their emotions.

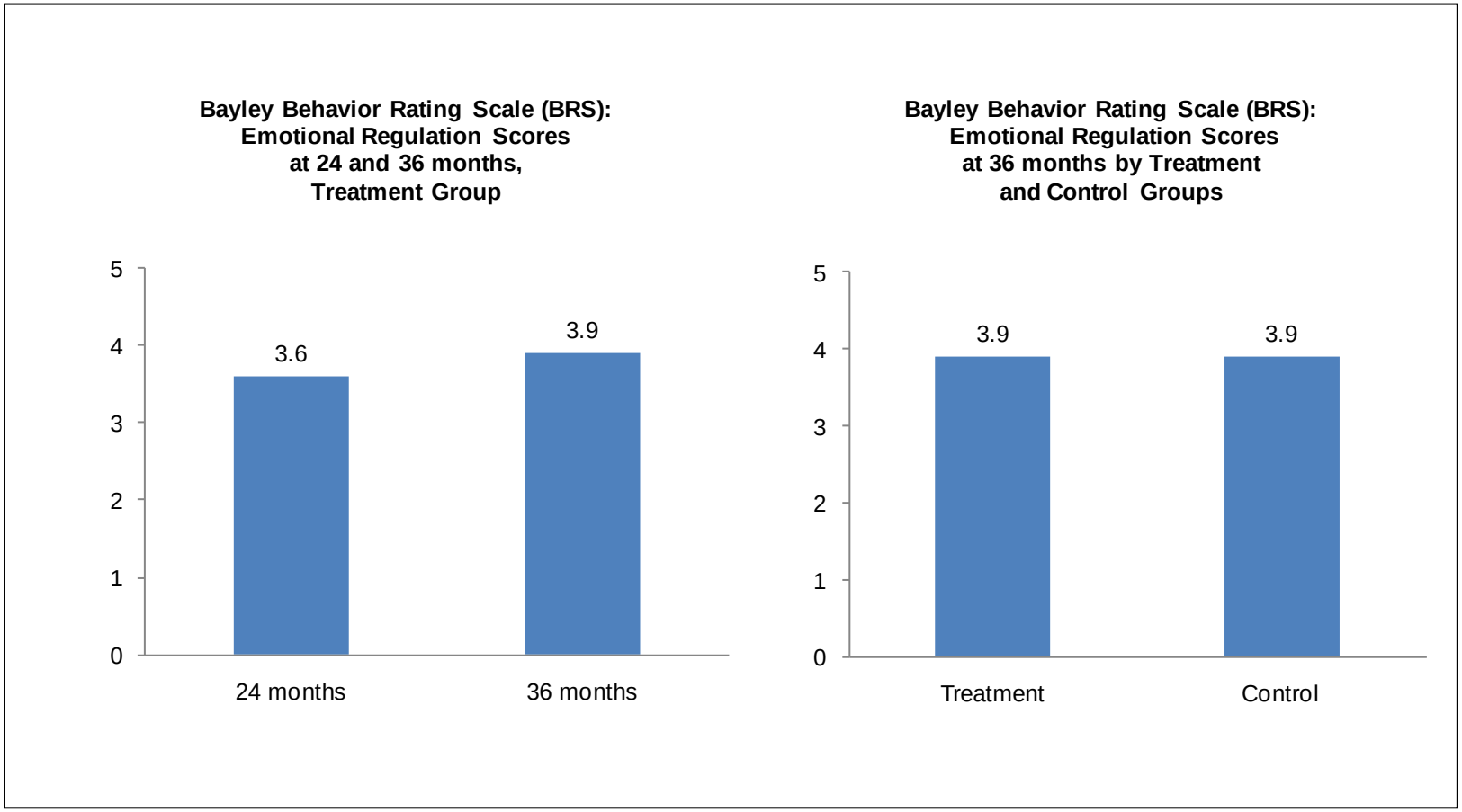
As explained earlier in this chapter, in order to separate the effects of the evaluated program from other effects and conclude with confidence that the Early Head Start program improves

children's emotional regulation skills, it is necessary to compare the post-test emotional regulation scores of children who were randomly assigned to receiving Early Head Start services (the treatment group) with the post-test scores of children who were randomly assigned to not receive the services (the control group). That is because the random assignment process guarantees that the only difference between the two groups is that one receives the services and the other does not. The second panel of Figure III.1 presents the average emotional regulation post-test scores (36 months) of children in the treatment and the control groups. The average scores at post-test are the same (3.9) for children in both groups, which means that participation in the Early Head Start program did not lead to an increase in children's emotional regulation skills since on average there is no difference in the emotional regulation outcomes of children who received Early Head Start services (treatment group) and the outcomes of children who did not receive the intervention (the control group). Thus, the comparison of pre- and post-test scores presented in the first panel cannot be interpreted as evidence of the effectiveness of the program to improve social-emotional outcomes of children.

What research questions can be answered with this approach? The following are examples of the research questions that can be answered with this approach:

- How does the distribution of QRIS-assigned quality ratings *before* the introduction of a new coaching method (that is, at pre-test) compare with the distribution of quality ratings *after* implementation of that method (that is, at post-test)?
- What are the QRIS ratings of providers *before* the introduction of curriculum requirements in the QRIS quality standards (at pre-test)? What are the ratings *after* introducing those requirements (at post-test)?
- What are the language and social-emotional outcomes of children served by programs in the QRIS in the fall of academic year Z (pre-test)? What are the language and social-emotional outcomes of those children in the spring of the same academic year (post-test)?

Figure III.1. Pre-Post Differences vs. Treatment and Control Group Differences



Source: Data is from Love, J.M., Kisker, E.E., Ross, C.M., Schochet, P.Z., Brooks-Gunn, J., Paulsell, D., Boller, K., Constantine, J., Vogel, C., Fuligni, A.S., & Brady-Smith, C. (June, 2002). *Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start. Vol. II: Final Technical Report Appendixes.* Princeton, NJ: Mathematica Policy Research.

The *Missouri Quality Rating System School Readiness Study* (Thornburg et al., 2009) employed a pre-post test approach to examine outcomes of children in child care programs that received different quality ratings (low, medium, and high ratings) by the Missouri QRS. Child outcomes were assessed in the fall of 2008 and in the spring of 2009. The outcomes assessed were social and behavioral skills (motivation, self-control, positive adult relationships, and incidence of emotional and behavioral problems); language outcomes (receptive vocabulary, knowledge of letters and sounds, early literacy, and print awareness); math skills (counting, addition, and subtraction); and motor skills (fine and gross motor skills). In particular, the *Missouri Quality Rating System School Readiness Study* asked the following research question:

- Do preschool children who attend higher quality programs as measured by the Missouri QRS show greater gains in school readiness than their peers who attend lower quality programs?

That study found that the social/behavioral skills and vocabulary outcomes of children in medium and high quality programs were higher in the spring of 2009 (post-test) than in the fall of 2008 (pre-test), but the same outcomes for children in low quality programs were higher in the fall of 2008 than in the spring of 2009. Also, that study found that the gross motor skills and knowledge of letters and sounds outcomes of children in all types of programs were higher at post-test than at pre-test. Although some greater gains were demonstrated among children who attended higher quality programs versus those in lower quality programs, the QRS (and the higher quality programs) cannot be fully attributed for *causing* this difference.

What are the key factors to consider when using this approach? A examination of a QRIS component, intervention, or service that is based on a pre-post test approach would yield valuable information on the characteristics and outcomes of evaluation participants (subjects) before and after the QRIS component, intervention, or service is implemented. However, when planning and implementing an evaluation of a QRIS initiative using a pre-post design it is important to consider

that any observed changes in outcomes from pre- to post-test cannot be attributed with confidence to the QRIS component, intervention, or service being examined. That is, causal inference cannot be drawn from the findings of an evaluation based on a pre-post test design because a comparison group does not always exist, or if one does, it is not based on random assignment and baseline equivalence between the groups cannot be established.

e. Descriptive Approach

What is this approach? A descriptive approach assesses a particular group of subjects at one point in time (that is, it looks at a cross-section of data or information on that group of subjects) or different groups of subjects at different point in time (that is, it looks at different cross-sections of data for each group of subjects) and identifies patterns or trends for the different characteristics and outcomes of the group or groups of subjects. In other words, a descriptive approach aims to find out *what is* and so this approach can be based on either quantitative or qualitative data. For example, a descriptive approach to examine outcomes for child care programs participating in a QRIS would look at the programs (family care and center-based) that have enrolled in the QRIS by a specific point in time (for example, the end of the first year of operations of the QRIS) and find out what are the qualifications of the staff in those programs, the amount and type of financial incentives received by those programs, the staff-to-children ratio, and other measures of program quality such as the quality of the space and furnishings and the quality of the interactions with children attending those programs. Alternatively, a descriptive approach to examine child care program outcomes can involve looking at the characteristics of programs that enroll in the QRIS in a given year and comparing them with the characteristics of other programs that enroll in the QRIS in a subsequent year. Finally, a descriptive approach to investigating outcomes for child care programs in the QRIS could also look at the outcomes described earlier by geographic area or location of the program, by characteristics of the children and/or communities served by those programs (for example, socio-

economic status and racial/ethnic composition), and by the quality rating level assigned by the QRIS at a specific point in time.

What can be learned with this approach? This type of approach provides answers to questions about *what is* for a particular situation or subject (or group of subjects). Thus, studies based on a descriptive approach provide information about what *outcomes are* for children, families, child care programs, or the community served by the QRIS. That information on outcomes can be reported in tables that contain summary statistics (including the mean, median, mode, standard deviation, and correlation between variables) for the sample of subjects participating in the evaluation (see Table III.9 for an example of a table shell that could be used to present summary statistics about a sample being examined),. Rich descriptions of those outcomes can also emerge from qualitative information such as observations, open-ended interviews, and focus groups. Studies based on a descriptive approach yield information (data) about what happens in child care programs and in the QRIS initiative that can be useful to determine the prevalence of certain characteristics of programs that receive high (or low) QRIS ratings, in planning how to allocate time and financial resources of the QRIS, and to identify areas for further research.

Table III.9. Example of A Table Shell Presenting Summary Statistics on the Sample Being Examined

Characteristic	Mean	Median	Mode	Standard Deviation
Sample of Child Care Programs				
Number of Children Enrolled in Program				
Number of Teaching Staff				
Number of Administrative Staff				
QRIS Rating				
Sample of Children				
Age				
Family Income				
Learning Assessment Score				

Although a descriptive approach can provide information that might not otherwise be noticed or encountered (for example, that child care programs using a standardized early learning curriculum have higher quality, as defined by the QRIS quality ratings, than child care programs not using a

standardized curriculum), it does not allow for making causal conclusions or for evaluating the effectiveness of the QRIS as a whole or of one of its components. That is, a descriptive approach could allow evaluators to find that child care programs using a particular curriculum obtain higher QRIS ratings than other programs not using that curriculum, however, a descriptive approach would not allow the evaluators to conclude that children receiving care in programs using that particular curriculum exhibit better outcomes than children attending programs that do not use that curriculum. A descriptive approach would only allow evaluators to find a correlation (or association) between the QRIS ratings (which reflect the quality of the curriculum used in the child care programs) and child outcomes.

What research questions can be answered with this approach? The following are examples of research questions that can be addressed with a descriptive approach:

- What are the QRIS quality ratings of child care providers enrolled in the QRIS at the end of year X? What are the quality ratings by geographic area? By characteristics of the children served?
- What are the qualifications of the staff in child care programs participating in the QRIS at time Y? What are the staff qualifications by type of program?
- How does the number of child care programs enrolled in the QRIS by the end of the second year of operations compare with the number of programs enrolled in the QRIS at the end of the first year of operations?
- What amount and types of financial incentives were provided in year A? How do those compare with the amounts and types of financial incentives provided in year B?
- What are the characteristics of the children (for example, age, race/ethnicity, socio-economic status, special needs status) served by the child care programs enrolled in the QRIS at time X? What are the characteristics of the children served by the child care programs enrolled in the QRIS at time Y?
- What are the language and socio-emotional outcomes of children in child care programs participating in the QRIS at time Z?

The *Evaluation of Pennsylvania's Keystone STARS Quality Rating System in Child Care Settings* (Barnard et al., 2006) examined the 2005–2006 scores of 572 child care providers (356 child care centers, 81 group child care homes, and 135 family care homes) on the Early Childhood Environment Rating

Scale Revised (ECERS-R) (Harms et al., 2005) and the Family Day Care Rating Scale (FDCRS) (Harms & Clifford, 1989). This study also examined whether child care programs in the QRIS used a curriculum, the qualifications and experience of the programs' staff, and the professional development requirements in these programs in 2006. The results from this study indicate that child care centers that received the highest QRIS ratings (STAR 3 and STAR 4 levels), on average, obtained higher scores on the ECERS-R than the child care centers that received the lowest QRIS rating or did not participate in the QRIS. The study also found that (1) child care programs using a defined curriculum obtained higher scores, on average, on the ECERS-R (child care centers) and the FDCRS (family care practitioners) than programs not using a curriculum; and (2) programs with staff holding college degrees and with more than five years of experience obtained higher ECERS-R and FDCRS scores than programs with staff with fewer qualifications.

What are the key factors to consider when using this approach? Although this type of approach can provide rich descriptions of the components of the QRIS, it does not provide information that allows for making causal conclusions regarding the effects of the QRIS components it described because this design does not involve using a control group (as do experimental designs) or a comparison group (as do quasi-experimental designs). For example, from a description of the quality levels or ratings that the QRIS assigned to child care programs it is not possible to conclude that participation in the QRIS resulted in several programs obtaining high quality ratings because the description of the quality levels does not provide information on what would have happened with the quality of child care programs had they not participated in the QRIS—which is the information that a control or a comparison group would provide. Another issue to consider is whether some of the information presented in a descriptive study is redundant. For example, if the quality ratings assigned by the QRIS are based on scores obtained on one of the ERS measures, then it is not necessary to present both the QRIS ratings and the ERS scores to

illustrate the quality of the programs in the community of interest. In that case, it would be enough to present either the QRIS ratings or the ERS scores.

D. Selecting Data Collection Methods and Developing a Data Collection Plan

After identifying the research questions and selecting an evaluation design for the QRIS evaluation, it is necessary to determine what data are required to answer those research questions, implement the evaluation design, and identify from whom and how those data can be obtained. Making these decisions early in the planning process for the QRIS evaluation can help streamline the data collection process and reduce the risk of collecting too much or irrelevant data.

Next, we provide information about different types of data collection methods and an overview of approaches to collecting the required data efficiently. This information might be helpful for making decisions about what methods to use, which data to collect from whom, and how to collect the required data.

1. Methods to Collect Quantitative and Qualitative Data

Quantitative data refers to information that can be counted or expressed numerically. In contrast, **qualitative** data refers to information that is not numerical and that can be represented in terms of qualities or categorizations. A data collection method is a specific tool or technique that is used to obtain data on implementation and outcome measures within the parameters determined by the evaluation design. Therefore, methods to collect **quantitative** data refer to techniques that use a structured set of questions that yield information in the form of numbers. For example, a survey of the staff in child care programs participating in the QRIS will provide information on the number of staff who participated in professional development activities in the past year and the number of hours spent in such professional development activities. Methods to collect **qualitative** data refer to techniques used to gather information on how a program or intervention works, why decisions were made in a particular way, and what those decisions might mean to those involved in the program or

intervention. For example, open-ended interviews or focus groups with child care program administrators can provide information about their experiences receiving the services provided by the QRIS or about the process they followed to decide how to use the QI support from the QRIS toward obtaining a higher quality rating. Because systematic qualitative data collection focuses on gathering information about how a program was planned to function and how it actually operates, this data collection method is employed in implementation evaluations.

We should note that although the format of the information (data) gathered with quantitative data collection methods is different from that of the information gathered with qualitative methods, both types of methods can serve to collect a specific piece of information (for example, either method could be used to obtain data on the types and frequency of QI technical assistance received from the QRIS), and some qualitative reports (data) can be converted to a number format (that is, they can be converted to a numerical measure).

The following list presents some of the most common types of **quantitative** methods:

1. Surveys or questionnaires (for example, surveys on child care programs' administrators and other staff, surveys on parents)
2. Extracting information from existing administrative data (for example, information from public databases [such as the National Child Care Information and Technical Assistance Center (NCCIC) Child Care Database][14] on the number and types of child care programs in the area served by the QRIS, information from QRIS administrative records on the number of children enrolled in the child care programs participating in the QRIS)
3. Structured observation that produces data that can be used to create numeric rating scales (for example, the ECERS-R [Harms, Clifford, & Cryer, 2005] protocol)
4. Standardized assessment instruments and tests (for example, tests to assess the language skills of children enrolled in the child care programs participating in the QRIS)

The following are examples of some of the most common types of **qualitative** methods:

1. Focus groups with child care program administrators, child care program staff, or parents

[14]See http://nccic.acf.hhs.gov/state-territory/index.cfm.

2. Open-ended and semistructured interviews, conducted with child care program administrators, child care program staff, or parents
3. Field notes and/or journal entries made during observation of the care environment, child care program staff meetings, or on-site assistance sessions provided by a QRIS-funded coach or mentor
4. Review of documents such as notes made by child care providers on skill and/or behavior of the children under their care, feedback (provided in an "opinion box") from parents regarding how useful they found the information provided by the QRIS, or procedures and training manuals used by the QRIS program staff

Using both quantitative and qualitative methods to collect data could be a desirable approach to produce the information necessary to document and explain the complexity of the program or intervention being evaluated. And, in deciding which data collection methods to use, the QRIS evaluation should remain focused on the adequacy of these methods to answer the research questions of interest. Chapter IV further discusses implementation and outcome measures and methods of data collection (qualitative and quantitative).

2. Developing a Data Collection Plan

After identifying research questions and determining the research methods that will be used in the evaluation, the next step in the evaluation planning process is to develop a data collection plan. Details about selecting measures and using data collection best practices are presented in Chapter IV, but key activities to develop a data collection plan are presented in Table III.10.

When developing a data collection plan, consider that an optimal approach to planning data collection activities involves, first, determining what types of data about the QRIS (or a particular component of the QRIS) and the program target population are already available and whether links between those existing sources of data can be made. QRISs conduct activities (for example, enrolling child care programs, conducting assessments of the QRIS at different points in time, providing training and on-site assistance to providers, and reaching out to families and the community to inform them about the QRIS ratings) from which information can be collected and used to adapt or build a database on subjects related to the QRIS (the community, child care

providers, parents, and children). That information can be used in the evaluation (for example, as a baseline or initial measure, or to track changes in implementation and/or outcomes over time). Therefore, the data collection plan can involve devising ways to use and examine existing tracking systems, investigating why certain data on the QRIS are collected and how they are used, and thinking critically about the types of data that the QRIS and its stakeholders need but have not collected in consistent ways.[15] Thinking about data collected in-house also encourages discussions about what relevant data should be collected on the different components of the QRIS across different stages of implementation so that it can be used to inform the QRIS operation and progress toward goals

Table III.10. Key Activities to Develop a Data Collection Plan

Activity	Considerations
Identify data elements and sources for these data	The data elements that have to be collected should be appropriate to answer the evaluation research questions
Select measures and create instruments	Measures and instruments should be appropriate to answer the evaluation research questions
Plan a schedule for all data collection activities	The schedule should meet the budget and time constraints of the evaluation. Data collection activities include obtaining IRB approval, assembling a data collection team, providing necessary training to data collection staff, and data collection (in the field, by phone, by mail, or by other methods, for example, online)
Make sure that data are complete	Have a plan for reaching and obtaining data from all of the targeted evaluation subjects. If the evaluation will use administrative data, then the data collection plan can establish steps to check that these data include all the subjects and information that are relevant to answer the research questions. If surveys will be used in the evaluation, the data collection plan can identify strategies to clear minimum targets for response rates, set rules about when efforts to collect data from individuals will be terminated, and what constitutes a refusal to participate overall or during the current data collection period

In addition, efficient data collection plans do not focus on collecting all types of data simply because it seems feasible to do so. Instead, good data collection plans make clear how each piece of

[15]Data from a QRIS monitoring or tracking system can contain valuable information on implementation and on the progress of QRIS activities in reaching goals (for example, the number of child care providers that have been assessed and rated), so using those data can help in reducing the costs of the data collection efforts.

data will be used in the evaluation, how it fits with other pieces of data, and how it will contribute to answering the evaluation questions.

In summary, the optimal data collection plans are focused on obtaining only the data that are relevant to answer the research questions of the QRIS evaluation and build on using data that have already been collected as part of any existing QRIS activities (that is, data from the QRIS tracking system).

E. Considerations for an Adequate and Credible QRIS Evaluation

In this section, we present key factors to consider when verifying that the evaluation design for your QRIS will adequately address the research questions of interest to the QRIS stakeholders and that can be helpful in ensuring that the evaluation will produce results that are credible. The adequacy of the evaluation design and the credibility of the findings depend on several factors, including the following (Bamberger & Rugh, 2008):

1. **The evaluation focus, approach, and methods are aligned with the types of information needed.** If state administrators, QRIS staff, or other stakeholders have information needs that cannot be addressed by the selected design evaluation, then that design would not be adequate.
2. **Data and data sources are available.** If the data needed to implement a certain evaluation design do not exist or cannot be generated, or if the appropriate data sources are not available to the evaluators, then that evaluation design would not be adequate.
3. **The data will support interpretations about the performance of the QRIS and its influence on child care providers, families, and child outcomes.** If the data needed to implement a given evaluation design do not provide information on achievement of the QRIS objectives, the extent of delivery of QRIS services, or the cost-effectiveness of QRIS activities, then that evaluation design and the findings it could yield would not be adequate.
4. **The evaluation team possesses the necessary qualifications.** If the evaluators do not have the technical expertise, knowledge of the ECE field, and resources necessary to implement the selected QRIS evaluation design, then the findings from the evaluation would not be adequate.

Checking that the methodology used in the evaluation is sound is also a key factor in verifying that the findings of the evaluation have not been compromised. A well-designed evaluation has a

strong approach to the following issues (Bamberger & Rugh, 2008; Shadish, Cook, & Campbell, 2002):

1. **The evaluation is objective.** The conclusions of the evaluation are drawn from the available evidence and the research is free from bias.
2. **The evaluation is reliable.** The research process for the evaluation is being carried out faithfully and consistently.
3. **The conclusions are valid.** Correct conclusions about the program are drawn from the findings of the evaluation.
4. **Constructs are valid.** Key constructs or concepts (for example, concepts such as quality of the environment in child care settings, highly skilled child care provider, and child cognitive outcomes) are defined and measured with precision and clarity.
5. **The evaluation has internal validity.** Given other plausible explanations, the effects found by the evaluation can be attributed with confidence to the program or intervention being examined.
6. **The evaluation has external validity.** The findings of the evaluation can be applied to different subjects, settings, or periods of time, that is, to the extent possible, the findings of the evaluation can serve to inform the evaluations of other QRISs.

F. Communicating the Findings of the Evaluation

As with other aspects of the evaluation (that is, the planning, data collection, and analysis stages), optimal strategies to communicate the findings of the evaluation are focused on providing answers to the questions that were identified as the key issues to be addressed by the QRIS evaluation. That is, when communicating evaluation results to stakeholders, the messages should be based on the conclusions about the program that are supported by the data and are in line with the methods used in the evaluation. For example, when communicating evaluation results, it is recommended to acknowledge whether the findings of the evaluation are purely descriptive or whether they might suggest patterns in outcomes but cannot attribute such patterns to the QRIS. Outside forces (for example, specific stakeholders) could also push for answers to questions about the QRIS that cannot be answered by the evaluation approach taken. In those cases, it is important to keep the message aligned to conclusions that the research design can draw.

Additionally, an earnest approach to communicating the evaluation findings involves (1) reporting *all* the findings, even when they were not the desired findings for the QRIS initiative (for example, the findings demonstrate implementation gaps); (2) being clear whenever some results are not conclusive and they warrant additional investigation; and (3) if potential explanations for the findings are presented, they should provide all plausible alternatives and it should be made clear that these potential explanations were not explored in the evaluation and should be the subject for future research.

Finally, in communicating the evaluation findings, it is good to keep the target audience in mind. For instance, state administrators, funders and potential funders (federal and local agencies, national and local foundations, other organizations), and/or QRIS administrators and other staff might be particularly interested in the findings regarding the implementation of the QRIS initiative (that is, findings on *how* the initiative works) and the extent to which the initiative can be linked to changes in outcomes for the evaluation subjects. Federal agencies, funders or potential funders, and the ECE research community might also be interested in learning about specific characteristics of the initiative being evaluated, the evaluation design and methods, and the analyses that were conducted. Community agencies and families might be interested in learning about how the QRIS initiative works and whether it has influenced outcomes for children and the community. Different products for different audiences might be necessary to present the same evaluation results in ways that are most meaningful.

