



EDUCATION

THE ARTS
CHILD POLICY
CIVIL JUSTICE
EDUCATION
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INTERNATIONAL AFFAIRS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
SUBSTANCE ABUSE
TERRORISM AND
HOMELAND SECURITY
TRANSPORTATION AND
INFRASTRUCTURE
WORKFORCE AND WORKPLACE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

[Jump down to document](#) ▼

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

Support RAND

[Purchase this document](#)

[Browse Books & Publications](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore [RAND Education](#)

View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND PDFs to a non-RAND Web site is prohibited. RAND PDFs are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation monograph series. RAND monographs present major research findings that address the challenges facing the public and private sectors. All RAND monographs undergo rigorous peer review to ensure high standards for research quality and objectivity.

Child-Care Quality Rating and Improvement Systems in Five Pioneer States

Implementation Issues and Lessons Learned

Gail L. Zellman, Michal Perlman

Prepared for the Annie E. Casey Foundation, the Spencer Foundation,
and United Way America



RAND

EDUCATION

The research described in this report was prepared for the Annie E. Casey Foundation, the Spencer Foundation, and United Way America, and was conducted by RAND Education, a unit of the RAND Corporation.

Library of Congress Cataloging-in-Publication Data

Zellman, Gail.

Child care Quality Rating and Improvement Systems in five pioneer states : implementation issues and lessons learned / Gail L. Zellman, Michal Perlman.

p. cm.

Includes bibliographical references.

ISBN 978-0-8330-4551-5 (pbk. : alk. paper)

1. Child care—United States—Evaluation. 2 Child care services—United States—Evaluation. I. Perlman, Michal. II. Title.

HQ778.63.Z44 2008

353.53'283—dc22

2008033800

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2008 RAND Corporation

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from RAND.

Published 2008 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: <http://www.rand.org>
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Preface

As demand for child care in the United States has grown, so have calls for improving its quality. One approach to quality improvement that has been gaining momentum involves the development and implementation of quality rating and improvement systems (QRISs): multi-component assessments designed to make child-care quality transparent to child-care providers, parents, and policymakers. By providing public ratings of child-care quality along with feedback, technical assistance, and improvement incentives, QRISs are posited to both motivate and support quality improvements.

In this report, we summarize the QRISs of five “early adopter” states: Oklahoma, Colorado, North Carolina, Pennsylvania, and Ohio. We then present results from in-depth interviews with key stakeholders in each of these states, focusing on major implementation issues and lessons learned. The goal of this report is to provide useful input for states and localities that are considering initiating or revising child-care QRISs.

This work represents a first product of the Quality Rating and Improvement System (QRIS) Consortium, a stakeholder group whose goal is to promote child-care quality through research and technical assistance. The work was funded by the Annie E. Casey Foundation, the Spencer Foundation, and United Way America. This study was carried out by RAND Education, a unit of the RAND Corporation. The study reflects RAND Education’s mission to bring accurate data and careful, objective analysis to the national discussion on early child care and education (ECCE). Any opinions, findings, and conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the funders or the QRIS Consortium.

Contents

Preface	iii
Figures	vii
Tables	ix
Summary	xi
Acknowledgments	xvii
Abbreviations	xix
CHAPTER ONE	
Introduction	1
Background	1
Quality Rating and Improvement Systems	3
Setting Goals, Expectations, and Standards	3
Establishing Incentives and Supports	5
Monitoring Performance Through Ratings	5
Assessing Compliance with Quality Standards	6
Encouraging Provider Improvement Through QI Support	6
QRIS Theory	7
QRISs in Practice	7
Limitations in Our Understanding of QRISs	9
Lack of Data	9
Limited Understanding of QRISs as Systems	9
Dearth of Practical Knowledge	10
QRIS Stakeholder Consortium	10
Study Limitations	11
Organization of This Report	11
CHAPTER TWO	
Methods	13
Sampling of States	13
Interviewee Selection	14
Interview Guide	15
Data Collection, Management, and Analysis	16
Other Research Informing This Study	16
CHAPTER THREE	
The Pioneer QRISs and How They Were Developed	19
Elements of a Rating System	19

Rating Components	19
How Components Are Weighted	20
Rating Systems in the Five Targeted States	21
Oklahoma: Reaching for the Stars.....	21
Colorado: Qualistar Rating System.....	22
North Carolina: Star-Rated License.....	23
Pennsylvania: Keystone STARS.....	25
Ohio: Step Up to Quality.....	26
Summary.....	27
QRIS Design and Implementation Processes.....	29
Goal-Setting and Feasibility Assessment.....	29
System Design.....	31
System Implementation	39
System Outputs	41
CHAPTER FOUR	
Lessons Learned	45
State Self-Assessments	45
States Generally Believe That Their QRISs Have Had a Positive Impact.....	45
Factors That Contribute to the Success of a QRIS.....	46
Challenges to Success.....	47
Impediments to Success.....	47
Setting Goals, Expectations, and Standards.....	47
Establishing Incentives and Supports	50
Monitoring Performance Through Ratings	52
Encouraging Provider Improvement Through QI Support.....	53
Dissemination of Information About the QRIS and Provider Ratings	54
QRIS Components and Their Relationships to Each Other	54
Recommendations	55
Precursors to a Successful QRIS	55
System-Development Process.....	56
What Should QRISs Include?.....	56
Quality Improvement	57
Evaluate the Effectiveness of the QRIS	58
APPENDIXES	
A. Interview Guide	59
B. Unpublished Mani Paper on QRISs	63
References	67

Figures

1.1. A Logic Model for QRISs	4
3.1. QRIS Design and Implementation Processes	30

Tables

2.1.	Study State Characteristics	13
2.2.	Interviews by State and Interviewee Category	15
3.1.	Oklahoma: Reaching for the Stars.....	21
3.2.	Colorado: Qualistar Early Learning.....	23
3.3.	North Carolina: Star-Rated License.....	24
3.4.	Pennsylvania: Keystone STARS	25
3.5.	Ohio: Step Up to Quality.....	27
3.6.	Summary of the Five Systems	28

Summary

Introduction

The generally low quality of child care in the United States, documented in a number of studies (e.g., Karoly et al., 2008), has led to calls for improvement. One approach that has been gaining momentum involves the development and implementation of quality rating systems (QRSs): multicomponent assessments designed to make child-care quality transparent and easily understood. Participating providers are assessed on each of the system components and receive a summary rating that they are encouraged to display and that may be made public in other ways as well. In theory, these simple ratings (often 0 to 5 stars or a rating of 1 to 4), enable parents, funders, and other stakeholders to make more informed choices about which providers to use and support, and they encourage providers to improve the quality of care that their program provides. Quality rating and *improvement* systems (QRISs) include feedback, technical assistance, and other supports to motivate and support quality improvements.

A systems perspective provides a useful framework for examining QRISs. Systems analyses posit a set of fundamental activities that, if carefully linked and aligned, will promote system goals. These activities include (1) setting goals, expectations, and standards for the system, (2) establishing incentives for participation and consequences for meeting (or failing to meet) expectations and standards, (3) monitoring the performance of key system entities (in the case of QRISs, program quality levels), and (4) evaluating how well expectations are being met, encouraging improved performance through quality-improvement (QI) support, and distributing performance incentives and other rewards.

Study Questions

In this report, we summarize the QRISs of five states that were early adopters of such systems. We then present results from in-depth interviews with key stakeholders in each of these states; the interviews focused on identifying major implementation issues and lessons learned.

The work attempts to answer four questions:

1. What is the theory of action underlying these systems?
2. What do these pioneer QRISs look like? Which aspects of quality are included as components in these QRISs?
3. How were they developed?
4. What challenges have system designers faced? What lessons may be learned from these early systems?

Methods

The five states included in the study were selected from among the 14 states that had a statewide QRIS in place as of January 2007. The states we chose were QRIS pioneers—they had longer experience designing and implementing a QRIS—and they represented a range of different approaches to QRIS design. We selected states that reflected diversity in terms of geography and population size because we thought that the presence or absence of large rural areas and wide dispersion of programs might significantly affect QRIS implementation. For example, if programs were widely dispersed and there were few programs in an area, parents might be less likely to use ratings as a program selection criterion.

Using these criteria, we selected Oklahoma, Colorado, North Carolina, Pennsylvania, and Ohio for study. We conducted a total of 20 in-depth telephone interviews from February 2007 to May 2007 with four key stakeholders in each state, using a semi-structured interview guide developed for the project. Interviewees included employees at state departments that oversaw or regulated early childhood programs, child care, or education; QRIS administrators; child-care providers; and representatives of key organizations involved in child care, such as local child-care resource and referral agencies, advisory group representatives, funders, and child-care advocates. Interview notes were transcribed, and coded. We then reviewed the interviews, identifying overarching themes and extracting key lessons learned.

Once our draft of the state QRISs was completed, we sent each interviewee our write-up of his or her state's QRIS for review. We then revised and updated our descriptions based on their feedback, incorporating changes that had been made to the systems after the interviews were conducted. In July 2008, one interviewee in each state was asked to review the entire manuscript. These reviews resulted in additional revisions, so that the information on each QRIS presented in this report is current as of July 2008.

Findings

QRISs generally adhere to a model similar to the one we developed and display in Chapter One. Key to the model are ratings of participating provider quality. The theory underlying the model posits that as parents learn about ratings, they will use them in making child care choices, selecting the highest-quality care they can afford. As the ratings are used, more programs will volunteer for ratings so as not to be excluded from parents' ratings-based choices. In the longer term, parents will have more higher-quality choices and more children will receive high-quality care. Ultimately, the logic model posits that this will result in better cognitive and emotional outcomes for children, including improved school readiness.

Across the five systems, there was considerable consensus concerning the key components of quality that belong in a QRIS. Each system includes measures of (1) staff training and education and (2) classroom or learning environment (although the latter is only measured at higher levels of quality in some states). States differ on whether they include parent-involvement assessments, child-staff ratios, or national accreditation status. Those states that include accreditation relied primarily but not exclusively on accreditation by the National Association for the Education of Young Children (NAEYC).

Cost issues strongly affected the choice of components and the use of particular component measures in most states. In a number of these pioneer states, environmental rating scales

(ERSs) are a particular subject of debate because of their high cost. An ERS evaluation requires an in-person visit by a trained observer, who evaluates such factors as the physical environment, health and safety procedures, and the quality of staff-child interactions. The ways in which the various quality components are summed and weighted to produce a rating differ across states. States also differ in the level of autonomy afforded providers in earning a rating. In point systems, in which summary ratings are based on total points across components, providers may focus their improvement efforts on those components they believe they can most easily improve (or those that are most important to them); in block systems, where providers must improve in all areas, improvement efforts are more prescribed.

The five states tended to follow similar processes in developing and implementing their QRISs. Each state set goals, assessed feasibility, and designed and implemented its system. In implementing a system, assessments must be conducted, ratings determined, and QI efforts begun. States devised a variety of ways to accomplish these tasks and used different combinations of staff to carry them out. The lack of piloting in most of these states and the relatively fast implementation of their QRISs led to early reassessments and numerous revisions, for example, in the role of accreditation and the number of rating levels.

Most interviewees reported increases in provider and parent interest in QRISs over time. They noted that more providers are volunteering to be rated, and more parents are asking resource and referral agencies about program ratings. Most interviewees believed that their QRIS had been helpful in raising awareness of quality standards for child care. They attributed success to political support, adequate financing of provider incentives, provider buy-in, public-awareness campaigns, and QI support for providers.

These states faced numerous challenges in implementing QRISs. First, a number of states struggled with standard-setting. Some states initially set standards low, because average quality of care was poor and designers worried that overly high standards would discourage provider participation. As programs improved over time, administrators increased standards, which programs resented. Second, states made different decisions concerning minimum standards that programs must meet to receive a rating. Three states require programs to be licensed before they can be rated. The other two states require some level of QRIS participation from all providers by assigning the lowest level of rating to licensed providers; to raise their rating, providers must agree to undergo a full QRIS rating. Several interviewees told us that this latter practice was confusing to parents because it was not clear whether a program received the lowest rating because it was licensed and chose not to participate in the QRIS or because it was part of the QRIS and had earned a low rating. At the same time, this practice brings licensing and the QRIS together and may encourage more providers to be rated so that they can attain a rating higher than the lowest one. States also faced challenges in making QI increments between ratings comparable. In one state, this issue led to significant changes in rating levels.

States also had to decide which components to include. Decisions about which components to include or omit are critical because they send a message to providers, parents, and policymakers about what is important in child care. Several programs struggled in particular about a parent-involvement component. Measures of this concept are not well developed, and the inclusion of additional components generally has nontrivial cost implications. At the same time, unmeasured components are likely to be ignored in favor of the measured ones.

The states we studied have invested substantial resources in their QRISs and have developed a range of financial incentives for system participation and quality improvement, including, for example, professional development support for staff in centers that attained a specified

rating and reimbursements for subsidy-eligible children that increased with provider rating. But funding remains an issue in most states. In some states, low reimbursement rates for children receiving child-care subsidies make it impossible for programs serving these children to attain the highest quality levels because these levels require low child-staff ratios and relatively well-educated providers, two very costly aspects of quality.

Providers are often understandably wary of the rating process and tend to view these ratings as they do licensing: something to “pass.” QRIS designers would like programs to replace this view with a culture of continuous quality improvement, but are unsure about how to effect this cultural change.

Recommendations

Based on our interviews and interpretation, we came up with the following recommendations for developing and refining QRISs.

Precursors to a Successful QRIS

1. Obtain adequate funding in advance and decide how it will be spent. QRISs require money to be effective. It is important to develop realistic cost estimates and to design the QRIS so that sufficient funds are available for key activities and are used in the most effective way.

2. Garner maximum political support for a QRIS. Such support does not require legislation, but lack of support from government, funding agencies, and other organizations that influence the child-care sector can be a major barrier to the ramping up of a QRIS in a timely manner and its continuing fiscal health. The need for broader public support, particularly from parents, is also important, as discussed below.

System Development Process

1. Conduct pilot work if possible and make revisions to the system before it is adopted statewide. If at all possible, significant time and effort should be devoted to an iterative revision process in response to a system pilot. Without a pilot phase, states were forced to make many changes after implementation was underway, which led to confusion and resentment. If pilot work is not possible, recognize that revisions are likely and both prepare participants and design the system to accommodate changes to the extent possible.

2. Limit changes to the system after it is implemented. Setting up a system of continuous quality improvement with clear incentives for improvement and a substantial number of rungs to climb may be the best way to encourage continuous quality improvement without imposing new requirements. Constant changes, including raising the bar to prevent provider complacency, create confusion for parents and may undermine their trust in the system. A strategy should be put in place as well to avoid the “provider fatigue” that may result from frequent changes.

What Should QRSs Include?

1. Minimize use of self-reported data as part of the QRS. Such data may bias ratings because providers have strong incentives to be rated well in these increasingly high-stakes systems where there may be significant consequences attached to ratings. However, such data can be helpful as part of QI efforts.

2. Licensing should ideally be integrated into the system. To the extent possible, rating systems should be integrated. One way to do this is to assign all licensed providers a star rating of “1” unless they volunteer for a rating and are rated higher.

3. Use ERSs flexibly by incorporating both self-assessments and independent assessments at different levels of the QRS. ERSs have substantial value. At least some of this value may be captured by using ERSs in more creative—and economical—ways.

4. Do not include accreditation as a mandatory system component. Accreditation based on the former NAEYC system imposed high costs (although limited scholarship dollars were available through NAEYC) and sometimes caused delays in completing ratings due to involvement of another entity. The new NAEYC system may obviate these problems but that is not yet clear. Using accreditation as an alternative pathway to higher ratings may be feasible but requires that decisions be made about equivalence.

5. The rating system should have multiple levels. Including many rungs makes progress more attainable at the lower quality levels, thereby facilitating provider engagement. It also allows for improvement at the higher end, preventing providers from shifting to a “maintenance” mode in which they no longer strive to improve.

Quality Improvement

1. Create a robust QI process. Without resources and support, few programs will be able to change. To effect change, a QRIS needs to provide some mix of staff development, financial incentives, and QI support.

2. Separate raters and QI support personnel. The rating and coaching tasks should be conducted by different individuals so as to avoid creating conflicts of interest that may bias the assessment process.

3. Public-awareness campaigns are important but should start after the system is in place; these campaigns need to be ongoing. Parents only need information about child-care quality for a relatively brief window of time while their children are young. To be useful, public-awareness campaigns need to be big enough to reach many parents and available on an ongoing basis. Such campaigns should be initiated once the system is fully developed, so that the system can deliver on its promises.

Evaluate the Effectiveness of the QRIS

1. Support research on systems and system components. Research that identifies best practices in QRISs is needed so that these practices can be shared. States would benefit from empirical work on key measurement issues, including how best to assess important components and how to combine ratings across components to provide reliable and valid ratings. Research on optimal QI practices and ways to reach parents is also needed. Establishing a QRIS Consortium is one way to accomplish this research.

Acknowledgments

This work was funded by the Annie E. Casey Foundation, the Spencer Foundation, and United Way America, all of which are also supporting the development of the QRIS Consortium. We are particularly grateful to the members of the Advisory Committee of the QRIS Consortium, including Garrison Kurtz, formerly of Thrive by Five Washington and now at Dovetailing; Meera Mani, Early Care and Education Consultant and now at Preschool California; Marlo Nash, formerly of United Way of America and now at Voices for America's Children; Doug Price, Founding Chairman of Qualistar Early Learning; Linda Smith of the National Association of Child Care Resource and Referral Agencies; and Gerrit Westervelt of The BUILD Initiative for their guidance and wisdom concerning QRISs and the political and bureaucratic contexts in which they operate. Special thanks to Meera Mani for allowing us to include her report of her QRIS interviews as an appendix to this report.

We are grateful to Emre Erkut of the Pardee RAND Graduate School and Lynda Fernyhough of the University of Toronto for their many contributions to this work. We also thank Cate Gulyas of the University of Toronto for her skilled work with the interview notes. We are most grateful to Christopher Dirks of RAND for his indispensable assistance on this report and throughout the work. The report benefited from reviews provided by Alison Clarke-Stewart of the University of California, Irvine, and Brian Stecher of RAND.

Finally, we are indebted to the interviewees who so willingly shared their time, insights, and hard-earned wisdom and accomplishments. They went beyond the call of duty in subsequently reviewing our write-ups of their state systems. We especially thank those who were asked to read the entire report for accuracy and graciously provided comments within a tight deadline.

Abbreviations

AA	associate of arts degree
CCDF	Child Care and Development Fund
CCR&R	child-care resource and referral agencies
CDA	associate's degree in child development
DHS	department of human services
ECCE	early child care and education
ECERS-R	Early Childhood Environmental Rating Scale–Revised
ERS	environmental rating scale
FDCRS	Family Day Care Rating Scale
ITERS-R	Infant/Toddler Environment Rating Scale–Revised
NACCRRA	National Association of Child Care Resource and Referral Agencies
NAEYC	National Association for the Education of Young Children
NCCIC	National Child Care Information Center
NICHHD	National Institute of Child Health and Human Development
OCDEL	Office of Child Development and Early Learning (Pennsylvania)
PD	professional development
QI	quality improvement
QRIS	quality rating and improvement system
QRS	quality rating system
R&R	resource and referral agency
STARS	Standards, Training/Professional Development, Assistance, Resources, and Support
TANF	Temporary Assistance to Needy Families

Introduction

Background

Research findings point to the importance of the preschool period for children's development and focus attention on the quality of care young children receive (Lamb, 1998; Scarr, 1998; Vandell and Wolfe, 2000). Numerous studies have demonstrated that higher-quality care, defined in various ways, predicts positive developmental outcomes for children, including improved language development, cognitive functioning, social competence, and emotional adjustment (e.g., Howes, 1988; National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network [ECCRN], 2000; Peisner-Feinberg et al., 2001; Burchinal et al., 1996; Clarke-Stewart et al., 2002). However, the care received by many children is not of high quality (NICHD ECCRN, 2003; Duncan, 2003; Karoly et al., 2008), and much preschool care is mediocre at best (Peisner-Feinberg and Burchinal, 1997; National Association of Child Care Resource and Referral Agencies [NACCRRA], 2006). A primary reason for the low levels of quality is the limited public funding for child care.

Concerns about poor-quality care have been exacerbated by a policy focus in recent years on children's academic achievement and the degree to which preschool care promotes school readiness and improves children's academic performance. The No Child Left Behind Act of 2001 (P. L. 107-110) is one policy that increased scrutiny of children entering kindergarten and drew attention to the social and cognitive skills children need to build successful careers at school. In some states, such as California and Oklahoma, concerns about child-care quality and children's readiness for school have led to increased support for publicly funded pre-kindergarten programs. In other states, such as Tennessee, government-funded Pre-K programs focused on young children who are most at risk of entering school without the skills necessary to succeed. These at-risk children are likely to be found in lower-quality care, since some of the most frequently assessed child-care quality indicators (e.g., favorable child-staff ratios and well-educated staff) are costly to achieve.¹

The generally low quality of child care has led to calls for improvement, amid recognition that the current child care system in the United States, if it can be called a system at all, does little to promote quality (National Early Childhood Accountability Task Force, 2007). Indeed, the United States' "system" of child care has been described as "a nonsystem of micro-enterprises" (Kagan, 2008). Most providers are underfunded and only loosely regulated.

¹ A significant exception to the association between cost and quality may be found at Head Start centers and at Child Development Centers sponsored by the Department of Defense for military dependents. In both of these settings, substantial subsidies enable low-income children to receive care of high quality at very low cost (Zellman and Gates, 2002; U.S. Department of Health and Human Services, 2004).

Child care is delivered by a variety of providers, including center-based programs (such as Head Start), Pre-K programs, and public and private centers, as well as home-based family child-care programs and friend-and-neighbor care. Centers and family child-care homes are the most likely to be licensed; they are also the types of care settings that are the focus of quality rating and improvement systems (QRISs).

Quality standards are largely defined by licensing requirements, which are set by states and vary widely in their scope and rigor. For example, while states generally require centers to be licensed, and most (34) require child care homes serving four or more children to be licensed,² seven states do not impose licensing requirements unless a program serves more than five children, and three states do not require any license for family child-care homes (NACCRRA, 2008).

Although much care is licensed, licensing represents a fairly low quality bar, focused as it is on the adequacy and safety of the physical environment. Licensing requirements focus on such things as fencing, square footage, and protecting children's health and well-being by covering plugs and locking up cleaning supplies. They essentially ignore other aspects of program quality, although some states may require minimal caregiver training (NACCRRA, 2006). Moreover, in its focus on easily assessed environmental features, the licensing process creates a "check-box" mentality among providers. Licensing is poorly understood by parents: 62 percent believe that all child-care programs must be licensed, and 58 percent believe that the government inspects all child-care programs. Many believe that licensing includes scrutiny of the program quality and that licensure indicates that a program is of high quality (NACCRRA, 2006).

But even if parents better understood licensing and quality more generally, the limited availability of care in many locations and for key age groups (particularly infants) provides ready clients for most providers, even those who do not offer quality care. This strong demand limits incentives for providers to take often-costly steps to improve. In some cases, providers may not know how to improve, even if they are motivated to do so. In addition, there are few empirical data available that providers can use to help them select the best ways to invest limited funds to maximize improvements in quality. Another constraint on quality improvement is parents' inability to recognize high-quality care and distinguish it from care of moderate or mediocre quality. Although some people believe that quality is obvious and parents will "know it when they see it," research suggests that this is not always the case; parents may not know what to look for, and even if they do, they may make care decisions based on other, more pressing considerations, such as cost and convenience.

The growing scrutiny of child care settings, the lack of market incentives to improve, and the lack of quality-improvement (QI) skills and knowledge among well-meaning providers have fueled concerns about the level of child-care quality and have focused attention on ways to improve it. Increasingly, quality rating systems (QRSs) are being promoted as a mechanism to improve quality in localities and states. These systems represent a mechanism to improve child-care inputs and outcomes through increased accountability, an approach that is consistent with policy efforts in K–12 education. Advocates for improved quality are quite enthusiastic about the potential of these systems, largely because of their scope, the infusion of public funds into them, and their focus on improving quality at all starting levels. They are

² Centers under religious aegis are license-exempt in a few states. Exclusions are made in six states for family child care that is limited to serving children from a single family.

sometimes contrasted with accreditation in this latter respect. Accreditation, generally associated with the National Association for the Education of Young Children (NAEYC) (although other organizations also accredit child care programs), is designed to help child-care providers improve the care they provide by engaging staff in a self-study process followed by a validation visit. However, because of the rigor and cost of the process, accreditation, which is entirely voluntary, has been taken up by very few providers: Less than 10 percent of nonmilitary programs are accredited (see Zellman et al. [2008] for further discussion of accreditation in quality rating systems).

Quality Rating and Improvement Systems

Quality rating systems are an increasingly popular tool for improving child-care quality. They are implemented statewide in a growing number of states, but are also being implemented at other levels (e.g., counties). QRSs are multicomponent assessments 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). The programs described in this report were all QRISs.

QRISs are essentially accountability systems centered around quality ratings that are designed to improve child-care quality by defining quality standards, making program quality transparent to consumers and providers, and providing incentives and supports for quality improvement. Ideally, these systems promote awareness of quality and encourage programs to engage in a process of continuous quality improvement. While QRISs ultimately are expected to promote improved child outcomes (see the QRIS logic model in Figure 1.1), the systems focus more immediately on assessing and improving program inputs and processes.³

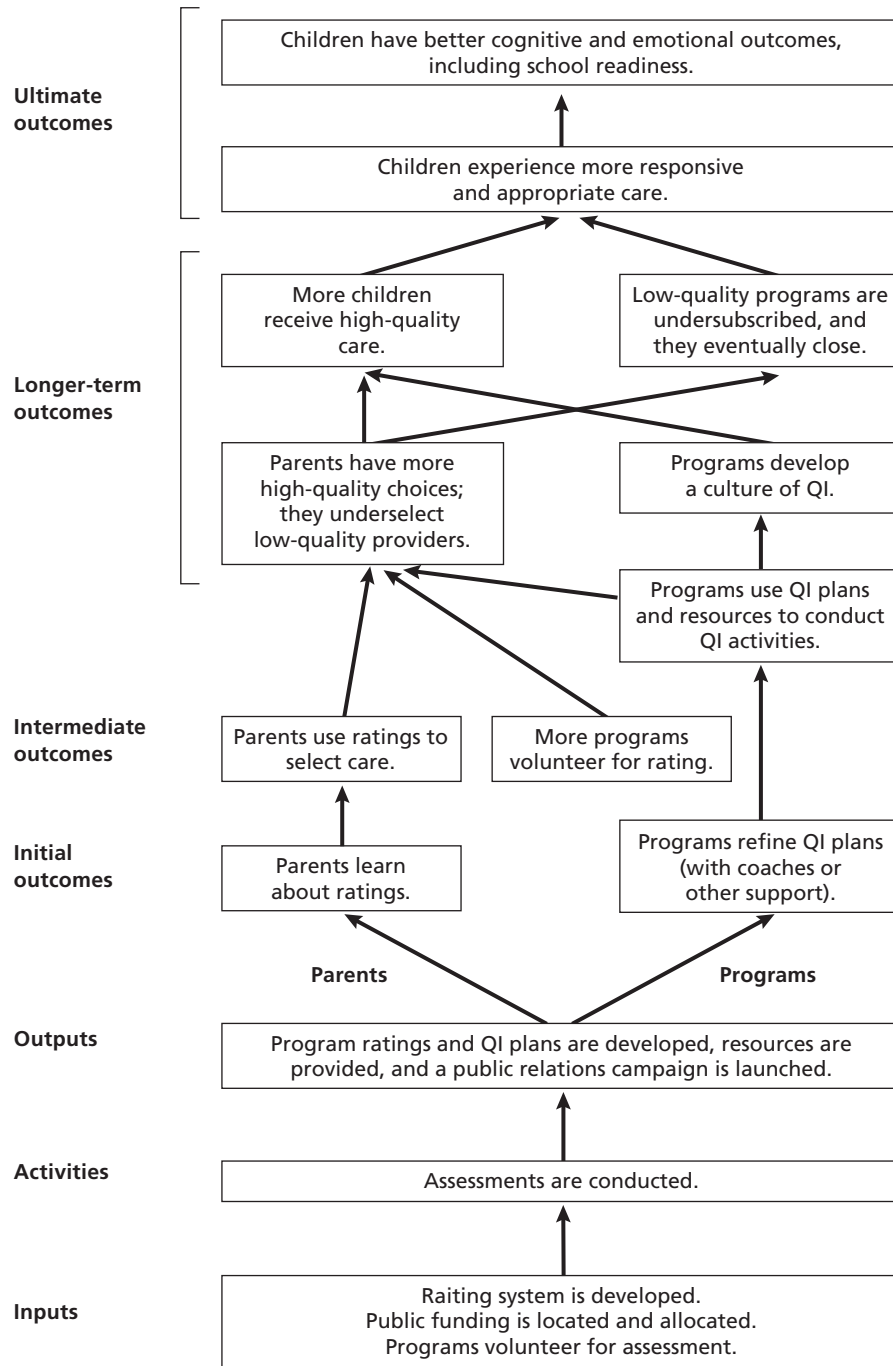
As accountability systems, QRISs can and should be assessed from a systems perspective. Systems analyses suggest a set of fundamental activities that, if carefully linked and aligned, will promote system goals. These activities include (1) setting goals, expectations, and standards for the system, (2) establishing incentives for participation and consequences for meeting (or failing to meet) expectations and standards, (3) monitoring the performance of key system entities (in the case of QRISs, program quality levels), (4) assessing compliance with standards and encouraging improved performance through (QI) support.

Setting Goals, Expectations, and Standards

Although it may seem that QRISs' goals are obvious by their very name, in fact, states (and other localities implementing a QRIS) may have different goals for their systems. For example, some states view ratings as primarily a mechanism for improving quality, whereas other states view these ratings as more of a consumer-education tool. Clarifying system goals is important in ensuring that system components align to support them.

³ In their focus on inputs and processes, QRISs differ from K–12 accountability initiatives, which focus on child outcomes. While most systems assert a link between improved quality and child outcomes, particularly school readiness, focusing on child outcomes in preschool accountability systems has raised concerns because of the mixed findings concerning linkages between child-care quality and child outcomes (see the National Early Childhood Accountability Task Force [2007] and Zellman et al. [2008]).

Figure 1.1
A Logic Model for QRISs



RAND MG795-1.1

Rating systems essentially define quality by identifying which program components will be assessed to determine program quality. States include a variety of components in their ratings, including teacher and director training, teacher credentials, weekly lesson plans, activity “interest areas” in the classroom, daily reading programs, parent involvement, self-assessments,

group size and child-staff ratios, environmental rating scales, and accreditation. However, there is considerable consensus concerning the key components of quality. These include child-staff ratios, group size, staff training and education, and some assessment of the classroom or learning environment. States differ in whether to include and how to weight parent involvement, child-staff ratios, and national accreditation.

A QRIS's highest rating indicates the level of care that the state or locality would like to see in all programs, even if it seems unattainable. By setting the standard high and rewarding progress through a set of ratings, everyone is clear about what it takes to reach the top.

Establishing Incentives and Supports

Incentives can take many forms. One of the key motivating factors in education systems is the quest for prestige (Brewer, Gates, and Goldman, 2001). Individuals generally want to be associated with organizations that are viewed positively by others. But prestige may not be a sufficiently strong motivator, because significant quality improvements, such as reduced child-staff ratios and improved staff education and training, are costly to implement.

Financial incentives, if they are sufficient, can support costly quality improvements. They also provide an additional revenue source for providers, many of which are small businesses, which can help to stabilize the operation and improve its functioning, especially if business-assistance support is also provided. Providing more funds to higher-rated programs can reward providers for higher quality and help them cover the higher costs of providing higher-quality care. States may also provide staff scholarships or other professional development programs for which eligibility depends on a program's rating.

Incentives may also occur in the form of hands-on QI support. Often, this support begins with detailed feedback on the rating results. In many systems, this feedback is accompanied by a QI plan that is much more specific than an overall quality rating. In many systems, coaches provide specific technical assistance concerning which areas to tackle and how. This package of support can be very motivating for providers, who often don't know how best to spend the limited QI funds they receive through their participation in the QRIS process or how to initiate QI efforts.

Monitoring Performance Through Ratings

The rating process and the quality of ratings that result represent the major QRIS monitoring activities. A number of issues surround these ratings. A key issue is cost: Conducting ratings requires monitoring and observations. These activities may be labor-intensive and therefore costly, particularly when they involve prolonged classroom observations, as is required to administer environmental rating scales. Frequent ratings arguably encourage programs to improve quickly, but they increase the portion of the budget that must be used for ratings. High rating costs reduce the funds available for other system activities, such as QI efforts and incentives.

Another important aspect of program-performance monitoring concerns the integrity of the assessment process itself. The integrity of the rating process is particularly at issue in the increasingly high-stakes contexts in which many QRISs operate. If the QRIS includes tiered reimbursement, a program's rating will affect its subsidy level and the length of its waiting list, so the measures must be reliable and valid indicators of quality. But the empirical support for measures of child-care quality is inadequate in general (e.g., Zellman et al., 2008); some con-

cepts, such as parent involvement, simply have not received significant attention. States that wish to include these concepts cannot employ generally accepted measures in their QRISs.

Another issue in the rating process concerns who will do the ratings. In most systems, raters are specially trained and conduct ratings on a regular basis. This keeps up their skills and reduces the likelihood that observer ratings will diverge over time. Some systems have tried to give raters dual roles, such as raters and coaches, usually to reduce costs (combining roles means that fewer individuals have to go out and visit programs). In most instances, these dual roles create problems. For example, rater-coach suggestions are often taken as requirements. (See Chapter Three for further discussion of this issue in specific QRISs.)

Finally, the rating process represents an opportunity to help programs embrace quality concerns and QI efforts on an ongoing basis. In this sense, the rating process ideally serves as a socialization tool in addition to its assessment function. Detailed feedback, in the form of a QI plan and coaching sessions, may help programs to view quality improvement as an ongoing process that is best done on a continuing basis. This view represents a major change from the approach that most programs take to licensing. There, easily quantifiable requirements, such as square footage and the height of fences, encourage programs to adopt a “check-box” mentality that is not conducive to a focus on continuous quality improvement. For those programs that already wanted to improve but lacked the revenue to do so, the rating process and its attendant QI support make it possible to deliver a higher-quality product.

Assessing Compliance with Quality Standards

System designers must determine how compliance with QRIS standards will be monitored and how closely providers must conform to standards to qualify for a given rating. States may give programs more or less autonomy in meeting standards by choosing to combine rating data through point or block systems. Point systems afford providers autonomy because they aggregate points across components, which allows programs to target specific components for improvement efforts. Block systems require programs to improve quality *within* each component to increase their rating. (See Chapter Four for further discussion of point and block systems.)

Encouraging Provider Improvement Through QI Support

States also must decide how they will deliver QI support. How detailed a plan will be developed? How will QI resources be allocated? Will programs be provided hands-on technical assistance in implementing the plan? Can programs choose which sorts of support they want? States must disseminate their ratings to maximize system effectiveness, although the timing of such dissemination must be carefully considered. If ratings are made public too soon, it may discourage provider participation and increase expectations too fast for a system that is being rolled out over time. Public-information campaigns and Web sites help to increase system effectiveness; some states have spent substantial funds to inform the public of the system. But the ratings are not always as transparent as one might hope. Setting licensing as the criterion at the lowest star level allows providers to opt out of higher ratings, which may serve some system purposes but can confuse the meaning of the lowest rating.

QRIS Theory

The premise underlying QRISs is that child-care quality is difficult to ascertain. Creating an assessment system that produces a single, easy-to-understand rating for each provider allows parents, providers, funders, and other stakeholders to more easily determine a provider's quality.

QRISs generally adhere to a model similar to one we developed, which is shown in Figure 1.1. This model describes the steps leading from the development of an assessment tool and the engagement of providers through several sets of outcomes. It illustrates how inputs, activities, and outputs relate to each other and to the outcomes.

QRIS activities focus on assessments of participating programs. In some states, these assessments are completely voluntary; in others, participation is required only at the lowest level, as it is equivalent to licensing.

The outputs of these assessments always include a program rating. Systems with a QI focus also may produce a QI plan based on the assessment. Some systems provide coaches or other technical assistance to help programs refine the QI plan and come up with plans for implementing selected strategies. Based on this QI activity, programs are expected to improve. Some systems also use the ratings in a public-awareness campaign to inform parents and other stakeholders of the ratings.

The logic model posits that as parents (on the left side of the figure) learn about ratings, they will use them in making child-care choices, selecting the highest-quality care they can afford. As the ratings are used, more programs (on the right side of the figure) are expected to volunteer for ratings, because they do not want to be left out as parents make rating-based choices.⁴ Participating providers are expected to improve their quality through QI support tied to their rating.

In the longer term, to the extent that ratings drive parental choice and programs improve, parents will have more high-quality choices. If parents can afford to choose higher-quality care (usually because subsidies make higher-quality care affordable), lower-quality programs will be undersubscribed and will either close or improve. The longer-term effect will be that more children receive high-quality care. Ultimately, the logic model posits that this will result in better cognitive and emotional outcomes, including improved school readiness.

QRISs in Practice

The idea behind QRISs is compelling. However, there are significant concerns about whether such systems can actually work, given the realities of U.S. child care. In many locations, there is not enough care to meet the need; this is particularly true if the need is for a specific kind of care, such as infant care. Moreover, high-quality care generally costs more than low-quality care, because it requires more favorable child-staff ratios and well-trained caregivers. Unless subsidies are available that lower the costs that parents must pay, high-quality care may not be affordable to parents.

⁴ In Pennsylvania, consumer demand is considered a longer-term process; the state encourages provider involvement by conditioning receipt of other public resources (e.g., participation in the state's preschool program) on QRIS participation and achievement of specified quality levels.

Once a QRIS is designed, providers are asked to volunteer to be rated.⁵ Participating providers are assessed on each of the system components (typically 4 to 7 components) and receive a summary rating that they are encouraged to display. These simple, readily understood ratings (often 0 to 5 stars or a rating of 1 to 4), convey information about the quality of care provided and theoretically enable parents, funders, and other stakeholders to make informed choices about which providers to use or support.⁶

Providers have a number of incentives to participate in quality ratings. Those who believe that they are delivering high-quality care may volunteer to participate in order to make their high-quality care more widely known. Providers who are not so confident may volunteer to participate in a QRIS because of the QI support that they will receive to improve their quality and their rating. Providers may also volunteer because of the incentives provided. In nearly all systems, planners have created financial incentives, such as staff-training scholarships, QI funds, and tiered reimbursement, to encourage participation and to help defray the nontrivial costs associated with quality improvements (see Chapter Four of this report, Stoney [2004], and Mitchell [2005] for further discussion of financial incentives in QRISs). These incentives are important because funding levels in many child-care programs are inadequate to enable significant quality improvements (National Early Childhood Accountability Task Force, 2007).

Parents play their part in the system by making child-care choices; widely available ratings enable them to make these choices based on program quality.⁷ According to market models, parents drive child-care quality by choosing better care for their children. According to Barraclough and Smith (1996), for example, “Neo-liberal policies of instrumental rationality . . . assume that parents seek out high-quality centers which will therefore flourish while those of lower quality will not survive because parents will not choose them” (p. 7). But there is widespread concern that parents are not good evaluators of child-care quality. For one thing, quality is not obvious to the untrained eye (Helburn, Morris, and Modigliani, 2002). In addition, parents have limited information about child-care providers (Grace and O’Cass, 2003; Helburn and Howes, 1996; Long et al., 1996). Many consider only a single center before enrolling their child (Van Horn et al., 2001). Certainly, most lack the background or training to identify or assess key quality indicators on their own (Fuqua and Labensohn, 1986). When researchers have examined whether parents’ child-care quality ratings are consistent with those of expert observers using accepted ratings schemes, they have found that parent and observer quality ratings are generally unrelated, suggesting that parent ratings are not driven by actual child-care quality (Barraclough and Smith, 1996; Cryer and Burchinal, 1997; Cryer, Tietze, and Wesels, 2002). Thus, providing parents with valid and reliable information about provider quality enables them to make the sorts of informed choices that should ultimately improve quality of care.

Another key element in QRISs is the QI plan, which is derived from a provider’s rating. The plan identifies areas of weakness and suggests ways to make improvements. A key aspect

⁵ The National Early Childhood Accountability Task Force (2007) argues that programs that receive public funds should be required to participate in such ratings.

⁶ Morris and Helburn (2000) found that suppliers sometimes supply lower-quality services at the same price as higher-quality services and can get away with it because of parent ignorance. (See also Helburn and Bergmann [2002].)

⁷ It may be that some parents will choose care based on other criteria, such as cultural or language consonance. However, QRISs are increasingly trying to include cultural competence in their systems and include a broader range of providers. If they are successful, parents may not need to choose between quality and other care attributes.

of these plans is the availability of resources that allow providers to implement some or all of the quality-improvement activities described in the QI plan. With hands-on support for quality improvement and resources to make needed improvements, the overall level of quality in the system theoretically will increase. Public funding tied to quality might, over time, constitute another incentive for providers to devote time and resources to quality improvement. This combination of widespread use of a QRIS; targeted QI efforts, including funding to providers to implement a QRIS-based QI plan; widespread dissemination of quality information; and public funding linked to quality constitutes an accountability system that will theoretically raise the overall quality of care.

QRISs have proved popular with state legislatures in recent years because they represent a conceptually straightforward way to improve child-care quality that does not require immediate investment of large amounts of capital. They are also consistent with a general trend toward demanding accountability in government-funded programs. The number of states implementing some form of rating system has increased from 14 in early 2006 to 36 at the beginning of 2008 (see the Mani paper in Appendix B).

Limitations in Our Understanding of QRISs

Despite the widespread appeal and rapid adoption of QRISs, their effectiveness may be limited by a lack of data and understanding concerning key aspects of their functioning. Here we discuss some of the major limitations.

Lack of Data

Although QRSs and QRISs are theoretically appealing, we do not know how well they measure what they purport to measure, whether providers who participate actually improve the quality of the care they provide, or whether children benefit from the improved care they receive as their providers receive QI support. Many of the existing systems are based on consensual ideas about what components of quality matter most. Many of the measures used to assess the components were developed in low-stakes settings, such as research studies or self-assessments, where there were few, if any, consequences attached to a particular score. These measures may not be appropriate in high-stakes settings, where scores could substantially affect a program's bottom line (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999). At the very least, such studies must be conducted; they may show that new measures need to be developed. Some quality components, such as parent involvement, have not been subjected to careful empirical assessment. Nor has the way in which components are weighted and combined into summary measures been studied.

Limited Understanding of QRISs as Systems

Little has been written about these QRISs as *systems*. Examining and understanding QRISs from a systems perspective holds considerable promise for improving these systems through alignment of key activities. These approaches (e.g., Zellman et al., forthcoming) define key system activities and focus on the ways to align them to promote system goals. For example, they stress the need to develop clear standards (in the case of QRISs, quality standards) and to ensure that all system players are afforded incentives to meet those standards.