IV. CHOOSING MEASURES AND DATA COLLECTION BEST PRACTICES

If the research questions and evaluation design serve as the foundation for a QRIS evaluation, the data sources and the measurement tools and strategies are the critical building materials for the evaluation structure. The measures and information collected in a QRIS evaluation will be used to answer the key research questions and will thus determine the ultimate conclusions that are drawn about how the QRIS works and the extent to which the initiative has reached its goals.

This chapter provides details about the measures used in a QRIS evaluation, including an overview of how measures are linked to the key domains and constructs in the evaluation, the process of selecting appropriate measures and measurement strategies, the specific types of measurement tools that can be used, and guidelines for data collection.

CHAPTER CONTENTS

A. Identifying Constructs and Data Elements

A first step in selecting measures to use in an evaluation is to develop a set of measurable constructs that relate to the research questions being asked. A *construct* is a characteristic or concept that can be measured using one of the qualitative or quantitative measurement strategies described in Chapter III. For example, if a research question asks "What are the characteristics of providers participating in the QRIS and how are they different from providers who don't participate?" the constructs needed to answer the research question will relate to provider characteristics. These constructs may include characteristics such as type and geographic location of provider and proportion of low-income children served. Together, this group of related constructs makes up a

domain. Because there are a wide range of constructs that could be examined in the domain of provider characteristics, the inputs and activities sections of the logic model will provide important guidance about which constructs should be selected to address this research question (see Chapter II for more details on building a logic model for your QRIS).

A similar process should be used for each of the research questions in the evaluation to create a set of concise constructs that can guide measures selection. The constructs selected for inclusion should be tightly linked to research questions and represent a manageable set of data elements that can be obtained feasibly and reliably. *Data elements* are the individual units of information that pertain to each construct.

After delineating the broad constructs that link to the research questions, the next step is to develop a measurement strategy. This involves further specification of the constructs so that a data source can be identified and the data elements within the construct can be measured. In the example provided earlier, the three constructs related to provider characteristics have to be further specified to clarify those dimensions that are important to measure. For example, provider type could be classified as center-based programs (including Head Start and state pre-kindergarten programs), licensed family child care programs, or other programs such as license-exempt providers. Geographic location of providers could be classified as urban, suburban, or rural. Proportion of low-income children served could be specified as the percentage of enrolled children in a program that receive child care subsidies. Additional details might be needed for each of these constructs to specify more clearly how they will be measured. For example, further definition would be needed to specify how the percentage of children receiving a subsidy is measured and calculated. It could be measured as an average over a quarter, over six months, or over a year, and details such as these would be needed to specify the measure for data collection and analysis.

Table IV.1 provides examples of common research questions and the domains, constructs, and data elements that can be identified for each. These examples are intended to be illustrative, not exhaustive or prescriptive, of how data elements should be defined.

Table IV.1. Examples of Research Questions and Corresponding Domains, Constructs, and Data Elements

Research Questions	Domains and Constructs	Specific Data Elements
What do key QRIS stakeholders view as the major implementation challenges and successes during the pilot and scale-up stage?	Stakeholders' perceptions: • Stakeholders' perceptions of challenges • Stakeholders' perceptions of successes	• (Specify *who* stakeholders are: QRIS administrators, program staff, contractors, staff from support agencies, legislators) • Stakeholders' perceptions of challenges and or success related to recruitment, rating, supporting quality improvement, distribution of incentives
What are the characteristics of providers participating in the QRIS?	Program/provider characteristics: • Provider type • Geographic location • Proportion of low-income children served	• Provider type = center-based programs (including Head Start and pre-kindergarten), licensed family child care programs, and legally non-licensed providers • Geographic location = urban, suburban, or rural • Proportion of low-income children served = average percentage of children enrolled over past 6 months who receive a child care subsidy
What was the focus and intensity of the trainings and on-site assistance provided by the QRIS?	Focus of training and on-site assistance: • Goals • Content Intensity of training and on-site assistance: • Frequency • Duration • Length	• Goals = overall improvement, specific quality standards • Content = curriculum, family partnerships, environment • Frequency = number of visits/training sessions • Duration = hours per session • Length = number of weeks/months
Do child care programs/providers receiving QRIS QI support achieve high ratings? Are high quality ratings for providers sustained over time?	Receipt of QI services: • Low • Medium • High Rating of provider over time	• Low QI = less than one month • Medium QI = two to six months • High QI= more than seven months Provider rating at Years 1, 2, and 3
Does the QRIS increase parents' access to information about quality of child care providers?	Parent access to QRIS: • Awareness of service • Access to written documents • Access to website Change over time	• Awareness of service = percentage of parents reporting that they have heard of the QRIS • Access to written documents = number of brochures that are distributed to parents • Access to website = monthly web traffic by new users • Change over time = examination of data elements at Years 1, 2, and 3

Research Questions	Domains and Constructs	Specific Data Elements
Has attending child care programs participating in the QRIS resulted in improved developmental outcomes for at-risk children?	Children's developmental outcomes: • General cognition and knowledge • Language • Approaches toward learning • Social and emotional development • Physical and motor development Attendance in QRIS program At-risk children	• Cognition = problem solving skills • Language = vocabulary and print awareness • Approaches toward learning = persistence • Social and emotional development = competence with peers, emotion regulatory skills • Physical development = gross and fine motor skills • Attendance in QRIS program = documented attendance over nine or more months • At-risk children = children receiving subsidy in past six months who attend QRIS-rated programs

B. Linking Constructs and Data Elements to Data Sources and Measurement Strategies

After constructs and sets of data elements have been specified, they must be linked to data sources and measurement strategies. For example, to measure the set of provider characteristics described earlier, the evaluator must identify a source of the information and a strategy for collecting information from that source. The data elements related to provider type, geographic location, and proportion of low-income children served might have been captured already during programs' enrollment in the QRIS if they filled out an application form, and these records can be shared with the evaluators. If not, evaluators might have to distribute a self-administered questionnaire to programs to collect this information. In general, measuring data elements in a QRIS evaluation will involve the use of existing data sources or new data collection using the qualitative and quantitative research methods described in Chapter III.

Evaluators can consider multiple existing data sources for measuring QRIS data elements. Table IV.2 provides an overview of possible data sources and an assessment of key advantages and challenges that should be recognized when using data from each source.

Table IV.2. Possible QRIS Evaluation Data Sources

Data Source	Advantages	Challenges
Child Care Licensing Records	• Contain information on programs' licensing status and violations, including those that involve child maltreatment	• Need for a unique identifier for each facility so that licensing data can be linked to QRIS facility-level data
QRIS Application/Enrollment Forms/Records	• Provide information on basic provider-level characteristics and dates of QRIS enrollment	• Facility-level data (name, address, and other characteristics) might not align precisely with data in the licensing or subsidy databases • Must ensure standardization of data elements that providers self-report (e.g., use of category selection on the application) for reliability and consistency
QRIS Ratings Records	• Provide details on the components used to calculate ratings and the rating assigned in each rating period	• Might contain aggregate data elements, not individual components (e.g., the score or points assigned for staff qualifications rather than the actual qualifications of individual staff members) • Historical data might not be available
Child Care Quality Observation Records	• Provide item-level detail from observational measures	• Need for a unique identifier for each facility so that classroom data can be linked to QRIS program data • Historical data might not be available • Observations might be done only at higher rating levels • Observations might not be done on all classrooms and/or caregiving staff • Observations might be done less frequently than annually
Technical Assistance/Quality Improvement Records	• Provide record of services provided or incentives that have been distributed	• Data might not be entered consistently across regions or agencies that are responsible for technical assistance • Database might not be centralized; rather, data elements might reside across different databases and in different formats (e.g., ACCESS database, Excel spreadsheets, etc.)
Provider Registry and Career Lattice Records	• Provide data about preservice and ongoing education and training; in some registries, data are verified by checking transcripts and attendance records	• Might not contain all individual caregivers (e.g., all caregiving staff at a center or every family child care provider) • Need for a unique identifier for each individual caregiver and facility to enable links between individual staff and the facility-level data that reside in the QRIS or licensing systems
Child Care Resource and Referral Records (program data)	• Possible to examine the characteristics of child care programs/providers across communities or by provider type	• Might not contain information for all child care providers and programs in the area serviced by the QRIS • Might be updated only at scheduled intervals

Data Source	Advantages	Challenges
Child Care Resource and Referral Records (parent data)	• Provide information on number and types of parent inquiries (potentially how they heard about the QRIS)	• Records represent a subset of parents who use referral services
Child Care Subsidy Payment Records	• Provide a direct count of how many children are paid for by subsidy in each program	• Need for a unique identifier at the facility level to link the number of subsidized children and amount of subsidy paid to the program with the QRIS facility-level data • All fields in the subsidy data might not be used regularly and might not be reliable
K–12 Education Records	• Provide data to link children's experiences in QRIS programs to educational outcomes	• Need for a unique identifier at the child level to identify children across systems to link their experience in the QRIS with outcomes in K-12
Early Learning/School Readiness Data	• Provide data to link children's experiences in QRIS programs to early learning outcomes	• Need for a unique identifier at the child level to identify children across systems to link their experience in the QRIS with assessment data • Use of early assessment tools might be voluntary or required only at the highest QRIS rating levels, representing a select subset of children

When using existing data sources, it is important to understand fully the applicability, coverage, and reliability of the data. Doing this before using the data in analyses or reporting will help to avoid making incorrect assumptions about how the data can be used in the evaluation, drawing inappropriate conclusions about the findings, and spending limited time and financial resources cleaning and preparing for analysis poorly specified or inconsistent data. Table IV.3 provides a list of questions to ask about the data you are considering including in your QRIS evaluation, organized by data principles that will support a successful analysis. Answering these questions can help determine if your data are appropriate for and ready to be used in the analysis.

Table IV.3. Questions to Ask About Potential Data Sources for QRIS Evaluation

Questions About the Data Source	Your Answer
Coverage—How Comprehensive Are the Data Across the Population of Interest?	
• What populations and/or geographic areas are covered in the data? • Is inclusion in the database mandatory or voluntary? What other inclusion or exclusion criteria exist?	
Reliability—Are Procedures in Place to Ensure Consistency and Quality in Data Entry?	
• What process is used for entering data? Does a written data entry manual exist? • What process is used to check and verify the accuracy of the data?	
Applicability—Does the Unit of Analysis Address the Research Question?	
• At what level will the data be provided? Can the data be aggregated at the level of child care facilities? Can data be available about individual classrooms? Can data be available about individual staff within classrooms? • If data are provided from structured measures (such as observational measures of global quality or measures of child development), are the data available at the level of individual items, by subscales, or by a total score? • What options exist for linking data at the level of facility, individual caregiver, family, or child?	
Feasibility—Are the Data Accessible and in a Usable Format for Analysis?	
• Are the data publicly available? If not, how is permission granted to access the data? Is IRB approval required? • What type of file can be produced from the data? • Is there a codebook to provide information about the data elements and the data structure?	
Confidentiality—How Will the Data Be Protected?	
• Can the data be de-identified? If not, what procedures are in place to protect the privacy of the data?	

IRB = institutional review board.

C. Types of Measurement Tools

This section describes in more detail the types of measures and measurement tools that can be used to tap particular data elements in an evaluation plan. The measures are described in two broad categories: *implementation* (or process) and *outcome* measures. *Implementation measures* describe or assess features of the QRIS and the extent to which key QRIS activities are being carried out as intended. *Outcome measures* assess the goals achieved with the QRIS through child care program outcomes or family or child outcomes. Each set of measures is linked to potential research questions described in

Chapter III to serve as examples. Appendix C provides a summary of implementation and outcome measures used in evaluations of several existing QRISs.

1. Implementation Measures

A QRIS can involve multiple activities, such as collection of documentation from programs, designation of ratings, provision of quality improvement services and financial incentives, and promotion of the QRIS to parents and programs. Some of these activities will build on existing structures and processes and others will be new activities. QRIS implementation is often launched with tight time lines and expectations from funders and policymakers for quick scale-up (Zellman & Perlman 2008). Therefore, measures are needed to assess the delivery and coordination of services and whether intended features are included as planned with fidelity and in accord with lessons from the field of implementation science.[16] Sample research questions that are important for implementation of a QRIS, their corresponding constructs and data elements, and possible data sources and measurement strategies are discussed later in this chapter. These implementation questions are meant to be illustrative of the range of QRIS evaluation strategies, but represent only one set of possible approaches. The implementation issues addressed here are fidelity of the quality rating process, providers' experiences in the QRIS, the provision of QI services, the provision of incentives and outreach, and dissemination of information to parents and the public.

In some cases (as noted later), the research questions about implementation/process and the corresponding constructs, data elements, and measures might best be framed as performance management and ideally would be conducted by the QRIS program management or internal evaluator as part of an ongoing review and monitoring of program performance (Walker & Moore 2011).

[16] Throughout this section, recommendations will be drawn from the synthesis of implementation science literature completed by Fixsen and colleagues (2005) and other resources from the National Implementation Research Network.

Fidelity of the Quality Rating Process. An important implementation question to assess in the pilot and scale-up stage is the fidelity of the quality rating process ("How are the QRIS standards and indicators measured and are ratings calculated with precision and integrity?"). Possible constructs for this question could include methods of data collection and sources of evidence for each quality rating component from programs, rating accuracy, and rating reliability.

Data collection from programs could be further specified by defining the tasks that are expected to ensure an efficient process that functions to protect the integrity of the ratings through the methods for documenting that specific components of the rating have been met. For example, such tasks might include the following:

- QRIS staff roles and responsibilities for data collection and rating are clearly defined
- QRIS staff receive initial and ongoing training on data collection processes and allowable sources of evidence for each component
- Expectations for documentation that child care programs complete and other data collection (observations, provider registry data) are clearly described for programs and for QRIS staff
- Programs understand the QRIS forms and document requirements
- A supervisor is available to monitor QRIS staff as they review program documents and assign ratings

Rating accuracy examines the strategies used to ensure that the data produced in the measurement of QRIS standards and used in the rating are accurate. Further specification of this construct could include the following:

- A manual or written guidance exists to explain scoring of documentation and decision rules for QRIS staff (especially rules for exceptions to scoring procedures)
- A process is in place to ensure the accuracy of data entry (for example, double keying or automatic reminders for fields that are left blank)
- Contractors or outside agencies that submit data to the QRIS follow the same data procedures for ensuring accuracy

Similarly, *rating reliability* includes activities to systematically check for the consistency with which individual raters score documentation or other materials submitted by QRIS programs. The constructs to assess rating reliability could include the following:

- A process and schedule exists to assess the inter-rater reliability of QRIS staff who score program documentation and materials or who conduct observations in the field
- A process exists to establish consensus on difficult rating items and relay decisions to other staff
- Contractors or outside agencies that submit data to the QRIS follow the same data procedures for ensuring reliability

Evaluators (and QRIS managers) can assess these process components using the following qualitative measurement strategies:

- Systematic review of program documents including training materials and data manuals
- Self-administered questionnaires, structured or semistructured interviews, or focus groups with QRIS staff
- Self-administered questionnaires, structured surveys, or semistructured interviews or focus groups with programs participating in the QRIS

When selecting among self-administered questionnaires, individual interview methods (either structured or semistructured), and focus group methods, evaluators can consider the type of information that would be most useful for addressing the research questions and the resources available for evaluation. Questionnaires can be administered and scored relatively inexpensively using web-based methods (or mail) and can reach a large number of providers. In contrast, interviews will produce in-depth data on the decision-making processes or experiences of an individual, but the process is more time and labor intensive than distributing self-administered questionnaires. Likewise, focus groups will provide in-depth information across a range of participants and the experience of participating in a discussion with other participants could spark contributions or experiences that participants might not have recalled in an individual interview. Each of these methods of collecting information requires either a trained interviewer (structured or

semistructured interviews) or a skilled facilitator (focus groups) and well-planned protocols to elicit the information that will be most relevant.

Providers' Perceptions of the Rating Process. Early in implementation of a QRIS (during the pilot, scale–up, or early operation stage), it is useful to understand providers' experiences with the rating process ("Do providers understand the rating system and feel that the process is fair and accurate?"), because these experiences have implications for providers' willingness to participate in the QRIS and to maintain participation over time. QRIS staff can use the information obtained to modify interactions with providers that are participating in the QRIS and to determine new strategies for recruiting and enrolling providers. When sampling providers, evaluators might consider selecting a sample of providers that has chosen to participate in the QRIS as well as providers that have chosen not to enroll in the QRIS. Possible data elements for these questions include providers' understanding of the rating system and their perceptions of the fairness of the rating process.

Both qualitative and quantitative strategies could be used to assess understanding and perceptions. Providers could answer questions using a Likert scale (so that an average understanding or perception of fairness could be calculated) and/or open-ended items on a telephone survey (which can be coded to describe the most and least common responses). Focus groups of providers would also provide insights into providers' decision-making processes related to participation in the QRIS and the points in the rating or QI processes that work smoothly or that challenge providers.

When selecting providers to participate in evaluation efforts, sampling of providers enrolled in the QRIS could target all participating providers or a subset of providers based on particular criteria (for example, star level or timing of enrollment in the QRIS). For providers who have not yet enrolled, it might be important to distinguish among those who do not know about the QRIS compared with those who have made a deliberate decision not to enroll. Sampling strategies of

providers who have chosen not to enroll in the QRIS could target providers who have requested enrollment materials or attended orientation sessions but who deliberately decided not to participate.

Quality Improvement. Implementation questions about quality improvement activities assess the staffing and implementation of the QI services being delivered by the QRIS. The evaluation can provide critical information about the degree to which QI services are being delivered as intended in accord with the service model used as well as QRIS staff and providers' experiences with QI ("How do QI specialists perceive the effectiveness of their services?" and "What are providers' experiences with QI services?").

These questions can be tapped by measuring constructs such as (1) the QI process, (2) staff perceptions of the QRIS QI model, (3) providers' understanding of QRIS QI services, and (4) providers' perceptions of the value of QI supports. We provide more details on measuring these constructs later in this chapter.

The *QI process* could be further defined and documented by measuring the tasks necessary to deliver effective quality improvement services:

- QRIS QI staff are hired with sufficient skills and competencies to deliver QI services
- QRIS staff receive initial and ongoing training in the QI service model
- QRIS staff roles and responsibilities for QI services are clearly defined
- QRIS staff deliver the intended dosage and intensity of QI services
- QRIS staff and supervisors can track and monitor QI service delivery to programs
- QRIS supervisors are available to assist with questions, provide resources, and promote reflection on practice

Staff perceptions of the QRIS QI model could be measured by assessing perceptions related to adequacy of training, match between QI approach and providers' needs, caseload management, and access to supports and resources as needed. *Providers' understanding* and *providers' perceptions of QI supports* could be measured by assessing awareness of QI services and perceptions of adequacy and usefulness of the services.

Data sources and methods to address constructs related to QI services might include training materials (which can be reviewed to assess the process used to train QI service staff), manuals for QI specialists (which can be used to assess the resources and supports available to QI staff), materials and forms distributed to providers, a case management data system to track dosage and intensity of services, self-administered questionnaires (with staff or with providers), and interviews and focus groups (also with staff or providers).

Financial Incentives. Implementation questions about financial incentives assess the types of payments made to providers (such as to support capital improvements, to reward movement up the rating levels, or to encourage enrollment of children from low-income families); the efficiency and accuracy of delivery of incentives; and providers' experiences with the payments ("What incentives are offered to providers?" and "What are their perceptions of the incentives?"). Examples of constructs to specify for financial incentives include timely and accurate delivery of incentives and providers' perceptions of incentives.

Timely and accurate delivery of incentives could be tapped by a review of program documents to assess whether incentives—such as tiered reimbursement, grants, materials, scholarships and bonuses—are delivered as intended:

- Written documentation outlines procedures for processing and delivery of QRIS incentives
- QRIS staff receive training on financial incentives
- QRIS staff roles and responsibilities for distribution of financial incentives are clear
- QRIS staff and supervisors can track and monitor distribution of financial incentives to programs

Providers' perceptions of financial incentives could be measured by their understanding of rules and procedures for financial incentives and perceptions of fairness and effective distribution of incentives.

Outreach and Dissemination. Implementation questions about dissemination focus on the degree to which information about the QRIS is developed and disseminated as planned to providers and to parents/consumers ("How are marketing materials about the QRIS disseminated to providers and to the public?" and "Do parents and the public understand the rating system and know how to interpret the ratings?"). Constructs to assess dissemination could include the dissemination process, parents' QRIS understanding, and public perception of the QRIS.

The *dissemination process* could be further specified by defining and measuring or documenting the tasks needed to ensure adequate promotion and marketing of the QRIS:

- QRIS staff roles and responsibilities for dissemination are clearly defined
- A strategic plan outlines marketing and website efforts and time lines
- Tools such as web analytics are in place to monitor website traffic and downloads of QRIS materials

Parents' understanding of the QRIS could be defined further to specify different aspects of awareness and perception of the QRIS and the ratings. For example, the following aspects could be defined:

- Parents know of an organization from which they can obtain information about child care quality
- Parents know the QRIS by name
- Parents understand the rating categories and what overall star ratings mean

Public perception of the QRIS could be defined to specify perception of the QRIS among key stakeholders, such as legislators, staff from early childhood support organizations (such as provider associations and professional organizations), and organizations such as child care resource and referral agencies that offer training and other support to practitioners.

Evaluators could assess these constructs by reviewing program documents and through interviews and/or focus groups with parents or key QRIS stakeholders.

With questions related to the implementation of outreach and dissemination strategies, as with all implementation questions, it is important to align the research questions and the implementation measures that are assessed with the stage of the QRIS and the activities that are conducted. It is important to ensure that measurement does not occur too early in the process, before implementation activities are fully in place.

2. Outcome Measures

The second category of measure that a QRIS evaluation can include is outcome measures. Depending on the explicit goals of a QRIS, multiple levels can be targeted by QRIS activities, including communities (or markets), programs/providers, parents, and children. *Outcome measures* assess the degree to which QRISs are linked to positive outcomes across these levels (with only certain evaluation designs permitting attribution of causality, as discussed in Chapter III). As noted in the description of process measures, the selection of outcome measures in a QRIS evaluation should directly link to the QRIS logic model and the research questions for the evaluation. For example, if there are no specific activities (such as marketing or outreach efforts) aimed at increasing parents' use of the QRIS, using evaluation resources to investigate outcomes for parents would not make sense. Assessment of outcomes should be conducted only for those outcomes explicitly targeted by the QRIS and/or to understand potential unintended consequences of QRIS activities.

Communities/Markets. Research questions related to community- or market-level outcomes in a QRIS might assess programs' uptake of the QRIS and the extent to which vulnerable children receive care in rated programs ("What proportion of the programs in a community participate in the QRIS? How many vulnerable children are served in QRIS-rated programs?"). Further specification of the data elements related to community-level outcomes could include the following:

- Number and proportion of community programs participating in the QRIS
- Number/percentage of slots available in QRIS-rated programs
- Number of children receiving subsidies served in QRIS-rated programs

- Percentage of slots in QRIS-rated programs that are filled by children receiving subsidies

Data sources to assess these data elements are most effective if they contain information that can be aggregated at a community level. For example, child care resource and referral databases typically contain facility information (including number of slots) on licensed programs across a state or in a specific geographical area. Records from a licensing database could also provide this level of detail (including number of available slots). Data from resource and referral or licensing databases typically contain addresses so that data can be aggregated at the level of city, zip code, county, or other designation (such as economic development region). These data can provide the denominator when calculating program participation or percentage of slots available in QRIS-rated programs.

Subsidy payment data also offer a means of examining community-level outcomes. These data must be linked with the facility-level QRIS data to estimate the average number and proportion of subsidy-receiving children participating in QRIS-rated programs. It is important to determine the time frame over which the averages are calculated and to know if the subsidy payment data represent hours authorized for care or actual attendance at a QRIS-rated program.

Programs/Providers. Research questions related to outcomes for programs and providers can address a wide variety of constructs depending on the QRIS goals. These goals are likely to address *overall movement up the QRIS rating levels*, but they might also specify particular desired outcomes connected with specific components on which the QRIS ratings are based. Such goals might also be related to the QI activities being implemented by the QRIS. For example, desired outcomes might include increased percentage of staff with certificates, credentials, or training hours in a particular topical area; increased use of approved curriculum or assessment tools; increased scores on the environment rating scales or other observational measures; increased frequency of practices to promote relationships with families; improved business and administrative practices; and attainment of professional accreditation (for example through the National Association for the Education of Young Children [NAEYC]).

It is likely that existing QRIS ratings data can be used to assess the constructs for research questions related to program and provider outcomes. The outcomes targeted by the QRIS typically are those that are assessed and rated as part of the QRIS. Thus, data on these outcomes is likely available in the existing QRIS data. However, if these data are not already collected by the QRIS, measures and a data collection processes will have to be identified. Depending on the resources available for evaluation, the strategies for gathering data could include the following:

- Self-administered questionnaires with staff and directors/family child care providers to assess professional development outcomes or accessing records available in a provider registry. Outcomes to consider include the following:
  - Participation in training linked to the QRIS
  - Attainment of certificates or credentials
  - Enrollment in credit-bearing courses
  - Preparation of a professional development plan
  - Participation in QI coaching, consultation, or mentoring (number of contact hours)
- Interviews with staff to gather information about changes in program practices:
  - Use of a new curriculum or assessment tools
  - Participation in goal setting for QI
  - Purchase of new materials
- Observations of practices using the environment rating scales or other observational tools, depending on the practices targeted by the QRIS quality improvement services. Information about available observation tools and the constructs they assess can be found in Appendix C. In general, observation tools may be used to assess the following:
  - Global quality (looking across routines, the environment, materials, and interactions)
  - Specific features of interactions between teachers and children, such as sensitivity, emotional support, and instructional practices
  - Practices to support specific developmental domains, such as language and literacy, social-emotional development, math and science, and health
- Self-administered questionnaires to gather information about practices with families and administrative practices. Practices assessed might include the following:
  - Communication strategies
  - Provision of resources and information
  - Opportunities to support interactions between staff and families

  - Beliefs and attitudes about families
- Accessing national databases to document and verify attainment of accreditation status

Because of the large number of constructs that are possible to measure in the domains described earlier, it is important to clearly specify a concise set of desired program outcomes that align with key research questions and QRIS activities.

Families. Research questions targeting QRIS outcomes for families ask about the extent to which the QRIS serves as a resource and decision-making tool when selecting early care and education arrangements for their children ("Do families use the QRIS to select care and education arrangements for their children?"). Questions can also address the potential of the QRIS to influence families' preferences, such that they prefer and seek arrangements that are higher quality ("Do families shift their priorities to seek higher-rated arrangements?"). Constructs to address these outcomes might include the following:

- Families' use of the QRIS website
- Families' preferences for higher-rated programs
- Patterns of enrollment in QRIS-rated programs

There are numerous challenges with the assessment of QRIS outcomes related to families. First, data are not typically collected from families as part of the QRIS. Thus, answering research questions about families usually requires investing in new data collection, either through self-administered questionnaires or through telephone interviews. Second, identifying and assessing the population of interest is difficult. Understanding whether and how the QRIS is used by all families with young children in a geographic area might require a household-based survey with screening to determine those households with eligible children that use nonparental care. This approach would provide helpful estimates of family QRIS use and preferences, but it is expensive to conduct. Alternatively, focus groups could be conducted with populations of special interest to the QRIS, and participants could be identified through connections with community services such as pediatricians'

offices or social service agencies. This method is also labor intensive because of the need for recruitment, logistics planning, expert facilitation, and data analysis.

As noted earlier, it is also important to assess the extent to which activities in the QRIS have been conducted to target families' decision-making process and to increase their familiarity with the QRIS quality indicators. If intentional efforts have not been undertaken, it would not be wise to pursue extensive data collection with families (other than to conduct a needs assessment with families).

Children. The inclusion of QRIS outcomes for children in an evaluation design requires careful planning and a sound measurement strategy. Elicker & Thornburg (2011) recommend that children's outcomes be assessed in a QRIS evaluation only in the mature operation stage and only if clear, conceptual links have been made between activities in the QRIS and expected outcomes for children. Ideally, the assessed child outcomes align closely with the QI activities that have been implemented. In addition, procedures should be in place to collect demographic data and other characteristics that are known to relate to children's developmental outcomes so that these can be controlled in analyses. Recommendations for using child assessments appropriately can be found in a volume published by the National Academy of Sciences (Snow & Van Hemel 2008).

For evaluations that focus on preschool-aged children, it is recommended that a set of measures be selected to tap the multi-dimensional domain of school readiness as defined by the National Education Goals Panel (Kagan, Moore, & Bredekamp, 1995) and subsequent conceptualizations (Love, Aber, & Brooks-Gunn, 1994; Love, 1999; Snow & Van Hemel 2008; U.S. Department of Health and Human Services 2010d). A comprehensive school readiness framework includes five dimensions: (1) cognition and general knowledge, (2) language development, (3) approaches toward learning, (4) social and emotional development, and (5) physical well-being and motor development. Including measures to assess the five constructs in this framework will produce a more complete picture of children's knowledge, skills, and abilities.

Within the framework of school readiness, there are three primary strategies for assessing children's development. The first is a comprehensive approach to combine direct assessments and teacher/parent ratings of developmental domains. In this approach, direct assessments and/or observations are conducted by trained research assistants using standardized tools that measure language and literacy skills, premath and numeracy skills, and general cognitive skills. Teacher and/or parent ratings are used to collect information about other developmental domains—such as approaches to learning, social-emotional development, and health outcomes—that are not easily assessed using direct standardized assessments. It is most effective to combine the direct assessments with ratings from teachers and parents because it offers a more complete picture of children's development across multiple domains and a picture of individual strengths and needs.

A second approach is to use teacher or caregiver ratings of children's development that are based on portfolios or collections of children's work. A portfolio can be used for assessment if the purpose of assessment has been defined, the criteria and methods for determining what is included in the portfolio are clear, and the criteria for using the items in the portfolio to make judgments about performance are specified. These assessment tools require specific training for teachers and caregivers and are based on intentional observation and collection of documentation of children's skills and competencies on an ongoing basis.

A third approach is to access children's school records when they enter kindergarten. Standardized assessments are typically not available for children until grade three. However, if school districts use the same instrument to assess children upon entry into kindergarten, these records could be available to assess children's skills and abilities at school entry.

Each approach to assessment of child outcomes has advantages and challenges, which are outlined in Table IV.4.

Table IV.4. Advantages and Challenges of Using Different Approaches to Assessment of Children's Development

Approach	Advantages	Challenges
Combination of direct assessment and teacher and parent ratings of developmental domains	• Uses multiple modes of collecting information that might provide a more accurate representation of children's abilities • Uses multiple reporters which allows for an assessment of children across different settings • The same tools can be administered across different programs	• Direct assessments are expensive to collect because they require a research assistant to spend one-on-one time with individual children • Direct assessments are not available in multiple languages (although some tools are available in Spanish) and might not have been validated across diverse cultural groups • Direct assessments rely on performance at one time point that might be influenced by other contextual factors • Reliability of parent and teacher ratings is not typically assessed and it is likely that there are inconsistencies in how different parents and teachers rate young children on different skills and abilities
Portfolio assessments	• Use information collected across time and activities to assess children's skills and abilities, which might increase the ability of the assessment to measure the skills and abilities adequately	• Assessment is based on performance in one setting (program/school) and might underestimate children's abilities • Requires extensive training to ensure that assessments are being used consistently • Might burden teachers/caregivers if the assessment for the evaluation is different from a tool they are using already
Standardized assessments at school entry or later grades	• Assessments provide information about children's skills and abilities in a school setting • The cost of obtaining data that are collected routinely in school settings will be lower than new data collection	• School districts might not use the same tools, which would make comparisons difficult • Linking children's performance in school to their participation in QRIS settings might be difficult if school identification numbers were not assigned before school entry

In addition to the considerations outlined in Table IV.4, additional criteria should be used to guide the selection of measures of child outcomes:

- Age of the children to be assessed; assessment of outcomes for infants and toddlers is more challenging than assessment for preschool children given the rapid and inconsistent pace of development of children under age three
- The focus and activities of the programs; measures should be related to the skills and abilities and outcomes targeted by the QRIS
- Appropriateness of measures for children from diverse racial, ethnic, and socioeconomic backgrounds

- Languages of the children to be assessed; most direct assessments are not available in languages other than English and Spanish
- Adequate psychometric properties, including reliability and validity with children birth through 8 years old[17]
- Established validity and reliability for the intended mode of administration
- Prior use in other state evaluations and large-scale early childhood surveys
- The cost and requirements associated with purchasing and administering the assessments

D. Additional Measures Selection and Data Collection Guidelines

As noted throughout this chapter, a variety of guidelines should be attended to when identifying measures for a QRIS evaluation and developing data collection strategies. Table IV.5 provides an overview of key considerations and best practices across the different types of measures and data collection strategies (as well as general design issues) described in the chapter.

[17]Reliability refers to the consistency of the measurements yielded by the assessment or test. That is, the results of an assessment have reliability if they are stable (the measurements are not subject to random or unsystematic variation when the assessment is administered in repeated occasions), meaningful (the measure actually reflects children's skills and abilities), and can be reproduced. Validity refers to the extent to which the assessment measures what it is supposed to measure and thus it refers to the suitability with which inferences can be made based on the assessment results (Sattler 1992).

Table IV.5. Considerations and Best Practices in Measures Selection, Data Collection, and Overall Design of QRIS Evaluation

Topic	Key Considerations	Best Practices
Measures Selection		
Implementation	• Who are the QRIS staff responsible for implementation? • What activities are being conducted by QRIS staff? • Where is the biggest investment of QRIS staff time and resources? • What early feedback from stakeholders can be used to guide implementation revisions/modifications?	• Develop measures that provide information about staff characteristics, their qualifications, and their training to conduct QRIS activities • In addition to perceptions of staff and QRIS participants, review staff resource documents and training materials to measure key aspects of implementation • Develop feedback loops with the program to ensure that data can be used immediately to make revisions/modifications to processes
Communities/ Markets	• What measures can be collected at the level of the community/market and at what frequency should these measures be reported? • What unintended consequences should be examined (e.g., programs exiting the market)?	• Develop indicators of communities and markets that can be tracked over time at yearly, semiannual, or quarterly intervals • Consider tracking other economic indicators that can be used as context for the findings
Providers/ Program	• What new information about programs must be collected for the evaluation? • What is the burden of data collection on programs/providers participating in the QRIS?	• To reduce the burden on programs and providers, incorporate measures needed for the evaluation directly into QRIS application and enrollment processes
Parents/Families	• Will the sample of parents include only those who are using QRIS programs? • Are particular populations of families (income, language, culture) targeted by the QRIS? How will they be involved in the evaluation?	• Convene an advisory group to provide input on recruitment and measures selection • Use strategies to identify and collect data from a broad group of parents and/or particular subpopulations of interest
Children	• What domains of development will be measured? • Are tools appropriate for use with children in different settings?	• Use a multidimensional school readiness framework to determine the outcomes that will be assessed • Share aggregate results with community partners to encourage use of the findings and engagement in the evaluation
Data Collection Methods		
Existing Databases	• Which external data sources can be used for the evaluation? (see Table IV.2)	• Use the worksheet provided in Table IV.3 to determine which existing data sources can be most useful for the evaluation
Self-Administered Questionnaires	• What needs, interests, or feedback on services in the QRIS can the evaluation assess?	• Use self-administered questionnaires to gather information across a large number of respondents (e.g., feedback on marketing initiatives, awareness of the QRIS)
Structured Surveys	• What characteristics of participants and patterns of service use can the QRIS evaluation assess?	• Use structured surveys to gather targeted information from an identified population (e.g., parents in QRIS-rated program; directors or family child care providers in QRIS programs)

Topic	Key Considerations	Best Practices
Semistructured Interviews	• What processes and practices in the QRIS have to be understood?	• Use semistructured interviews to gather in-depth information about processes and practices (e.g., experiences with QI services; perception of provider needs)
Focus Groups	• What groups of respondents (and their experiences) are of interest to the evaluation? • How will the focus groups be recruited? • Are skilled facilitators available to lead the groups?	• Use focus groups to gather information about decision making and perceptions of processes and practices across identified groups (e.g., QI staff, family child care providers) • Develop clear protocols for recruitment and conduct
General Design and Method Issues		
Stage of Implementation	• Are measures being collected at the appropriate time in the development of the QRIS?	• Align research questions with stages of the QRIS implementation • Create a time line of data collection to ensure appropriate timing with QRIS maturity
Cultural Relevance	• Are measures and strategies for data collection appropriate for the population?	• Include members of cultural communities who can review materials in a technical advisory group for the evaluation or the QRIS in general • Ensure that protocols and surveys/interviews are translated and reviewed with a cultural lens
Protection of Human Subjects	• Have study procedures for data collection and measures been reviewed by an IRB? • Has permission/consent been provided for all data and study procedures?	• Submit evaluation protocol to an IRB for review
Training and Supervision of Evaluation Staff	• Have all research staff received adequate training to administer study tools and measures? • Do research staff meet regularly with supervisors and other staff to discuss evaluation progress?	• Establish a process for training staff and ensuring consistency and reliability in administration of tools • Maintain supervision of staff to monitor adherence to evaluation processes over time

This page has been left blank for double-sided copying.

V. EVALUATION COSTS AND FUNDING SOURCES

Evaluation is critical to developing and maintaining a successful QRIS because of its role in continuous program improvement and in identifying program benefits that in turn can strengthen public support for the program. However, evaluation research will require staff time and resources that are scarce when state, local, and non-profit budgets are strained. Meeting the resource challenge is a critical part of evaluation planning so that stakeholders can ensure that the QRIS is effectively managed and improved over time.

QRIS partners can use evaluation resources efficiently by balancing their evaluation objectives against their research and evaluation budgets. They can also seek outside funding for research and evaluation activities. In this chapter we discuss how the evaluation designs and data collection needs to address questions of interest lead to different levels of evaluation cost and provide examples of funding levels for actual QRIS evaluations to help in estimating what an evaluation is likely to require in terms of resources. We also discuss the sources of funding that have been used for QRIS evaluations across the country.

CHAPTER CONTENTS

A. Design Components that Influence Evaluation Costs

Evaluation research can have a large or moderate price tag, depending on the type of evaluation (implementation or outcome evaluation), the comprehensiveness of the research questions, and the extent of new data collection required to address the questions well. Chapter 3 discussed potential evaluation designs, including examples of research questions that could be addressed, and what conclusions can be drawn from each design. Chapter 4 discussed the measures that could be used in

evaluations. Here, we discuss how choices of evaluation design, measurement, and data collection strategy influence the cost of evaluation.

To illustrate sources of variation in the costs associated with different evaluation designs, Table V.1 provides examples of research questions and the types of data that could inform each question. The table presents implementation evaluation designs first, and then outcome evaluation designs. Within each evaluation design (for example, a process study) we provide one or two research questions as examples. Many of these questions have been the focus of QRIS evaluations in other states. The next columns describe the data that might be needed to address each research question. This is divided into existing data (which is clearly less expensive to use) and new data collection that might be needed to fully address the question. Finally, we note the design factors, generally associated with data collection of various types, which will increase evaluation costs.

Next, we discuss the cost considerations for each of the major evaluation types (implementation and outcome evaluations).

1. Implementation Evaluation

The cost of implementation evaluations will vary depending on (a) the types of existing data that might be used in the evaluation, and (b) the types of new data that need to be collected specifically for the implementation evaluation.

a. Existing data that might be used for the evaluation.

Existing data sources are the first place to look for information that will address the research questions if costs are to be controlled. Data for many types of implementation evaluations can come from the QRIS program's management information system (MIS), from resource and referral agency partners, from licensing agencies, and other state agencies with critical data items on the population of interest (as discussed in Chapter IV). For example, MIS data could address questions about the pace of enrollment of providers in the QRIS and the frequency and variety of technical assistance used by providers in the system.

Table V.1. Examples of Major Evaluation Design Components that Influence Evaluation Cost

				Design Factors that Could Increase Evaluation Costs					
	Examples of Research Questions that Could Be Addressed	Existing Data that Might Be Used in the Evaluation	New Data Collection	New Data Collection	Large Represent-ative Sample	Comparison Group Outside QRIS	Multiple Data Collection Waves	Observational Measures of Quality	Direct Child Assess-ments
Implementation Evaluation									
Process Study	Did child care programs enroll in the QRIS at expected rates?	Management information system data on QRIS enrollees compared with Resource and Referral data and licensing data on all providers	Interview QRIS staff about enrollment process	X					
	Did child care providers receive technical assistance from the QRIS at expected intensity?	Management information system data on QRIS technical assistance; number of TA encounters per provider during a year	Interview QRIS staff about technical assistance offered, how topics were selected, perceptions of provider engagement.	X					
Comparative Study of a QRIS Element	What approaches have been taken to informing parents about the QRIS ratings, and what are the challenges in parent understanding and use of the ratings in choosing care?		Interview QRIS staff about parent information approaches; focus groups of parents on understanding and use of the ratings in choosing care; parent survey on understanding and use of ratings	X	X				
	How do the QRIS ratings correlate with independent measures of quality, and what accounts for differences in these correlations across QRIS programs?	MIS data on QRIS ratings	Observational measure of quality					X	
Cost Study	What is the cost of rating a child care program and keeping the rating up to date?	MIS information on number of child care programs rated over a year	Interviews with program staff on steps to rating a child care program and time required for each step; time frames for updating rating	X					

	Examples of Research Questions that Could Be Addressed	Existing Data that Might Be Used in the Evaluation	New Data Collection	Design Factors that Could Increase Evaluation Costs					
				New Data Collection	Large Represent-ative Sample	Comparison Group Outside QRIS	Multiple Data Collection Waves	Observational Measures of Quality	Direct Child Assess-ments
	What is the cost of offering technical assistance to caregivers to improve quality, including advertising the TA, engaging caregivers, and providing TA?	MIS information on number of caregivers receiving TA during a year	Interviews with program staff on advertising, TA provided, and efforts to keep caregivers engaged	X					
Validation Study	How is the quality rating constructed, what elements of the quality rating are most variable and which contribute most to the final rating? How do the ratings correlate with an independent measure of quality?	Management information system data on elements of quality rating by provider	Observational measure of quality for each provider included in the rating study					X	
Outcome Evaluation									
Experimental Design	Does a program of coaching combined with financial incentives increase the quality of care compared with the prevailing approaches to TA?		Observational measure of quality of care; caregiver survey	X	X			X	
Quasi-experimental Design	Do communities included in the QRIS service area experience a greater increase in the quality of caregiving than comparable communities not served by a QRIS?		Observational measure of quality of care; caregiver survey	X	X	X		X	
	Do children receiving care in programs participating in QRIS have improved language and social-emotional outcomes compared with comparable children in programs not participating in QRIS?	MIS to identify children in programs participating in QRIS; resource and referral data to identify other caregivers	Measures of children's outcomes in QRIS and non-QRIS programs	X	X	X			X
Time Series Approach	How has the quality of center-based care changed over time as the QRIS was introduced and expanded?		Observational measure of quality of care; caregiver survey	X	X		X	X	

	Examples of Research Questions that Could Be Addressed	Existing Data that Might Be Used in the Evaluation	New Data Collection	Design Factors that Could Increase Evaluation Costs					
				New Data Collection	Large Represent-ative Sample	Comparison Group Outside QRIS	Multiple Data Collection Waves	Observational Measures of Quality	Direct Child Assess-ments
	How has parent knowledge of the QRIS rating system improved as the QRIS was introduced and expanded?		Parent surveys	X	X		X		
Pre-Post Test Approach	How did child care program ratings change from the initial rating at the start of the program to two years later, when the programs were rated again?	MIS data on indicators comprising the QRIS rating and the overall rating							
Descriptive Approach	What are the characteristics of families and children in programs participating in the QRIS?	Possibly MIS data if QRIS providers are required to collect demographic data	Survey of families using caregivers who participate in the QRIS		X				
	What are the characteristics of child care providers participating in the QRIS?	MIS data on characteristics of participating child care providers	Survey of QRIS providers for broader information on caregiver training, motivations, and intention to remain in the field		X				

Over time, the MIS and other partner data systems could be expanded to routinely collect information that can help QRIS stakeholders to evaluate participation, service use, intensity, and costs of the program. Expansions can include additional data elements, clearer definitions of existing data elements, and greater accuracy and higher completion of data elements. In addition, links between administrative data systems facilitate analyses of many important questions about a QRIS. For example, by associating children with their providers, the stability and quality of care can be measured for QRIS providers. By associating parents with their providers, the stability of employment can be measured among parents using QRIS providers.

b. New data collection

The types of new data that need to be collected to answer the research questions on implementation will also influence the evaluation costs. Next, we discuss how collecting new data using three different methods—interviews, observations, and surveys—influence the costs of implementation evaluations.

Interviews. Implementation evaluations often require some information from program operators to provide context for the data obtained from the MIS and other agency sources. This information can typically be obtained through semi-structured interviews with staff members, adding modestly to evaluation costs since it entails a limited number of interviews that can be completed over the telephone or by visiting staff offices for a day or two. For example, to address the question, "Did child care providers receive technical assistance from the QRIS at the expected intensity," the data on the number of technical assistance visits and number of providers served by those visits should be interpreted based on interviews with QRIS staff. Staff interviews can illuminate the types of technical assistance that was offered, how the topics were selected, efforts made to publicize the technical assistance, how technical assistance workshops or visits were scheduled, feedback from providers, and perceptions of the level of provider engagement and interest.

Observations. Some research questions—for example, validation of the QRIS ratings—require an independent observation of the child care environments that can be compared with the overall ratings and with specific component measures within the overall ratings. Observational ratings of child care environments are expensive. They entail training observers to code an observation measure with high reliability, scheduling visits to providers, spending hours at the child care setting making the observation, making multiple visits in the case of broken appointments, and managing the observation staff.

Surveys. Some research questions focus on how well information about the QRIS program has been communicated to parents in the community. To provide a valid answer to such a question, a survey could be fielded of parents in the community. The parents surveyed should be representative of the parent groups of interest; for example, parents in a particular geographic area in which the QRIS is particularly active or parents of children in a particular age group. The group of interest should be carefully defined – for example, location of residence, parental work status, and ages of children – so that the responses to the survey are meaningful. Finding a suitable source of contact information on the group of interest – the sampling frame – can be challenging. The sampling frame needs to provide the necessary contact information on households and not include large numbers of people who are not eligible for the survey. Parent surveys will be more expensive when the number of parents to be surveyed is larger, when the available sampling frame includes a larger proportion of people who are not eligible for the survey, and when the potential respondents are unwilling to respond to the survey. The length of the survey itself can also contribute to the cost of the survey, both because of the time required to obtain responses to the questions and because a longer survey can make it more difficult to gain respondents' cooperation.

2. Outcome Evaluation

Outcome evaluations are often more expensive than the implementation evaluations because essential features of their designs are costly to implement. For example, an experimental or quasi-experimental design will require data collection from a comparison group, and a time series design will require data collection at multiple points over time. As in the case of implementation evaluations, the cost of outcome evaluations will be influenced by whether existing data sources can provide information for the evaluation or new data have to be collected. Data to measure how a QRIS component, intervention, or service affects outcomes could be obtained from administrative data—such as the MIS or other administrative data—only if appropriate data elements can be found in the data and if the timing of measurement is appropriate. However, administrative data might not include the key outcomes measured at the right time. In that case, follow-up data collection—observational measures of quality, perhaps coupled with a brief provider questionnaire—would need to be completed at the end of the professional development period. Additionally, concerns about the comparability of the data obtained through administrative data systems and those obtained separately by trained data collectors often lead the researcher to recommend that data from both groups be collected in the same way. Administrative data may be helpful to identify providers, families, and children for the QRIS sample, but data on outcomes of evaluation subjects would need to be obtained through caregiver or parent surveys, observations of care arrangements, and direct child assessments.

Some factors that will affect the cost of outcome evaluations are specific to the design, while others apply across outcome evaluations (though to varying degrees). We discuss these design features and how they contribute to evaluation costs next.