Dearth of Practical Knowledge

Also missing from the literature is a practical knowledge base to which policymakers can refer in crafting QRIS legislation, designing QRISs, or implementing QRIS components. Although the National Child Care Information Center (NCCIC) has produced several useful publications (e.g., Stoney, 2004; Mitchell, 2005), individuals who must design a QRIS often rely on colleagues in other states to share the lessons they have learned. This information is conveyed informally and unsystematically.

The widespread availability of practical knowledge is extremely valuable to people in the field charged with designing, implementing, and refining QRISs. They must make many complex decisions as they develop their plans. Many of the decisions involve costly tradeoffs (e.g., whether to focus resources on quality assessments or limit assessments so that more funds are available for QI support). Such decisions have profound effects on the system, yet there is little organized information about how to make these decisions.

Such knowledge is important because designing effective QRISs is a challenging task. QRIS are complex structures that involve multiple goals, public and private sectors, and multiple stakeholders of different backgrounds. In addition, QRISs generally operate with limited resources, so that resource allocation decisions that misdirect resources can have negative and long-term implications for the system and its goals.

QRIS Stakeholder Consortium

In January 2006, the RAND Corporation hosted a meeting that brought together state policymakers and implementers representing eight states, staff of child-focused foundations, child advocates, and researchers to explore whether there was interest in forming a QRIS Stakeholder Consortium that would harness existing knowledge, support new research, and create joint research and information-sharing efforts. Interest was high.

One outcome of the meeting was a decision to go forward with the design of a QRIS Consortium. A steering committee was identified that included all of the individuals and organizations that attended the RAND conference; others who were interested but were unable to attend were later added. A far smaller advisory committee was also established that would meet monthly to try to bring the QRIS Consortium into being. This was recognized to be a relatively long-term task, as money would need to be raised, an agenda agreed on, and organizational governance established. In the meantime, some of the work that had been determined in the meeting to be important for the field was initiated. One piece of that work was a report on lessons learned by a small number of states that were among the first to design and implement QRISs.

This is that report. Its goal is to provide useful input for states and localities initiating or revising child-care QRISs. In this report, funded by the Annie E. Casey Foundation, the Spencer Foundation, and United Way America, we summarize the QRISs of five early adopters of such systems: Oklahoma, Colorado, North Carolina, Pennsylvania, and Ohio. We present results from in-depth interviews with key stakeholders in each of these states, focused on identifying major implementation issues and lessons learned.

1. What is the theory of action underlying these systems?

2. What do these pioneer QRISs look like? Which aspects of quality are included as components in these QRISs?
3. How were they developed?
4. What challenges have system designers faced? What lessons may be learned from these early systems?

Study Limitations

The study conclusions and recommendations are based on the experiences of only five states. Although we selected these states to be representative, their experiences surely differ from those of other states, especially states that adopted QRISs later. Within states, we spoke to individuals who were involved in and very knowledgeable about their state's QRIS. However, since we conducted only four interviews in each state, it is possible that other equally knowledgeable actors would have provided different insights and conclusions.

We did not systematically collect the same information from every interviewee or state. Instead, we asked a set of basic questions but encouraged interviewees to focus on those aspects of these complex systems with which they were most familiar. We did ask for recommendations about who might provide information they lacked; in some cases, we followed up with these individuals. The end result was a rich set of information, but that information was not entirely consistent across states.

Organization of This Report

In Chapter Two, we describe study methods. In Chapter Three, we present a brief summary of the QRIS in each of the five states we studied, then we describe the QRIS development and implementation process in these states. In Chapter Four, we describe key lessons identified by the interviewees and extracted by comparing the experiences of the five states. We conclude Chapter Four with a set of concrete recommendations. We also discuss the implications of this work for the development of the QRIS Stakeholder Consortium. Appendix material includes the interview guide and an unpublished report of QRIS Consortium efforts to engage stakeholders.

Methods

This study relied on in-depth interviews of key stakeholders involved with QRISs in five pioneer states. Notes from these discussions were analyzed to understand each system and to identify lessons learned about the design, implementation, and refinement of QRISs.

Sampling of States

States were selected from among the 14 states (as of January 2007) that had a statewide QRIS in place. We chose five states that had had longer experience in designing and implementing a QRIS and that jointly represented a range of approaches to designing and implementing their QRISs.

The primary selection criterion was being a QRIS pioneer. This meant that efforts to design a statewide QRIS had begun before 2002. However, we also sought to ensure some level of diversity in terms of geography and population size, because it seemed possible that geography, particularly the presence of large rural areas or a wide dispersion of programs, could significantly affect QRIS implementation. QRIS stakeholders to whom we spoke held quite different views on this matter. Some suggested that the political history and context of each state had a substantial effect on what a QRIS needed to look like and that dispersion of programs substantially influenced implementation. Others argued that children's developmental needs did not differ by geography and therefore quality measurement and QRISs should not differ by geography. Given the lack of clarity on the role of geography and political tradition, we chose five states that differed in population size and that represented different parts of the country.

The study sample is shown in Table 2.1.

Table 2.1
Study State Characteristics

State	Year QRS Started	Region	Population (in millions)
Oklahoma	1998	South	3.6
Colorado	1999	West	4.8
North Carolina	1999	Atlantic/South	8.9
Pennsylvania	2002	Northeast	12.4
Ohio	2004	Midwest	11.5

The QRIS launch dates in Table 2.1 clarify the pioneer status of most of the states. Although the launch of Ohio's QRIS did not occur until 2004, the state actually began to develop its QRIS in 1999. Although political and financing issues delayed the launch of a pilot until 2004, much of the decisionmaking around the QRIS occurred much earlier. Ohio was selected because we felt that its early design, later implementation, and the substantial period between the two might provide unique insights into issues in QRIS design and implementation.

Interviewee Selection

The selection of interviewees in each state was based on general assumptions about the key stakeholders who are likely to be involved in QRIS development or to be affected by a QRIS. These include state-level departments responsible for early childhood programs, QRIS administrators, child-care providers, parents, advocacy groups, and funders. This assumption did not hold completely in every state; in Colorado, for example, the QRIS is not state-administered, so we did not interview a state-level person there. Within each category, discussed in more detail below, we endeavored to find one or two individuals who had filled these roles or who best represented each category. We began in all cases by asking members of the QRIS Consortium Advisory Committee for nominations of a key person to contact in each state. Within states, when we asked for names, there was considerable consensus about the key players in each QRIS category. We sought to interview each of these individuals.

In each state, interviewees were sought from the following categories, although the categories represent guides rather than requirements. As noted above, in some states, the QRIS process differed, so that we interviewed a set of individuals representing slightly different groups:

- state-level departments that oversee or regulate early childhood programs, child care, or education
- administrators of QRISs (these people might be part of or outside government)
- child-care providers (for-profit and not-for-profit alike). We looked for child-care provider organizations and interviewed their heads whenever possible.
- key organizations involved in the delivery or oversight of child care and early education initiatives, such as resource and referral agencies and Smart Start
- child-care or early-education advocacy groups
- private funders of early child care and education (ECCE) initiatives
- child-care “champions”: individuals devoted to improving child care who don't fit into any of the other stakeholder categories.

To find interviewees, we used a combination of Internet searches (to identify institutions) and snowball sampling (asking already-identified interviewees to nominate people in other groups). We found that it was relatively easy to locate the institutions, since they tend to be limited to just a few per state. Each state has at least one office in charge of child care or child safety. QRISs are managed by an umbrella organization in each state. In states where there also was a county-based or regionalized rating organization, we talked with an administrator at its headquarters. In some states, there were several associations representing providers; we selected the one that was described as most involved in the QRIS. In most states, there was at

least one foundation that focuses on early childhood development; we picked the foundation most often identified by interviewees as being most engaged in the design or implementation of the QRIS. In one state, we found a child-care champion who was not affiliated with any of the above groups.

Our biggest challenge came in finding organizations representing parents who had been involved in some way in the QRIS; indeed, we were unable to locate such an organization in any state. Given the lack of parent organizations, we chose not to interview individual parents, believing they would not be representative of anyone. Of course, most of our interviewees were parents; we made a point of discussing parental views and the lack of parent organizations with them.

The 20 interviews we conducted covered a wide range of individuals, as shown in Table 2.2. Interviewees participated in the study under an assurance of anonymity.

Interview Guide

The research team developed an interview guide to organize the collection of information in the context of semi-structured telephone interviews. The interview guide content was based on the team's experience in researching QRISs. A draft was shared with members of the QRIS Consortium Advisory Committee, and their input was incorporated. The instrument was then informally piloted with two interviewees who were involved with or knowledgeable about at least two QRISs; the question list and ordering of questions were revised based on their feedback.

The interview guide includes almost 60 questions organized around the following topic areas. These areas roughly parallel the process by which a QRIS is conceived, developed, and implemented:

- impetus for a QRIS—what problems would it address, goals for the system
- QRIS design and planning—including, among other issues, key system components, the QRIS's place within the broader early ECCE ecosystem in the state, relations to licensing and accreditation
- financing—including who pays for ratings, incentive structure, tiered reimbursement

Table 2.2
Interviews by State and Interviewee Category

State	State Department	QRIS Administration	Provider Group	Child-Care Initiative	Private Funder	Child-Care Champion	Totals
Oklahoma	1	1		2			4
North Carolina	1	1	1	1			4
Colorado		1	1		1	1	4
Pennsylvania	1	1	1		1		4
Ohio	1	1	1	1			4
TOTAL	4	5	4	4	2	1	20

- implementation—including, among other issues, any piloting of the QRIS, scope of the system, the rating process, roles of key actors, and QI efforts, if any
- modifications to the QRIS during the implementation process
- overall perspective—including, among other issues, perceived needs for research and lessons learned.

The interview guide was designed to be only a guide and not a survey form. The senior researchers who conducted the interviews emphasized different question subsets depending on what type of stakeholder was being interviewed, the nature of each interviewee's experience, and the issues that arose in each state and in each interview. The interview guide may be found in Appendix A.

Data Collection, Management, and Analysis

In-depth, semi-structured telephone interviews of selected interviewees were conducted by members of the research team. Interviews generally lasted about 90 minutes, with a range from 45 to 95 minutes. Interviews were conducted from February to May 2007.

Each interview was conducted by one or two senior researchers; a junior researcher listened and took detailed notes. A research assistant integrated the notes for each question by state and by interviewee type. The study authors then reviewed all the comments on a particular topic and synthesized the material, identifying overarching themes on a state or topic basis and extracting lessons learned by comparing and contrasting the experiences of individual states.

We generated a description of each state's QRIS based on the interviews and on reports sent to us by the interviewees and available on the Web. We sent each his or her state's description for comment. Almost all interviewees responded. We updated the text to reflect changes that had been made to the systems after the interviews were conducted. In July 2008, we asked one interviewee from each state to review the entire manuscript and provide us with additional updates and corrections. Therefore, information in this report on each state's QRIS is accurate as of July 2008.

Other Research Informing This Study

This study has benefited from three other sources of information in addition to the completed interviews:

- Web sites of state systems
- Web sites of ECCE organizations with a nationwide mission
- an unpublished paper written by Meera Mani for the QRIS Consortium, which was the product of Mani's efforts to engage people involved in QRIS design and implementation with the QRIS Consortium. (That paper, which provides updated figures on statewide QRISs and information about additional states, may be found in Appendix B.)

The QRIS Web site for each participating state was accessed before interviews in each state began. These Web sites provided useful factual information in advance of interviews. This allowed the interviewer to ask more system-specific questions and freed up interview time for questions tailored to each individual's experiences and perceptions. The amount and complexity of information provided on a Web site, intended for providers and parents alike, provided an interesting comparison across states as well; state Web sites varied in the level and sophistication of the information provided.

Web sites of ECCE organizations, such as the National Child Care Information Center and the National Association of Child Care Resource and Referral Agencies, including prior studies conducted by these organizations, were also useful for obtaining nationwide information to put the subject states in perspective and for selecting our state sample.

We asked interviewees in each state included in the study whether there had been any evaluations of the QRIS or its component parts. Several states referred us to Web sites of universities that had conducted such studies. We incorporated relevant findings and insights from those studies when appropriate.

The Pioneer QRISs and How They Were Developed

One of the goals of the study was to examine and compare the states' QRISs and their experiences and choices in developing them. This chapter incorporates data from the interviews we conducted in the five study states. The first part of this chapter focuses on what the systems look like. We begin with a brief description of rating system elements and then summarize the key characteristics of the rating systems in each of the five pioneer states.

The second part of this chapter focuses on processes. We first present a figure that we created based on what we learned about the states' QRIS development processes and experiences. In discussing this figure, we highlight similarities and differences across the states in terms of the processes they followed and the decisions they made about the design and implementation of their systems.

Elements of a Rating System

Rating systems are designed to provide information about program quality in a format that, ideally, can be understood at a glance. All states include a number of components of quality, discussed below, which are combined in different ways to produce an overall quality rating. The pioneer states in our study use either a "star" system (in which the lowest ranking is 0 or 1 star) or, in the case of Ohio, a step system. The intention is to create an easily understood ranking system to help both parents and providers quickly distinguish among programs in terms of the quality of care they were providing at the time of the rating.

Rating Components

The quality components, and the weight assigned to each component, vary in important ways from state to state. In assigning a rating, states may decide to evaluate a wide variety of factors, as noted in Chapter One. But, in general, there is considerable consensus concerning the key components of quality. They include child-staff ratios, group size, staff training and education, and some assessment of the classroom or learning environment. States differ in whether to include and how to weight parent involvement and national accreditation. Most states that include accreditation focus on accreditation by the National Association for the Education of Young Children, although they may permit accreditation by other groups (see Mitchell [2005, p. 24] for a list of other accrediting organizations). NAEYC established rigorous standards in 10 specific child-care areas, including curriculum, progress assessment, teacher qualification,

health, and safety.¹ ECCE programs volunteer to be measured against NAEYC's standards. Candidates for accreditation must meet NAEYC's criteria and are judged on the basis of written submissions and a site visit from one or more members of the NAEYC Academy.

Cost and measurement issues strongly affect the choice of components and the use of particular component measures in QRISs, according to our interviewees. Environmental Rating Scales (ERSs) have been the subject of considerable debate in some states. An ERS evaluation of a classroom requires an in-person visit by a trained observer, who evaluates such factors as the physical environment, health and safety procedures, and the quality of child-staff interactions and instruction. To administer an ERS, the rater must spend several hours in each rated classroom. Consequently, such evaluations are costly. Moreover, several interviewees argued that ERSs place too much emphasis on physical attributes of the setting and on hygiene issues, such as hand-washing, and not enough on processes, such as adult-child interactions. There has been considerable debate in some states about whether or not to include a separate measure of parent involvement in the QRIS.² While most agree that parent involvement is an important component of program quality and something that QRISs should promote, there are few available measures of this construct that have even the appearance of validity (see Zellman and Perlman [2006] for an in-depth discussion of the goals of parent involvement in child care settings, and Zellman et al. [2008] for some analysis of a new measure of parent involvement, the Family Partnership measure).

The issues surrounding rating components are discussed in more detail under the "Content Decisions" subhead later in this chapter.

How Components Are Weighted

The way in which the various quality components are weighted and summed has received little empirical or policy attention, but this process and the assumptions underlying it are quite important and vary across states. Some states simply decide to assign equal points to each component, because there is no empirical basis for differentiation. Once points are assigned, some system designers come up with "best guesses" concerning where to cut scores in assigning the rating. Other states choose a maximum number of points and divide them among the components.

States also differ in terms of the autonomy afforded providers in earning a rating through choosing a point or block system. Point systems aggregate across components, which allows programs to target areas for improvement based on ease of change, costs, or other program-determined considerations. Block systems require programs to achieve quality levels *within* categories. This approach imposes more consistency across programs, but it may make it much more difficult and costly for programs to improve their rating level. Point and block systems are discussed in more detail below, in the context of the five QRISs.

¹ NAEYC introduced a revised accreditation process in September 2006. The revised process is based on new Early Childhood Program Standards. The process has also changed: There are now four steps and new forms, terminology, and deadlines. The process still requires annual reports, involves random unannounced visits, and requires programs to report major status and program changes. The accreditation criteria have been field-tested by the Center for Improving Child Care Quality at the University of California, Los Angeles (NAEYC, 2005).

² The ERS includes a Parents and Staff subscale that relies on questions asked of care providers.

Rating Systems in the Five Targeted States

The states' rating systems are presented in the order of the year that their system was first launched, as a pilot or on a statewide basis. A summary table of each state's QRIS characteristics is presented at the end of this section for easy cross-state comparison.

Oklahoma: Reaching for the Stars

Oklahoma's goals were to improve the quality of child care by increasing the training and education of providers, providing parents with a simple tool to evaluate quality of care, and increasing the level of provider reimbursements, which would help to increase the number of spaces available to low-income families. Indeed, this latter goal drove the development of the system: System designers had learned that the state legislature was not willing to increase reimbursements given the low quality of many programs, and it was hoped that a QRIS would make legislators more willing to allocate additional reimbursement funds.

Oklahoma began its QRIS design process in 1997 and launched Reaching for the Stars in 1998, making Oklahoma's the fastest system-development process among the five states we examined. Table 3.1 summarizes Oklahoma's system.

Oklahoma's rating process includes four levels: 1 star, "1 star plus," 2 stars, and 3 stars. Assignment to star levels occurs via a block system in which each star has a fixed set of requirements in terms of quality components.

State licensing has been brought into the QRIS: State licensing directly confers 1-star status. "One star plus" is a temporary program rating that requires a move up to 2 stars within two years; if this doesn't happen, the program reverts to 1-star status. To earn 2 stars, a program must meet the criteria or be nationally accredited. To earn 3 stars, a program must meet the criteria and be nationally accredited. Licensing visits occur three times a year; star reviews occur annually.

Reaching for the Stars heavily emphasizes the quality of staff and the learning program. The rating components include compliance with licensing requirements, teacher and director training, teacher credentials, weekly lesson plans, activity interest areas, implementation of a

Table 3.1
Oklahoma: Reaching for the Stars

Level	Voluntary/ Mandatory	Requirements	Components Rated	Rating System	Frequency
*	Mandatory for license	1 star is automatic with license	Minimum Licensing Requirements	Block	3 times yearly
*+	Within two years, must move up to 2 stars or drop back to 1 star	Apply and meet criteria	Above plus teacher and director training, weekly lesson plans, activity interest areas, daily reading program, parent involvement	Block	Annual
* *	Voluntary	Apply and meet criteria or national accreditation	Above plus teacher credentials, salary compensation, program evaluation including ERS	Block	Annual, plus ERS every 4 years
* * *	Voluntary	Apply and meet criteria and national accreditation	Above	Block	Annual, plus ERS every 4 years

daily reading program, parent involvement, staff compensation, and program evaluation. Providers are expected to use ERS scales for program evaluation, but they do not constitute an input to the 1-star and 1-star plus ratings. Two- and 3-star programs must have an ERS assessment conducted by the state every four years. Group size and child-staff ratios are not part of the system.

Oklahoma did not pilot its rating system, but has changed its system regularly since its initial rollout. Oklahoma began with a two-level system. One star was awarded automatically with licensing. A second star required that a program meet internal quality criteria or achieve NAEYC accreditation. In 1999, a third star was added to its two-level system. The following year, the “1-star plus” level was added because so few programs could reach the 2-star level.

Infrastructure is now in place to encourage programs to raise quality. Rated providers receive state-provided tiered reimbursements (from \$14 to \$31 per child per day) depending on county, provider type, child’s age, and star level. To support this process, designers recognize they need to provide more technical assistance; this responsibility is met by resource and referral agencies affiliated with NACCRRRA, and a few staff at the state department. In 2003, the system started requiring centers that provide care to children whose families qualify for a child-care subsidy to have earned at least 1-star-plus status. Child-care homes must meet 1-star-plus requirements until they are fully licensed to care for these children. Once fully licensed, they may drop back to 1 star (at significant cost), but few do so. As of the time of our interviews, Oklahoma was also contemplating a reduction in the frequency of assessments, now done yearly.

Colorado: Qualistar Rating System

Colorado’s system had multiple goals. It was designed to provide parents with accessible information about quality that would improve their ability to make good child-care choices. The system is also intended to create incentives for providers to improve the quality of their care. The rating process was designed to readily translate into a QI plan that would provide direction for quality improvements that were motivated by differential reimbursement and by public ratings. In addition, it was hoped that differential reimbursement would promote a better-educated, better-paid workforce.

Colorado started taking QRIS-design steps in 1997 and launched its system in 1999. Although Colorado did not formally pilot the QRIS, the gradual implementation of the privately developed rating system, which began in a few limited geographic areas, created a *de facto* pilot. Table 3.2 summarizes Colorado’s system.

The voluntary Qualistar rating includes five components: the learning environment, family partnerships (a measure developed to replace a previous parent involvement measure), staff training and education, child-staff ratios and group size, and accreditation status. The weight accorded to accreditation in the overall score is very low (2 out of 42 possible points; other components are worth 10 points each). Colorado’s QRIS is the most comprehensive of the five states in our study, although it has been criticized for under-weighting accreditation and ignoring administrative practices. Qualistar used its own data and the limited published work available as a basis for component weighting. But research on weighting and combining of components is limited, and this system, like all those we examined, was hampered in basing these decisions on strong empirical data.