Large representative sample. In all outcome evaluations that require new data collection the cost of the evaluation will depend on the size of the sample (see Table V.1). As the number of subjects from whom data need to be collected increases—the number of caregiving staff that will be

surveyed, the number of settings that will be observed, the number of children that will be assessed—the cost of data collection and therefore the cost of the evaluation will also increase. Having a large representative sample, as discussed in Chapter III, allows evaluators to achieve results that are more statistically precise than if smaller samples are used. And in particular, for experimental and quasi-experimental designs a large enough sample is required to detect the impacts of the intervention being studied. For example, the question of whether an enhanced professional development approach has an effect on quality of caregiving requires a sufficient number of providers in the "treatment" and the "comparison" groups to detect effects on caregiving practices and quality of the environment. The number of providers required for the study depends on a careful analysis of the variability of the outcome measures, the expected change in the quality of caregiving and the smallest change that would be considered meaningful to QRIS stakeholders. Also important to consider is the likelihood that providers would engage in the professional development offered and participate in the data collection effort once training is completed.

Comparison group outside QRIS. Experimental and quasi-experimental evaluations require that data are collected on a control or comparison group. In an experimental design, the evaluator randomly assigns evaluation participants to either a treatment or a control group and collects information on all members of both groups. In quasi-experimental evaluations, the evaluator might be able to collect data directly from the comparison group or might have to obtain the data from secondary sources such as surveys from databases on child care programs, families, and children compiled by other organizations.

Multiple data collection waves. The number of data collection waves and their timing are determined based on the research questions, but these decisions will also affect evaluation costs. An experimental design may require one follow-up data collection after the program being tested is completed, while a time series study will require multiple follow-up measures over the period of change. Each additional wave of data collection has significant data collection costs associated with

it. Thus, a cross-sectional descriptive study will cost less than a pre-post study of the same sample size and involving the same types of data collection, which must collect two waves of data. The length of time between the beginning of the study and the follow-up data collection also affects the evaluation cost. Because there is so much movement of staff and children in and out of programs, longer follow-up periods make it more difficult to locate and complete data collection on a larger proportion of the sample. Thus, the cost of the evaluation increases with the length of follow-up because of the complexities involved with tracking the sample and finding sample members who are no longer associated with the child care programs in the evaluation.

Direct child assessments. Measuring child outcomes such as cognitive and language development, health, and behavior can be costly to implement. Before including children in a study, the researcher must have procedures formally reviewed by an institution that safeguards the interests of research subjects (an institutional review board) and obtain informed consent from parents. The research should focus on children within a specific age group because children develop quickly and so the measures of child outcomes change over the age span, making analysis across wide age groups very difficult. Widening the age span will require a larger sample so that analyses of separate age groups of children (necessary because of changes in the measures over age groups) will have sufficient sample for analysis.

B. QRIS Evaluation Cost Examples

In light of the cost of evaluation activities, states can take one of two different approaches to developing their evaluation plans. They can develop an evaluation plan that fits the budget they are able to set aside for it. Alternatively, they can think expansively about the research that should be done, set priorities among them, and then seek outside funding to meet their evaluation goals over time. Knowing what evaluation activities can be afforded for the level of funding that is available to QRIS partners is not easy to determine without careful thought and planning. An initial design phase

(discussed in the introduction to Chapter VI) may be needed to be sure the budget and purpose align.

A few illustrative examples of QRIS evaluations and details about their evaluation budgets and activities are provided in Table V.2. As described earlier, costs will vary based on the overarching design, as well as the decisions about such elements as sample size, data collection, and type of measures that support the goals of the evaluation.

C. Common Sources of Funding

Recent evaluations of QRISs have been funded largely by organizations that were already invested in the development or maintenance of the QRIS. These include public/private partnerships, private foundations, and state and local government entities.

Internal sources. Evaluations funded by QRIS partners include the evaluation of the Qualistar Early Learning QRIS, which was funded by Qualistar Early Learning, a nonprofit organization. Similarly, Prime Time Palm Beach County, Inc. funded the evaluation of Prime Time in Florida. Private foundations supported the evaluations of Parent Aware in Minnesota and the Missouri Quality Rating System.

Government sources. State and local governments are often involved in financing evaluations, especially when the QRIS is a statewide system. Indiana's Paths to QUALITY is a statewide system whose evaluation was funded by an agency within state government. Internal partners with financing from state government, nonprofits, and private business donors evaluated Delaware Stars for Early Success. The federal government has provided funding for research in the area of child care and quality rating systems through the Child Care and Development Fund block grant to states. The evaluations of QRISs in Indiana, Kentucky, and Ohio were supported in part by the Child Care and Development Fund through the set-aside designated for quality improvement.

Table V.2. Examples of QRIS Evaluations with Funding Levels and Evaluation Activities

State or County and QRIS	Evaluator	Funding Level	Evaluation Design	Evaluation Activities
Minnesota, Parent Aware	Child Trends	$431,000	Implementation evaluation (process and validation)	Assessed implementation after 3 years and trends in outcomes. Also identified areas for program improvement to aid in statewide implementation.
Miami-Dade, Quality Counts	University of North Carolina	$350,000	Implementation evaluation (in-depth), and Outcome evaluation (descriptive)	First 2 years focus on implementation, award levels, and the best mix of QI supports. Year 3 will examine associations between child care quality and child outcomes.
Kentucky, KIDS NOW Initiative	University of Kentucky	$250,000 (most recent year)	Outcome evaluation (descriptive)	Examines associations between participation in QRIS to child and child care quality outcomes. Data collected using surveys, child assessments, parent questionnaires, and classroom observations. Yearly executive summaries released, as well as research-to-practice briefs.
Pennsylvania, Keystone STARS	University of Pittsburgh	$1,400,000	Outcome evaluation (time series; descriptive)	Analysis of trends in quality among child care programs participating in QRIS and nonparticipating providers. Also examining the components that predict higher quality, such as curricula and teacher qualifications.
QRS Assessment Project; multiple QRS	Mathematica; Child Trends; Christian & Tvedt Consulting	$1,336,786	Implementation evaluation (comparative; validation); Outcome evaluation (experimental; descriptive)	Multi-faceted approach: summative of the key elements of 26 QRIS; comparative study of quality measurement across 5 QRIS to include secondary data analysis of associations between quality components, observed quality and rating levels; comparative study of 2 QRIS on the role of QRIS in ECE system integration; pilot experimental study of classroom observation measures

Source: U.S. Department of Health and Human Services (2010b) and Tout, Starr, Soli, Moodie, Kirby, and Boller 2010.

Other opportunities. When seeking federal support for evaluations, states often identify and apply for these types of grants jointly with academic research partners. A federal research grant partially supported the evaluation of Parent Aware in Minnesota. Keeping current and monitoring the potential sources for evaluation funding across the federal government (for example, from the U.S. Department of Education, the National Institutes of Health, and the Administration for Children and Families) is important when seeking funding opportunities In particular, the U.S. Administration for Children and Families, Office of Planning, Research and Evaluation

(ACF/OPRE), has supported a large number of evaluations and other research studies of QRIS, for example:

1. A mixed methods study of the ways in which QRISs target the various dimensions of quality in child care programs[18]
2. An evaluation of Maine's Quality Rating System[19]
3. An investigation of how low-income families in Minnesota make child care decisions [20]
4. A study that pilots and validates four linked state child care QRSs; measures the effects of training on child care quality; and) assesses provider attitudes about professional development, and parent attitudes about quality ratings[21]

A complete list of these studies and its descriptions is provided in the ACF/OPRE Child Care & Early Education Research Connections website: http://www.researchconnections.org/childcare/resources/?topic=quality-rating-systems.In addition, ACF/OPRE supports the Quality Initiatives Research and Evaluation Consortium (INQUIRE), which is designed to facilitate the identification of issues and the development and exchange of information and resources related to research and evaluation of QRIS and other quality initiatives. INQUIRE consists of leading QRIS researchers convened by OPRE to develop useful resources about evaluation methods and findings for state child care administrators, program operators, and other stakeholders. INQUIRE members also provide technical assistance to states on evaluation and data issues as requested.

18http://www.researchconnections.org/childcare/resources/16288?topic=quality-rating-systems&publicationYear=2008&paging.startRow=1

19http://www.researchconnections.org/childcare/resources/16284?topic=quality-rating-systems&publicationYear=2008&paging.startRow=1

20http://www.researchconnections.org/childcare/resources/13219?topic=quality-rating-systems&publicationYear=2007&paging.startRow=1

21http://www.researchconnections.org/childcare/resources/12418?topic=quality-rating-systems&publicationYear=2004&paging.startRow=1&author=Raikes%2C+Helen

VI. GUIDELINES FOR SELECTING AND WORKING WITH AN EVALUATOR

An important part of the evaluation planning process is to select an individual or organization to conduct the evaluation and report the findings. The evaluator will be responsible for keeping the evaluation on schedule and within budget; managing the expectations of the evaluation partnership (which includes the evaluation funder, the administrators of the QRIS, academic and state advisors, local child care resource and referral agencies, and other stakeholders); and working closely with participating child care providers. Accordingly, it is important to choose an evaluator or a team with the qualifications, resources, and experience to complete the evaluation effectively. This chapter discusses the steps to choosing the right evaluator for the job and how to work effectively with an evaluator.

Before approaching the task of recruiting an evaluator, however, it may be useful to set aside a block of time to think about the various design options and what they would cost, and to seek the guidance of an experienced consultant in this matter. An experienced evaluator or consultant could help to outline the basic framework for the evaluation, including the logic model and key research questions. They could also lay out potential design options and estimate what costs would be involved. This initial design phase would enable the evaluation partnership to release a strong RFP that is based on a thorough consideration of the partnership's needs and which will ensure the highest quality proposals. The partnership should also draw on the knowledge of their partners and of other stakeholders they work with during this phase to build the strongest possible foundation for the evaluation.

CHAPTER CONTENTS

A. Considering Internal or External Evaluators

As we discussed in Chapter III, evaluation can be done at different levels, from a descriptive evaluation to a more complex evaluation seeking to identify impacts. Each comes with different costs and, similarly, with regard to an evaluator, each level requires a different capacity and expertise.

QRIS partnerships are typically made up of several organizations with varying expertise. In deciding who should conduct an evaluation, partners should first consider their internal capacity to conduct the evaluation. They will have to consider the research questions they want to answer and the data collection activities that will be necessary to answer those questions. A QRIS that already maintains an MIS with in-depth data on providers and other aspects of the program could easily answer questions about implementation progress and program improvement by assigning in-house staff with sufficient qualifications or research experience to do the job. Evaluations do not necessarily have to be completed by an external organization, and an internal evaluation is sometimes the best choice for what is desired from the evaluation. An internal evaluation could be carried out by individuals within the QRIS, such as researchers from an academic or other research institution that is part of the QRIS partnership.

When the evaluation requires methods and a time commitment that are beyond the capacity of the members of the partnership, it might be necessary to seek external expertise. This expertise can be provided in two ways: (1) as the primary evaluator, or (2) as a consultant on an as-needed basis. This choice will depend on the needs of the partnership. If the partnership has the staff time to conduct the evaluation but lacks the methodological expertise to construct indicators or measure outcomes, the evaluation could benefit from the advice of a seasoned child care researcher. If the

partnership lacks both expertise in evaluation and the staff resources to collect data or write reports, then it might be best to hire an external organization that can offer both staff and evaluation knowledge and experience.

In addition to capacity, there are additional factors to consider in putting together an evaluation team:

- **Budget constraints.** Using in-house staff for an evaluation might make it easier to control costs or identify efficiencies to save on expenses. A critical issue to consider is whether the internal evaluator has sufficient time to devote to the project. On the other hand, strategically involving a consultant who has years of experience with evaluation could cut down on the time spent planning and executing the evaluation, as a seasoned researcher will have the tools to conduct the work more efficiently than QRIS partners attempting evaluation for the first time.
- **Credibility.** External evaluators lend higher credibility and a disinterested and unbiased eye to evaluations, which an internal evaluator does not. To outside observers such as parents or potential replicators of the QRIS model, an evaluation conducted by an external evaluator is more likely to be unbiased in its estimates of what the model can achieve or its assessment of the validity of the rating system in classifying different child care providers.
- **Objectivity**. Although they might know more than anyone else about how the QRIS works, internal staff might find it harder to distance themselves from the day-to-day operations of the QRIS and maintain an objective perspective. They could also be more generous in judging the quality of care provided or management of the system because they are loyal to their own employer. An external evaluator will have a fresh perspective to identify areas for improvement. External evaluators should have no stake in the outcome of the evaluation and should reveal any potential conflict of interest before being hired. For example, the evaluator should have no financial interests in any aspect of the QRIS and funders of the evaluation should require that evaluators certify this in writing.

B. Developing a Request for Proposals (RFP)

If the decision is made to hire an external evaluator to evaluate the QRIS, information should be released in the public domain to notify potential evaluators of the opportunity. This is typically done through a solicitation or the release of a request for proposals (RFP). An RFP is essentially a job description detailing what services an evaluator would provide to the commissioner of an evaluation. Responses to the RFP would describe applicants' expertise in evaluation and/or early childhood education and child care and the approaches they would use to evaluate the QRIS. An

evaluator would be chosen from among those that submitted a proposal, taking into consideration qualifications of the team, merits of the evaluation design, and costs.

A well-written RFP is the key to bringing in strong proposals. Doing the necessary groundwork in fleshing out the logic model and research questions as discussed in the introduction to this chapter are the first step in writing a good RFP, but there are also content decisions that can influence how potential evaluators respond to the request. Table VI.1 provides some suggestions for what should and should not be in an RFP.

Table VI.1. The DOs and DON'Ts of a Good RFP

DO	DON'T
Specify your key research question(s)	Provide a laundry list of questions that will lead down too many paths; focus on the key questions of interest
Specify an expected level of effort	Set expectations beyond available funding; you will only get what you can afford. A high quality, focused effort is more useful than a thin or unsystematic exploration across a broad spectrum
Use clear, consistent language and terminology to convey research goals and expectations	Mix or use terms that suggest methods outside of the scope of the proposed work; for example, don't use "impacts" if it is a descriptive study
Consider the need for an advisory group or technical work group within the RFP; what skills are necessary, what organizations or entities should be represented, what balance should exist between internal staff and external experts	Require such a group within the RFP if it does not need to be organized and managed by the evaluator; consider if the need for a group is such that the partnership might convene it separately (and for other purposes), or if it is specific to the evaluation itself

1. Components of an RFP

An RFP can provide as little or as much detail as desired, but some basic components are typically found in an RFP. The more information and guidance given to potential applicants, the more likely that proposals submitted will align with the objectives of the organization commissioning the evaluation. On the other hand, if the RFP gives less guidance, potential evaluators can use their own expertise and creativity in proposing the best design for the QRIS evaluation based on their knowledge. Table VI.2 describes the types of information that could be provided in an RFP for the evaluation of a QRIS.

Table VI.2. What to Include in a Request for Proposals

1. **Name** and **contact information** for the organization sponsoring the evaluation as well as a point person to contact for any questions related to the RFP.
2. **Due date** for asking questions about the RFP and submission of proposals, including the time and time zone
3. **Instructions** for submitting proposals, including the preferred method(s) of delivery, acceptable and unacceptable formats, and number of copies.
4. **Proposal guidelines** such as preferred font and font size, page limits, and additional attachments such as resumes for proposed team members and business references.
5. **Background** on the QRIS. This could describe how the QRIS was founded and the stage and breadth of implementation, as well as the stakeholders involved and how many child care providers are currently participating. This summary could also describe the structure of the rating system and any QI programs that are part of the system. A logic model for the QRIS should be included here. Links to and sources of any previous evaluation results should be included and accessible to applicants to assist in understanding the context for the current evaluation.
6. **Purpose** of the evaluation, including the research questions that the evaluation will answer, what products are required and how the results of the evaluation will be used.
7. **Design** of the evaluation or what components are absolutely required, as well as a description of each task the partners plan to fund. Such tasks include coordination with the funders/stakeholders; development of final evaluation plans (design, data collection, analysis, and dissemination); measures development and testing; data collection and analysis; and reporting and dissemination.
8. **Time line** for the RFP process and the evaluation, including when the evaluation contract will be awarded and when major deliverables (for example, design report and final report) are desired.
9. **Available budget** for the evaluation or a proposed level of effort, stated as a dollar amount, in person hours, or in full-time equivalents (FTEs).
10. **Proposal requirements**, such as a description of the proposed evaluator's qualifications, experience with evaluation, content knowledge in the area of early childhood education and child care, the proposed design for the evaluation, data collection and analysis plans, evaluation budget, resumes for members of the evaluation team, references from previous funders of projects of similar size and scope, and any other desired elements.
11. **Criteria** for evaluating and scoring proposals and the selection process.
12. **Terms and conditions** of contracting with the evaluation sponsor. For example, if there are restrictions on contracting with a for-profit firm or on overhead, these should be clearly described in the RFP. Any restrictions on publication or dissemination should also be described. This section should describe the type of funding mechanism (whether it is a grant, firm fixed price contract, contract plus fixed fee, or other type). Institutional review board (IRB) requirements should be included (often IRB activities are included in the RFP as an evaluation task).

The components listed in Table VI.2 are suggestions for what to include in an RFP or how one might be organized. The QRIS partnership can decide to include more or less information according to its needs. If the QRIS is managed by a government agency, such as the state's department of education, there might be additional requirements for the RFP laid out in the state's procurement guidelines. Developing a high quality RFP that clearly describes state needs and requirements takes time. A committee review or external review of the RFP is useful for ensuring that applicants will be able to provide the best proposals possible.

2. Developing Scoring Criteria

A set of scoring criteria can help in evaluating each proposal on its own merits and ranking them in terms of quality. The criteria used should reflect what the partnership wants from an evaluator who will work with it for an extended period, will be closely involved with partnership staff, and will have access to knowledge of the day-to-day activities of the QRIS. The criteria should apply to the proposed team's (1) qualifications for carrying out the work, (2) its approach to the evaluation, and (3) its cost proposal. The cost proposal is usually evaluated separately from the other two components, depending on the priorities of the partnership, to prevent any bias in assessing the merit of the evaluation itself.

Each of these components can have an overall score associated with it, but can also be broken down into smaller segments to make it easier to award points. General items such as submission of the proposal on time, staying within the page limit, and providing references for previous work of a similar nature should be prerequisites to consideration of a proposal and do not require points. An example of how points might be awarded in evaluating a proposal can be found in Table VI.3.

3. Dissemination

The RFP should be disseminated to a wide range of individuals and organizations to ensure that there is a good sample of proposals to select from at the end of the RFP process. The QRIS partnership might already have potential evaluators in mind, and should seek out additional candidates, including academics, research institutes at higher education facilities, private contractors, nonprofits, and think tanks. Local universities or community colleges provide a wealth of expertise in research and have the advantage of physical proximity. It might help to look for individuals with degrees in the social sciences who have specialized in statistics or evaluation methods, as well those who have conducted similar evaluations in the past (W. K. Kellogg Foundation 2004a). University department chairs in the early childhood or education areas are another source of references to skilled evaluators and can also share the RFP with their colleagues.

Table VI.3. A Sample Proposal Scoring Guide

A. Proposal Preparation – Required for Consideration	
Proposal submitted by deadline	
Contact information provided	
Stayed within page limit (not including resumes)	
References from prior funders of related work included (three references at minimum)	
B. Technical Approach – 100 points	
Evaluation design meets requirements laid out in RFP	X points
Research questions are in line with RFP	X points
Design reflects understanding of QRIS and child care environment	X points
Proposes to engage program and provider staff	X points
Allows for feedback into program improvement	X points
Proposed time line aligns with desired time frame	X points
Data collection minimizes burden on QRIS and providers	X points
Reflects team's ability to remain flexible and establish and meet deadlines	X points
Contains structured plan for data collection, analysis, and deliverables	X points
Team will maintain sufficient confidentiality of identifiable data	X points
C. Organization and/or Team Qualifications – 100 points	
Proposed director has successfully led similar projects for at least 5 years	X points
Reflects capacity and organizational support for completing the work on time and within budget	X points
Significant combined years of relevant methodological and technical expertise and experience	X points
Key staff have successfully conducted similar work in the past	X points
Reflects the capacity to conduct data collection across multiple sites	X points
Capacity to conduct interviews in multiple languages as needed	X points
Ability to present information concisely and in approachable language	X points
Experience working with complex systems and partnerships	X points
D. Cost Proposal – 100 points	
Proposed services represent a good value	X points
Cost is reasonable	X points
Total Score Possible	**300 points**

Inviting potential applicants to participate in a conference call or webinar about the RFP in which the partners answer questions provides an opportunity to check for understanding of the RFP. Following up in writing ensures that all potential applicants have access to the same information as background for writing their proposals. Setting a clear deadline for questions and providing quick turnaround responses that go to all potential applicants or are posted on a public website also help to ensure the quality of the proposals. Transparency in the solicitation process is critical because if potential applicants have a sense that the competition is not fair or there is a favored candidate, the funder runs the risk of not getting a range of strong applications from different institutions.

C. Working with an Evaluator

Whether the chosen evaluator is internal staff or an external team, the partnership should establish clear lines of communication, common understanding, and accountability before the evaluation begins. This is especially important for an external evaluation team that has its own organizational styles of working and collaborating.

1. Avoid Surprises

Establish from the outset a regular meeting with the evaluation team, either the project director alone or the director and other key evaluation staff. The meetings can be in person or over the telephone as is convenient; it will be an opportunity to discuss progress on the evaluation and key deliverables, address roadblocks and challenges, and answer questions from the evaluation team. These regular meetings will also help to keep parties on task and manage any problems before they become major obstacles for the evaluation. The QRIS partnership might consider asking for monthly or bimonthly updates in written form from the evaluator in addition to or in lieu of meetings.

Setting a schedule for major milestones in the evaluation will also prevent surprises and keep the evaluation on the proposed time line. Milestones include deliverables and major data collection activities, such as interviews with parents and observations of provider facilities.

2. Establish Roles

It can be helpful to clarify any expectations of the QRIS partnership before the work begins. If any documents will need approval—such as surveys of parents or staff, observation protocols for provider visits, or recruitment letters—make sure the evaluator is aware of that need and can budget sufficient time to have documents approved by the partnership before they are used in the evaluation. If the staff member overseeing the evaluation would prefer to be the point of contact for providers or parents interested in the evaluation, that should also be made clear. All assumptions

and a clear communications plan should be discussed as soon as possible to ensure a smooth start to the evaluation.

3. Engage Program Staff

Any evaluation will need the cooperation of QRIS and provider staff to progress smoothly. To this end, the evaluator should agree on a protocol for notifying staff and families about evaluation activities and updates as necessary. The evaluators could be asked to attend a few system-wide meetings or visit providers early in the evaluation to introduce themselves to the various stakeholders, introduce the evaluation approach, and answer questions. A staff person from the QRIS could be designated as a liaison to the evaluator in case the evaluator has questions or concerns during the course of the evaluation.

4. Absorbing Feedback

Engaging community child care leaders and staff early can also facilitate the acceptance of evaluation findings that recommend changes at the implementation or provider levels. Even before the evaluation has been completed, a mechanism for feedback from the evaluator should be established that gathers useful findings that can inform program improvement. A QRIS is a complex initiative and an evaluation is a prime opportunity to check on the health of the system and identify where it is weak. The evaluation task and the system as a whole can benefit from timely feedback without having to wait for the release of a final evaluation report at some future time (for example, two or three years after the start of evaluation efforts).

D. Other Resources

A number of useful resources are available that provide additional guidance for creating an RFP and selecting an evaluator. ACF recently published a second edition of *the Program Manager's Guide to Evaluation* (U.S. Department of Health and Human Services 2010c), which includes advice on how to hire an outside evaluator and manage the selected evaluator. The W. K. Kellogg Foundation's

Evaluation Handbook (2004) discusses the skills evaluators need to carry out an effective evaluation and has helpful tools to select the right evaluation team, including worksheets and illustrative examples. These are two well-regarded sources on the subject, but numerous other resources are available in the public domain.

REFERENCES

Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth. Olympia, WA: Washington State Institute for Public Policy.* http://www.wsipp.wa.gov/rptfiles/04-07-3901.pdf.

Bamberger, M., & Rugh, J. (2008, November). *RealWorld evaluation: Working under budget, time, data and political constraints.* Summary chapter prepared for the American Evaluation Association Professional Development Workshop, Session 21. Retrieved October 26, 2010, from http://www.realworldevaluation.org/RealWorld_Evaluation_resour.html.

Barnard, W., Smith, W. E., Fiene, R., & Swanson, K. (2006). *Evaluation of Pennsylvania's Keystone STARS quality rating system in child care settings.* University of Pittsburgh, Office of Child Development; Pennsylvania State University, Prevention Research Center. Retrieved from http://www.pakeys.org/docs/Keystone%20STARS%20Evaluation.pdf.