Colorado’s rating process yields five levels: 0 stars through 4 stars. Colorado has changed its QRIS rating mechanism several times in response to empirical work conducted by RAND,

Table 3.2
Colorado: Qualistar Early Learning

Level	Voluntary/ Mandatory	Requirements for this Level	Components Rated	Rating System	Frequency
Zero	Voluntary	0-9 points If a program earns 0 points on learning environment, it cannot get higher than Star 0, equal to "provisional" status.	Learning environment; family partnership; staff training and education; group size and child-staff ratio; accreditation (2 points)	Point (highest score is 42)	Every 2 years
★	Voluntary	10–17 points	Above	Point	Same
★★	Voluntary	18–25 points	Above	Point	Same
★★★	Voluntary	26–33 points	Above	Point	Same
★★★★	Voluntary	34–42 points	Above	Point	Same

NOTE: See the Qualistar Early Learning Web site (Qualistar Early Learning, no date) for discussion of the criteria required to attain different numbers of points.

feedback from the provider community, and data on ratings costs (e.g., Zellman et al., 2008; Le et al., 2006; Zellman and Perlman, 2006). Although criteria definitions, weighting of the components, and method of data collection have changed, the basic structure in terms of component areas (inputs) and number of levels (outputs) has not. For instance, the RAND evaluation of the Qualistar QRIS demonstrated that the initial parent involvement measure produced no variation: Almost all parents were happy with their care. This lack of variation meant that the parent involvement measure was making no contribution to program ratings. As a result, Colorado changed the way it measured parent involvement, focusing on ratings by providers and parents on the degree to which the program partners with parents to help each child (see Zellman and Perlman [2006] for more detail on this measure).

North Carolina: Star-Rated License

North Carolina's goals were to (1) raise the quality of child care, (2) create a licensing system that differentiated quality above minimum licensing requirements and incentivized volunteer providers to engage in QI, and (3) provide parents with better information about quality of care. Reimbursement rates reward higher-quality providers. In addition, the system is designed to help providers intentionally plan for and invest in improvements that result in higher quality and higher star ratings.

North Carolina issued the first star ratings in 1999 to a test group of providers. It launched its system, the Star-Rated License, in 2000, when the Division of Child Development issued Star-Rated Licenses to all eligible providers. Table 3.3 summarizes North Carolina's system.

The three-component licensing system began to transition to a two-component system in 2006, when compliance history (level of conformity with licensing requirements) was dropped as a component for star ratings because it was not a good quality indicator. Seventy-five percent compliance with licensing requirements became a prerequisite for all regulated programs, including religious programs.

The Star-Rated License includes two major rating components: program standards and staff education standards. As noted above, a third component, compliance with licensing requirements, was dropped. Compliance is now considered the baseline level, and all programs

Table 3.3
North Carolina: Star-Rated License

Level	Voluntary/ Mandatory	Requirements	Components Rated	Rating System	Frequency
★	One star is automatic with license	Must have history of at least 75% conformance with licensing requirements	Program environment; child-staff ratios; child-child and adult-child interactions; staff education; teacher and administrator credentials	Point	
★★	Voluntary	4–6 points	Above plus increased square footage and/or reduced child-staff ratios; staff selection and training requirements; parent participation opportunities; operational and fiscal management policies; objective program evaluation; increased staff qualifications; and additional activity requirements	Point	Every 3 years
★★★	Voluntary	7–9 points	Above plus possible ERS	Point	Every 3 years
★★★★	Voluntary	10–12 points	Above plus ERS and further reduced child-staff ratios	Point	Every 3 years
★★★★★	Voluntary	13–15 points	Above plus ERS	Point	Every 3 years

NOTE: See North Carolina Division of Child Development (no date) for information and discussion of the criteria required to attain different numbers of points.

must maintain 75 percent compliance by law. Program standards include program environment (including square footage, activity areas), child-staff ratios, and child-child and adult-child interactions. ERS scores are required only at the 5-star level; however, the majority of programs that have a 4-star license have an ERS done in order to earn the points needed for 4 stars (90 percent of programs). With the change to a two-component system, it is anticipated that a higher percentage of 3-star programs will also have an ERS completed (about 30 percent of 3-star programs have done so). The absence of parent involvement and administrative practices as rating inputs at the 1-star level is worth noting; both are included in the enhanced standards that first apply at the 2-star level.

The rating process includes five levels. Stars are based on a point system; the star rating is determined by the total number of points received out of a possible 15. Program and education standards are weighted equally; an additional quality point is awarded for meeting additional education or program enhancements. The system is voluntary at the higher star levels, but 1 star denotes licensing and is required. Ratings remain in effect for three years, unless the provider requests a reassessment or there is reason to believe that the quality level may have changed. This can happen, for instance, when providers have excessive staff turnover (e.g., more than 50 percent). If the provider asks for a reassessment, it must cover all costs.

Pennsylvania: Keystone STARS

Pennsylvania's system was designed to encourage and support programs and practitioners to improve child outcomes through provision of higher-quality care. Pennsylvania launched its Standards, Training/Professional Development, Assistance, Resources, and Support (STARS) system in 2002 as a pilot. Table 3.4 summarizes Pennsylvania's system.

Keystone STARS includes four components: director and staff qualifications and professional development; early learning program; partnership with family and community; and leadership and management. Each component contains two or more quality dimensions, such as curriculum, community resources, and employee compensation practices. An ERS is used at most levels, but in different ways depending on star level. At the highest levels (3 and 4), a program must achieve a specified ERS score based on external assessor rating. At lower levels, the ERS is used as a self-assessment tool. Ratios and group sizes are not included in quality ratings, because ratios required for licensing are considered to be good and licensing is a mandatory part of the system.

Table 3.4
Pennsylvania: Keystone STARS

Level	Voluntary/ Mandatory	Requirements	Components Rated	Rating System	Frequency
Start with STARS	Voluntary	None; this is an education and readiness level			Annual
*	Licensing is a prerequisite, but provider must meet additional requirements	Program meets performance standards for learning program, family and community partnerships and leadership management	Director and staff qualifications; early learning program; partnership with family and community; and leadership and management; ERS used only for self-assessment	Block	Annual
**	Voluntary	Above, plus staff qualifications and professional development performance standards	Above	Block	Annual
***	Voluntary	Above, plus average facility ERS score across classrooms of at least 4.25; no classroom ERS below 3.5	Above plus ERS score	Block	Annual
****	Voluntary	Above, plus average facility ERS score across classrooms of at least 5.25 and no classroom ERS below 4.25; NAEYC accreditation accepted for a 4-star rating, but ERS required in Years 2 and 4 of accreditation cycle	Above plus ERS score, NAEYC accreditation (not required), learning time, and goal setting	Block	Annual

NOTE: See Pennsylvania Early Learning (2006) for more information on the Keystone STARS program.

The rating process yields one of five levels: “Start with STARS,” 1-star, 2-star, 3-star, and 4-star. Start with STARS serves as an introduction to the concept of continuous quality improvement. As with Oklahoma, assignment to star levels occurs via a block system, with a fixed set of expectations for each level. System participation is voluntary. Licensing is a prerequisite for star 1, but not sufficient to earn that rating.

Pennsylvania recently has made some minor changes to its standards, based on new research relating standards to outcomes. The changes include more focus on curriculum, assessment, and outcomes for children. Accredited programs now require ERS scores (prior to 2008, accredited programs could automatically earn 4 stars) and are assessed in terms of professional development hours for staff and annual goal setting. Next steps include a strengthening of the continuous QI requirements in the standards and focus in professional development on crafting an annual program improvement plan with benchmarks based on sources of evidence.

Ohio: Step Up to Quality

Ohio’s QRIS, Step Up to Quality, was designed to improve the quality of licensed early childhood programs to support optimal child development and learning and to provide parents with a tool to assist in selecting quality programs.

Step Up to Quality design work began in 1999 as a way to enhance the child-care licensing system. A first draft of a QRIS was completed in 2000, but the system was shelved until 2005 because of the imposition of cost-containment measures in the child-care subsidy program.

In 2005, Step Up to Quality was piloted in nine counties. In 2007, the governor took the program statewide before leaving office. As of summer 2008, the governor has supported expansion efforts across the state. The system is fully publicly funded, primarily out of Ohio’s quality set-aside and the federal Temporary Assistance to Needy Families (TANF) program. Table 3.5 summarizes Ohio’s system.

Step Up to Quality was designed to be a three-step system, but the pilot revealed that a number of programs could not activate or maintain compliance with key licensing requirements; an additional status, “Getting Ready,” was created to provide support to these programs. When Step Up to Quality went statewide, the “Getting Ready” status became known as “Emerging Star.” Programs eligible for “Emerging Star” status have 12 months to become “star rated.” During this 12-month period, the program receives intensive technical assistance based on ERS scores in all age groups as well as support in meeting the Step Up benchmark indicators. After an Emerging Star program becomes star rated, it receives quarterly technical assistance visits during the next year until the program’s star rating is renewed.

Ratings are conducted annually. Provider participation is encouraged through provision of quality payments based on rating and on percentage of subsidized children. These payments may be used for critical repairs, classroom supports, early learning resources, staff training and education, and compensation.

Step Up to Quality covers five broad-based components: ratios and group size, staff education and qualifications, specialized training, administrative practices, and early learning. As in Oklahoma, programs are expected to use ERS for self-assessment, but an ERS measure is not a part of the actual rating. Parent involvement is not included. NAEYC accreditation does not play a formal role in Step Up; however, NAEYC accreditation may be used as an alternate pathway for meeting ratio and group size requirements for Steps 2 and 3.

Ohio’s system is voluntary; licensing compliance is required to be eligible for a rating. At the time of our interviews, Ohio was considering making participation mandatory for pro-

Table 3.5
Ohio: Step Up to Quality

Level	Voluntary/ Mandatory	Requirements	Components Rated	Rating System	Frequency
Emerging Star	Voluntary	License is a prerequisite. These are programs with a serious risk of licensing noncompliance and are not eligible for star rating.	Programs receive technical assistance; ERS assessments result in a Quality Improvement Plan	Block	Up to one year of technical assistance
Step 1	Voluntary	Administrator has CDA or equivalent; one lead teacher with ECE, AA (associate of arts degree), or equivalent; 5 hrs. annual specialized training; administrator self-assessment and PD plans; staff benefits; content standards available	Ratios and group size; staff qualifications; training; administrative practices; early learning (ERS for self-assessment)		Annual
Step 2	Voluntary	Administrator has an associate's degree in early childhood education or equivalent; 10 hrs. annual specialized training; program action plans and PD plans implemented; more staff benefits; curriculum aligned with standards; parent communication; classroom self-assessments; children screened	Above		Annual
Step 3	Voluntary	Administrator has an associate's degree in early childhood education or equivalent; each classroom has lead teacher who has an AA; 50% have assistant teachers with CDA or equivalent; 15 hrs. annual specialized training; program action and PD plans implemented; more staff benefits; curriculum aligned with standards using child assessment data; developmental screenings; parent communication; classroom action plans	Above, alignment with NAEYC criteria		Annual

NOTES: CDA = associate's degree in child development; PD = professional development.

grams funded through the Early Learning Initiative; it is now mandatory. Funding has not been tiered as of summer 2008 and instead is based on enrollment of subsidized children.

Summary

There are significant similarities in the QRISs adopted in these five pioneer states. For example, all include QI as a goal. However, there are also substantial differences, particularly in the components included in the rating systems, the incentives attached to different ratings, and their relationships to licensing and accreditation. Key similarities and differences are summarized in Table 3.6.

Table 3.6
Summary of the Five Systems

QRS Information	Oklahoma	Colorado	North Carolina	Pennsylvania	Ohio
System name	Reaching for the Stars	Qualistar Rating System	Star Rated License	Keystone STARS	Step Up to Quality
Year launched	1998 (with two levels)	1999	September 2000	2002	2006 (statewide)
Did the QRIS development process include a pilot phase?	No	Yes (in practice)	No, but gradual introduction	Yes	Yes; Pilot in nine counties from September 2005 through November 29, 2006.
Has there been executive or legislative action to support the QRIS?	No	State legislation targeted some providers; city legislation for Denver	Yes	No	Yes
Is provider participation in the QRIS voluntary or mandatory?	Providers with children receiving public child-care subsidy must have more than 1 star	Yes	1 star (licensing) mandatory; additional stars voluntary	Yes	Yes
Participation rates ^a	97% of centers	10%	All licensed providers	68%	Unknown
Role of licensing	1 star	License prerequisite for rating	1 star	License required for rating	100% compliance with key licensing standards required for rating
Role of accreditation	Yes: 2 stars require assessment or accreditation; 3 stars require assessment and accreditation	Yes: Accredited programs receive 2 points out of 42	No	Yes: National accreditation confers 4-star status	No, but tried to align Step 3 with national accreditation criteria. Accreditation can meet ratio/group size benchmark in Steps 2 and 3
Number of levels	4 (1–3 stars and “1 star plus”)	5 (0–4 stars)	5 (1–5 stars)	5 (1–4 stars and “Start with STARS”)	4 (Steps 1–3 and “Emerging Star”)
Number of components	9	5	2	4	5
Does the QRIS use a point or block system?	Block	Point	Point	Block	Block
Does the QRIS include an ERS evaluation?	For self-assessment and monitoring only, no input to Star rating except 2 Star Plus programs get ERS every four years	Yes	Only required for highest level	Score thresholds exist for 3 and 4 stars; self-assessment only for 1 and 2 stars	For internal self-assessment (no rating input to QRIS), although overall score equal to or greater than 5 with no subscale score less than 4.5 can meet ratio/group size benchmark in Steps 2 and 3

Table 3.6—Continued

QRS Information	Oklahoma	Colorado	North Carolina	Pennsylvania	Ohio
Has the QRIS been revised since its inception?	Yes	Yes	Yes	Yes	No
Rating frequency	Three times yearly	Every 2 years	Every 3 years	Yearly	Yearly
Components	Teacher and director training; teacher credentials; weekly lesson plans; activity interest areas; daily reading program; parent involvement; self-assessment	Ratios; ERS; family partnership; staff education and training; accreditation	Program standards (program environment; child-staff ratios); education standards (interactions; staff/administration credentials)	Staff qualifications and professional development; early learning program; parent and community partnership; leadership and management	Ratio and group size; staff education and qualifications; administrative practices; early learning

^a We were not able to obtain QRIS participation rates from all states.

QRIS Design and Implementation Processes

Our interviews revealed that the five states tended to follow similar processes in developing and implementing their QRISs. Each state set goals, assessed feasibility, and designed, implemented, and evaluated their system. The process is continuous; the outputs from the “final” stage (evaluation) are, in turn, used to reassess feasibility and make further design and implementation changes.

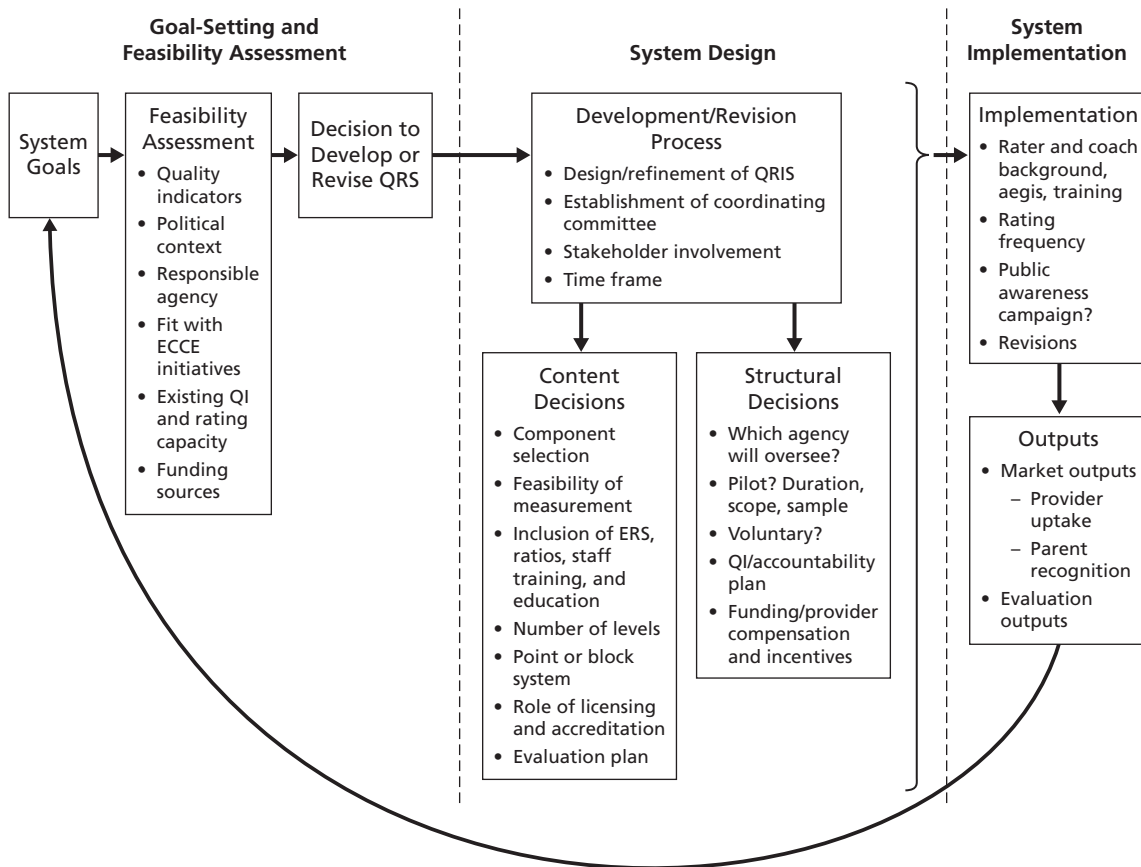
The schematic presented in Figure 3.1 captures the processes followed by these five QRIS pioneer states as they designed, implemented, and refined their QRISs. Below, we discuss some of the key decisions that states made, noting similarities and differences. In Chapter Four, we discuss some of the implications of these decisions for states that are deciding whether and how to launch a quality rating system.

Goal-Setting and Feasibility Assessment

Typically, a QRIS development process begins with a recognition of low child-care quality and a decision to pursue a rating system as a way to improve it. In most cases, the need is apparent: Many or most child care settings are known or understood or suspected to be of unacceptably low quality. In every one of the five pioneer states, the overriding goal for the QRIS is to improve the quality of care provided. Additionally, some states hope to use a rating system and the attention and resources it can bring to the child care arena to raise subsidy levels for low-income children and thereby increase the number of subsidized spaces, professionalize child care, or bring child care into alignment with Pre-K and K–3 education efforts. (See Mitchell [2005] for further discussion of the many goals that underlie these systems.)

Once goals are established, the feasibility assessment process focuses on the *indicators of quality*—what program features indicate quality. Detailed decisions in this area will be made in the design phase, but during the feasibility assessment, it is important to begin developing consensus about what constitutes high-quality child care.

Figure 3.1
QRIS Design and Implementation Processes



RAND MG795-3.1

The *political context* is also evaluated, and, if possible, the support of leading politicians is obtained. In Colorado and North Carolina, for instance, the governor not only lent critical political support but also facilitated funding.

The *agency* that will be responsible for developing and overseeing the QRIS is designated or, if necessary, steps are taken to create it. QRISs may be established by the legislature (through statute) or by the executive branch (through the applicable regulatory agency). While agency regulation allows for far greater flexibility if and when the system needs to be revised, the advantage of statute is its permanence. Other key issues that are generally addressed in this phase include the compatibility of a QRIS with other state *early child-care and education* initiatives and *current quality improvement and rating capacity*.

For the most part, QRISs have benefited from support from the public and private sectors. In all five states, system designers were able to enlist significant political and often *financial* support from the governor for the rating system. Most systems rely on public funding, usually a combination of federal and state money. Funding in Colorado was initially raised through United Way and other private sources. (See Stoney [2004] for further funding details.)

At this point, there is a decision made about whether to move forward with a rating system. Since all the states represented in this report decided to move forward, the development process then began under the aegis of the entity designated during the feasibility assessment.

With one exception, respondents did not distinguish between the planning and design phases of the process; Colorado interviewees specifically identified a six-month planning phase that included about 40–50 meetings. The duration of the design phase generally took 1–2 years.

System Design

Development/Revision Process. In all but one state, *oversight of the design and planning phase* was the responsibility of *coordinating committees or boards* that were created specifically for the project. In Colorado, Qualistar Early Learning (then known as Educare) created an oversight committee to coordinate the design process. In North Carolina, the state Child Care Commission, mandated by statute, coordinated the design process with 16 appointed members drawn from providers, parents, and citizens. Ohio created a broad-based group that designed the system at the outset. This leadership team included members from the resource and referral agency (R&R) and provider associations.³ In addition, two facilitators were hired: one to determine the processes for accomplishing specific tasks, and another (an early childhood expert) to provide content direction. There was a several-year hiatus after the original design was completed. When resources were secured for a pilot, the Ohio Department of Job and Family Services became the lead agency. It worked in conjunction with the Ohio Child Care Resource and Referral Association and the local R&Rs to implement Step Up to Quality. In Pennsylvania, the Child Care Committee's function was to determine what each star level would represent. The committee developed standards then went to the state department for approval and action. Currently, a Keystone Stars advisory committee provides input on program design and operations to the State Office of Child Development. In contrast, Oklahoma interviewees described an absence of real coordination except by the Department of Human Services, where staff were hired to assist with training and administration.

A key issue in this process includes how many and which *stakeholders* to involve. Although bringing many stakeholders to the table was viewed by a number of our interviewees as critical to ensuring buy-in and ultimate support, others noted that including too many participants slows the process and reduces momentum. Key outputs of this process are consensus on QRIS goals and some sense of a *time frame* for moving forward.

Oklahoma involved many stakeholders from both the public and private sectors, including NACCRRRA, Native American tribes, child-care centers, nonprofit and professional organizations, proprietary providers, NAEYC, the state Department of Education, the state Health Department, and colleges. Many business leaders also made significant contributions. Similarly, in Pennsylvania, the design committee, chaired by the United Way, included researchers, practitioners, and association representatives. However, support did vary across states. For instance, in Colorado, QRIS development was a process that many in the child care field did not want, suggesting that providers likely had little to do with creating support for its implementation. In contrast, providers in Ohio and Oklahoma had a significant role in creating support for the QRIS. Indeed, providers in Ohio advertised their participation in the QRIS publicly. As a result of these efforts, families began questioning which centers had good ratings, thereby encouraging provider competition and QRIS uptake. Some of the opposition to QRIS in these states came from people who argued that competition among providers without QI

³ R&Rs work with parents to familiarize them with available programs; in many states, R&Rs have won contracts to provide a range of services to the QRIS, including QI assistance.

support puts some providers at a distinct disadvantage. They contended that low quality occurs because providers who serve lower-income families cannot raise their rates, not because people are not motivated to improve the quality of care they provide.

A broad group of stakeholders was involved in the system design process in Colorado, Ohio, and Pennsylvania. In Colorado, approximately 50 people were involved in the selection of the quality domains, assigned values to indicators, and developed a mission statement with support from child-care experts. In Ohio, the system was designed by a 60-person planning team comprising nonprofit and for-profit provider groups, Montessori and school-age programs, and programs from both public and private schools. Oklahoma solicited input from provider groups, but a few experts and advocates designed the system. This was also the case in North Carolina. There, child-care organizations were more involved in the revision process than in the original design.

Parents were notably missing from the design phase in all states. According to respondents in Colorado, parents were hard to recruit because of busy schedules at work or with their children. After some perseverance, a few parents did become involved later in the process. However, many design-team members were parents themselves and were able to contribute parental sensibilities to the process. Interviewees from both Colorado and North Carolina noted the importance of engaging the provider community in this process as a way of reducing potential resistance to the QRIS.

Content Decisions. The design process in each state focused on content and structural decisions.

One of the most important decisions in QRIS design is *component selection*. As one interviewee astutely noted, “What matters is to include what matters.” In selecting the components for their systems, all states turned to the research literature and precedents to determine which care components were associated with better child outcomes. Interviewees in Colorado and Ohio specifically mentioned relying on the results of the Cost, Quality, and Child Outcomes Study (Peisner-Feinberg et al., 1999; Peisner-Feinberg and Burchinal, 1997; Helburn, Culkin, and Morris, 1995). Oklahoma also included components that developers believed needed attention because they either were either absent or set at low levels in licensing requirements.

These reviews were then subjected to discussions about which components could be well measured and which were *feasible to measure* given wide variation in the cost of their administration. For example, Ohio interviewees noted that parent involvement was originally included in their rating system, but was later dropped when the measures they were using to assess this component, such as number of parents who attend meetings, began to be seen as not credible.

In general, the component choices that the five states made were fairly similar, which is not surprising, as they looked to the same literature and, in some cases, the same studies as the basis for their choices. At the same time, each state had to carefully weigh empirical data against their own values and goals. In all states, interviewees who discussed component choice noted that what gets included is attended to, whereas what is excluded is likely to be ignored. This reality led to different component decisions across states and less than strict reliance on research findings in some cases. For example, several interviewees mentioned a lack of data to support the inclusion of parent involvement and the lack of a clear way to operationalize the construct. Nonetheless, a decision was made in three states to include parent involvement because, in the words of one interviewee, including a parent involvement measure is likely to increase programs’ efforts to promote it.

Use of ERS scales. All states include ERS scales somewhere in their rating systems, although most don't include them as part of the rating at all quality levels. These are labor-intensive observational instruments that focus on the physical environment, health, safety, and the quality of interactions and instruction. ERSs, including the Early Childhood Environmental Scale–Revised (ECERS-R), and its infant and toddler version (the Infant/Toddler Environment Rating Scale–Revised [ITERS-R]), family child-care version (the Family Day Care Rating Scale [FDCRS]), and school-age care version (the School Age Care Rating System [SACRS]) have received special scrutiny because they are the most costly components to assess.

Moreover, questions are increasingly being raised about the validity of these measures (e.g., Perlman, Zellman, and Le, 2004; Scarr, Eisenberg, and Deater-Deckard, 1994). Ohio decided not to include an ERS as a rating component; Oklahoma began with an ERS, then dropped it for cost reasons (except at the higher-rating levels); Pennsylvania and Oklahoma use ERSs as self-assessment tools for the lower ratings. Ohio dropped the ERS when their initial evaluation indicated that ERS scores were not significantly different across Steps, which suggested that the ERS components “do not capture it all,” in the words of one interviewee. Ohio still uses an ERS as a self-assessment tool. Oklahoma chose to use an ERS as an improvement tool rather than as a determinant of high-stakes funding. In Pennsylvania, an ERS became the proxy for quality at the higher levels; designers decided that the use of a standardized ERS would offer providers a well-defined framework of expectations and provide comparability across providers within and across states.

Another issue that must be decided if the ERS is used is how to combine scores across classrooms. In North Carolina, centers were assigned the lowest ERS score received by any of their classrooms under the original three-component system. In the two-component system, centers receive an ERS score based on a classroom average, assuming no classroom falls below a specified floor. One interviewee commented that as a result of this practice, the behavior of a single staff member can have a significant effect on the provider's overall score. In Pennsylvania, centers must have an average total score that is greater than 4.5 to achieve a 3-star rating, provided no classroom receives a score less than 3.5. To achieve the highest (4-star) level, a center must achieve an average facility score of 5.25, with no classroom below 4.25.

Several interviewees noted that ERSs are cumbersome and felt that they overemphasize health and safety. One interviewee noted that centers want more feedback (specifically, item-level feedback) from the ECERS-R/ITERS-R and FDCRS, but Teachers College Press, which holds a copyright on the scoring process and mechanisms, will not allow assessors to provide such feedback. Other interviewees commented on inconsistency in ratings among ERS raters, on raters' disruptiveness in the classroom, and on ERSs being the greatest source of complaints about the QRIS rating process. At the same time, some states, Pennsylvania, reported no ERS problems.

Finally, several respondents questioned the reliability of ERS ratings. They noted instances in which scores notably varied across raters who were rating the same program within the same timeframe. As a result, providers may be penalized (or, for that matter, benefit from) scores that don't reflect their average performance.

The role of child-staff ratios. Some states have opted to include child-staff ratios in their ratings, assigning better ratings to providers that have fewer children per staff member. However, since staff salaries are by far the biggest expense incurred in child care, inclusion of ratios runs the risk of discouraging providers who may not be able to improve their ratios due to cost constraints. Improving on the ratios established through licensing is often impossible for

resource-limited programs, and this can result in little variance in ratios across providers. This reduces the value of ratios in distinguishing gradations of quality among providers. Oklahoma and Pennsylvania do not include child-staff ratios in their QRISs. Interviewees in these states noted that ratios are a licensing requirement, with certification essentially ensuring that these standards are met. One respondent from Ohio argued that improving ratios did not improve quality.

Staff training, education, and turnover. All states include some measure of staff training and education in their QRIS. This practice, although supported by research, raises many issues. Several interviewees pointed out that current practice, which rewards higher education, pushes out otherwise good staff. One interviewee said that insisting on formal education and not allowing staff to substitute experience in the field runs the risk of pushing out the “warm, wise grandmas.” Another interviewee noted that tracking course equivalents is a huge undertaking and determining course equivalents earned in other states or countries has made hiring staff trained elsewhere difficult. Another issue in assessing staff training and education concerns which staff are included. In North Carolina, only full-time staff are counted. This raises issues; for example, the head teacher may only work part-time and therefore her training would not be factored into her center’s rating.

North Carolina is unique in including a measure of staff turnover in its QRIS. Staff turnover is generally considered an important indicator of quality, so it is interesting that North Carolina is the only state to include it.

Number of rating levels to include. Most systems have settled on 4–5 levels, largely based on a sense that fewer are too few and more are too many. Oklahoma’s experience was helpful to other states in this regard. It began with two levels and quickly realized that this was not enough, but the three-level system that came next was also judged to be inadequate: The lowest star was very low, the highest very high. This put most providers at the same rating, defeating the purpose of adding an additional one.

Ideally, interviewees noted, programs will distribute themselves across the levels, moving upward over time. Some states monitor these distributions and tweak their levels to promote a better distribution across participating providers. States also strive to equalize the difficulty of moving from one star level to the next. This has proved quite challenging in several systems. Oklahoma, for example, created a “1-star-plus” rating when it became clear that the leap from 1 star to 2 stars was extremely difficult compared with other transitions.

Point versus block system. Each of these approaches has merits and drawbacks. A point system aggregates across components, which allows programs to target areas for improvement based on program-determined limitations and considerations. For example, a typical resource-constrained program that may not have the funds to improve a component that is quite costly, such as ratios, or that is slow to improve, such as staff education and training, may gain points in other ways that may be less costly or that can be changed more quickly, e.g., instituting a policy of daily written feedback to parents. Point systems also provide benefits to QRIS designers. In particular, states can weigh the point levels attached to each of the components to address their own sense of what is important. A particularly noteworthy example is Colorado’s decision to allot only two points to NAEYC accreditation in their 42-point rating system, while all other components are allotted ten points. Another advantage to point systems from the perspective of system designers is that point weightings can be fairly easily altered.

Block systems require programs to achieve quality levels *within* component categories to improve their ratings. As a result, all the system components must be considered. Advocates

of this approach argue that block systems increase consistency across programs with the same rating, and arguably within star levels, although one interviewee contended that her state's block system produced large quality differences within a given star level. Compared with point systems, block systems reduce provider autonomy. For example, child-staff ratios may need to be under certain specified levels by age group in order for a program to qualify for a particular star rating. A program cannot choose to ignore its ratios and still achieve a particular rating, as they might be able to do in a point system. As Mitchell (2005) states, this disadvantage of point systems can be reduced by requiring that some points be earned in each component category or requiring that the points necessary to achieve a high-level rating exceed the total in one or two components. Interviewees in Pennsylvania, which has a block system, argued that the benefit of the block system is that providers clearly understand what is expected of them and that the block system conveys the idea that improving quality requires work in multiple areas. Block systems also assure parents that certain criteria are met. The drawback, they conceded, is that their block system may create a checkbox mentality—the provider is told exactly what to do to raise quality. As a result, providers are less likely to engage in a process that considers what it takes to raise quality; block systems are less likely to lead to an environment within child-care programs that is focused on quality improvement and how best to achieve it.

The role of licensing and accreditation. Licensing standards are generally understood by all interviewees to be the minimum acceptable standard for child-care programs. What varies is whether the minimal standard must be reached *before* a program may be rated or whether it represents the lowest level of the rating system. In Colorado, Pennsylvania and Ohio, a decision was made that licensing was a prerequisite for a rating. In North Carolina and Oklahoma, licensing represents the lowest level of their respective QRISs—1 star. The decision to include licensing within the rating system has complicated these systems, because the inclusion of licensing means that participation at this level is not voluntary, whereas participation at higher levels is. As a result, it is not possible to know whether a provider with a 1-star rating simply is not interested in participating in the rating system beyond the minimum level and might in fact be a high-quality program, or whether it has tried for a higher rating but failed to achieve it. At the same time, when licensing is part of the QRIS, the QRIS involves the vast majority of centers and many home care providers (depending on a state's licensing policies—see discussion in Chapter One). Some argue that this wide net serves two useful purposes. By making the QRIS nearly universal, the argument for public funding is substantially strengthened. In addition, at least some programs that are given a 1-star rating because they decided not to fully participate in the system may be motivated to do so in order to earn a rating that reflects their actual quality rather than being forced to display an ambiguous 1-star rating.

The place of NAEYC accreditation in states' QRISs also varies; all QRISs made these decisions on the basis of the previous NAEYC system. Like the ERS, accreditation is a complicated rating system component. Requiring accreditation or accepting it as evidence of having reached a particular quality level forces providers to become involved with another system, which has its own rules and timetables. Some states have decided that accreditation rules are not sufficiently consistent with their own views about the components of quality and have chosen not to include it. Moreover, accreditation is costly and time-consuming for programs. However, if accreditation is determined to represent the top rating level, a *state* avoids the cost of rating those programs itself. In some states, there has been considerable discussion about whether to simply make the highest QRIS level equal to accreditation. Oklahoma firmly embedded accreditation in its rating system; accreditation is an option in Pennsylvania. Colorado,

North Carolina, and Ohio have minimized its influence. In these latter states, accreditation has been a contentious issue between providers and QRIS proponents. In Colorado, providers are unhappy that only two of the possible 42 QRIS points are granted for accreditation; they argue that national accreditation deserves more weight in the rating system. Other states, such as North Carolina, have ignored accreditation entirely. Many who oppose including accreditation in QRISs believe that it is not a practical or realistic goal for many providers; many parents cannot afford to support the highly credentialed staff that accreditation demands.

Ohio also does not include accreditation in its QRIS, on the grounds that state-level recognition might be more important than national-level accreditation in changing the culture of programs. One Ohio interviewee did note that providers who have been through the Step Up program find that it is easier to attain national accreditation (if they choose to undergo the process).

Oklahoma has also debated the role of NAEYC accreditation. Currently, a 3-star rating is equivalent to accreditation, based on the argument that accreditation is sufficient and should stand alone as the highest possible rating. Those opposed maintain that accreditation standards are not always maintained after accreditation is obtained. NAEYC has responded to this complaint by instituting a program of random, unannounced visits as part of its revised accreditation system.

Pennsylvania includes accreditation in its QRIS by allowing accreditation status to be one of several pathways to a 4-star rating. Interviewees suggest that some programs would never obtain accreditation if they attempted it on their own. Yet, they do progress using the QRIS, eventually obtaining 4 stars.

The place of accreditation is, in many instances, part of a larger discussion about how to appropriately balance the goal of high quality against reality: Many programs cannot hope to provide care of the highest quality because of the high cost of two of the major drivers of quality—low child-staff ratios and well-educated staff. Programs that serve low-income families in particular often cannot raise their fees sufficiently to pay for better ratios and better-educated staff; they would need a substantial increase in their subsidy level even to pursue accreditation.

This creates a dilemma for system designers, particularly when provider participation is voluntary. Very high standards may discourage providers from agreeing to participate in ratings, yet high standards are the whole point. Some pioneer states chose relatively low standards as a starting point, in an effort to engage lower-quality providers in QI initiatives. North Carolina is an example of a state where standards have been raised several times. But changing standards over time carries its own risks: The public may become confused, and providers may resent achieving a goal only to be told the bar has been raised. This issue is discussed further in Chapter Four.

Evaluation plans. All states recognized the importance of system assessment and committed themselves to conducting ongoing evaluations of their QRISs. To accomplish this, they contracted with local universities. However, as discussed in Chapter Four, the evaluations are limited. They generally focus on the relation between system outputs (ratings) and ERS scores and generally do not address the longer-term outcomes described in the logic model in Chapter One.

Structural Decisions. Structural issues include *which agency will oversee the system* and *how the system will be implemented*. Will there be a pilot? Will rollout be statewide or sequenced? Is provider participation voluntary or required? Decisions also had to be made concerning fund-

ing and provider incentives. Funding availability, along with the desire to begin to improve quality as quickly as possible, drove implementation decisions in several states. For example, when system funding became available, some system designers were reluctant to conduct a pilot, fearing that funding might not be there by the time the pilot phase was completed. Funders also dictated scope in some cases. For example, in Colorado, local funding availability determined the location of the first pilot sites.

All the states we studied decided to make *provider participation voluntary*, although, as discussed above, in states where licensing constitutes the lowest rating level, all licensed providers are assigned the lowest rating level unless they earn a higher rating through a formal rating process.

Funding—for ratings and providers—had to be accessed. As Stoney (2004) notes, some funding already may be available in the system; system designers need to focus on how to draw on existing resources to support system activities. Tapping new funds, both public and private, helped some systems ensure adequate funds. Allocation of funds within the system requires careful decisionmaking. For example, keeping rating costs low may enable the system to provide richer incentives for provider participation and quality improvements.

Another key structural decision concerns whether the system will include a *quality-improvement process*. In other words, will the system be a QRS or a QRIS? If improvement is part of the system, decisions need to be made about who will provide QI, and how the system will make resources available for such efforts. This is a significant decision. In order to create a QI process, ratings—which have intentionally been simplified for parents—must be capable of producing detailed feedback on which a QI plan can be built. Excluding a QI process saves a great deal of money, but without one, the system must assume that providers will be motivated to improve *and* know how to improve on their own once they receive their quality rating. The five states examined in this study determined that a rating alone would not produce improvement; each decided that it was necessary to build infrastructure to support providers who are striving to improve. Without sufficient timely technical assistance, training, and resources, these designers concluded, the system is unlikely to succeed over time.

At the same time, given the resource constraints that characterize the child-care arena, it is not surprising that QI support generally falls short. Funding levels for children eligible to receive a public child-care subsidy are often set at 75 percent of the average market rate for care in the area. Such funding levels are not adequate to cover the costs of quality improvement. And states have not allocated sufficient funds to cover the costs of QI support. For example, Oklahoma, through its Department of Human Services (DHS) Division of Child Care, acknowledged that it did not factor in enough resources for technical assistance. Colorado and Ohio both focused their efforts on creating a broader accountability system as part of their QRIS. Indeed, an underlying principle of the Colorado QRIS was to create accountability. Providers, once rated, were given comprehensive documentation about their quality rating, including recommendations on how to improve quality as well as resources to do so. Subsequent assessments would determine whether quality had improved.

Pennsylvania continues to refine its QI activities, which include technical assistance, case management, education support, and retention awards. In Ohio's pilot phase, resource and referral agencies were responsible for technical assistance. In Oklahoma, the R&R agencies spent six months informing the field about the rating system and provided technical assistance and training to providers.

Colorado has a decentralized system under which a number of local agencies provide ratings, QI support, and incentives. The state's Early Learning Fund, which includes public and private monies, supports R&R staff. Almost every county in Colorado has an early childhood council that can offer grants for professional development (e.g., scholarships and college credits). Colorado school-readiness programs also offer funding for ratings, professional development, and new equipment. Local R&R agencies and local school-readiness coordinators offer technical assistance to providers. Qualistar's training tool, "Getting Ready for Ratings," also helps providers understand what is involved in the rating process. (See the Qualistar Early Learning Web site [no date] for more information.)

In North Carolina, Smart Start—North Carolina's early childhood initiative designed to ensure that young children enter school healthy and ready to succeed—provides technical assistance services. These efforts are guided by very specific performance indicators that must be met, such as a specified percentage of the children in child care in each Smart Start service area are to be in 3-to-5-star programs. North Carolina's R&R agencies receive state funds to work with programs to assess quality (by conducting mock ERS assessments), develop QI plans, define success (star rating plan), and implement QI plans. Support is offered at the board of directors, provider, classroom, and staff levels. The process can take up to a full year, depending on where the program is at inception and what its goals are. For efficiency, some initiatives are pooled across programs. For instance, the R&R agency has developed the "Stars Guard" project, which essentially helps providers guard against reduced ratings (some of which may result from the revision of the assessment and process).

Pennsylvania, through its Office of Child Development and Early Learning (OCDEL) reported that the STARS manager organizes a meeting between the provider and the technical assistant to provide support in specific areas. QI efforts begin after the technical assistant develops a service plan with the provider.

Provider compensation and incentives. In any QI system, incentives are key to promoting provider participation and encouraging quality improvements, especially those that are costly and slow to achieve, such as staff education and training. Care of high quality simply costs more than mediocre care, and even the most motivated providers may not be able to improve unless there is financial support for improvement. In voluntary systems, these incentives are critical to win cooperation. Providers don't want to risk a bad rating; knowing that they can get help and resources to improve may make that risk more acceptable. Some systems link ratings to reimbursement levels to motivate provider improvement, as discussed below.

Oklahoma (through its DHS Child Care Division) offers a variety of incentives to encourage system participation. Salary supplements are provided to help cover the costs of better-trained staff and to decrease staff turnover. Centers are offered \$200–\$2,000 every six months, depending on the teachers' education level, with holders of bachelor's degrees qualifying the center to receive the highest amount. Tiered reimbursement, an incentive structure in which higher-rated programs are paid more, encourages providers to improve quality. For instance, the cash incentive per infant varied from \$14 at level 1 to \$29 in level 3 in metropolitan areas.

Pennsylvania also offers a variety of incentives. For instance, support grants are allocated at the beginning of each year to participating programs. If programs are able to increase their star rating by the end of the year, they receive a merit award. Staff members are also rewarded for getting more education via the Education and Retention Award—money is provided to programs for distribution to staff with specified job tenure who have attained a new ECE

degree. Training is also offered to child-care staff, along with reward points and per diem reimbursements for attending training sessions. Centers are reimbursed for the release time of all enrolling teachers. The governor has a new initiative to provide \$75 million under the Pre-K Initiative to those programs that obtain a rating of 2 stars or higher.⁴ Previously, providers had to be serving at least 25 percent subsidized children to qualify for cash awards. That figure has been restructured: If subsidized children make up between 5 percent and 34 percent, providers are eligible for one level of subsidy; if they serve 34 percent or more subsidized children, they are eligible for another higher level of subsidy.

In North Carolina, differential reimbursement rates for subsidized children are based on star ratings. For some Smart Start subsidies, a provider must obtain at least a 3-star rating. One issue that has resulted from tying some subsidy spots to ratings is that children receiving these subsidies have to move if their provider's quality drops below 3 stars. Subsidies do not cover the full cost of improving quality. In addition to the differential reimbursement rates, many staff receive salary bonuses if the center obtains higher star ratings. Smart Start dollars, United Way contributions, and other foundation money is used for this purpose. Incentives for meeting transition deadlines (e.g., transitioning between star 2 and 3) are also offered.