Berry, D., Bridges, L., & Zaslow, M. (2003). *Early childhood measures profiles.* Washington, DC: Child Trends. Retrieved from http://www.researchconnections.org/childcare/resources/8634/pdf.

Blase, D. (2008). *Cascading Logic Model.* Chapel Hill, NC: National Implementation Research Network, Frank Porter Graham Child Development Institute, University of North Carolina at Chapel Hill.

Boller, K., Atkins-Burnett, S., Malone, L. M., Baxter, G. P., & West, J. (2010). *Compendium of student, teacher and classroom measures used in NCEE evaluations of educational interventions. Volume I. Measures selection approaches and compendium development methods* (NCEE 2010-4012). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/pubs/20104012/pdf/20104012.pdf.

Brandon, R. N., & Stutman, T. J. (2009). *Potential improvements to Ohio's Step Up to Quality program: Quality-based costs to providers, families and funding agencies.* University of Washington, Evan School of Public Affairs, Human Services Policy Center. Retrieved October 26, 2010, from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.170.953.

Breitner, L., Brandon, R., & Lalic, N. (2010). *Budgeting as a tool for policy development.* Training materials prepared for the UNICEF Social Protection and Inclusion Project in Bosnia and Herzegovina. Retrieved August 25, 2010, from http://www.hspc.org/topics/SPIS/EnTrainingModule_1.aspx.

Bronte-Tinkew, J., & Calkins, J. (2001). *Logic models and outcomes for early childhood programs.* Report to the DC Children and Youth Investment Trust Corporation. Washington, DC: Child Trends. Retrieved from http://www.childtrends.org/Files/LogicModelsandOutcomesforEarlyCH.pdf.

Centers for Disease Control and Prevention. (1999). *Framework for program evaluation in public health* (MMWR, Report No. 48 - RR-11). Washington, DC: Author. Retrieved October 26, 2010, from http://www.cdc.gov/eval/framework.htm.

Centers for Disease Control and Prevention. (2008). *Logic model basics* (Evaluation Brief). Retrieved August 25, 2010, from http://www.cdc.gov/healthyyouth/evaluation/pdf/brief2.pdf.

Centers for Disease Control and Prevention Evaluation Working Group. (2005). *Evaluation Working Group homepage.* Retrieved from http://www.cdc.gov/eval/index.htm.

Cheatham, J., Pope, B., & Myers, G. (2005). *Evaluating quality in state child care licensing: The Tennessee Report Card and Star-Quality Child Care Program.* University of Tennessee College of Social Work, Office of Research and Public Service. Retrieved October 26, 2010, from http://www.csw.utk.edu/about/stimulus/enhanced/2005_fall/childcarepaper.pdf.

Child Care and Early Education Research Connections. *Child care and early education glossary.* Retrieved from http://www.researchconnections.org/childcare /childcare-glossary.

Child Care and Early Education Research Connections. *Research glossary.* Retrieved from http://www.researchconnections.org/childcare/research-glossary.

Child Trends. (2009). *Logic models for the evaluation of Minnesota's early childhood and school-age professional development system Summary.* Washington, DC: Author. Retrieved August 19, 2010, from http://www.dhs.state.mn.us/main/groups/children/documents/pub/dhs16_144670.pdf.

Clegg & Associates, Inc. (2007). *Clegg & Associates Logic Model Game.* Retrieved August 19, 2010, from http://www.cleggassociates.com/html/modules.php?name=Content&pa=showpage&pid=23&cid=5.

Coffman, J. (2007). *A framework for evaluating systems initiatives.* Build Initiative: Build Strong Foundations For Our Youngest Children, Early Childhood Funders' Collaborative. Retrieved from http://www.buildinitiative.org/files/BuildInitiativefullreport.pdf.

Corporation for National and Community Service Resource Center. (n.d.). *AmeriCorps logic model resources.* Retrieved from http://www.nationalserviceresources.org/star/ac-logic.

Corso, P. S., & Lutzker, J. R. (2006). The need for economic analysis in research on child maltreatment. *Child Abuse & Neglect*, *30*, 727–738.

Delaware Department of Education. (2007). *Delaware Stars for Early Success - Phase 1 report.* Retrieved from http://www.doe.k12.de.us/infosuites/students_family/earlychildhood/files/DE%20Stars%20for%20Early%20Success%202007%20report.pdf.

Digital Library for Earth System Education. (2010). *Evaluation and assessment help for educators.* Retrieved October 26, 2010, from http://www.dlese.org/educators/eval.php.

Doran, G. T. (1981). There's a SMART way to write management's goals and objectives. *Management Review*, *70*(11), 35–36.

Elicker, J., Langill, C. C., Ruprecht, K., & Kwon, K. (2007). *Paths to QUALITY: A child care quality rating system for Indiana: What is its scientific basis?* Purdue University, Child Development and Family Studies. Retrieved October 26, 2010, from http://www.cfs.purdue.edu/cff/documents/project_reports/07_paths_to_quality.pdf.

Elicker, J., & Thornburg, K. (2011). *Evaluation of quality rating and improvement systems for early childhood programs and school-age care: Measuring children's development* (Research-to-Policy, Research-to-Practice Brief OPRE 2011-11b). Washington, DC: U.S. Department of Health and Human

Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Farell, K., Kratzmann, M., McWilliam, S., Robinson, N., Saunders, S., Ticknor, J., & White, K. (2002). *Evaluation made very easy, accessible and logical.* Halifax, Nova Scotia, Canada: Atlantic Centre of Excellence for Women's Health. Retrieved October 26, 2010, from http://www.acewh.dal.ca/eng/reports/EVAL.pdf.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication #231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network.

Florida Alcohol and Drug Abuse Association and Florida Department of Children and Families. (n.d.). *How to design substance abuse intervention logic model.* Retrieved August 19, 2010, from http://www.fadaa.org/services/resource_center/PD/WebEx/InterventionLogicModel.pdf.

Frechtling, J. (2002, January). *The 2002 user friendly handbook for project evaluation.* National Science Foundation, Evaluation and Communication, Division of Research. Retrieved January 21, 2011, from http://www.nsf.gov/pubs/2002/nsf02057/start.htm.

FRIENDS National Resource Center for Community-Based Child Abuse Prevention and Child Welfare Information Gateway. (n.d.). *Evaluation toolkit and logic model builder.* U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau. Retrieved from http://www.childwelfare.gov/preventing/developing/toolkit/.

Halle, T., Vick, J., & Anderson, R. (2010). *Quality in early childhood care and education settings: A compendium of measures (second edition).* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Retrieved from http://www.researchconnections.org/childcare/resources/18804/pdf.

Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early Childhood Environment Rating Scale, Revised Edition, updated.* New York, NY: Teachers College Press.

Hepburn, K. S., Kaufman, R. K., Perry, D. F., Allen, M. D., Brennan, E. M., & Green, B. L. (2007). *Early childhood mental health consultation: An evaluation toolkit.* Washington, DC: Georgetown University, Technical Assistance Center for Children's Mental Health; Johns Hopkins University, Women's and Children's Health Policy Center; Portland State University, Research and Training Center on Family Support and Children's Mental Health. Retrieved October 26, 2010, from http://gucchd.georgetown.edu/products/ECMHCToolkit.pdf.

Howes, C., Pianta, R., Bryant, D., Hamre, B., Downer, J., & Soliday-Hong, S. (2008). *Ensuring effective teaching in early childhood education through linked professional development systems, quality rating systems and state competencies: The role of research in an evidence-driven system* (National Center for Research in Early Childhood Education White Paper). Retrieved from http://www.ncrece.org/wordpress/wp-content/uploads/2008/09/ncrecewhitepaper2008.pdf.

Innovation Network, Inc. (n.d.). *Logic model workbook.* Retrieved August 19, 2010, from http://www.innonet.org/client_docs/File/logic_model_workbook.pdf.

Isakson, E. A., Davidson, L. L., Higgins, L. B., & Cooper, J. L. (2011, February). *State-level indicators for social-emotional development: Building better systems.* Columbia University, Mailman School of Public Health, National Center for Children in Poverty. Retrieved from http://nccp.org/publications/pub_997.html.

James Bell Associates. (2009, September). *Selecting an evaluation approach* (Evaluation Brief). Arlington, VA: Author. Retrieved from http://www.jbassoc.com/reports/documents/evaluation%20brief%20-%20selecting%20an%20evaluation%20approach%20v5.pdf.

Kagan, S. L., Moore, E., & Bredekamp, S. (Eds.) (1995). *Reconsidering children's early learning and development: Toward shared beliefs and vocabulary.* Washington, DC: National Education Goals Panel.

Karoly, L. A., Greenwood, P. N., Everingham, S. S., Hoube, J., Kilburn, M. R., Rydell, C. P., Sanders, M., & Chiesa, J. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions* (pp.73–102). Santa Monica, CA: The Rand Corporation. http://www.rand.org/pubs/monograph_reports/MR898.html.

Love, J. M., Aber, J. L., & Brooks-Gunn, J. (1994). *Assessing community progress toward achieving the first national educational goal.* Princeton, NJ: Mathematica Policy Research, Inc.

Love, J. M. (1999). *A Framework for developing research designs for the national Head Start impact study: what design(s) will answer the question: what strategies will allow us to make the strongest causal inferences about Head Start's effectiveness?* Princeton, NJ: Mathematica Policy Research.

Marek, L., Peterson, D., Mancini, J., Betts, S., & McDonald, D. (n.d.). *Sustaining community projects: Logic model construction and implementation.* Children, Youth, and Families Education Research Network. Retrieved August 19, 2010, from http://ag.arizona.edu/sfcs/cyfernet/evaluation/BREEZE%20training_rev2.htm.

McCawley, P. F. (2001). *The logic model for program planning and evaluation.* University of Idaho Extension. Retrieved from http://www.uiweb.uidaho.edu/extension/LogicModel.pdf.

Mitchell, A. (2005, July). *Stair steps to quality: A guide for states and communities developing quality rating systems for early care and education.* Alexandria, VA: United Way Success By Six. Retrieved from http://www.earlychildhoodfinance.org/downloads/2005/MitchStairSteps_2005.pdf.

Mitchell, A., & Ghazvini, A. (2007, July). *Florida's quality rating system: A conceptual model for estimating cost assumptions, explanations and supporting data.* Policy Group for Florida's Family and Children, publication 2007-Q001a. Retrieved from http://www.policygroup.org/downloads/policy%20brief%20series/2007-Q001a.pdf.

Nagle, G., Bronfin, M., & Lagarde, G. P. (2009). *BrightStart progress report.* Louisiana Department of Health and Hospitals, Office of Public Health, Maternal and Child Health Program. Retrieved from http://www.dhh.louisiana.gov/offices/publications/pubs-1/SCR%2083%20report.pdf.

National Child Care Information Center (NCCIC). (n.d.). NCCIC QRIS Cost Estimation Model. Accessed at http://qriscostmodel.nccic.acf.hhs.gov.

National Child Care Information and Technical Assistance Center. (2008, March). *Early childhood education workforce personnel and training registry systems.* Retrieved November 24, 2010, from http://seed.alaska.edu/support%20docs/National%20Early%20Childhood%20Registry.pdf.

National Science Foundation. (n.d.). *Online evaluation resource library homepage*. Division of Research, Evaluation and Communication, Directorate for Education and Human Resources. Retrieved August 27, 2010, from http://oerl.sri.com/home.html.

Norris, D. & Dunn, L. (2004). *Reaching for the Stars family child care home validation study*. Early Childhood Collaborative of Oklahoma. Retrieved October 26, 2010, from http://www.oklahomachildcare.org/system/files/ECCO%2520FamilyHome%2520Full%2520Report.pdf.

Organization for Economic Co-Operation and Development. (2008). *Best practice guidelines for evaluation* (PUMA Policy Brief No. 5). Retrieved October 26, 2010, from http://www.oecd.org/dataoecd/56/17/35060864.pdf.

Ovretveit, J., & Gustafson, D. (2002). Evaluation of quality improvement programs. *Quality and Safety in Health Care, 11*(3), 270–275. Retrieved October 26, 2010, from http://qshc.bmj.com/content/11/3/270.full.pdf.

Parsons, B. (1999). *Making logic models more systemic: An activity*. Boulder, CO: InSites. Retrieved August 25, 2010, from http://www.insites.org/documents/lmact.pdf.

Paulsell, D., Porter, T., & Kirby, G. (2010, March). *Supporting quality in home-based child care* (Final Brief). Submitted to the Administration for Children and Families, Office of Planning, Research, and Evaluation. Princeton, NJ: Mathematica Policy Research. Retrieved from http://www.mathematica-mpr.com/publications/pdfs/earlychildhood/HBCC_supporting_brief.pdf.

Paulsell, D., Porter, T., Kirby, G., Boller, K., Martin, E. S., Burwick, A., Ross, C., and Begnoche, C. (2010, March). *Supporting quality in home-based child care: Initiative design and evaluation options*. Report submitted to the Administration for Children and Families, Office of Planning, Research, and Evaluation. Princeton, NJ: Mathematica Policy Research. Retrieved from http://www.mathematica-mpr.com/publications/pdfs/earlychildhood/HBCC_supporting_options.pdf

Pennsylvania Keys to Professional Development. (2008). *Career lattice information sheet*. Retrieved from http://www.pakeys.org/docs/SQ03%20Career%20Lattice%20Info%20Sheet%20_05.01.08.pdf.

Porter, T., Mabon, S., Kearns, S., Robertson, A., & Kreader, J. L. (2003). *A toolkit for evaluating initiatives to improve child care quality*. New York, NY: Bank Street College of Education, Institute for a Child Care Continuum. Retrieved October 26, 2010, from http://www.bankstreet.edu/gems/ICCC/QualityStudy03.pdf.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Sattler, J. M. (1992). *Assessment of Children*. La Mesa, CA: Jerome M. Sattler Publisher, Inc.

Schmitz, C., & Parsons, B. (n.d.). *Everything you wanted to know about logic models but were afraid to ask*. Paper funded by the W. K. Kellogg Foundation. Retrieved August 25, 2010, from http://www.insites.org/documents/logmod.htm.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference, second edition*. Boston, MA: Houghton Mifflin.

Spielberger, J., Lockaby, T., Mayers, L., & Guterman, K. (2009). *Ready for prime time: Implementing a formal afterschool quality improvement system by Prime Time Palm Beach County, Inc.* Chicago: Chapin Hall at the University of Chicago. Retrieved October 26, 2010, from http://www.chapinhall.org/sites/default/files/Prime_Time_Report_04_18_09.pdf.

Swenson-Klatt, D. & Tout, K. (2011). *Measuring and rating quality: A state perspective on the demands for quality measurement in a policy context.* In M. Zaslow, I. Martinez-Beck, K. Tout & T. Halle (Eds.), Quality measurement in early childhood settings. Baltimore: Paul H. Brookes Publishing.

Thornburg, K. R., Mayfield, W. A., Hawks, J. S., & Fuger, K. L. (2009). *The Missouri quality rating system school readiness study.* Columbia, MO: University of Missouri Center for Family Policy and Research. Retrieved from http://mucenter.missouri.edu/MOQRSreport.pdf.

Tout, K., Starr, R., & Cleveland, J. (2008). *Evaluation of Parent Aware: Minnesota's quality rating system pilot: Year 1 evaluation report.* Minneapolis, MN: Child Trends. Retrieved October 26, 2010, from http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7B61078E31-3393-49B1-B301-D5CBC9A0DAC3%7D.PDF.

Tout, K., Starr, R., Isner, T., Cleveland, J., Soli, M., & Quinn, K. (2010, March). *Evaluation of Parent Aware: Minnesota's quality rating and improvement system pilot: Year 2 evaluation report.* Minneapolis, MN: Child Trends. Retrieved from http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BE0D4D742-A334-4305-BAF1-8432F580856D%7D.PDF.

Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G. & Boller, K. (2010, April). *The Child Care Quality Rating System (QRS) Assessment: Compendium of Quality Rating Systems and Evaluations*, OPRE Report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Retrieved from http://www.childcareresearch.org/childcare/resources/18554.

Tout, K., Starr, R., Isner, T., Cleveland, J., Soli, M., & Quinn, K. (2010, November). *Evaluation of Parent Aware: Minnesota's quality rating and improvement system pilot: Year 3 evaluation report.* Child Trends. Retrieved from http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BB5ADD2AE-D080-4290-A698-A972B0A93B6A%7D.PDF.

Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). *Issues for the next decade of quality rating and improvement systems* (Issue Brief). Administration for Children and Families, Office of Planning, Research and Evaluation and Child Trends. Retrieved October 26, 2010, from http://www.childtrends.org/Files/Child_Trends-2009_5_19_RB_QualityRating.pdf.

University of Kentucky and University of Louisville. (2007). *KIDS NOW evaluation project. Executive summary 2007.* Retrieved from http://www.education.ky.gov/NR/rdonlyres/C0105037-187D-4874-AB25-B0A0D4989E78/0/KIDSNOWEvaluationProject.pdf.

University of South Florida. (2009). *Logic model team website.* Retrieved August 19, 2010, from http://logicmodel.fmhi.usf.edu/logic_models.html.

University of Wisconsin-Extension, Program Development and Evaluation Unit. (n.d.). *Logic Model.* Retrieved from http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html.

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. (2010a). *QRIS profiles*. Washington, DC: Child Trends, Mathematica Policy Research, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Retrieved November 2, 2010, from http://www.acf.hhs.gov/programs/opre/cc/childcare_quality/.

U.S. Department of Health and Human Services, Administration for Children and Families, National Child Care Information and Technical Assistance Center. (2010b). *Quality rating and improvement system resource guide: State information*. Retrieved January 19, 2011, from http://stage.nccic.acf.hhs.gov/qrisresourceguide/index.cfm?do=qrisstate.

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. (2010c). *The program manager's guide to evaluation: second edition*. Washington, DC. Retrieved October 26, 2010, from http://www.acf.hhs.gov/programs/opre/other_resrch/pm_guide_eval/reports/pmguide/program_managers_guide_to_eval2010.pdf.

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start. (2010d). *The Head Start Child Development and Early Learning. Framework*. Washington, DC. Retrieved June 15, 2011, from http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/eecd/Assessment/Child%20Outcomes/HS_Revised_Child_Outcomes_Framework.pdf

U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention. (n.d.). *Performance measures: Logic models*. Retrieved August 19, 2010, from http://www.ojjdp.ncjrs.gov/grantees/pm/logic_models.html.

W. K. Kellogg Foundation. (2004a). *Evaluation handbook*. Battle Creek, MI: Author. Retrieved from http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx.

W. K. Kellogg Foundation. (2004b). *Logic model development guide*. Retrieved from http://www.wkkf.org/knowledge-center/resources/2010/Logic-Model-Development-Guide.aspx.

Walker, K., & Moore, K. (2011). *Performance management and evaluation: What's the difference* (Research to Results Brief 2011-02). Washington, DC: Child Trends.

Westmoreland, H., Lopez, M. E., & Rosenberg, H. (2009). *How to develop a logic model for districtwide family engagement strategies*. Harvard Family Research Project. Retrieved August 19, 2010, from http://www.hfrp.org/publications-resources/browse-our-publications/how-to-develop-a-logic-model-for-districtwide-family-engagement-strategies.

World Bank Group. (n.d.). *Participatory monitoring and evaluation: Tools and methods*. Retrieved from http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALDEVELOPMENT/EXTPCENG/0,,contentMDK:20509330~menuPK:1278210~pagePK:148956~piPK:216618~theSitePK:410306,00.html.

World Health Organization. (2004, June). *Monitoring and evaluation toolkit: HIV/AIDs, tuberculosis, malaria.* Retrieved from http://www.who.int/hiv/pub/epidemiology/en/me_toolkit_en.pdf.

Zellman, G., Perlman, M., Le, V. N., & Setodji, C. M. (2008). *Assessing the validity of the Qualistar early learning quality rating and improvement system as a tool for improving child-care quality.* Santa Monica, CA: RAND Corporation. Retrieved October 26, 2010, from http://www.rand.org/pubs/monographs/2008/RAND_MG650.pdf.

Zellman, G. L., Brandon, R. N., Boller, K., & Kreader, J. L. (2011). *Effective evaluation of quality rating and improvement systems for early care and education and school-age care* (Research-to-Policy, Research-to-Practice Brief, OPRE 2011-11a). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Forthcoming online.

Zellman, G. L., & Perlman, M. (2008). *Child-care quality rating and improvement systems in five pioneer states: Implementation issues and lessons learned.* RAND Corporation. Retrieved from http://www.rand.org/pubs/monographs/2008/RAND_MG795.pdf.

GLOSSARY

Accreditation. A process through which child care programs voluntarily meet specific standards to receive endorsement from a professional agency. The National Association for the Education of Young Children (NAEYC) and the National Accreditation Commission for Early Care and Education Programs (NAC) are among the organizations that offer accreditation programs for child care.

Adult-Child Ratio. A ratio of the qualified caregivers to children in a child care program.

Assessment. Measurement of a child's cognitive, language, knowledge and, psychomotor skills in order to evaluate development compared to children of the same chronological age.

Baseline. Measurement taken before an intervention is implemented to establish a comparison point for outcomes measured at a later time period.

Bonuses. QRISs can offer a one-time merit or achievement payment upon receipt of the rating. These awards are generally modest (between $250 and $2,500), depending on the type of program and its star level. QRISs can also offer similar-size awards for achievement or maintenance of quality on an annual (or biennial) basis.

Building Block System. Quality rating structure for child care providers in which all the quality standards in one level must be met before moving on to the next higher rating level.

Career lattice. Framework outlining multiple pathways for professional growth and development. It shows how practitioners can move vertically, horizontally, or diagonally across different occupational settings in the child care and early education system to advance their careers. Practitioners can view the minimum skills requirements of positions across licensed care, family learning centers, schools, or other settings within the system so they can plan their education in relation to the position they are seeking.

Center-based child care. Programs that are licensed or otherwise authorized to provide child care services in a nonresidential setting.

Child care provider. An institution or person who provides child care services.

Child Care Resource and Referral (CCR&R) agencies. Local and statewide services, including (1) guidance and referrals for parents seeking child care; (2) the collection of information about the local supply of child care; and (3) provider training and support. Some CCR&R agencies also administer child care subsidies.

Construct Validity. The degree to which key concepts (or constructs) are defined and measured with precision and clarity. Construct validity refers to the extent to which what needs to be measured is actually measured.

Descriptive Evaluation. An evaluation that answers questions about implementation of an intervention, how the intervention functions, or how characteristics of the intervention or participants may have changed over time.

Dissemination. Providing information about the QRIS, its standards, and ratings to the general public. The information provides parents and other stakeholders with the tools to judge absolute and relative quality of child care providers and settings.

Effectiveness. The ability to cause or produce desired outcomes.

External Validity. The ability to generalize the conclusions of the evaluation to future or different conditions.

Financial incentives. Awards or subsidies offered to providers, provider staff, or other caregivers to encourage quality improvement and participation in the QRIS or related activities. Incentives can include staff scholarships, retention bonuses, provider grants, subsidies contingent on performance (see "Tiered reimbursement system"), and funds for quality improvement initiatives.

Home-based child care. Child care provided for a group of children in a home setting. Most states have regulatory guidelines for family child care homes if they serve a number of children or families over a specified threshold or it they operate more than a specified number of hours each month.

Hypothesis testing. Using a statistical method to make decisions based on data.

Improvement Awards. Awarded to child care providers participating in a QRIS for improvements in quality. The award amount can be provided (1) in a matrix, with amounts differing by quality level, type or size of program, and the density of at-risk children served; or (2) as a standard amount or an upper threshold for grant amounts. The grant amount can be left unspecified but should indicate that it will align with items in the program's quality improvement plan.

Indicator. An observation assumed to be evidence of the attributes or properties of some phenomenon. Indicators allow assessment of progress toward the achievement of intended outputs, outcomes, goals, and objectives.

Internal validity. The ability to assert accurately that a program or intervention has caused the effects found in the evaluation given other plausible explanations.

Intervention. A program or initiative that is expected to change outcomes.

Licensing. Requirement necessary for a provider to legally operate child care services in a state or locality, including registration requirements established under state, local, or Tribal law.

Logic model. A systematic and visual way to present the relationships that are expected to exist among the resources available to the effort or program, the activities or policies that are to be put in place, and the changes or results that are expected to follow.

Monitoring. Oversight of provider and caregiver performance and compliance with quality standards.

Null hypothesis. The hypothesis (within the context of hypothesis testing) that is being falsified by a specific statistical test based on data.