The Colorado QRIS is unique in that half of its funding consists of foundation and other private funds. These funds are combined into the Early Learning Fund. Moreover, much of the work of the QRIS is carried out by local entities. As a result, the structure and governance of QRIS activities varies, sometimes substantially, across cities and counties. Most providers who have been rated have used funds provided through school-readiness programs to help pay for evaluation and improvements. However, in order to access these funds providers must meet certain other criteria, which basically limited ratings to those in poor neighborhoods. Local early childhood councils in a number of counties have provided funds that have supported the expansion of the QRIS into their communities. The Early Learning Fund channels resources for ratings and QI through R&Rs. Qualistar attempts to align all these local efforts and ensure quality by monitoring and approving local R&R training protocols and QI operating standards and managing the statewide R&R network.

Ohio had no systematic payment bonuses (incentives) for improved quality during its pilot. However, QI grants of \$4,500 were allocated to all participating programs; technical assistance was also provided. These funds were not sufficient to support a scaling up to statewide implementation, nor was there a tiered reimbursement system. Funding was based on the number of subsidized children enrolled, as the focus was on increasing quality for low-income children. However, Ohio is now moving toward a focus on increasing quality for all children.

System Implementation

Once a system has been designed, assessments must be conducted, ratings determined, and QI efforts begun. States have devised a variety of ways to accomplish these tasks and use different combinations of staff to carry them out. As with so many other aspects of QRISs, system implementation involves difficult decisions.

Raters' and Coaches' Background, Responsibilities, and Training. Decisions must be made that balance rater efficiency and cost against independence and expertise. For example, in Ohio, a licensing person visits all participating programs at regular intervals. Might this

⁴ The standard is to increase to a 3-star rating next year.

person also do quality ratings, or is their mindset (and perhaps their training) too focused on licensing and therefore too compliance-oriented? The person who conducts the rating assessments knows the programs well and understands how to improve ratings. This person could serve in a QI function, but would program personnel feel free to openly discuss and even challenge a QI recommendation if it were delivered by someone who was likely to rate the program in the future? A rater who also provides technical assistance may experience a conflict of interest when rating a provider who has not improved, as she may feel that this would reflect badly on her coaching abilities.

The pioneer states we studied considered these issues and adopted different staffing patterns and different schedules for making their ratings.

In North Carolina, the University of North Carolina–Greensboro has a contract with the state to conduct the ERS assessment. A licensing consultant, separate from the assessor, meets with the center director to provide feedback and suggestions for improvements.

In Ohio, Step Up specialists verify everything on site. Licensing staff perform annual inspections mandated by licensing regulations. The QRIS process is completely separate from the licensing process.

In Oklahoma, once the provider's application is received, it is reviewed to evaluate rating eligibility, and a rating is assigned based on paperwork. This approach is unique to Oklahoma. After that, the monitoring of the program is assigned to a licensing specialist. Thirty-six licensing specialists conduct three visits a year to programs. During these visits, accuracy of ratings is verified with a checklist.

Pennsylvania's OCDEL described two separate roles in their rating process, performed by separate individuals. In a coaching role, the STARS manager helps the provider understand the program and its requirements. Although that individual may be an expert in the field, her role is to provide support. The "STARS designator" is an analysis expert who reviews sources of evidence for each rating component using standard worksheets. A separate group conducts ERS and undergoes regular reliability checks.

In Colorado, a number of local agencies support QI activities, which reflects Qualistar's goal of strengthening local communities' capacity to conduct ratings and support QI. Local R&R staff, trained by Qualistar, conduct ratings and, independently, coach programs toward improvement based on Qualistar QI standards. School-readiness grants also support these efforts. In Denver, the Denver Preschool Program also engages providers in these activities.

Consultants in North Carolina, Ohio, and Oklahoma also teach directors how to conduct ERS self-assessments so that they have a better grasp of the system and potential ways to improve. Oklahoma contracts with an outside agency, supervised by the licensing coordinator in the Child Care Division, to conduct ECERS-R observations. ECERS-R assessments are conducted every four years when a provider has obtained a rating of 2 or more stars (this amounts to roughly 2,000 facilities).

Rating Frequency. As shown in Table 3.6, two of the five pioneer states have chosen to rate programs yearly. North Carolina, with a three-year rating cycle, is unique in length of time between ratings. However, all programs are monitored annually; if certain indicators are present (e.g., new director, high staff turnover, serious licensing violations), a rating is conducted. Interviewees in the two annual-rating states reported that a yearly cycle seemed a good compromise: frequent enough so programs did not feel they had to wait an unreasonably long time for a new rating, but far enough apart to allow for serious QI efforts. North Carolina chose less frequent assessments to save money; it has implemented a process to reassess pro-

grams that request a re-rating during the period between ratings to accommodate the longer between-rating period. Colorado moved from an annual to a biannual cycle to save resources and to allow time for program improvements.

Public-Information Campaigns. Another aspect of implementation concerns communication with parents and the community about the rating system. Such communication is critical because parental choice is a key accountability mechanism in a QRIS. Interviewees agreed that a good, ongoing marketing and public-awareness campaign is important. They suggested that these campaigns must focus on publicizing the system and conveying to parents what components are considered in achieving a particular star level. This allows parents to compute their own ratings if they choose. For example, a particular family might value child-staff ratios above provider education levels. Indeed, more than one interviewee stressed that, ideally, parents will consider the star rating to be just one indicator of quality, not the sole indicator. A few interviewees noted a downside to high ratings: In their states, parents have begun to equate stars with cost and limit their search process to lower-rated programs. They argued that while there is a relation between cost and quality, subsidies and other factors blur this association. Public-information campaigns need to address these issues so that parents have as much choice as possible. In North Carolina, interviewees attributed the success of their system (at least in part) to a strong marketing campaign. Interviewees also agreed that it takes a substantial, costly campaign to inform parents about what a rating system is and how it can help them select a quality center. Many aspects of any QRIS can be confusing to parents; lack of consistency across states adds to the confusion. Different states administer the ERS in different ways, for example. One interviewee noted, for instance, that the ratio requirement for the highest star in her current state does not match that of the lowest star in her previous state.

Communication strategies are complicated, of course, by implementation decisions. A small statewide system pilot made a communication campaign particularly challenging in Pennsylvania because ratings began to be published before all participating providers were rated. In another state, some people wanted to hold off on marketing until there was a critical mass of providers at higher star levels, but not everyone agreed. Moreover, as one interviewee noted, a campaign cannot just end at some point after the system launch. Such a campaign would only reach parents who are looking for child care at that point in time, since parents generally attend to information about child care only when they are making child-care decisions; they are less likely to pay attention at other times. Continued communication with providers and other stakeholders is also critical as a means of encouraging participation in a voluntary system. A high level of provider participation is an important tool for engaging parents.

QRIS Revisions. As discussed above, the lack of piloting (especially given the limited empirical knowledge base about QRISs) and the relatively fast implementation of these QRISs led to early reassessments and numerous mid-course corrections in nonpilot states (Ohio used its pilot experience to make some changes to its system before the statewide launch). Some corrections were based on lessons learned in the implementation process: For example, several states revised the role of accreditation in their system. Other corrections reflect the maturing of the QRIS. As we will discuss in Chapter Four, states that set low initial standards (such as Oklahoma) have taken steps to raise the standards over time.

System Outputs

As shown in the QRIS logic model (Figure 1.1 in Chapter One), QRISs produce a number of outputs and outcomes that might be examined, ranging from initial outcomes (more-informed

parents and programs that are developing QI plans), intermediate outcomes (parents using ratings in selecting care and more programs volunteering to be rated), and longer-term outcomes (programs conducting QI activities, more children in higher-quality care). All of the states we included in our study have examined one or more of these outcomes, although the focus of their efforts to date has been on earlier outcomes, with the exception of Colorado.

Provider Uptake and Parent Awareness. Most interviewees report growing provider and parent interest in QRISs over time. Providers are volunteering for ratings in increasing numbers, and parents are increasingly asking about programs' ratings in interactions with referral agencies. In Oklahoma, for example, 2- and 3-star programs rarely have vacancies because parents are eager to acquire a space in these programs. In other states, parents are beginning to question programs that don't participate in the system.

States are investing substantial resources in their QRISs. Assessing in a formal way how well these systems are working is obviously important, particularly so in these pioneer states, where there was limited precedent on which to base their system.

Evaluation Outputs. Evaluations may focus on a number of different issues that are presented in the logic model in Chapter One. For example, evaluations might focus on

- program assessments and their outputs, particularly whether assessment procedures and tools are efficient and effective
- whether the rating system is a reliable and valid measure of quality
- which components are the best predictors of quality
- whether the mechanisms used to weight and combine component measures to produce a single rating are valid
- whether parents know about and use ratings
- whether programs implement QI activities
- whether there are significant differences in quality between star levels
- whether the QRIS is improving program quality
- whether more children receive high-quality care
- whether that care creates better learning environments
- whether children attending high-quality settings (e.g., those with higher quality ratings) have better outcomes (see Stoney [2004] for further detail on evaluation of QRISs).

Oklahoma has been evaluating its system since its beginning, contracting with a joint team from the University of Oklahoma and Oklahoma State University. Much of the work has focused on whether there are differences between star levels on key system criteria, such as ratios, parent involvement indicators, and staff education. The evaluations indicate that there are significant quality differences by star status. Evaluations have also examined whether the system has improved average program quality over time. Using a subset of centers measured soon after implementation and three years later, improvement over time was found. For example, 75 percent of centers at 1-star status in 1999 had reached 2-star status by 2003. Twenty-seven percent of centers at 2-star status in 1999 had reached 3-star status by 2003 (Norris, Dunn, and Dykstra, 2003).

Oklahoma continues to engage a range of stakeholders in their evaluation efforts. Evaluators have conducted an online survey of parent recognition and satisfaction (in which 650 parents participated). Further, to solicit input about the QRIS, the professional development unit of the state department hosts quarterly Reaching for the Stars focus groups consisting of

licensing staff, Reaching for the Stars consultants, and providers. In addition, Reaching for the Stars outreach specialists host focus groups in rural areas, post surveys to parents, and gather information from parents through R&R agencies. R&R agencies also conduct needs assessments of their local areas.

Colorado was committed to rigorous evaluation of its QRIS from the beginning, and therefore data collection has been an integral part of the rating process. Colorado engaged the RAND Corporation to evaluate its system early on. More than one Colorado interviewee stated that, in retrospect, evaluation had begun too soon, before the system components had been tested and refined. The RAND evaluation of the QRIS therefore focused more on the system components than had been planned; two components on which research attention was focused—ratios and family partnership, a parent involvement measure—were improved. The study found that participating providers improved their quality, although the researchers could not determine if this improvement was a function of Qualistar’s QI efforts or the result of self-selection or simple participation. The evaluation is unique in examining the effects on child outcomes of changes in quality over time; no effects were found. Interim findings were provided to Qualistar, which led to a number of modifications to the measurement process, as noted above (see Zellman et al. [2008] for a report of the evaluation).

In North Carolina, the Frank Porter Graham Center at the University of North Carolina, Chapel Hill examined the relationship between star rating level and other indicators of program quality (ERS, teacher education, staff wages, and turnover) in the first group of centers to be licensed under the new Five-Star Child Care Licensing System. Statistically significant associations between star rating level and each of the indicators validated the new system, according to the evaluators (Peisner-Feinberg, 2000). An ongoing assessment effort conducted by the North Carolina Rated License Assessment Project continues to monitor assessments, focusing on average ERS subscale scores and highlighting those that are lowest (see, e.g., Cassidy et al., 2005).

The Pennsylvania Office of Child Development commissioned a study carried out by the University of Pittsburgh Office of Child Development and the Pennsylvania State University Prevention Research Center (Barnard et al., 2006) that closely followed the design used in North Carolina. Researchers examined the association between star rating level and other indicators of program quality (ERS, teacher education, teacher experience, and professional development compliance). Unlike in North Carolina, the study included homes and some randomly selected programs that were not participating in Keystone STARS or were at the low end of the star ratings.

The findings were similar to those in North Carolina; star ratings were related to ERS levels. In addition, teachers with an AA degree or higher had higher ERS scores, as did teachers with five or more years of experience. Finally, use of a standardized curriculum was associated with significantly higher ECERS-R scores on nearly all scales. The authors concluded that the Keystone STARS QRIS is improving quality in participating child-care programs and that it is a reliable indicator of quality.

Ohio has only recently moved to statewide implementation of its QRIS. A presentation of pilot data in 2006 understandably focused on process data. Administrators were asked about the ease or difficulty of several aspects of the application process. Results showed that some aspects, such as the mandatory orientation, were considered simple, while other components, such as the development of an evidence binder, were considered challenging by most. Administrators also generally endorsed the importance of the key system benchmarks: ratios,

accreditation, staff education, staff training, workplace characteristics, and early learning. Most administrators asserted that the system has helped staff, board members, and parents better understand quality. More than two-thirds valued QI and early learning resource grants.

Like North Carolina and Pennsylvania, Ohio also assessed the association between rating level and ECERS-R score using pilot data and, like the other states, found some association between ECERS-R score and rating level. In this case, ECERS-R scores differed significantly between centers achieving any step level and “Getting Ready” centers that had not been able to achieve any step level; differences across levels in ECERS-R scores were not significantly different (Buettner, 2007).

These evaluation efforts focus on testing the validity of these systems and ask basic questions appropriate to this task: Do the star ratings relate to other measures of quality? Are the QI efforts resulting in improvements in participating provider equality?

For the most part, these efforts suggest that the systems measure some aspects of quality and that QI efforts seem to be improving provider quality. However, the limited sample sizes and circumscribed goals of these evaluations cannot speak to the many outstanding questions about QRISs, many of which we have raised in this report. As QRISs mature, it may be useful to conduct cross-state research on key issues that have not been addressed. By pooling data and evaluation expertise, research may be able to provide answers to many questions.

In Chapter Four, we examine lessons learned to date and present a set of recommendations for these systems going forward. We end the paper with a recommendation to greatly expand the research base for these systems and suggest an approach for accomplishing this.

Lessons Learned

QRISs are systems designed to improve child-care quality. By creating and aligning system attributes, QRISs motivate and help to produce improvements in the quality of care. In Chapter One, we briefly described a generic QRIS and the logic model underlying it, and we identified its key attributes: setting goals, expectations, and standards through a set of ratings; establishing incentives for system participation; monitoring performance through assessments of programs and reporting of ratings; evaluating how well programs meet expectations; and encouraging improved performance through QI plans and resources. By ensuring that these attributes are carefully considered and are aligned with each other, a QRIS has a much better chance of fulfilling its goals and engaging key stakeholders in the process.

In this chapter, we extract lessons from the experiences of the five pioneer states, both those insights that interviewees noted themselves and others that we gained during our analysis of their QRIS experiences. We organize these lessons in terms of the key systems attributes listed above. Based on these lessons, we make a number of recommendations concerning how to approach the development, implementation, and refinement of a QRIS.

State Self-Assessments

States Generally Believe That Their QRISs Have Had a Positive Impact

Below we summarize what our interviewees saw as evidence of the success of their state's QRIS. Most interviewees thought that their QRIS has been very helpful in raising awareness of quality standards for child care. A Colorado interviewee reported that the QRIS has reinforced the importance of professional development in improving quality. Rated programs in some states, such as Pennsylvania, are reaching critical mass, which has resulted in greater stakeholder awareness of quality. Certainly, providers are now more aware of what it takes to get a good rating.

In Colorado, the QRIS has, for the first time, created accountability, which alone might be "reason to celebrate," according to one interviewee. The QRIS gives providers the necessary stepping stones for quality improvement. North Carolina's QRIS is generally considered to be successful because it is the first initiative that provides objective information to parents about the quality of programs. Increased awareness of quality reportedly has resulted in significant improvement in staff-child interactions. Creating greater accountability has also led to greater investment in quality. Requests for ratings by child-care providers in North Carolina have been high. During the transition to the QRIS, it was especially difficult for the state to keep up with the demand. One interviewee noted that improvements have been difficult to sustain

because the market cannot bear the necessary corresponding wage increases. For instance, North Carolina offered education incentives to teaching staff: The TEACH initiative offered scholarships and a modest raise to teachers who completed additional education. This initiative “hit a ceiling,” according to North Carolina interviewees, because parents could not afford the fee increases required for increased salaries for better-educated staff, and public subsidies were not sufficient to cover these costs.

Ohio interviewees generally expressed satisfaction with their QRIS. According to one interviewee, parents have begun to ask about which centers have good ratings, which has encouraged competition among providers and ultimately has improved quality. Furthermore, the program is successful because “parents are accessing the Web site,” and there has been “lots of momentum” around the program. One interviewee pointed to a study (Buettner, 2007) that found that higher star levels correlated with higher ERS scores. The same interviewee also pointed out that the Ohio child-care industry had long focused on maintaining minimum standards; the QRIS has shown providers that more is possible.

According to Oklahoma interviewees, the state’s Department of Human Services considered its QRIS to be successful, given that 65 percent of participants had a rating of “1-star plus” or higher (recall that in Oklahoma, all licensed providers are included in the state QRIS). In addition, 87 percent of the 660 parents who completed an online parent survey knew about the QRIS, another sign of success. Furthermore, more children are enrolled in higher-quality facilities: All 2- and 3-star programs are full. In the past, many programs would not take subsidized children because the fees were too low. Now, there are some programs that will *only* take subsidized children because the income is guaranteed—a real turnaround. One interviewee emphasized that Oklahoma started out with three goals: (1) to improve the training and education levels of providers, (2) to increase Department of Human Services reimbursement rates so that the number of subsidized slots of reasonable quality would grow, and (3) to give parents a way to access quality. She believes that the state has reached all three goals and is now in a position to explore what to do next.

Interviewees from Pennsylvania think that its QRIS has been successful in several respects. One interviewee reported that there is a focus on improvement; many programs at the “Start with STARS” or 1-star levels are thinking about how to achieve a 2-star rating. Pennsylvania’s OCDEL has sufficient funding this year to help programs increase their quality. In the last quarter, two of 1,200 programs went down in their ratings due to director/staff turnover, while a little over 100 (8 percent) of the programs went up a level. Almost 90 percent of the members of the leading provider association are involved in the QRIS, and almost 70 percent of all programs are participating.

Factors That Contribute to the Success of a QRIS

Interviewees from the pioneer states point to a number of factors that were important in promoting the success of their QRISs.

Funding. In Colorado and Pennsylvania, funding was described as very important. In Pennsylvania, adequate financing is believed to have made the QRIS a success, particularly given the critical nature of incentives to providers. Pennsylvania has combined funding from separate governmental streams within the OCDEL. This means that various ECCE initiatives are coming together under a single agency. North Carolina locked in funds to support QI efforts, particularly the hiring of qualified staff.

Public Awareness. Interviewees from several states pointed to the value of public awareness. One Ohio interviewee attributed a fair share of the system's success to providers who helped raise public awareness of the QRIS. North Carolina interviewees also described their state's marketing of the program as extremely useful. Information is available on Web sites, and handouts are distributed to parents and providers. In Oklahoma, a considerable amount of information was easily disseminated through group and individual meetings and mail-outs. In addition, because of the state's relatively small size, key people know and trust each other, which facilitated cooperation.

Support. Interviewees in Colorado pointed to the importance of including QI and technical assistance support with the QRIS ratings as a motivator to providers and an important lever in increasing provider quality and the effectiveness of the QRIS as a whole. In Oklahoma, the fact that staff members were specifically assigned to the Reaching for the Stars program, which meant that no one was asked to expand their own job to work on Reaching for the Stars, ensured that proper attention could be dedicated to the QRIS.

Structure. Interviewees in several states noted the importance of the structure of oversight of their QRIS. In Colorado, having a board of directors from the child-care community as well as the business sector was key. In Pennsylvania, bringing together QI and technical assistance people in the Keystone STARS program has helped a great deal.

Challenges to Success

In analyzing the interviews from the five pioneer states using the system attributes discussed above, we noted a number of challenges in implementing and aligning these components.

Impediments to Success

All interviewees were asked to discuss what they might do differently if they were starting to design their QRIS now. Interviewees in Colorado indicated that they would spend more time thinking about how to reduce rating costs and that they would use the freed-up funds to provide more QI support. In Ohio, more time would be devoted to getting rating infrastructure in place. Without it, enthusiastic providers quickly became frustrated. More time also would be devoted to rolling out a training system. The initial training infrastructure was inadequate, and people with motivation and monetary incentives to be trained could not be accommodated.

In Pennsylvania, one interviewee would have changed the pilot design. Instead of randomly selecting providers, she would have selected one or two geographic areas on which to focus, which would have allowed a targeted marketing strategy and a more effective pilot of the QRIS. Some providers were upset when a local newspaper published the ratings before everyone had a chance to be rated; a geographically focused pilot could have obviated this problem.

Setting Goals, Expectations, and Standards

Rating systems embody the standards that drive QRISs.¹ The highest rating indicates the level of care that the state would like to see in all programs, even if it seems unattainable. By setting

¹ States are increasingly focusing on the development of early learning standards, which specify what children should know and be able to do at different ages. These standards enable states to examine and align standards across child care,

high standards and rewarding specific achievements along the way, QRISs clearly articulate what it takes to reach the top. However, this apparently simple idea has not been easy to implement in practice. Interviewees from some of the pioneer states told us that they are uncomfortable with such a system because they want all programs to continuously improve; providers that reach the highest rating may stop trying to improve if the system does not build in higher standards over time. Other states were reluctant to set initial standards too high, because most programs were not delivering high-quality care at the time the system was designed. Designers feared that very high standards might discourage providers from even trying. In two states, national accreditation was adopted as the highest standard, but over time, QRIS planners began to question this decision. An interviewee in one of these states argued that accreditation did not capture every aspect of quality; another thought that more could be accomplished by making the highest rung locally determined. Yet another state resisted turning over important aspects of their state's QRIS to NAEYC or any other accrediting agency, which would have its own procedures and timetables. Establishing equivalence between accreditation and QRISs will require significant effort, which may not be warranted given the small number of providers that are currently accredited. A revised version of the NAEYC accreditation system was implemented in 2006 in an effort to address concerns with the earlier version, some of which are described above. Whether the revised version alleviates these concerns remains to be seen.