Points System. Quality rating structure for child care providers in which points are earned for each quality standard. The points can be added together for a combined points score that is equated

to a corresponding rating level. In addition, the rating system can establish minimum requirements for the number of points achieved in different categories of standards to ensure broader coverage.

Professional development. In the child care field, this term refers to opportunities for child care providers to receive ongoing training to increase their preparation for, and skill in, caring for children. These opportunities include mentoring programs, credentialing programs, in-service training, and degree programs.

Quality improvement. Initiatives designed to increase the quality or availability of child care programs or to provide parents with information and support to enhance their ability to select child care arrangements most suited to their family and child's needs. Common quality initiatives include child care resource and referral services for parents, training and professional development and wage enhancement for staff, and facility improvement and accreditation for child care programs.

Quality Rating and Improvement System (QRIS). Quality Rating Systems (QRSs) are multicomponent assessments with public ratings of child care quality designed to make child care quality transparent and easily understood. Some QRSs explicitly include feedback and technical assistance and provide incentives to motivate and support quality improvement; these are Quality Rating and Improvement Systems (QRISs). A QRIS rates providers by levels of quality according to established criteria to encourage high-quality service delivery. Other elements of a QRIS include financial incentives, monitoring, and dissemination of the ratings to the public.

Quality Rating System (QRS). See Quality Rating and Improvement System (QRIS).

Registries. Databases that store and track a variety of professional development information. Personnel and training registries track practitioners' completed education, training, and other experience and professional activities. A registry can also produce records that validate qualifications or ongoing professional development for accreditation, a QRIS, wage incentives, and credentials. In some states, the registry is the system used to verify provider qualifications to meet state licensing requirements.

Reliability. Refers to the degree to which a measure (or set of measures) is consistent in measuring what it is supposed to measure.

Retention bonuses. A monetary payment to child care practitioners who remain employed with their provider for a specified period of time. These incentives help providers retain quality child care professionals, particular those who have completed training or professional development programs.

Standards. The components of child care services upon which quality ratings are based. The standards establish the criteria or requirements for being awarded a level of quality. The standards used in a QRIS can vary widely from state to state.

Technical assistance. Support or training provided to child care providers and staff by experts in relevant areas of programmatic, quality improvement, or professional development topics.

Tiered reimbursement system. A subsidy payment system that offers higher payments for child care that meets higher-quality standards or for child care that is in short supply.

Validity of ratings. The degree to which ratings are accurate reflections of the relative quality of child care services.

Wage enhancements. Grants provided by state or local government to supplement the wages of child care staff and improve employee retention.

The following resources were used to construct the glossary:

Career Lattice Information Sheet, Pennsylvania Keys to Professional Development. http://www.pakeys.org/docs/SQ03%20Career%20Lattice%20Info%20Sheet%2005.01.08.pdf.

Child Care and Early Education Glossary, Child Care and Early Education Research Connections. http://www.researchconnections.org/childcare/childcare-glossary.

Child-Care Quality Rating and Improvement Systems in Five Pioneer States: Implementation Issues and Lessons Learned. Zellman, G. L., & Perlman, M., RAND Corporation. http://www.rand.org/pubs/monographs/2008/RAND_MG795.pdf.

Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G. & Boller, K. (2010, April). *The Child Care Quality Rating System (QRS) Assessment: Compendium of Quality Rating Systems and Evaluations*, OPRE Report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. http://www.childcare research.org/childcare/resources/18554.

Early Childhood Education Workforce Personnel and Training Registry Systems, National Child Care Information and Technical Assistance Center. http://seed.alaska.edu/support %20docs/National%20Early%20Childhood%20Registry.pdf.

Logic Model Development Guide, W. K. Kellogg Foundation. http://www.wkkf.org/know ledge-center/resources/2010/Logic-Model-Development-Guide.aspx.

Research Glossary, Child Care and Early Education Research Connections. http://www. researchconnections.org/childcare/research-glossary.

APPENDIX A

EXTERNAL RESOURCES ON LOGIC MODELS, EVALUATION DESIGN, MEASURES, AND GUIDELINES ON SELECTING AND WORKING WITH AN EVALUATOR

This page has been left blank for double-sided copying.

To create the Toolkit, the QRS Assessment study team searched for and compiled a large selection of resources on evaluation and child care rating and improvement systems. This appendix showcases these resources and provides a brief overview of the unique features of each one. The resources are organized according to the chapters for which they were referenced, but each resource may have provided material for more than one chapter and, therefore, may appear more than once. Appendix B describes the methodology for searching for relevant resources.

A. Resources for Building Logic Models

We conducted a systematic search for resources on logic models (Appendix B details the methods used to find these resources). To organize the resources, we classified them into three categories:

1. **QRIS-Specific Resources.** These resources include information on at least one of four topics: (1) elements of logic models for the Quality Rating and Improvement System (QRIS); (2) guidelines for creating a logic model for QRIS; (3) examples of QRIS-specific logic models; and (4) information on how to use the logic model for QRIS planning, implementation, or evaluation.
2. **QI Intervention/Other Intervention Resources.** These resources include information on the elements, planning, and use of logic models for QI (quality improvement) interventions and other programs that are not QRISs. They also include information on how the logic models were used to inform the planning, implementation, and evaluation of those QI interventions and other programs.
3. **General Resources.** These resources provide information useful for understanding, creating, and interpreting logic models in general, without referencing specific interventions or programs.

1. QRIS-Specific Resources

General Information on QRIS Logic Models

Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). *Issues for the next decade of quality rating and improvement systems.* (Issue Brief). Washington, DC: Administration for Children and Families, Office of Planning, Research and Evaluation (OPRE) and Child Trends. http://www.childtrends.org/Files/Child_Trends-2009_5_19_RB_QualityRating.pdf.

- Summarizes issues and suggestions for QRIS evaluation improvement from recent OPRE-led meetings

- Discusses importance of logic models as part of QRIS evaluations and provides a logic model template
- Notes that few states have logic models for their QRISs and accompanying evaluations

Information on the Elements (Inputs, Activities, Outputs, Outcomes) of QRIS Logic Models

Howes, C., Pianta, R., Bryant, D., Hamre, B., Downer, J., & Soliday-Hong, S. (2008). *Ensuring effective teaching in early childhood education through linked professional development systems, quality rating systems and state competencies: The role of research in an evidence-driven system.* National Center for Research in Early Childhood Education White Paper. http://www.ncrece.org/wordpress/wp-content/uploads/2008/09/ncrecewhitepaper2008.pdf.

- Makes the case that state professional development systems must be better integrated with quality rating systems and Early Childhood Educator Competencies to improve the delivery of early childhood education services
- Discusses importance of creating a logic model to describe the integration process and its goals

Mitchell, A. (2005, July). *Stair steps to quality: A guide for states and communities developing quality rating systems for early care and education.* Alexandria, VA: United Way Success By Six. http://www.earlychildhoodfinance.org/downloads/2005/MitchStairSteps_2005 .pdf.

- Provides guidelines, tools, and resources on developing Quality Rating Systems (QRS)
- Describes the five components of QRSs: (1) standards, (2) accountability, (3) quality improvement, (4) incentives, and (5) dissemination
- Provides QRIS-specific examples of logic model elements

Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G. & Boller, K. (2010, April). *The Child Care Quality Rating System (QRS) Assessment: Compendium of Quality Rating Systems and Evaluations*, OPRE Report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. http://www.childcareresearch.org/childcare/resources/18554.

- Comprehensive information on the characteristics of 26 QRISs and on the evaluations they have undertaken
- Describes the inputs, activities, and outputs broadly across the 26 included QRISs

Information on Using the Logic Model as a QRIS Evaluation Tool

Zellman, G. L., Brandon, R. N., Boller, K., & Kreader, J. L. (2011). *Effective evaluation of quality rating and improvement systems for early care and education and school-age care.* (Research-to-Policy, Research-to-Practice Brief, OPRE 2011-11a). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Forthcoming online.

- Provides information on using logic models to plan a QRIS evaluation
- Links elements of a QRIS-specific logic model to research questions

Examples of QRIS Logic Models

Zellman, G. L. & Perlman, M. (2008). *Child-care quality rating and improvement systems in five pioneer states: Implementation issues and lessons learned.* Santa Monica, CA: RAND Corporation. http://www.rand.org/pubs/monographs/2008/RAND_MG795.pdf.

- Provides a general QRIS logic model
- Describes how a logic model is particularly helpful for QRIS evaluation planning as it helps stakeholders distinguish between intermediate program goals, such as improving program inputs and processes, and long-term outcomes, such as improved child outcomes

2. QI Interventions/Other Intervention Resources

General Information on Logic Models for QI Interventions/Other Interventions

Centers for Disease Control and Prevention Evaluation Working Group. (2005). *Evaluation Working Group homepage.* http://www.cdc.gov/eval/index.htm.

- Website containing resources for designing evaluations of public health programs
- Includes links to resources on using logic models to plan and execute a program evaluation
- Contains a bibliography with more than nine pages of "logic models in program evaluation" references [http://www.cdc.gov/eval/logic%20model%20bibliography.PDF]

University of South Florida. (2009). *Logic Model Team website.* http://logicmodel.fmhi.usf.edu/logic_models.html.

- Technical assistance website targeting Comprehensive Community Mental Health Services Program for Children and Their Families grant recipients
- Focuses on using theory-of-change logic models to plan programs and monitor results

U.S. Department of Justice (DOJ), Office of Juvenile Justice and Delinquency Prevention. (n.d.). *Performance measures: logic models.* http://www.ojjdp.ncjrs.gov/grantees/pm/logic_models .html.

- Explains DOJ's use of logic models in grant solicitations, emphasizing importance of logic models depicting key program features, outcomes, and indicators in grant applications
- Includes tools and resources for DOJ grant applicants, such as a list of DOJ performance measures, examples of generic logic models, and logic model templates

Information on the Elements of Logic Models for QI Interventions/Other Interventions

Bronte-Tinkew, J. & Calkins, J. (2001). *Logic models and outcomes for early childhood programs.* Report to the DC Children and Youth Investment Trust Corporation. Washington, DC: Child Trends. http://www.childtrends.org/Files/LogicModelsandOutcomesforEarlyCH.pdf.

- Presents logic models and measurable outcomes for the early childhood programs of the DC Children and Youth Investment Trust Corporation

Child Trends. (2009). *Logic models for the evaluation of Minnesota's early childhood and school-age professional development system – summary.* Washington, DC: Author. http://www.dhs.state.mn.us/main/groups/children/documents/pub/dhs16_144670.pdf.

- Discusses the development, components, and use of logic models for this evaluation
- Includes the development of five component-level logic models and one comprehensive logic model
- Highlights how using logic models can be useful for identifying program outcome measures and building consensus/common vocabulary among stakeholders

Marek, L., Peterson, D., Mancini, J., Betts, S. & McDonald, D. (n.d.). *Sustaining community projects: Logic model construction and implementation.* Children, Youth, and Families Education Research Network. http://ag.arizona.edu/sfcs/cyfernet/evaluation/BREEZE%20training_rev2.htm.

- Provides information on elements of a logic model and logic model worksheets
- Discusses the importance of logic models and consensus building among all program stakeholders

Westmoreland, H., Lopez, M. E., & Rosenberg, H. (2009). *How to develop a logic model for districtwide family engagement strategies.* Harvard Family Research Project. http://www.hfrp.org/publications-resources/browse-our-publications/how-to-develop-a-logic-model-for-districtwide-family-engagement-strategies.

- Provides step-by-step instructions on how to develop a logic model, especially to describe districtwide family engagement strategies

Information on Using the Logic Model as an Evaluation Tool for QI Interventions/Other Interventions

FRIENDS National Resource Center for Community-Based Child Abuse Prevention and Child Welfare Information Gateway. (n.d.) *Evaluation toolkit and logic model builder.* U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau. http://www.childwelfare.gov/preventing/developing/toolkit/.

- Information on using logic models to evaluate the effectiveness of prevention and family support programs, including child abuse and neglect prevention and parenting programs

- Features a hands-on web-based Logic Model Builder, which includes a step-by-step interactive guide to creating a logic model for child abuse or post-adoption services/programs; definitions and examples of key logic model features; and examples of outcomes, indicators, and assessment tools

Examples of Logic Models for QI Interventions/Other Interventions

Breitner, L., Brandon, R., & Lalic, N. (2010). *Budgeting as a tool for policy development.* Training materials prepared for the UNICEF Social Protection and Inclusion Project in Bosnia and Herzegovina. http://www.hspc.org/topics/SPIS/EnTrainingModule_1.aspx.

- Provides logic model instruction and development aid sheets

Corporation for National and Community Service Resource Center. (n.d.) *AmeriCorps logic model resources.* http://www.nationalserviceresources.org/star/ac-logic.

- Provides hands-on logic model development tools for AmeriCorps applicants and volunteers
- Includes information on basic logic models, a logic model worksheet, and logic model puzzles for practice
- Includes examples of logic models for Corporation environmental programs

Florida Alcohol and Drug Abuse Association and Florida Department of Children and Families. (n.d.). *How to design substance abuse intervention logic model.* http://www.fadaa.org/services/resource_center/PD/WebEx/InterventionLogicModel.pdf.

- Defines a logic model and its core components (inputs, outputs, impacts, and outcomes)
- Includes an example logic model

Program Development and Evaluation Unit, University of Wisconsin-Extension. (n.d.). *Logic model.* http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html.

- Introduces logic models and their core components
- Provides templates for creating a logic model and examples of logic models
- Provides information and materials on teaching and training about logic models

3. General Resources

General Information on Logic Models

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. (2010c). *The program manager's guide to evaluation: Second edition.* Washington, DC: Author. http://www.acf.hhs.gov/programs/opre/other_resrch/pm_guide_eval/reports/pmguide/program_managers_guide_to_eval2010.pdf.

- Provides an example of a logic model for a child abuse prevention program and includes a logic model worksheet

Clegg & Associates, Inc. (2007). *Clegg & Associates logic model game.* http://www.clegg associates.com/html/modules.php?name=Content&pa=showpage&pid=23&cid=5.

- Provides information and key definitions on building and designing logic models for outcome-based evaluation systems
- Includes an interactive game in which users identify elements of the logic model

Schmitz, C., & Parsons, B. (n.d.) *Everything you wanted to know about logic models but were afraid to ask.* Funded by the W. K. Kellogg Foundation. http://www.insites.org/documents/logmod.htm.

- Discusses differences between logic models and action plans and describes the purposes of logic models
- Provides information on the appropriate level of detail to include in a logic model, who should create the logic model, and how often the model should be revised

W. K. Kellogg Foundation. (2004b). *Logic model development guide.* http://www.wkkf.org/knowledge-center/resources/2010/Logic-Model-Development-Guide.aspx.

- Comprehensive guide on building logic models
- Discusses how to use logic models in program design and planning, and the importance of logic models to identify appropriate program success indicators

Information on the Elements of Logic Models

Centers for Disease Control and Prevention. (2008). *Logic model basics.* (Evaluation Brief). http://www.cdc.gov/healthyyouth/evaluation/pdf/brief2.pdf.

- Part of CDC's program evaluation briefs and tutorials online resource library
- Describes logic model components and provides a logic model worksheet

McCawley, P. F. (2001). *The logic model for program planning and evaluation.* University of Idaho Extension. http://www.uiweb.uidaho.edu/extension/LogicModel.pdf.

- Provides definition of logic model and its components and information on developing logic models for program

Information on Using the Logic Model as an Evaluation Tool

Innovation Network, Inc. (n.d.) *Logic model workbook.* http://www.innonet.org/client_docs /File/logic_model_workbook.pdf.

- Introduction to logic models and discussion of the role they can play in evaluations

- Step-by-step workbook activities to be used when creating a logic model

Farell, K., Kratzmann, M., McWilliam, S., Robinson, N., Saunders, S., Ticknor, J., & White, K. (2002). *Evaluation made very easy, accessible and logical.* Halifax, Nova Scotia, Canada: Atlantic Centre of Excellence for Women's Health. http://www.acewh.dal.ca/eng/reports/EVAL .pdf.

- Provides key definitions related to program evaluation and highlights logic models as an evaluation tool

B. Resources on Evaluation Design

We conducted a systematic search for resources on evaluation design (Appendix B details the methods used to find these resources). To organize the resources, we classified them into four categories:

1. **Evaluations of Existing QRIS.** These resources include links and summaries of the information presented in 14 reports published to date on evaluations of 12 existing QRISs.
2. **Resources on QRIS Evaluation Design.** These resources include information on at least one of the following four topics: (1) profiles of existing QRIS and their evaluation efforts; (2) information on QRIS planning, development, and research; (3) discussions of the status of research on QRIS; and (4) information on the challenges of evaluating QRIS and tools to address these challenges.
3. **Resources on Evaluation Design for Other QI Initiatives.** These resources include information on designing evaluations of initiatives aimed at improving quality of child care and initiatives to improve health outcomes.
4. **General Resources on Evaluation Design.** These resources provide information useful to plan, design, and implement evaluations in general, without referencing specific interventions or programs.

1. Evaluations of Existing QRS

Colorado

Zellman, G., Perlman, M., Le, V. N., & Setodji, C. M. (2008). *Assessing the validity of the Qualistar Early Learning quality rating and improvement system as a tool for improving child-care quality*. Santa Monica, CA: RAND Corporation. http://www.rand.org/pubs /monographs/2008/RAND_MG650.pdf.

- Describes implementation and outcomes evaluation of Colorado's Qualistar Early Learning program
- Collected data using the Caregiver Interaction Scale, the Pre-Kindergarten Snapshot, teacher surveys, student examinations, and family background information collected from parents
- Conducted waves of data collection, each 12 months apart
- Analyzed the program's quality measures and suggested improvements
- Analyzed data over time and within subgroups of children

Delaware

Delaware Department of Education. (2007). *Delaware Stars for Early Success - Phase 1 report*. http://www.doe.k12.de.us/infosuites/students_family/earlychildhood/files/DE%20Stars%20for%20Early%20Success%202007%20report.pdf.

- Describes how Delaware's QRIS, the Delaware Stars for Early Success initiative, was tested on 15 child care programs in preparation for offering Delaware Stars to more programs (Phase 1).
- Researchers gathered information from Delaware Stars database, surveys, interviews and meeting discussions with programs and technical assistants, and from advisory groups. Child care providers were assessed using the Environmental Rating Scale (ERS).
- Findings suggest revising center standards to have one version for early care and one for school-age-only programs, making qualification requirements at the Star Level 2 better interface with licensing, and moving several standards to a higher level. Findings also suggest revising grant funding to offer different amounts to small and large centers and providing additional support to care providers for navigating the QRIS and implementing plans to improve quality.

Florida, Palm Beach

Spielberger, J., Lockaby, T., Mayers, L., & Guterman, K. (2009). *Ready for prime time: Implementing a formal afterschool quality improvement system by Prime Time Palm Beach County, Inc.* Chicago: Chapin Hall at the University of Chicago. http://www.chapinhall.org/sites/default/files/Prime_Time_Report_04_18_09.pdf.

- Analyzes the process evaluation of the Quality Improvement System (QIS) operated by Prime Time, Inc., an organization entrusted with improving the quality of afterschool programs in Palm Beach County, Florida
- Discusses the implementation experience from the perspectives of 43 elementary and middle afterschool program directors, in addition to agency directors and Prime Time, Inc. staff members
- Describes the Quality Advising and Peer Coaching component of the QIS, which provides short- and long-term, on-site technical assistance to program directors and staff

Indiana

Elicker, J., Langill, C. C., Ruprecht, K., & Kwon, K. (2007). *Paths to QUALITY—Child care quality rating system for Indiana: What is its scientific basis?* Child Development and Family Studies, Purdue University. http://www.cfs.purdue.edu/cff/documents/project_reports_/07_paths_to_quality.pdf.

- Report presents the findings from an evaluation of the scientific validity of the quality standards contained in Paths to Quality (PTQ), Indiana's child care QRS.
- Standards identified 10 "key quality indicators": (1) regulation, (2) teacher education and specialized training, (3) structural quality, (4) process quality, (5) assessment, (6) provisions for children with special needs, (7) program policies and procedures, (8) director professional development, (9) parent-teacher communication and involvement, and (10) Accreditation by the National Association for the Education of Young Children (NAEYC) or other organizations.
- The review of the 10 main quality indicators (and 12 additional subindicators) within the PTQ standards revealed substantial evidence for the validity of the PTQ criteria:

Seventy-five percent of the quality indicators examined had "substantial evidence" for their validity, and 61 percent had significant evidence that they support children's development, learning, or well-being in child care.

Kentucky

University of Kentucky and University of Louisville. (2007). *KIDS NOW evaluation project. Executive summary 2007*. http://www.education.ky.gov/NR/rdonlyres/C0105037-187D-4874-AB25-B0A0D4989E78/0/KIDSNOWEvaluationProject.pdf.

- Presents the results from the most recent evaluation (2006–2007) of the STARS for KIDS NOW Initiative, which is Kentucky's QRS
- Examines child care providers participating in the STARS for KIDS NOW Initiative using surveys, child assessments, parent questionnaires, and classroom observations
- Offers five key findings: (1) preschool Head Start classrooms scored significantly higher than classrooms in nonprofit and for-profit programs in the Early Childhood Environment Rating Scale-Revised (ECERS-R) and the Early Language and Literacy Classroom Observation (ELLCO). For infant/toddler classrooms, nonprofit programs scored significantly higher than for-profit programs in the Infant/Toddler Environment Rating Scale (ITERS). Children in nonprofit centers scored significantly higher than children in for-profit and Head Start centers on the Peabody Picture Vocabulary Test (PPVT), Woodcock Johnson Test of Achievement-Applied Problems, Woodcock Johnson Test of Achievement-Dictation, and comprehension measures; (2) programs serving children who receive subsidy scored lower on the ECERS-R and ELLCO than programs who do not, and programs serving more minority children scored lower on the ECERS-R than programs serving fewer minority children; (3) participation in STARS for KIDS NOW and Healthy Start training leads to higher ITERS scores; (4) moderate increases from 2006 to 2007 in the overall score on the ECERS-R and ELLCO were found for all providers; and (5) centers performing in the top 25th percentile in overall quality had the highest percentages of children with disabilities, the lowest turnover rates, and the highest participation in STARS for KIDS NOW

Louisiana

Nagle, G., Bronfin, M., & Lagarde, G. P. (2009). *BrightStart progress report*. Louisiana Department of Health and Hospitals, Office of Public Health, Maternal and Child Health Program. http://www.dhh.louisiana.gov/offices/publications/pubs-1/SCR%2083%20report.pdf.

- Reports on the progress of the BrightStart initiative in 2008–2009, which includes promoting the implementation of a voluntary QRS for Louisiana, called Quality Start.
- Quality Start aims to increase the quality of child care and early learning and to educate parents and consumers in understanding, assessing, and demanding higher quality care.
- Six hundred child care centers applied to participate in Quality Start by January 1, 2009, with 495 achieving their first start. Some centers (97) applied for higher stars, with 75 centers achieving higher ratings.

Minnesota

Tout, K., Starr, R., & Cleveland, J. (2008). *Evaluation of Parent Aware: Minnesota's quality rating system pilot: Year 1 evaluation report.* Minneapolis, MN: Child Trends, 2008. http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7B61078E31-3393-49B1-B301-D5CBC9A0DAC3%7D.PDF.

- Outlines the Minnesota Parent Aware program, including quality standards, monitoring, incentives, and dissemination of ratings
- Reports results of an implementation (process) evaluation for the program's first year and notes that outcome evaluation results are forthcoming
- Describes data collection through semistructured discussions with various program stakeholders; provides discussion protocols in an appendix
- Analyzes quantitative Parent Aware program data
- Discusses success and challenges of program implementation, including those related to recruitment and enrollment, promotion and marketing, the rating process, and provider observations

Tout, K., Starr, R., Isner, T., Cleveland, J., Soli, M., & Quinn, K. (2010, March). *Evaluation of Parent Aware: Minnesota's quality rating and improvement system pilot: Year 2 evaluation report.* Minneapolis, MN: Child Trends. http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BE0D4D742-A334-4305-BAF1-8432F580856D%7D.PDF.

- Reports results of an implementation (process) evaluation for the program's second year
- Presents descriptive analyses of the characteristics of, and the quality ratings obtained by, child care programs participating in Parent Aware
- Describes data collection through semistructured discussions with various program stakeholders; provides discussion protocols in an appendix
- Presents preliminary findings on the scores from observation assessments obtained by child care programs participating in Parent Aware

Tout, K., Starr, R., Isner, T., Cleveland, J., Soli, M., & Quinn, K. (2010, November). *Evaluation of Parent Aware: Minnesota's quality rating and improvement system pilot: Year 3 evaluation report.* Minneapolis, MN: Child Trends. http://www.melf.us/vertical/Sites/%7B3D4B6DDA-94F7-44A4-899D-3267CBEB798B%7D/uploads/%7BB5ADD2AE-D080-4290-A698-A972B0A93B6A%7D.PDF.