Pioneer states have struggled with standard-setting because of the large gap between where designers would like programs to go and where they are when the QRIS is designed. In the earliest states (Oklahoma in particular), program planners couldn't themselves envision programs attaining really high quality levels, and therefore they focused their efforts on raising standards above where they currently were. Yet in states that set a modest standard as its highest level, problems occurred when designers raised the standard over time. Providers rightly argued that they had signed on to participate in a system defined by a given set of standards, and then, after they strove to reach them, and often succeeded, those standards had been raised. Providers that had received a particular rating were especially unhappy that they were suddenly at risk of receiving a lower rating for the same standard of care. As discussed in the recommendation section at the end of this chapter, one possible solution is to build in a larger number of "rungs" in the rating system. This will ensure that providers who are initially at the low end of the scale can experience some success and improvements in the context of a system with sufficiently high standards that the system does not need to be changed as providers improve over time.

The lowest rating defines the lowest acceptable quality standard. In three of our pioneer states, licensing is the entry ticket to the QRIS; programs must be licensed to participate in the system and undergo a rating. In these pioneer states, compliance with key licensing standards is regarded as a necessary condition for the delivery of a safe program; licensing requirements ensure only that the program meets minimal standards. Licensing as a prerequisite then allows the QRIS to focus on quality, not basic safety. Two states automatically assign the lowest QRIS rating to all licensed providers, whether or not they choose to participate in or opt out of system ratings, which are voluntary in all of these systems. An advantage of this approach is that it ensures that virtually all providers participate in the QRIS at some level (depending

Pre-K, Head Start, and K-3 programs. At this time, these standards have not found their way directly into QRISs in most states, although these standards have been reconciled with and, in some states, have driven program standards and staff professional development criteria.

on licensing requirements in a given state, as discussed in Chapter One). It may also motivate providers who receive the lowest rating to agree to be rated. However, including licensed programs and assigning them the lowest star rating has confused some parents, as it is not clear to them whether a program has attained a one-star rating because it is licensed and the director chose not to participate in a QRIS rating or because it was evaluated and received the lowest rating. At the same time, bringing licensing into the QRIS means that parents have one fewer system to understand.

Another key issue in setting standards and ratings concerns efforts to make the QI increments roughly comparable. Ideally, it should require equal amounts of effort to move from one rating to the next across the entire range of levels. Equal intervals reward providers for improving equally, regardless of where they fall on the rating scale. This too has been a challenge in at least one pioneer state. Oklahoma started with just three ratings: licensing at the low end, national accreditation at the high end, and everything else in between. It became clear rather quickly that moving from a rating of 1 star to 2 stars was very challenging, which was discouraging to many programs. As a result, Oklahoma added a “1 star plus” rating to bridge the gap. But this change, along with many others, created both confusion and resentment among parents and providers.

States also had to decide which standards to include. What is included in the rating systems really matters; through inclusion or omission, messages are sent to the provider community, to parents, and to policymakers about what is important in child care. Furthermore, in high-stakes settings such as QRISs, where substantial consequences often are attached to ratings, people pay close attention to what is included in the ratings, making content selection (and omission) especially important. Finally, some components, such as ERSs, are valuable because they are objective measures that generate the process information that is important to the development of QI plans. It is for this reason that North Carolina and Colorado included an ERS as one component of their system.

At the same time, the addition of each new component has nontrivial cost implications and may raise measurement issues as well. Many states have struggled with the inclusion of an ERS for these reasons. ERSs are expensive to administer and require costly validation and revalidation of raters (see Stoney [2004] for a discussion of how states handle ERS training and reliability). Oklahoma began with an ERS but decided to forgo the use of it as part of the quality rating, mainly because of cost concerns.² Pennsylvania has limited ERS use to programs that have already demonstrated a higher quality level. Ohio uses a self-assessment tool, which does not require costly training of objective observers.

Some components, such as parent involvement, present problems because there are no generally accepted measures that meet psychometric criteria. Yet, parent involvement is considered by many to be an important aspect of program quality. Ohio included a parent involvement component but then dropped it when their measure, which focused on numbers of parents attending meetings, seemed not to be capturing meaningful parent involvement. Other states have stayed with an assessment of parent involvement. Colorado used RAND’s empirical analyses to modify its measure; it moved from assessing parent satisfaction to examining a parent partnership measure (see Zellman and Perlman [2006] for further discussion of this measure).

² Oklahoma does require high-rated programs to be assessed using an ERS every four years.

Staff training and education are also difficult to rate; there is still disagreement concerning which aspects of staff background and training are most important, and little data about how to combine training and education characteristics across staff. Given that incentives are tied to performance in these increasingly high-stakes QRSs, use of self-reported data (for example, staff reports about their efforts to involve parents) should be minimized. However, while conflict of interest and self-presentation biases make the use of such data as part of the rating system problematic, self-assessment data can be extremely useful in QI efforts. Research on many of these questions, including reliance on self-reported data, is badly needed so that states can draw on the results of empirical studies rather than have to reinvent the QRIS “wheel” for themselves.

Establishing Incentives and Supports

All of the five states we examined provide financial incentives to support quality improvement. These incentives are given to providers and, in some cases, to the caregivers who work in participating centers. Current incentives include subsidy payments that increase with higher quality ratings (tiered reimbursement), staff scholarships, and other professional development programs where eligibility depends on a program’s rating, as well as QI resources that may be tied to a QI plan. Differential reimbursement by star rating rewards providers for higher quality and helps them cover the higher costs of providing higher-quality care. Oklahoma, North Carolina, and Ohio have such systems; in Colorado, provision of tiered reimbursement is at the discretion of local counties; such a system currently operates in Denver. States employ a number of approaches to encourage system participation. Funds are targeted in some cases to address problem areas, such as high staff turnover and low levels of education. Oklahoma and Pennsylvania provide a number of grants and scholarships to improve the capacity and compensation of child care workers. North Carolina, among other states, provides TEACH fellowships for this purpose. States also support staff with salary supplements and retention incentives. In the case of Pennsylvania, staff must be working in 2-, 3-, or 4-star programs to qualify; Oklahoma’s program requires that staff be working in programs above the 1-star level to qualify (see Mitchell [2006] for details on these programs). North Carolina has limited QI funds to the Star Rated License program, but since all providers must participate in the QRIS at least at the 1-star (licensing) level, this approach effectively includes all programs. In 2007, the Pennsylvania’s OCDEL submitted a proposal for a tiered reimbursement add-on (in addition to QRIS-linked awards already offered), which would acknowledge high-quality programs for taking on at-risk children. In addition, the governor has earmarked \$75 million to expand public Pre-K programs beyond at-risk districts so that there will be reimbursement for every child. As a result, there is now collaboration between school districts and providers. In this new initiative, Pre-K programs will be required to participate in the QRIS. A more detailed description of the various incentive structures in these and other states is provided by Mitchell (2005).

To motivate improvements in child-care quality, QRISs depend, to some degree, on prestige. By creating easily understood summary measures, parents and providers can see a given program’s quality standing relative to that of others. But prestige may not be a sufficient incentive in the child-care arena, because without additional resources, it may be impossible to improve one’s ranking, even if motivation to do so is high. Some improvements, particularly reduced ratios and better-trained staff, are extremely costly. QRISs theoretically link prestige indicators to parent choice. Specifically, the idea is that, given clear information about child-

care quality, parents will choose only high-quality care for their children. But parent choice is often illusory in the child-care arena, where many parents feel fortunate to find any care that is affordable and available during the hours they need it. In addition, when most care is mediocre at best, even parents who are aware of the rating system and want to make quality-driven choices often cannot do so because of lack of supply of high-quality care. For example, in Oklahoma, all programs at the 2- and 3-star levels (in Oklahoma's three-level system) are completely full. However, there are some examples of choice beginning to take hold. In Ohio, a public-information strategy supported by providers led parents to begin to ask which centers had good ratings. The result, according to one interviewee, was increased competition among providers. But as discussed below, these campaigns are themselves expensive to develop and maintain.

This suggests that QRISs cannot rely on prestige and parent choice as the sole motivation for improvement and that they must depend on financial incentives and QI support to motivate participation and improvement. These incentives must be built in a way that recognizes that, in a voluntary system, any provider who is willing to be rated accepts some risk that her program will not get a good rating. Further, submitting to a rating takes time and distracts staff from other tasks.

Funding for QRISs remains an issue in most states. Low reimbursement rates for children receiving child-care subsidies represent a barrier to success in most systems. Oklahoma's biggest obstacle to success was that incentives only applied to children receiving child-care subsidies. This meant that if centers did not have subsidized children, they had no incentive to participate. Money comes from a range of sources, including federal support through the Child Care and Development Fund, administered by the Child Care Bureau, money from welfare programs, particularly Temporary Assistance to Needy Families, and private sector contributions. As Stoney (2004) notes, effective financing may include not just raising or identifying new funds, but the ability to draw on existing resources. To the extent that a QRIS aligns with other programs, for example, Pre-K or Head Start, some of those resources may contribute to QRIS ratings and QI efforts.

In general, however, incentives are not sufficient (and often are not delivered reliably enough) to motivate providers to change their practices and improve the quality of care they provide. In Ohio, for example, low subsidy rates make achievement of high ratings impossible in centers that serve low-income children without third-party money. In Oklahoma, the salary supplement reaches only about 10 percent of staff.

Adequate incentives should increase the percentage of providers who participate in the system. This, in turn, should put pressure on nonparticipating providers to join the system as parents increasingly seek rating information as they search for care. In Colorado, incentives to providers dropped dramatically because of a shortfall in funding. One interviewee felt that this did not reduce provider motivation to change and that the higher funding level was not necessary. Research that examines the levels of incentives needed to motivate change is badly needed. States vary dramatically in the level of participation they have achieved. Some of the states with higher rates have achieved good provider participation through a combination of political support (e.g., from the governor) and a solid incentive package (e.g., both Oklahoma and North Carolina, which have high participation rates, provide tiered reimbursements). Colorado had the lowest reported participation rate (Stoney, 2004) and also the highest ratings costs. These high costs can be attributed, at least in part, to requiring ERS assessments at all star levels.

Monitoring Performance Through Ratings

Creating a Culture That Emphasizes Quality of Care in the Child-Care Sector. Quality ratings represent a cultural shift for child-care programs that have long focused on meeting licensing requirements. Not surprisingly, interviewees in several states reported that providers are often wary of the rating process, particularly those aspects that involve quality improvement. As one interviewee described it, licensing ratings are based on a check-box mentality: Requirements are generally unambiguous, and attainment is clear. Quality ratings differ from licensing in two important ways. First, at least some components (e.g., an ERS) are not entirely clear-cut, as they focus on staff interactions with children, relationships with parents, and other subjective aspects of the program. Second, in an effort to create a culture of quality improvement, the rating process, and particularly the QI process that follows, asks programs and their staff to become involved on a continuing basis in thinking about quality. This is particularly important because providers are used to “passing” their licensing inspection and then thinking little about it. Continuous quality improvement requires staff to take far more responsibility for the quality of care that they provide. This may involve thinking about what may be missing, how to best improve those aspects of the program found to be weak, and what additional actions might help, such as enrollment in training or a new way of interacting with parents.

Rating Frequency. The cost of ratings varies enormously depending on what is rated. As noted above, ERSs are a particularly labor-intensive, expensive aspect of quality that are sometimes included in QRSs. ERSs are costly because raters must be trained and their reliability assessed on a regular basis and because raters must observe in each rated classroom for several hours. The frequency of monitoring performance is another cost issue. In the five pioneer states, the frequency with which providers are rated varied widely and was generally related to cost. In Oklahoma, where an ERS is not part of the rating process, ratings are conducted three times per year. In contrast, in North Carolina, where ERS is used, they are conducted once every three years. Even with this long period between ratings, North Carolina has a waiting list for ERS assessments because there are simply not enough trained raters to get to programs in a timely manner. In Colorado, classrooms in each program are ERS-assessed every other year. The cost associated with this requirement may contribute to the relatively low uptake rate in that state, despite the fact that the vast majority of ratings, including ERS costs, have been paid for by public and private third parties.

Long lags between ratings are clearly problematic in several respects. They create disincentives to make immediate changes, and they contribute to a sense among caregivers and directors that the process is unfair because programs may be “stuck” for a long time with a rating that they feel is out of date.

Who conducts the ratings is another important issue. Real or perceived rater conflict of interest can undermine the credibility of a QRIS in the eyes of the public and may also reduce the value of the QI process. States have struggled with the assignment of the tasks of licensing inspections, quality ratings, and QI coaching. One reason for assigning more than one task to a single individual is that it can result in substantial costs savings. As Stoney (2004) notes, states often view these as strategic decisions that allow them to use existing funds to support the QRIS. For example, in Ohio, Step Up to Quality programs have visits from licensing specialists (with a health and safety focus) and from QRS specialists (focusing on quality) every two years. It is expensive to have two sets of people conducting assessments. Some have suggested that the two jobs be integrated, but others worry that the two types of work require different skills.

Combining roles sometimes creates problems. Raters may be more inclined to rate providers more positively as they get to know them through coaching efforts. Further, rater-coaches may feel that improvements (or lack of improvements) reflect their own coaching abilities, which may cause them to overlook flaws. A North Carolina interviewee stressed the importance of separating licensing and quality assessment, something North Carolina did not do. Indeed, QRIS tasks were given to licensing specialists so that no new funding was required. Even if it appears more “efficient” to have raters (who, after all, are already in classrooms and know the factors that increase quality) provide technical assistance, several interviewees questioned whether it is possible for raters to provide useful technical assistance. Recipients are always uncertain whether the advice is just advice or whether it must be followed.

Yet cost cannot be ignored. Several Colorado interviewees noted that lowering costs for ratings is key to the viability of its system. If too much money is devoted to ratings, there may not be enough to fund other critical aspects of the QRIS, such as provider incentives and QI coaching, two costly aspects of system functioning. As noted above, several states are dealing with the cost issue in the selection of components to be rated.

Performance monitoring also brings to the fore many of the measurement problems that were discussed during our interviews. A number of interviewees worried about rater reliability, particularly on the ERS. One interviewee talked about instances of ERS rater inconsistency. She said that a particular provider received two rather different ERS scores from two different raters although everything in the center was the same during both assessments. Regardless of who conducts the ratings, the reliability of their observations across classrooms, providers, and time must be tested regularly.

Assessing Provider Compliance with Quality Standards. System designers must determine how compliance with QRIS standards will be monitored and how closely providers must conform to these standards to qualify for a given rating. The pros and cons of the point and block systems were discussed in Chapter Three. Briefly, these systems vary in terms of the autonomy afforded providers in meeting standards. Point systems allow for greater discretion in how to improve. Providers who may not have the resources or motivation to improve a particular component can gain points in other ways that may be less costly or that can be changed more quickly. In block systems, all the system components must be attended to. Advocates of this approach argue that block systems increase consistency across programs with the same rating. The greater control and rigor afforded by the block system are appealing. However, use of a block system implies confidence in the relative weights of the various aspects of quality being measured. This confidence may not be warranted given gaps in the current empirical basis for making such decisions. As with many aspects of QRISs, research about how to combine what is measured to create summary scores is badly needed.

Encouraging Provider Improvement Through QI Support

All states represented in this study provided some form of QI support to participating child-care providers. Technical assistance or coaching uses information gathered as part of the assessment process to give providers feedback about how to improve their practices. The information from the quality ratings is often augmented by some form of self-assessment (often ERSs at lower levels of some QRISs) or the collection of additional data that are deemed useful for QI purposes.

Despite some commonalities in QI efforts across the five states, the extent and delivery methods for QI varied widely. Some professional development opportunities are provided to

staff in the form of scholarships and funds to cover their time. In Colorado, providers were awarded between \$3,000 and \$4,000 per classroom per year for QI efforts, although the distribution and the uses of the funds varied across the state. Some states include differential reimbursement to providers in their systems that may result in higher staff salaries. Others specifically link completion of training to salary increases. These efforts constitute incentives and are discussed under that heading earlier in this chapter. They also serve a QI purpose by facilitating recruitment and retention of staff that are likely to provide higher quality care.

Provision of sufficient support for improvement may be critical to achieving the goals of QRISs. In a cash-starved sector where supply is limited, quality improvement (along with ratings and the broader accountability system) may be required in order to bring about change. It is important that those who are considering the development of a QRIS consider what information beyond ratings-generated information may be needed to support providers. It is also critical to recognize the dearth of research on what strategies are most effective and how to best spend limited QI funds.

Dissemination of Information About the QRIS and Provider Ratings

Simply providing information about a new system, its goals, and the quality standards may, in the view of several interviewees, cause programs to begin to think about quality. All states have posted information about their systems on a Web site; Oklahoma used word of mouth and satellite radio as well. A high rating is something to be celebrated, and programs are happy to display the placard indicating their good evaluation. In most states, the rating is part of the information that R&Rs provide to parents, which of course increases the dissemination of rating information and does so at a point when parents are actively soliciting program information and making the choices that are key to the vitality of QRISs.

However, rating information is not always widely disseminated or entirely transparent. As noted above, in states where licensing represents the entry rating level, a 1-star rating does not indicate to parents whether the program chose to opt out of the rating process after licensing or whether it was evaluated and simply is not providing care of high quality. And in still other states, resource and referral staff are cautioned to tell parents that a rating is not the only measure of quality. While cautious and well intentioned, these messages are likely to be confusing to parents.

QRIS Components and Their Relationships to Each Other

It is critical that, as a QRIS is developed, the different components of the system and the relations among them are considered and aligned. For example, if rating standards are to be set high, incentives to participate have to be correspondingly higher to induce participation. Moreover, since quality improvement is costly, QI funds must be adequate to ensure that improvement can occur. It is also important to consider whether a point or block system is being used when setting incentives: Block systems allow for less maneuvering, so setting reimbursements and rewards consistent with QI costs should be more easily achieved. One interviewee suggested that states think seriously about goals and standards. In this interviewee's view, standards and participation work in opposite ways: High standards reduce participation; lower standards increase it. She stated that states have two options: (1) Set lower standards at the beginning, which will bring in more providers, then try to raise standards over time, or (2) set standards high at the beginning, which will lower initial participation, then try to increase participation over time. One possible solution is to design a rating system with more "rungs,"

enabling lower-quality providers to get on to the scale and experience improvement in their ratings at the low end while establishing appropriate standards of quality at the high end from the outset. The decision that is made will have important implications for scaling up the system.

Another interviewee noted the importance of aligning incentives with standards. For example, if more training is needed to attain a particular rating, attainment of the rating should include payment of training costs. As noted throughout this report, there is little excess money in the child-care system; providers may want to raise quality, but they cannot afford to do so unless the costs are covered. This is particularly true for providers that serve low-income families. Subsidies generally are insufficient to cover quality improvement, and providers in low-income communities cannot raise parent fees to cover quality improvements. Alignment also involves the relation between standards and training. For example, including a standard that requires directors to observe and provide teacher feedback implies a need to train directors on classroom observation and feedback. If that training is not widely available, then the requirement needs to be phased in or delayed until the training infrastructure is in place.

Examination of the system as a whole can and should be considered in a larger sense as well. The focus on improved quality that is at the heart of QRISs can be used as a way to bring together ECCE initiatives in the state. By aligning multiple initiatives, it may be possible to capitalize on funds available from other sources to promote quality, as discussed above. (See Stoney, 2004, Kagan and Kauerz, 2007, and the National Early Childhood Accountability Task Force, 2007, for further discussions of alignment in early education policy from different perspectives.)

Recommendations

We conclude with a list of recommendations drawn from our interviews and analyses of the QRISs in these five pioneer states. These recommendations are organized in terms of what needs to be in place prior to the inception of a QRIS, the system-development process, what the system should include, and assessment of QRIS outcomes. Given the qualitative nature of this study and the small number of states and interviewees involved, these recommendations should be regarded as potentially useful advice rather than guidelines that all states should follow.

Precursors to a Successful QRIS

1. Obtain adequate funding in advance and decide how it will be spent. QRISs require money to be effective. Money is necessary to do ratings, to provide incentives for participation, and to improve quality. In the child-care arena, money is very limited. It is important to design the QRIS so that available funds are used in the most effective way. This requires analysis of the costs of the various components, such as ratings and coaching. If, for example, the cost of ratings is so high that it compromises other QRIS activities, it may be necessary to rethink what is being measured or find other sources of support to fund them. It is critical that sufficient incentives are available to improve quality. While early rating system ideas in some states relied on the motivation of programs to improve, it has become increasingly apparent that improvement cannot occur without incentives and support. Providers accept risks in participating in rating systems. Those risks must be attached to real benefits. An underfunded system risks failure.

2. Garner maximum political support for a QRIS. In North Carolina and Colorado, the governors lent political support and facilitated funding for their QRISs; in Pennsylvania, gubernatorial support has contributed to a comprehensive approach with sufficient funding. Lack of such support can be a major barrier to the ramping up of a QRIS in a timely manner and its continuing fiscal health.

System-Development Process

1. Conduct pilot work and make revisions to the system before it is adopted statewide.

One of the most obvious recommendations gleaned from the interviews is that significant time and effort should be devoted to an iterative revision process in response to a system pilot. Piloting may become less critical once more research and information sharing between states about QRISs is available to guide decisionmaking. Without a pilot phase, states in this study were forced to make many changes after implementation was underway. This led to confusion and resentment. In a constantly changing policy environment, piloting may not always be possible, as people fear losing funds if they do not use them when they are available. In such cases, the ideal approach is to lock in funds for the longer term but still launch a pilot and revision process. Premature implementation of QRISs that are not developed carefully may undermine future public investment in such systems.