- Reports results of an implementation (process) evaluation for the program's third year
- Presents the findings from an in-depth analysis of the Parent Aware rating tool, including the four categories of program standards that form the framework for the rating tool, the individual indicators that make up each category, and how these indicators are scored
- Reports the findings on the scores from observation assessments obtained by child care programs participating in Parent Aware

- Presents a descriptive analysis of parent perceptions about quality early care and education and their understanding of Parent Aware
- Discusses a number of considerations for statewide implementation of Parent Aware

Missouri

Thornburg, K. R., Mayfield, W. A., Hawks, J. S., & Fuger, K. L. (2009). *The Missouri quality rating system school readiness study.* Columbia, MO: University of Missouri Center for Family Policy and Research. http://mucenter.missouri.edu/MOQRSreport.pdf.

- Studies the relationship between early childhood program quality ratings and child outcomes in Missouri.
- Includes 38 licensed early childhood programs and 350 children ages 3 to 5 that attended participating programs full-time.
- Researchers used a variety of assessments to measure school readiness for children, as well as documenting the learning environment and family involvement.

Ohio

Brandon, R. N., & Stutman, T. J. (2009). *Potential improvements to Ohio's Step Up to Quality program: Quality-based costs to providers, families and funding agencies.* University of Washington, Evan School of Public Affairs, Human Services Policy Center. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.170.953.

- Presents findings from an examination of (1) the costs of Ohio's Step Up to Quality (SUTQ) program to providers, family, and public agencies; (2) the financial implications of potential improvements to SUTQ; (3) the effectiveness of the Quality Achievement Awards (QAA) in offsetting the costs to providers of meeting standards; and (4) the increase in costs of assisting families to afford higher quality care that would result from updating the SUTQ standards.
- The findings suggest that, although the costs to providers of moving from meeting licensing standards to meeting the first step of SUTQ standards are minimal, the costs of moving above the first quality level are significant. Also, Ohio's QAAs offset only a small portion (10 to 20 percent) of what it costs providers to meet the standards. Finally, if staff compensation is maintained at current levels, middle-income parents would likely need assistance to meet the cost of higher quality levels. If staff compensation were increased to the level earned by public school teachers with comparable qualifications, only the wealthiest families could afford the highest level of quality.

Oklahoma

Norris, D., & Dunn, L. (2004). *Reaching for the Stars family child care home validation study.* Early Childhood Collaborative of Oklahoma. http://www.oklahomachildcare.org/system/files/ECCO%2520FamilyHome%2520Full%2520Report.pdf.

- Presents the findings from an examination of the differences in quality between family child care homes representing the various levels (one-star, one-star plus, two-star, and three-star) of the "Reaching for the Stars" program, Oklahoma's QRS
- Analyzes 189 family child care homes across the state, and data sources included provider surveys and classroom observation
- Describes differences in family care homes by Star category on measures of child care quality such as the Family Day Care Environment Scale (FDCES) and the sensitivity of provider-child interaction as measured by the Arnett Caregiver Interaction Scale
- Describes the differences in family care homes by Department of Human Services (DHS) reimbursement rate area

Pennsylvania

Barnard, W., Smith, W. E., Fiene, R., & Swanson, K. (2006). *Evaluation of Pennsylvania's Keystone STARS quality rating system in child care settings*. University of Pittsburgh Office of Child Development; Pennsylvania State University Prevention Research Center. http://www.pakeys.org/docs/Keystone%20STARS%20Evaluation.pdf.

- Presents the findings from an evaluation of the Pennsylvania Keystone STARS program
- Compares 356 Keystone STARS child care centers with 216 non-STARS programs, and also compared centers having higher STARS ratings with those having lower ratings, to examine the program's effect on child care program quality
- Measures program quality and improvement over time using the ECERS-R and the Family Day Care Rating Scale (FDCRS)

Tennessee

Cheatham, J., Pope, B. & Myers, G. (2005). *Evaluating quality in state child care licensing: The Tennessee Report Card and Star-Quality Child Care Program.* Knoxville, TN: University of Tennessee College of Social Work, Office of Research and Public Service. http://www.csw.utk.edu/about/stimulus/enhanced/2005_fall/childcarepaper.pdf.

- Describes the Tennessee Child Care Report Card and Star-Quality Child Care Program, which consists of (1) assessing (using observational tools such as the ECERS-R, the School Age Care Environment Rating Scale [SACERS] and the FDCRS) and rating the quality of care provided in centers and family and group homes; and (2) rating other characteristics of child care providers, such as qualifications of staff, compliance history, parent involvement, ratio and group size, and staff compensation. Providers receiving at least a one-star rating on the quality assessment component, a one-star rating on the compliance history component, and an overall score (average of all component scores) of at least one, are eligible to participate in the Star-Quality Program.
- Providers participating in the Star-Quality Program receive a Star-Quality Report Card with their overall rating and the rating on each component that was assessed. They also receive bonus payments above the base reimbursement rate according to their rating.

Providers post their report cards for parents to view and are expected to make changes to their programs to improve their quality.

- Findings from this report indicate that, by the program's third year, 76 percent of the providers that were assessed obtained an overall rating of at least one star, and 54 and 38 percent of the centers and homes that were assessed, respectively, earned a three-star rating.

2. Resources on QRIS Evaluation Design

Information on QRIS Planning, Development, and Research

Quality Rating and Improvement Systems National Learning Network, 2009–2010
http://www.qrisnetwork.org/

- QRIS information clearinghouse
- Provides access to a resource library covering topics related to QRIS, including standards and standards alignment, evaluation, data collection and data systems, and measuring quality
- Provides links to other online resources related to QRIS planning, development, and research

Discussions of the Status of Research on QRIS

Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). *Issues for the next decade of quality rating and improvement systems.* (Issue Brief). Washington, DC: Administration for Children and Families, Office of Planning, Research and Evaluation and Child Trends. http://www.childtrends.org/Files/Child_Trends-2009_5_19_RB_QualityRating.pdf.

- Presents a discussion of the status of research and evaluation of QRS
- Highlights importance of conducting research on the implications of the unique structure and design of each QRS for its development and success
- Describes limitations of existing QRS evaluations, including inability to draw causal conclusions
- Notes that future QRS research should examine (1) cost of QRS implementation; (2) how to measure critical components of quality in ways that are cost-effective and apply to different types of child care settings and to children who differ in age, language, ability, and culture; and (3) the unintended consequences of QRS, such as how low-income families choose between regulated and unregulated care when a QRS is implemented

Profiles of Existing QRIS and Their Evaluation Efforts

Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G. & Boller, K. (2010, April). *The Child Care Quality Rating System (QRS) Assessment: Compendium of Quality Rating Systems and Evaluations*, OPRE Report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and

Families, U.S. Department of Health and Human Services. http://www.childcare research.org/childcare/resources/18554.

- Provides in-depth profiles of 26 QRS programs, including detailed information on evaluations conducted by 18 of the programs
- In addition to evaluations, details the following information on each QRS: (1) implementation dates, (2) rating processes and quality standards, (3) observational measures, (4) incentives, (5) quality improvement processes, and (6) outreach and marketing
- Provides the names of external evaluators, including universities and consultant, used by 13 of the programs
- Notes which of the 18 evaluations were implementation, validation, quality improvement, or outcomes evaluations
- Lists the programs with published evaluation reports
- Notes the use of Environment Rating Scales to measure quality improvement over time of four programs, as well as to validate QRS provider measures

Zellman, G. L. & Perlman, M. (2008). *Child-care quality rating and improvement systems in five pioneer states: Implementation issues and lessons learned.* Santa Monica, CA: RAND Corporation. http://www. rand.org/pubs/monographs/2008/RAND_MG795.pdf.

- Summarizes the QRSs of five states, including their rating systems and components, design, implementation processes, and system outputs
- Discusses implementation issues based on in-depth interviews with QRS stakeholders in each state
- Recommends ways to improve QRS based on lessons learned in the five states
- Highlights effectiveness evaluations as one of five ways that states can improve their QRSs
- Suggests establishment of a national QRS consortium to pool research funds and collect comparable data across states

Information on the Challenges in Conducting Evaluations of QRIS and Tools to Address These Challenges

Zellman, G. L., Brandon, R. N., Boller, K., & Kreader, J. L. (2011). *Effective evaluation of quality rating and improvement systems for early care and education and school-age care.* (Research-to-Policy, Research-to-Practice Brief, OPRE 2011-11a). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Forthcoming online.

- Provides information on why it is important to evaluate QRISs
- Discusses challenges in conducting evaluation of QRISs and strategies to addressing these challenges

3. Resources on Evaluation Design for Other QI Initiatives

Information on Designing Evaluations of Initiatives Aimed at Improving Child Care Quality

Porter, T., Mabon, S., Kearns, S., Robertson, A., & Kreader, J. L. (2003). *A toolkit for evaluating initiatives to improve child care quality.* New York: Bank Street College of Education, Institute for a Child Care Continuum. http://www.bankstreet.edu/gems/ICCC /QualityStudy03.pdf.

- Discusses evaluation in the context of efforts to improve child care quality
- Uses examples of specific QI initiatives to illustrate evaluation process
- Provides methodology for selecting instruments and makes instruments available in a format users can tailor to their needs

Paulsell, D., Porter T., Kirby, G., Boller, K., Martin, E. S., Burwick, A., Ross, C., & Begnoche, C.. (2010, March). *Supporting quality in home-based child care: Initiative design and evaluation options.* Report submitted to Administration for Children and Families, Office of Planning, Research and Evaluation. Princeton, NJ: Mathematica Policy Research. http://www.mathematica-mpr.com/publications/pdfs/earlychildhood/HBCC_supporting_options.pdf

- Describes potential strategies for supporting quality in home-based child care settings
- Discusses considerations for decision making and ongoing evaluation of strategies for supporting quality in home-based child care settings

Paulsell, D., Porter, T., and Kirby, G. (2010, March). *Supporting quality in home-based child care.* Final brief submitted to Administration for Children and Families, Office of Planning, Research and Evaluation. Princeton, NJ: Mathematica Policy Research. http://www. mathematica-mpr.com/ publications/pdfs/earlychildhood/HBCC_supporting_brief.pdf

- Presents findings on the prevalence and quality of home-based child care
- Discusses the characteristics of caregivers, quality initiatives for home-based care, and evidence of effectiveness for home-based care initiatives
- Presents an agenda for program development and research designed to foster effective quality initiatives for home-based care

Karoly, L. A., Greenwood, P. N., Everingham, S. S., Hoube, J., Kilburn, M. R., Rydell, C. P., Sanders, M., & Chiesa, J. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions* (pp.73–102). Santa Monica, CA: The Rand Corporation. http://www.rand.org/pubs/monograph_reports/MR898.html.

- Synthesizes results from a number of previous evaluations of early childhood interventions.
- Concludes that early childhood interventions can benefit participants through emotional and cognitive development, education, economic well-being, and health.

Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia, WA: Washington State Institute for Public Policy. http://www.wsipp.wa.gov/rptfiles/04-07-3901.pdf.

- Presents findings from an analysis of whether prevention and early intervention programs for youth provide positive investment returns to the taxpayer.
- Concludes that some programs do have greater benefits than costs, but that research in this area is still relatively young.
- Recommends evaluating current programs that have not been evaluated yet in terms of their costs and benefits.

Corso, P. S., & Lutzker, J. R. (2006). The need for economic analysis in research on child maltreatment. *Child Abuse & Neglect*, *30*, 727–738.

- Provides a framework for conducting economic analysis (i.e. analysis of the impact that a program has on outcomes and costs) of interventions to prevent child maltreatment.
- Discusses the available research in this area, as well as the need for further research.

Information on Designing Evaluations of Initiatives Aimed at Improving Health Outcomes

Hepburn, K. S., Kaufman, R. K., Perry, D. F., Allen, M. D., Brennan, E. M., & Green, B. L. (2007). *Early childhood mental health consultation: An evaluation toolkit*. Washington, DC: Georgetown University, Technical Assistance Center for Children's Mental Health; Johns Hopkins University, Women's and Children's Health Policy Center; Portland State University, Research and Training Center on Family Support and Children's Mental Health. http://gucchd.georgetown.edu/products/ECMHC Toolkit.pdf.

- Provides information on designing and incorporating evaluation for early childhood mental health consultation services, identifying measures, and determining outcomes
- Provides guidance on using evaluation data to report service outcomes and to build the evidence base related to early childhood mental health consultation services

World Health Organization. (2004, June). *Monitoring and evaluation toolkit: HIV/AIDS, tuberculosis, malaria*. http://www.who.int/hiv/pub/epidemiology/en/me_toolkit_en.pdf.

- Provides an overview of issues to consider in formulating a participatory national monitoring and evaluation (M&E) strategy for HIV/AIDS, tuberculosis, and malaria
- Provides information on (1) designing M&E systems that are sustainable and can be used to report on results and impact during the implementation stages of programs, (2) implementation and quality control of M&E systems and reporting of progress, and (3) evaluation, review, and improvement of M&E systems over time

4. General Resources on Evaluation Design

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. (2010c). *The program manager's guide to evaluation: Second edition.* Washington, DC: Author. http://www.acf.hhs.gov/programs/opre/other_resrch/pm_guide_eval/reports/pmguide/program_managers_guide_to_eval2010.pdf.

- Provides information on the questions that an evaluation can answer and guidelines for conducting a successful evaluation
- Discusses the importance of program evaluation to (1) finding out if it is working, (2) giving results to funders and the community, and (3) identifying areas for improvement, among other things
- Responds to common concerns about program evaluation, such as its purported burden on staff and the fear of getting negative results
- Reviews evaluation costs, including those associated with hiring an outside evaluator
- Lists data elements, sources, and collection instruments
- Advises on ways to analyze and disseminate results, includes evaluation planning worksheets, and provides references for evaluation toolkits, online evaluation resources, and consultants

Digital Library for Earth System Education. (2010). *Evaluation and assessment help for educators.* http://www.dlese.org/educators/eval.php.

- Provides information on evaluation and assessment, such as evaluation planning, data collection, data analysis, and data reporting
- Provides links to other resources on evaluation and assessment

Bamberger, M., & Rugh, J. (2008, November) *RealWorld Evaluation: Working under budget, time, data and political constraints.* Summary Chapter prepared for the American Evaluation Association Professional Development Workshop Session 21. http://www.realworldevaluation.org/RealWorld_Evaluation_resour.html.

- *Introduces the "RealWorld Evaluation" approach, a framework of steps to address constraints in evaluation including* funding, time, availability of data, and clients' needs
- Encourages the use of mixed-method designs to minimize threats to validity
- *Recommends techniques to strengthen evaluation designs and improve the utility of findings*
- *Assesses the strengths and weaknesses of different evaluation designs*
- *Discusses ways to efficiently use evaluation resources, including funding, staff, and time, as well as evaluation data and findings*
- *Includes checklists for gauging the validity of baseline data and of the evaluation*

Centers for Disease Control and Prevention. (1999). *Framework for program evaluation in public health.* (MMWR, Report No. 48 - RR-11). Washington, DC: Author. http://www.cdc.gov/eval/framework.htm.

- *Provides a guide for public health professionals for using program evaluation*
- *Includes a step-by-step guide for program evaluation and presents standards for effective evaluation*
- *Provides information on qualitative and quantitative approaches and tools for meeting evaluation needs at all stages of program implementation*

Coffman, J. (2007). *A framework for evaluating systems initiatives.* Build Initiative: Build Strong Foundations For Our Youngest Children, Early Childhood Funders' Collaborative. http://www.buildinitiative.org/files/BuildInitiativefullreport.pdf.

- *Provides a framework for conducting evaluation of systems initiatives, including efforts to improve child well-being by building or reforming education and human service systems*
- *Emphasizes the importance of connecting evaluations with specific theories of change*
- *Discusses selecting an evaluation approach that is both feasible and rigorous*
- *Suggests assessing both interim and long-term impacts and outcomes*
- *Recommends disseminating results as part of continuous feedback loops*
- *Includes table with examples of evaluation methodologies as they align with particular types of research questions*

Farell, K., Kratzmann, M., McWilliam, S., Robinson, N., Saunders, S., Ticknor, J., & White, K. (2002). *Evaluation made very easy, accessible and logical.* Halifax, Nova Scotia, Canada: Atlantic Centre of Excellence for Women's Health. http://www.acewh.dal.ca/eng/reports/EVAL .pdf.

- Defines key terms such as evaluation, process evaluation, impact evaluation, and outcome evaluation
- Suggests conducting a needs assessment prior to an evaluation and provides guidelines for doing so
- Describes empowerment and participatory evaluations and discusses their utility
- Explains the CDC framework for planning and conducting an evaluation
- Provides a step-by-step plan for disseminating evaluation results

Frechtling, J. (2002, January). *The 2002 user friendly handbook for project evaluation.* National Science Foundation, Division of Research, Evaluation and Communication. http://www.nsf.gov/pubs/2002/nsf02057/start.htm.

- Focuses on evaluations funded by NSF, but notes the information is pertinent to the wider program evaluation audience
- Discusses formative and summative evaluations

- Discusses how to select an appropriate and rigorous evaluation design
- Outlines how to collect and analyze data, and disseminate results

James Bell Associates. (2009, September). *Selecting an evaluation approach.* (Evaluation Brief). Arlington, VA: Author. http://www.jbassoc.com/reports/documents/evaluation%20brief%20-%20selecting%20an%20evaluation%20approach%20v5.pdf.

- Provides a step-by-step guide to developing an evaluation of a program
- Presents examples of research questions for process and outcome evaluations
- Discusses evaluation designs such as experimental, waitlist/overflow, match case, comparison site, time series, pre-post test, and case studies designs

National Science Foundation. (n.d.). *Online evaluation resource library homepage.* Division of Research, Evaluation and Communication, Directorate for Education and Human Resources. http://oerl.sri.com/home.html.

- Presents a matrix of user scenarios that provides a roadmap for how to use the toolkit to meet specific user needs
- Provides access to evaluation plans, instruments, and reports for NSF projects that can be used as examples for designing, conducting, documenting, or reviewing program evaluations

Organization for Economic Co-Operation and Development. (2008). *Best practice guidelines for evaluation.* (PUMA Policy Brief No. 5). http://www.oecd.org/dataoecd/56/17/35060864.pdf.

- Defines "evaluation" and discusses its objectives, main actors, benefits, and costs
- Discusses practices related to strategic management of evaluation activities, enhancing credibility of the evaluation, and fostering an "evaluation culture"
- Discusses practices related to building effective evaluations, such as ensuring links with policymakers, choosing the right evaluator, and involving stakeholders and openly communicating findings

Ovretveit, J., & Gustafson, D. (2002). Evaluation of quality improvement programs. *Quality and Safety in Health Care*, *11*(3), 270–275. http://qshc.bmj.com/content/11/3/270 .full.pdf.

- Provides an overview of the status of the research on hospital quality improvement programs
- Discusses the challenges to evaluating hospital quality improvement programs and presents evaluation designs and methods to address those challenges
- Discusses strategies to improve future evaluations of hospital quality improvement programs

U.S. Department of Education, Institute for Education Sciences. (2010). *What Works Clearinghouse.* http://ies.ed.gov/ncee/wwc/.

- Provides standards for rigorous research in education, including randomized control trials and quasi-experimental studies, and rates the rigor of existing education research [http://ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf]
- Contains a registry of education researchers with experience conducting rigorous evaluations, including those of early childhood education programs
- Provides a checklist of "things to get right" when conducting rigorous education research [http://ies.ed.gov/ncee/wwc/pdf/guide_RCT.pdf]
- Assesses research on early childhood education programs for children ages 3 to 5 [http://ies.ed.gov/ncee/wwc/pdf/ece_protocol_v2.0.pdf]

W. K. Kellogg Foundation. (2004a). *Evaluation handbook.* Battle Creek, MI: Author. http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx.

- Provides a step-by-step guide to program evaluation, from planning and implementing the evaluation through communicating findings
- Details the three components of project evaluation: (1) context evaluation, (2) implementation evaluation, and (3) outcome evaluation
- Discusses research questions, reviews budgeting, and advises how to select an outside evaluator
- Outlines how to collect and analyze data
- Discusses the importance of, and techniques for, communicating evaluation findings, both formative and summative

World Bank Group. (n.d.). *Participatory monitoring and evaluation: Tools and methods.* http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALDEVELOPMENT/EXTPCENG/0,,contentMDK:20509330~menuPK:1278210~pagePK:148956~piPK:216618~theSitePK:410306,00.html.

- Provides descriptions and guidance on the utility and purpose of participatory monitoring and evaluation approaches
- Highlights the strengths and appeal of participatory monitoring and evaluation methods

C. Resources for Selecting Measures and Data Collection Strategies

We conducted a systematic search for resources on selecting measures and data collection strategies (Appendix B details the methods used to find these resources). To organize the resources, we classified them into three categories:

1. **QRIS-Specific Resources.** These resources include information on at least one of three topics: (1) measuring development of children attending child care programs participating in a QRIS, (2) constructs that could be measured in a QRIS evaluation, and (3) elements of QRIS that could be measured in evaluations.
2. **Child Care and Early Childhood Education Resources.** These resources include guidelines for selecting measures and information on measures used in evaluations and other research studies of child care and early childhood education.
3. **General Resources.** These resources provide information on identifying sources of information, selecting instruments, and using best practices data collection procedures.

1. QRIS-Specific Resources

Elicker, J., & Thornburg K. (2011). *Evaluation of quality rating and improvement systems for early childhood programs and school-age care: Measuring children's development.* (Research-to-Policy, Research-to-Practice Brief OPRE 2011-11b). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning. Research and Evaluation.

Tout, K., Zaslow, M., Halle, T., & Forry, N. (2009). *Issues for the next decade of quality rating and improvement systems.* (Issue Brief). Washington, DC: Administration for Children and Families, Office of Planning, Research and Evaluation and Child Trends. http://www.childtrends.org/Files/Child_Trends-2009_5_19_RB_QualityRating.pdf.

- Identifies possible constructs or data elements that could be measured in a QRIS evaluation (page 8)

Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G. & Boller, K. (2010, April). *The Child Care Quality Rating System (QRS) Assessment: Compendium of Quality Rating Systems and Evaluations*, OPRE Report. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. http://www.childcareresearch.org/childcare/resources/18554.

- Details QRIS evaluations conducted as of 2009 and the elements of QRIS that could be measured in a QRIS evaluation

Zellman, G. L., Brandon, R. N., Boller, K., & Kreader, J. L. (2011). *Effective evaluation of quality rating and improvement systems for early care and education and school-age care.* (Research-to-Policy, Research-to-Practice Brief, OPRE 2011-11a). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Forthcoming online.

- Includes possible constructs or data elements that could be measured in a QRIS evaluation
- Contains a resource list for further information on QRIS

2. Child Care and Early Childhood Education Resources

Berry, D., Bridges, L., & Zaslow, M. (2003). *Early childhood measures profiles.* Washington, DC: Child Trends. http://www.researchconnections.org/childcare/resources/8634/pdf.

- Provides profiles of measures to assess a variety of developmental domains in early childhood

Boller, K., Atkins-Burnett, S., Malone, L. M., Baxter, G. P., & West, J. *Compendium of student, teacher and classroom measures used in NCEE evaluations of educational interventions. Volume I. Measures selection approaches and compendium development methods.* (NCEE 2010-4012). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. http://ies.ed.gov/ncee/pubs/20104012/pdf/20104012.pdf.

- Provides criteria to guide measures selection
- Provides comparative information about outcome measures used in studies of educational effectiveness

Halle, T., Vick, J. & Anderson, R. (2010). *Quality in early childhood care and education settings: A compendium of measures (Second edition).* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. http://www.researchconnections.org/childcare/resources /18804/pdf.

- Provides profiles of measures used to assess dimensions of quality in early childhood settings. Profiles include an overview of constructs assessed and procedures for training and reliability.

Hepburn, K. S., Kaufman, R. K., Perry, D. F., Allen, M. D., Brennan, E. M., & Green, B. L. (2007). *Early childhood mental health consultation: An evaluation toolkit.* Washington, DC: Georgetown University, Technical Assistance Center for Children's Mental Health; Johns Hopkins University, Women's and Children's Health Policy Center; Portland State University, Research and Training Center on Family Support and Children's Mental Health. http://gucchd.georgetown.edu/products/ECMHCToolkit.pdf.

- Provides information on tools to measure process and outcomes
- Provides examples of instruments to assess various dimensions of an intervention

Isakson, E. A., Davidson, L. L., Higgins, L. B., & Cooper, J. L. (2011, February). *State-level indicators for social-emotional development: Building better systems.* Columbia University, Mailman School of Public Health, National Center for Children in Poverty. http://nccp.org/publications/pub_997.html.

- Proposes indicators along a spectrum of social-emotional well-being, providing information across the range of normal and at-risk conditions, as well as services that are part of the larger system
- Describes the experience of states in identifying indicators in the social-emotional domain and proposes a framework for developing indicators at the state level

3. General Resources

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. (2010c). *The program manager's guide to evaluation: Second edition.* Washington, DC: Author. http://www.acf.hhs.gov/programs/opre /other_resrch/pm_guide_eval/reports/pmguide/program_managers_guide_to_eval2010.pdf.

- Provides information on sources of information in an evaluation, effective instruments, and data collection procedures (Chapter 7)

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* (FMHI Publication #231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network.

- Provides tools and information to identify and assess effective implementation of evidence-based practices.

University of North Carolina, Chapel Hill. (2008). *National Implementation Research Network.* http://www.fpg.unc.edu/~nirn/.

- Website provides a variety of resources to learn about effective implementation

D. Resources for Selecting and Working with an Evaluator

We conducted a systematic search for resources on selecting and working with an evaluator (Appendix B details the methods used to find these resources). All the resources found provide general guidelines for how to select and work with an evaluator.

1. General Resources

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. (2010c). *The program manager's guide to evaluation: Second edition.* Washington, DC: Author. http://www.acf.hhs.gov/programs/opre /other_resrch/pm_guide_eval/reports/pmguide/program_managers_guide_to_eval2010.pdf.

- Provides guidelines for finding an outside evaluator, managing an evaluation headed by an outside evaluator, and creating a contract with an outside evaluator
- Advises on the challenges of hiring an evaluator and problems that arise in working with an outside evaluator
- Lists the potential responsibilities of the program manager and the evaluator

Organization for Economic Co-Operation and Development. (2008). *Best practice guidelines for evaluation.* (PUMA Policy Brief No. 5). http://www.oecd.org/dataoecd/56/17/35060864.pdf.

- Provides guidelines for the type of evaluator that is most appropriate given a project's needs and characteristics

W. K. Kellogg Foundation. (2004a). *Evaluation handbook.* Battle Creek, MI: Author. http://www.wkkf.org/knowledge-center/resources/2010/W-K-Kellogg-Foundation-Evaluation-Handbook.aspx.

- Describes characteristics of different types of evaluators (e.g., external and internal evaluators)
- Provides guidelines for selecting the right evaluator for a project
- Includes a worksheet on elements to check when selecting an evaluator

APPENDIX B

METHODS USED TO IDENTIFY EXTERNAL RESOURCES INCLUDED IN THE TOOLKIT

This page has been left blank for double-sided copying.

Table B.1 lists the search terms the study team used to compile the resources described in Appendix A referenced for this Toolkit. We used the search terms to find online, publicly available resources. We reviewed the resources the search yielded to verify that they provided relevant information that would aid in development of the Toolkit. Search results that did not include relevant information were not included in the list provided in Appendix A.

Table B.1. Methods Used to Compile External Resources in Appendix A

Topic	Search Terms	Relevant Criteria for Inclusion
Building Logic Models	• Logic model development • Creating logic models • Logic model elements • QRIS logic model • Logic model 'and' intervention • Logic model 'and' evaluation	• Definitions of the elements of a logic model • Guidelines on creating a logic model • Using logic models to guide planning, implementation, and/or evaluation of programs • Examples of logic models for specific interventions/programs • Materials (visual aids, training sheets, builders) developed to help users learn to build a logic model
Designing Evaluations	• Designing evaluations • Evaluation design • Evaluation best practices • Evaluation best practices quality • Evaluation quality rating system • Quality improvement evaluation • Program quality rating evaluation	• Planning, designing, and implementing evaluations • Identifying research questions • Options for evaluation design • Research methods • Guidelines on developing evaluation plans on data collection and analysis • Materials (training sheets, links to other resources) developed to help users learn topics on evaluation planning, design, and implementation
Choosing Measures and Data Collection Best Practices	• Measures QRIS • Child care quality measures • Child development measures • Selecting measures evaluation • Data collection process • Data collection QRIS • Best practices data collection	• Frameworks for selecting measures • Processes for developing measures or collecting data to feed into measures • Measures used in evaluations of QRIS • Measures used in child care quality studies
Selecting and Working with an Evaluator	• Selecting evaluator • Working with evaluator • Hiring an evaluator • Evaluation RFP	• Discussion of RFP process, elements of an RFP, and criteria for evaluating an RFP • Guidance on types of evaluators to look for in early childhood evaluations • Issues relating to working with an external evaluator • Discussion of internal and external evaluation

This page has been left blank for double-sided copying.

APPENDIX C

MEASURES USED IN QRIS EVALUATIONS

This page has been left blank for double-sided copying.

Table C.1 summarizes, and links to, measures used in recent QRIS evaluations. The measures are of varying formats, from protocols used in observations of child care environments to questionnaires sent to child care provider staff. The instruments also vary in their uses: some measure quality, while others document processes or implementation.

To create this list of measures, we contacted members of each QRIS evaluation team that conducted the QRIS evaluations mentioned throughout this Toolkit. The teams provided us with the measures they used to collect data, which we then categorized and summarized for this appendix. The table includes links to each measure, as well as a reference indicating the measure's source.

Table C.1. Measures Used in Evaluations of Quality Rating and Improvement Systems

QRIS	Measure	Measure Type	Description
Colorado Qualistar Early Learning	Family Questionnaire for Child Care Centers Qualistar Early Learning (2003). *Family partnership survey for center-based parents.* Denver, CO: QEL.	Questionnaire	Questionnaire for parents of center-based children. • Includes questions on the nature and frequency of information and activities offered to parents, as well as policies around child illness. Most responses use Likert-type scales or multiple-choice answers.
	Family Questionnaire for Family Homes Qualistar Early Learning (2003). *Family partnership survey for family home-based parents.* Denver, CO: QEL.	Questionnaire	Questionnaire for parents of children in family homes. • Includes questions on the nature and frequency of information and activities offered to parents by family home care facilities, as well as policies around child illness. Most responses use Likert-type scales or multiple-choice answers.
	Documentation Checklist Qualistar Early Learning (2003). *Family partnership documentation checklist.* Denver, CO: QEL.	Interview	Protocol for interview with providers and corresponding ratings. • Includes checklist for site visitors to assess the types of information and activities offered to parents by providers, and the methods providers use to gather information and feedback from parents and families. Identifies the types of documentation that should be reviewed with points corresponding to each assessment question.
Indiana Paths to QUALITY	Paths to QUALITY Classroom Teacher/Lead Caregiver Survey Langill, C., Elicker, J., Ruprecht, K., Kwon, K., & Guenin, J. M. (2009). *Classroom teacher/lead caregiver survey.* West Lafayette, IN: Purdue University, Center for Families and Department of Child Development and Family Studies.	Questionnaire	Questionnaire for lead caregiver or classroom teacher. • Asks about educational qualifications, membership in professional associations, attendance at trade conferences, and participation in professional development. Uses multiple-choice responses.

QRIS	Measure	Measure Type	Description
	Paths to QUALITY Initial Survey with Childcare Providers Langill, C., Elicker, J., Ruprecht, K., Kwon, K., & Guenin, J. M. (2009). *Initial survey with childcare providers.* West Lafayette, IN: Purdue University, Center for Families and Department of Child Development and Family Studies.	Questionnaire	Mail questionnaire for child care providers. • Gathers information on provider's experience with the rating system and perspective on fairness and quality of rating standards. Includes questions on the awareness of staff and parents about the rating system and benefits of participation. Also discusses technical assistance, licensing, and professional development activities of the director. Includes multiple-choice, Likert-type scale, and open-ended questions.
	Paths to QUALITY Follow-Up Telephone Surveys with the Original Sample of PTQ Providers Langill, C., Elicker, J., Ruprecht, K., Kwon, K., & Guenin, J. M. (2009). *Follow-up telephone surveys with the original sample of PTQ providers.* West Lafayette, IN: Purdue University, Center for Families and Department of Child Development and Family Studies.	Interview	Protocol for direct survey of participating providers. • Gathers perspective from providers on their experience with the rating system and changes in rating levels of their program. Also includes questions on technical assistance and program fees. Uses Yes/No and open-ended questions.
	Paths to QUALITY Telephone Survey with Parents Who Have Children in PTQ Classrooms/Family Child Care Homes Langill, C., Elicker, J., Ruprecht, K., Kwon, K., & Guenin, J. M. (2009). *Telephone survey with parents who have children in PTQ classrooms/family child care homes.* West Lafayette, IN: Purdue University, Center for Families and Department of Child Development and Family Studies.	Interview	Protocol for telephone survey of parents and guardians of PTQ children, including decision-making and demographic questions. • Gathers information on family's utilization of child care and the motivation behind their child care choices. Also includes questions on parents' knowledge of provider's licensing and quality rating activities, and demographic background of parent. Mostly multiple-choice responses.

QRIS	Measure	Measure Type	Description
	Paths to QUALITY General Public Survey Langill, C., Elicker, J., Ruprecht, K., Kwon, K., & Guenin, J. M. (2009). *Paths to QUALITY general public survey.* West Lafayette, IN: Purdue University, Center for Families and Department of Child Development and Family Studies.	Interview	Protocol for telephone survey with randomly sampled adults in the general public. • Collects information on demographics and child care arrangements for all children under age 6 in sampled households. Uses rating scales for questions on reasons for choosing child care providers, including location, referrals, and qualifications of the provider. Also includes questions on interviewee's knowledge of provider licensing and quality ratings, and the impact that a quality rating system would have on child care decisions.
Kentucky STARS for KIDS NOW	STARS Quality Rating System Questionnaire Rous, B., Schroeder, C., & Naber, P. (2007). *KY STARS for KIDS NOW evaluation questionnaire.* Training Into Practice Project.	Questionnaire	Mail questionnaire for targeted providers. • Used to determine provider's level of knowledge about the quality rating system. Asks about prior experience with rating system, barriers and incentives to participation, and personal background. Uses multiple-choice and open-ended questions.
	STARS Technical Assistance Follow-Up Survey Kentucky Quality Enhancement Initiative. (2006). *TA Follow-up survey*.	Interview	Protocol for direct survey of providers to gain perspective on technical assistance visits. • Short survey using rating scales to evaluate satisfaction with a recent technical assistance visit, and whether changes were made based on the visit.

QRIS	Measure	Measure Type	Description
Maine Quality for ME	Center-Based Teaching Staff Questionnaire ME QRIS Evaluation. (2010). *Center-based staff questionnaire.* Adapted with permission from Educare, Bounce Network national evaluation instrument. Portland, ME: University of Southern Maine.	Questionnaire	Questionnaire for center-based teaching staff on program and personal background. • Section I surveys staff using rating scales on how well prepared they feel to work with children with special needs, their experiences interacting with parents, job satisfaction, the amount of control they have over their job and classroom, and workplace climate. Also includes rating scales for beliefs about early care and education, what makes children successful in school, and important factors for high quality services, • Section II gathers background information on respondents, including race, education, credentials, professional development, job details, work experience, and languages used.
	Family Child Care Home Teaching Staff Questionnaire ME QRIS Evaluation. (2010). *Family child care home providers questionnaire.* Adapted from Educare, Bounce Network national evaluation. Portland, ME: University of Southern Maine.	Questionnaire	Questionnaire for home-based teaching staff on program and personal background. • Section I gathers background information on respondents, including race, education, credentials, professional development, job details, work experience, and languages used. • Section II surveys staff using rating scales on how well prepared they feel to work with children with special needs, their experiences dealing with parents, job satisfaction, the amount of control they have over their job and classroom, and workplace climate. Also includes rating scales for beliefs about early care and education, what makes children successful in school, and important factors for high quality services,

QRIS	Measure	Measure Type	Description
	Quality of Child Care Services - Parent Survey ME QRIS Evaluation. (2010). *Parent questionnaire.* Includes Emlen scale on parent perceptions of quality; able to use without permission. Portland, ME: University of Southern Maine.	Questionnaire	Questionnaire for parents on child care program. • Gathers background information on children in child care, including age, gender, special needs, and nature of care provided. Includes questions with rating scales on the types of information provided to parents by the child care provider, and opportunities to get involved and give feedback. Also asks parents to rate statements about their satisfaction with their care provider and the care environment.
Miami-Dade Quality Counts	Director Survey, QRIS Programs Miami-Dade County Quality Counts Evaluation Project Team. (2009, September). *Quality Counts director survey year 1.* Chapel Hill, NC: University of North Carolina at Chapel Hill, FPG Child Development Institute.	Questionnaire	Questionnaire for directors of QRIS programs on views about the QRIS. • Includes multiple-choice questions on experience and satisfaction with the quality rating system, standards, and technical assistance. Includes questions on parent awareness of the rating system, involvement of community partners, and background on the program and director.
	Director Survey, Non-QRIS Programs Miami-Dade County Quality Counts Evaluation Project Team. (2009, September). *Non-Quality Counts director survey year 1.* Chapel Hill, NC: University of North Carolina at Chapel Hill, FPG Child Development Institute.	Questionnaire	Questionnaire for directors of non-QRIS programs on views about QRIS. • Includes multiple-choice and open-ended questions about the provider, their knowledge of the rating system, reasons for non-participation, and involvement of community partners in the rating system.
	Parent Survey Miami-Dade County Quality Counts Evaluation Project Team. (2009, July). *Parent survey year 1.* Chapel Hill, NC: University of North Carolina at Chapel Hill, FPG Child Development Institute.	Questionnaire	Questionnaire for parents of children in QRIS programs. • Asks parents about access to child care in their community and their own utilization of and spending on child care. Also includes questions about awareness of the rating system and how it affects child care decisions, Responses are multiple-choice answers or rating scales.

QRIS	Measure	Measure Type	Description
	Lead Teacher Survey Miami-Dade County Quality Counts Evaluation Project Team. (2009, July). *Quality Counts teacher survey year 1.* Chapel Hill, NC: University of North Carolina at Chapel Hill, FPG Child Development Institute.	Questionnaire	Questionnaire for lead teaching staff on background and views about the QRIS. • Gathers information on teacher's work experience, professional development, credentials, and background. Includes rating scales for knowledge of and experience with quality rating system.
Minnesota Parent Aware	Y2 Report Telephone Interview Questions Child Trends. (2009). *2009 Parent Aware evaluation - stakeholder interviews*. Minneapolis, MN: Author.	Interview	Protocol for various interviewees, including curriculum review committee, providers, legislator, and resource specialist. • Open-ended questions on the implementation of the quality rating system, including the process for reviewing curricula and assessments, the experience of raters, the experience of quality assurance specialists, successes and challenges, the impact of the quality rating system on providers, parents, and children, and the potential effect of statewide implementation.
	Director Survey – Codebook Minnesota Early Learning Foundation Research Consortium. (2009). *2009 Parent Aware evaluation - director survey.* Minneapolis, MN: Minnesota Early Learning Foundation.	Questionnaire	Codebook for survey of child care provider directors. • Contains survey questions and codes for possible answers on the background of the provider, staff and qualifications, salaries and benefits, professional development, spending on quality improvement, types of children enrolled, services and curriculum provided, experience with the quality rating system, background on the director, business activities, and community support. Responses are multiple-choice, yes/no, or open-ended.

QRIS	Measure	Measure Type	Description
	Family Child Care Provider Interview – Codebook Minnesota Early Learning Foundation Research Consortium. (2009). *2009 Parent Aware evaluation - family child care provider survey.* Minneapolis, MN: Minnesota Early Learning Foundation.	Questionnaire	Codebook for survey of child care provider staff. • Contains survey questions and codes for answers for provider staff on work experience, job satisfaction, professional development, beliefs about early care and education, interaction with parents, services and curriculum provided, lesson planning, assessment practices, background and special needs of children in the classroom, quality improvement experiences, and business activities. Combination of rating scale, multiple-choice, and yes/no responses.
	Parent Telephone Survey 2008 Minnesota Early Learning Foundation Research Consortium (2009). *2008 Parent Aware evaluation - parent interview.* Minneapolis, MN: Minnesota Early Learning Foundation.	Interview	Protocol for telephone interview with parents of children in participating providers. • Gathers information on home and family activities, child health and development, parental support and health, provider services, child care arrangements and choices, satisfaction with child care provider, parental involvement, and demographics.
	Survey of Providers Not Currently Enrolled in PA Child Trends. (2010). *2010 Parent Aware evaluation - survey of non-Parent Aware programs.* Minneapolis, MN: Author.	Interview	Protocol for telephone interview with non-participating providers. • Gathers background information on the children served, feedback about a presentation on the quality rating system, as well as perceptions of the system.
Oregon Child Care Contribution Tax Credit (CCCTC) Projects	CCCTC Facility Director Baseline Survey NPC Research. (2007, August). *Child Care Contribution Tax Credit facility director baseline survey.* Portland, OR: Author.	Questionnaire	Baseline questionnaire for provider director. • Includes questions regarding enrollment, income, and facilities.

QRIS	Measure	Measure Type	Description
	CCCTC Facility Director Follow-Up Survey NPC Research. (2009, September). *Child Care Contribution Tax Credit facility director follow-up survey.* Portland, OR: Author.	Questionnaire	Follow-up questionnaire for provider director. • Includes questions about business practices and revenue.
	CCCTC Parent Baseline Survey NPC Research. (2009, April). *Child Care Contribution Tax Credit parent baseline survey.* Portland, OR: Author.	Questionnaire	Baseline questionnaire and consent form for parents of participating children. • Includes questions on parents' experience with child care provider, child care history, finances, and personal background,
	CCCTC Parent Follow-Up Survey NPC Research. (2009, September). *Child Care Contribution Tax Credit parent follow-up survey.* Portland, OR: Author.	Questionnaire	Follow-up questionnaire and consent form for parents of participating children. • Includes questions on parents' experience with child care provider, child care history, finances, and personal background,
	CCCTC Baseline Provider Interview NPC Research. (2005, December). *Child Care Contribution Tax Credit baseline provider interview.* Portland, OR: Author.	Interview	Protocol for baseline interview with providers regarding child care practices. • Includes questions about recent changes, plans for staying in the field, and feelings about being a child care provider.
	CCCTC Follow-Up Provider Interview NPC Research. (2006, May). *Child Care Contribution Tax Credit follow-up provider interview.* Portland, OR: Author.	Interview	Protocol for follow-up interview with providers regarding child care practices. • Includes questions about recent changes, plans for staying in the field, and feelings about being a child care provider.
	Provider Enrollment Survey Child Care Research Partnership and NPC Research. (2007, September). *Participant enrollment survey.* Portland, OR: NPC Research.	Questionnaire	Baseline questionnaire sent to provider staff regarding business practices and qualifications. • Includes questions regarding provider business practices, licensing, skill level, professional development, job satisfaction, and qualifications.

QRIS	Measure	Measure Type	Description
	Provider Follow-Up Survey Oregon Child Care Research Partnership and NPC Research. (2009, September). *Provider follow-up survey.* Portland, OR: NPC Research.	Questionnaire	Follow-up questionnaire sent to provider staff. • Includes questions regarding skills, professional development, job satisfaction, and experience with the rating system.
	Quality of Early Childhood Care Settings: Caregiver Rating Scale (QUEST) Goodson, B. D., Layzer, J. I., & Layzer, C. J. (2005). *Quality of Early Childhood Care Settings: Caregiver Rating Scale (QUEST).* Cambridge, MA: Abt Associates, Inc.	Observation	Observation instrument for visits with child care providers. • Instrument looks at child interaction, environment, materials in the classroom, and potential dangers.
Pennsylvania Keystone STARS	Child Development Evaluation Scale (CDPES) Fiene, R. (1984, August). *Child Development Evaluation Scale.* Harrisburg, PA: Office of Children, Youth and Families.	Observation	A tool for evaluating child care programs. • Includes (1) a scale to assess compliance with licensing requirements, (2) a scale for assessing program quality, and (3) an observation protocol for documenting caregiver-child interaction.
	Child Care Quality Indicators (CCQI) Scale Fiene, R., & Carl, B. (2010, August). *Child Care Quality Indicators Scale.* Middletown, PA: Early Childhood Research Institute.	Observation	An observation tool based on 13 key quality indicators. • Provides protocols to assess caregiver quality, as well as the quality of the child care environment.
Virginia Star Quality Initiative	QRIS Evaluation Form - Director's Version Virginia Department of Social Services. (2009). *Virginia Star Quality Initiative – evaluation form – director's version.* Richmond, VA: Author.	Questionnaire	Questionnaire for directors of child care centers regarding the QRIS rating process. • Includes questions regarding application, rating, site visits, and challenges. Mostly open-ended questions.
	Rater Site Visit Evaluation Form Virginia Department of Social Services. (2009). *Virginia Star Quality Initiative – rater site visit evaluation form.* Richmond, VA: Author.	Observation	Observation instrument for site visit by QRIS rater regarding process of visit. • Includes questions on the receptiveness of staff, any challenges, lessons learned, and time spent on various tasks.

QRIS	Measure	Measure Type	Description
	Assistant Teacher Education, Qualifications, and Training Form Virginia Department of Social Services. (2009). *Standard 1: Assistant teacher education, qualifications, and training form.* Richmond, VA: Author.	Questionnaire	Form for assistant teachers of preschool children and toddlers. • Form includes questions regarding teachers' educational experience, ongoing training, leadership and professional development, personal reflections, and publications.
	Director Education, Qualifications, and Training Form Virginia Department of Social Services. (2009). *Standard 1: Director education, qualifications, and training form.* Richmond, VA: Author.	Questionnaire	Form for program directors. • Form includes questions regarding program directors' educational experience, ongoing training, leadership and professional development, personal reflections, and publications.
	Teacher Education, Qualifications, and Training Form Virginia Department of Social Services. (2009). *Standard 1: Teacher education, qualifications, and training form.* Richmond, VA: Author.	Questionnaire	Form for preschool and toddler teachers. • Form includes questions regarding teachers' educational experience, ongoing training, leadership and professional development, personal reflections, and publications.
	Staff to Child Ratios and Group Size Virginia Department of Social Services. (2009). *Standard 3: Structure – staff to child ratios and group size.* Richmond, VA: Author.	Questionnaire	Data collection form for hourly group sizes in child care--includes children and teachers and staff-child ratios.
	Transition Practices Documentation Checklist Virginia Department of Social Services. (2009). *Standard 4: Environment and instruction – transition practices documentation checklist.* Richmond, VA: Author.	Questionnaire	Data collection form for documents required for child transition to another class or setting according to QRIS standard.

QRIS	Measure	Measure Type	Description
Washington Seeds to Success	Center Director Questionnaire Kovac, M., & Early Learning Initiative Evaluation Team. (2010). *The Seeds to Success modified field test center director questionnaire – winter 2010.* Princeton, NJ: Mathematica Policy Research.	Questionnaire	Questionnaire for directors of child care centers on quality improvement activities. • Paper survey with questions on staff recruitment and management processes, student funding and financial practices, teacher education, training, curriculum, family involvement, employment and personal background, Multiple-choice or rating-scale responses.
	Family Child Care Provider Questionnaire (English) Family Child Care Provider Questionnaire (Spanish) Kovac, M., & Early Learning Initiative Evaluation Team. (2010). *The Seeds to Success modified field test family child care provider questionnaire - winter 2010.* Princeton, NJ: Mathematica Policy Research.	Questionnaire	Questionnaire for home-based child care providers on child care services. • Paper survey with questions regarding child care setting and staffing, curricula and assessments, professional development, family involvement, management practices, and educational background. Multiple-choice or rating-scale responses.
	Teacher Questionnaire (English) Teacher Questionnaire (Spanish) Kovac, M., & Early Learning Initiative Evaluation Team. (2010). *The Seeds to Success modified field test teacher questionnaire - winter 2010.* Princeton, NJ: Mathematica Policy Research.	Questionnaire	Questionnaire for child care staff on center policies and practices. • Paper survey with questions regarding center organization, professional development and training, family relationships, management practices, and personal background. Multiple-choice or rating-scale responses.

This page has been left blank for double-sided copying.