2. Limit changes to the system after it is implemented. States worry a good deal about complacency among providers; those that get a reasonably good score in particular may move into what one interviewee described as “maintenance mode.” This is a prime motivator for changes to QRISs. Our sense, however, is that states don’t worry enough about the other side of this equation—risking provider fatigue and willingness to participate by making multiple changes to the system over time. This risk should be considered in deciding whether to make changes to the system. Setting up a system of continuous quality improvement with clear incentives for improvement and a substantial number of rungs to climb may be a better way to encourage continuous quality improvement without imposing new requirements. Consider the value of stability against continuous improvement. Our interviews suggest that constantly raising the bar creates confusion among parents, may undermine their trust in the system, and contributes to “provider fatigue.” If changes are made, moreover, informing and updating the public is a more challenging task; too many changes may lead to uncertainty about the stability of the system. This is likely to be an even more significant problem if changes lead to altered program ratings.

What Should QRISs Include?

1. Minimal self-reported data. In these high-stakes systems, where substantial funds may be tied to ratings, providers will be highly motivated to score well. Their need to be well rated makes them potentially biased informants. Self-reported information can be very useful for QI purposes, but it should be used sparingly, if at all, as part of the rating system.

2. Licensing should be integrated into the system. There are multiple assessment systems that speak to the quality of a child-care provider. This is confusing for parents who do not understand the difference between licensing, accreditation, certification, rating systems, etc. To the extent possible, these rating systems should be integrated. One way to do this is to assign all licensed providers the lowest QRIS rating unless they volunteer for a rating and are rated higher. One advantage to assigning licensed providers the lowest QRIS rating is that it results in virtually universal participation by providers, albeit at a low level (i.e., they may get

a rating of “1” by default for being licensed without actually being rated). This may motivate some of the 1-star providers to be evaluated and would reduce parental confusion because the vast majority of providers would be rated. However, states may not want to assign licensed providers the lowest QRIS rating if they believe that licensing represents too low a quality threshold.

3. Use ERSs flexibly by incorporating both self-assessments and independent assessments at different levels of the QRS. ERSs have substantial value, particularly with regard to quality improvement. At least some of this value may be captured by using ERSs in more creative—and economical—ways, as several states are doing. The ERS can be used as a self-assessment tool at lower quality levels, as is the practice in Ohio and Pennsylvania. Several interviewees thought that such use was quite helpful in a number of programs. An ERS can be used as part of a QRIS, but a substantial amount of money can be saved if it is administered only after a program shows evidence in other ways that it is likely to be highly rated (e.g., by achieving a certain level based on ratios, training, and education, etc.), as in Pennsylvania and North Carolina.

4. Do not include accreditation as a mandatory system component. Issues raised about accreditation by our interviewees included the high cost (although limited scholarship funds are available through NAEYC) and labor intensiveness of becoming accredited, delays in completing ratings due to waitlists for validation visits, and the lack of validation of the measure. Furthermore, discrepancies in standards between states and accreditation bodies were identified by some interviewees as a source of concern about including accreditation in the system. Given these concerns, it seems inadvisable to build accreditation into QRISs, particularly as a required element. States could choose to make accreditation an alternative path to ratings, but they must recognize that doing so imposes challenges in establishing equivalence—what star level would accreditation equal? It is worth noting that a new NAEYC accreditation system was recently implemented in response to some of the limitations associated with the earlier version. Data about the validity and equivalencies of the new version are needed before stronger recommendations about the role of NAEYC accreditation in QRISs can be made.

5. The rating system should have multiple levels. Having many rungs makes progress more attainable at the lower quality levels, thereby facilitating provider engagement. It also allows for improvement at the higher end, preventing providers from shifting to a “maintenance mode” in which they no longer strive to improve. We suggest that states aim high at the beginning, even if there are concerns that high standards are far above most providers. By providing incentives for progress (critical because fewer providers will join if they don’t think they can reach the top levels), child-care providers may be willing to take risks to improve the quality of the care they are able to provide even if their first ratings are likely to be low.

Quality Improvement

1. Create a robust QI process. Standards and accountability may help providers understand what quality is and may also underscore the parts of their programs most in need of change. But without resources and support, few programs will be able to improve. To effect change, a QRIS needs to provide some mix of staff development, financial incentives, and QI support. Some of this support may already be available; QRIS planners need to inventory potential support. These efforts should ensure that the support aligns with system requirements. For example, community colleges need to offer the ECCE coursework that is required to attain the credentials required in the QRIS. Once key system elements have been deter-

mined, it is critical to begin to examine the whole ECCE system in the state to look for places where similar activities may be occurring that could be aligned to provide continuing support for the QRIS.

2. Separate raters and QI support personnel. Raters know a great deal about programs and therefore seem ideally situated to provide QI input. Such an arrangement saves money, too: Raters already visit the program and could simply provide some technical assistance while they are there. But such dual roles create problems for programs; they often don't know whether they must comply because the person is a rater or if the advice may be considered and rejected. To avoid such problems, it is preferable to have distinct groups of individuals conduct assessments and assist in QI efforts.

3. Public-awareness campaigns are important but should start after the system is in place; these campaigns need to be ongoing. Parents only need information about child-care quality for a relatively brief window of time while their children are young. To be useful, public-awareness campaigns need to be big enough to reach many parents and available on an ongoing basis. Such campaigns should be initiated once the system is fully developed, so that the system can deliver on its promises. In cases where a pilot is geographically circumscribed, e.g., operating in a single county, a public-awareness campaign could be launched in that county once the system was implemented.

Evaluate the Effectiveness of the QRIS

1. Support research on systems and system components. Research that identifies best practices in QRISs is needed so that these practices can be shared among the states involved in the development of a QRIS. Research on QRISs has been limited in focus and depth. A large share has focused on process assessments, which is appropriate given that these systems are relatively new. Validation studies have generally been limited to assessing relations between ECERS-R scores and star ratings. However, since QRISs may include the ERS as part of their quality ratings, the magnitude of these relationships likely overstates the relationship between the QRIS and ERSs and may lead to erroneous conclusions about the validity of the QRIS.

States would benefit from empirical work on key measurement issues, including how best to assess important components, such as staff education or training. States also would benefit from research that examines how to combine ratings across components to provide reliable and valid ratings. Testing of the logic model underlying QRISs is important as well. Without such research, each state must essentially reinvent the wheel in terms of developing a QRIS. However, this research will be complex and costly. Moreover, sample sizes in individual states may be inadequate to conduct these complex studies.

Establishing a QRIS Consortium is one way to accomplish this research. Discussions between various stakeholders about such a consortium took place at RAND in January of 2006. The QRIS Consortium concept proposes that states pool research funds, agreeing to collect some comparable data using compatible platforms. By pooling these data, the QRIS Consortium could conduct research that these systems need but currently do not have. The QRIS Consortium idea represents a promising approach to overcoming some of the barriers to conducting important research on QRISs. As these systems proliferate across states and continue to change within them, the ability to draw on high-quality research in making decisions about components, measures, weighting, and other critical system issues will become even more important.

Interview Guide

QRIS Questionnaire – Cover sheet

Thanks again for agreeing to talk with us today.

To remind you of why we're taking your time, RAND is part of a consortium with broad national representation from many stakeholders, including child-care practitioners, accreditation agencies, philanthropists, and researchers. The consortium works on the prospect of an institute for quality rating and improvement in early child care and education. The institute would help states with no QRS system establish their own and help other states improve on their existing systems. As one of the initial steps of this effort, we're trying to develop lessons from experience.

We have picked [insert state] since it is one of the most experienced nationwide in child-care quality rating systems. We feel you are best positioned to inform us about your state's QRIS system from the [insert stakeholder type]'s point of view. Your name has been referred to us by [insert other contact] AND/OR we understand you are a key participant in your state's QRIS system based on our preliminary Internet research.

[HSPC Disclaimer]

Could you describe your current and former positions associated with QRS in your state?

In what capacity were you involved with QRS system development?

In what capacity were you involved with QRS system implementation?

Could you describe your background in ECCE and formal education?

Now, let's proceed to system-related questions.

QRIS Questionnaire – Questions

Impetus

1. What was the impetus for your state's QRIS?
 - a. Which groups/stakeholders played major roles in obtaining support for a QRIS?
 - State legislature
 - Governor's office
 - Key department/agency
 - Outside organization or individual
2. Did parents play a role in creating support for a QRIS?
3. Did providers play a role in creating support for a QRIS?

4. What was understood to be the primary objective of the system as the planning process began? Was there consensus about this?

System Design and Planning

5. Which group coordinated the design phase?
6. How long did the planning process take?
7. Which stakeholders were involved in system design? (Check all that apply)
 - a. State departments/agencies
 - b. Provider groups
 - c. Parent groups
 - d. Child advocacy groups
 - e. Research organizations
 - f. Legislators
 - g. Philanthropy
 - h. Other (describe)
8. How were components selected?
9. [If obvious component missing] Why doesn't your system include [missing component]?
10. How was the number of quality levels/tiers/steps determined?
11. How was the required frequency of ratings determined?
12. Did you focus on a broader accountability system, or just the QRS?
13. Is QI part of the accountability system? Why/why not?
14. How did you plan for technical assistance and training to support providers?
15. How is the QRS linked with licensing?
16. How is the QRS linked with accreditation?
17. Is the system voluntary? Why / why not?
18. What incentives were/are available for providers to participate (financial and other)?
19. Was any feature of the system specifically mandated by legislation or executive decision?
20. To what extent were these decisions based on research about child-care quality or quality measurement? How did you access that information (e.g., consultant, literature review)?
21. How much was your system influenced by what other states were doing? How did you get that information? How did you incorporate it?
22. What was the total budget for system design?
23. How did funding limitations affect the design process?

Implementation

24. Was the system piloted?
25. What were the most important things you learned from the pilot?
26. What organization administers the system? How was that organization selected?
27. How are raters trained?
28. How is a rating conducted?
29. How much does it cost to rate a classroom?

30. What are the largest cost drivers in a rating?
31. How are ratings funded? How much must providers pay to be rated?
32. Did you strive to minimize the cost per rating? How?
33. What rating components have been most problematic to implement? Why?
34. What rating components have been least problematic to implement? Why?
35. How many providers have been rated? What proportion of providers is that?
36. Are homes and centers equally interested in being rated?
37. What form of technical assistance is available (prior to/after rating)? Who provides QI technical assistance?
38. Are funding limitations affecting QRS implementation?
39. What were other key implementation challenges?

Modifications

40. Has your system been modified since its initial design? If so, when, why and how?
 - a. Components
 - b. Levels
 - c. Provider Incentives
 - d. Technical Assistance
 - e. Other
41. What data are being collected about the QRS—both for administrative and evaluation purposes?
 - a. Do you have follow-up on the impact of ratings?
 - b. Effects of ratings on parent behavior?
 - c. Other?
42. Has the system been changed based on your system's or other data?
 - a. Based on what other states are doing?
 - b. On what the literature on quality measurement has shown?

Overall Retrospective

Grand-tour Questions:

43. How successful would you say your state QRIS has been? Is it doing what you hoped it would do?
 - a. Probe: match with stated goals
44. What concerns you most?
45. What has contributed to the success of your effort?
46. What has impeded its success?
47. Are these factors (promoting and impeding success) unique to your state?
48. What do you wish you had known when you first started to design your system?
49. What would you do differently if you were to do it again?

Targeted Questions—ask as appropriate:

50. Would you raise/lower the number of rating levels? Why/why not?
51. Would you add/omit any component? Why/why not?

52. Would you increase/reduce rating costs? Why? How?
53. Would you increase/reduce rating frequency? Why/why not?
54. Would you modify provider incentives? Why/why not? How?
55. Would you modify parent education initiatives? Why/why not? How?
56. Would you explore additional funding sources? If so, which ones?

Concluding Questions:

57. Based on your state's experience, what would you tell another state newly embarking on a QRS to pay attention to?
58. What would you like to have known?
 - a. Research based
 - b. Experience based

Thank you for your time and contribution. Your input is instrumental to our research effort. We will be interviewing/have interviewed several other participants from your state, and we are looking at four other states. Your responses will be interpreted in conjunction with the responses of other contributors, and will help us distill broad-based lessons from experience for better quality child care nationwide.

Unpublished Mani Paper on QRISs

Quality Rating and Improvement Systems Multi-State Conversations Meera Mani, August 2007¹

Quality Rating Systems (QRS) have become a growing national phenomenon. Thirty-six states are implementing some form of rating, and a large number of them are tying it to a tiered reimbursement system.

The RAND Corporation hosted a meeting in Santa Monica in January 2006 to help state leaders explore ways they might collaborate on the development and implementation of QRIS systems. Representatives from eight states attended. The discussion focused on the issues states were grappling with as they implemented quality rating systems, and on the possibility of forming a consortium of states to help save money and speed up development of effective models. At least three areas of potential collaborative work were identified:

- technology and database systems
- quality improvement methods, including training of coaches and QI staff
- examination and validation of measures used in the ratings, such as classroom environment, ratios, and teacher qualifications.

Now the Build Initiative, which is working in seven states (NJ, PA, OH, IL, MI, MN, WA) to help develop early care systems, and United Way of America, which has a long-standing interest in this area, are trying to reinvigorate interest in the idea of a national consortium or collaborative technical assistance effort to help states improve their QRIS models. To that end, a sample of individuals who attended the January 2006 meeting, along with selected others, were contacted and interviewed. This interview sought to

- explore/reaffirm interest in a national collaboration
- ask questions regarding progress to date in the state including changes in the ratings landscape in the last 18 months
- ascertain desire and need for technical assistance needs that may be provided by a national organization such as the Build Initiative, United Way of America, and other potential partners.

¹ Reprinted with permission.

Seventeen individuals representing 13 states were interviewed, and the following report provides a summary of the views expressed regarding the formation of a national consortium as well as specific requests for technical assistance from such a body, followed by a state-by-state profile. The individuals interviewed, along with their affiliations, are named in each of the state profiles.

National Consortium

Do you think it would be beneficial to have a consortium to support state and local efforts either currently engaged in or contemplating the adoption of a QRIS system?

Without exception, all of the individuals interviewed responded in favor of such an entity. Each was able to identify benefits and share some issues for consideration. A summary of those responses is presented here.

Identified benefits:

- Share methodology to understand why some states are more successful than others
- Create an effective forum to encourage peer to peer cross-state dialogue
- Learn from other states about what worked and what did not
- Examine QRS/QRIS models in other states—specifically, design and cost, and learn how to ramp up to scale
- Provide much needed cross-state dialogue to create a national momentum that may influence decisions at the state level
- Provide direct and intentional support to states
- Bring credibility to rating systems nationally
- Encourage a more coherent dialogue on quality nationally
- Explore a shared research agenda and coordinate research efforts nationally
- National research agenda could give credibility to this movement
- Conduct national research on impact of QRS systems on children birth to five, including children with special needs and linguistic difficulties
- Engage in a national discussion on the wisdom and feasibility of tying child outcomes to QRS
- Serve as a forum to resolve compatibility issues between QRS/QRIS and NAEYC accreditation
- Build common language around advocacy nationally.
- Develop a shared advocacy agenda especially with regard to federal child care block grant reauthorization
- Create momentum around a national movement that embeds QRS/QRIS in federal child care legislation

However, note of caution:

- The consortium must honor existing expertise and not become a self-appointed expert
- There must be some thought given to how the consortium membership is constructed—
 - Should it be open to all states with some interest in QRS or should it initially bring together states who have reached a certain level of QRS implementation?

- Members should be compensated for participating
- Invite National Women’s Law Center and Voices for America’s Children to participate
- Participating states must have public and private representation
- Small group preferable to large group for truly constructive dialogue to occur
- Need a better definition of what is meant by a shared research agenda
- Would only consider a shared research agenda if it can deliver local data for reporting
- National research must be less expensive than individual state efforts for a shared research agenda to make sense

Possible leaders:

The Build Initiative and United Way of America were named as possible leaders. NAEYC, NACRRA, NCCIC and The Women’s Law Center were named as important partners.

Technical Assistance Requests:

If such a consortium were to exist on which issues could you use most help and support?

A summary of the recurring themes that emerged from the conversations is presented here.

QRS/QRIS Design and Development:

- Community Readiness Assessment
 - How can one assess the true readiness of a community to embrace QRS?
 - Who are and how can one engage the key stakeholders?
 - Is there a logical order to stakeholder engagement?
 - What are the essential elements of a community readiness plan?
- Calculation of systemic and unit cost. Need a cost template
 - What are the broad parameters and discrete elements of QRS development and implementation?
 - Is there a template that will guide individual state’s planning efforts while allowing customization for local context?
- Cost-effective QRS/QRIS and QI models
- True cost and capacity to support programs for quality improvement
 - What are the essential elements of an effective technical assistance program?
- Essential elements of an accountability system and how to design it

Rating Implementation:

- Training of raters
- Effective data collection techniques and tools to manage data
- Successful technical assistance and quality improvement models
- Proven best practices for technical assistance for quality improvement

Public Awareness and Advocacy:

- How to develop a public awareness and publications plan
- Tools to use with policymakers

Evaluation and data questions:

- Which QRS/QRIS domains work best in terms of evaluating quality?
- Which implementation methodologies have been most successful?
- How are states addressing child outcomes?
- Are there benefits to a shared research agenda?

References

- American Education Research Association, American Psychological Association, and the National Council on Measurement in Education, *Standards for Educational and Psychological Testing*, Washington, D.C.: American Educational Research Association, 1999.
- Barnard, W., W. E. Smith, R. Fiene, and K. Swanson, *Evaluation of Pennsylvania's Keystone STARS Quality Rating System in Child Care Settings*, Harrisburg, Pa.: Pennsylvania Office of Child Development, 2006.
- Barraclough, S. J., and A. B. Smith, "Do Parents Choose and Value Quality Child Care in New Zealand?" *International Journal of Early Years Education*, Vol. 4, No. 1, 1996, pp. 5–26.
- Brewer, D. J., S. M. Gates, and C. A. Goldman, *In Pursuit of Prestige: Strategy and Competition in U.S. Higher Education*, Santa Monica, Calif.: RAND Corporation, DRU-2541-EDU, 2001. As of July 28, 2008: <http://www.rand.org/pubs/drafts/DRU2541/>
- Buettner, Cynthia K., *Step Up to Quality Pilot Final Report*, Columbia, Ohio: The Ohio Collaborative: Research and Policy for Schools, Children, and Families, January 2007.
- Burchinal, M. R., J. E. Roberts, L. A. Nabors, and D. M. Bryant, "Quality of Center Child Care and Infant Cognitive and Language Development," *Child Development*, Vol. 67, 1996, pp. 606–620.
- Cassidy, D. J., L. L. Hestenes, A. Hegde, S. Hestenes, and S. Mims, "Measurement of Quality in Preschool Child Care Classrooms: an Exploratory and Confirmatory Factor Analysis of the Early Childhood Environment Rating Scale–Revised," *Early Childhood Research Quarterly*, Vol. 20, No. 3, 2005, pp. 345–360.
- Clarke-Stewart, K. A., D. L. Vandell, M. Burchinal, M. O'Brien, and K. McCartney, "Do Regulable Features of Child-Care Homes Affect Children's Development?" *Early Childhood Research Quarterly*, Vol. 17, No. 1, 2002, pp. 52–86.
- Cryer, D., and M. Burchinal, "Parents as Child Care Consumers," *Early Childhood Research Quarterly*, Vol. 12, No. 1, 1997, pp. 35–58.
- Cryer, D., W. Tietze, and H. Wessels, "Parents' Perceptions of Their Children's Child Care: A Cross-National Comparison," *Early Childhood Research Quarterly*, Vol. 17, No. 2, 2002, pp. 259–277.
- Duncan, G., "Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development," *Child Development*, Vol. 74, No. 5, September–October 2003, pp. 1454–1475.
- Fuqua, R. W., and D. Labensohn, "Parents as Consumers of Child Care," *Family Relations*, Vol. 35, No. 2, 1986, pp. 295–303.
- Grace, D., and A. O'Cass, "Child Care Services: An Exploratory Study of Choice, Switching and Search Behaviour," *European Journal of Marketing*, Vol. 27, Nos. 1–2, 2003, pp. 101–132.
- Helburn, S. W., and B. R. Bergmann, *America's Child Care Problem: The Way Out*, New York: Palgrave, 2002.
- Helburn, S. W., M. L. Culkin, and J. Morris, *Cost, Quality, and Child Outcomes in Child Care Centers: Public Report*, Denver, Colo.: University of Colorado at Denver, 1995.
- Helburn, S., and C. Howes, *Child Care Cost and Quality: The Future of Children: Financing Child Care*, Los Altos, Calif.: Center for the Future of Children, 1996.
- Helburn, S. W., J. Morris, and K. Modigliani, "Family Child Care Finances and Their Effect on Quality and Incentives," *Early Childhood Research Quarterly*, Vol. 17, 2002, pp. 512–538.

- Howes, C., "Relations Between Early Child Care and Schooling," *Developmental Psychology*, Vol. 24, 1988, pp. 53–57.
- Kagan, S. L., "Buckets, Banks, and Hearts: Aligning Early Childhood Standards and Systems," presentation at the Quality Rating and Improvement Systems: Creating the Next Generation of QRISs conference, St. Paul, Minn., June 4, 2008.
- Kagan, S. L., and K. Kauerz, "Reaching for the Whole: Integration and Alignment in Early Education Policy," in R. C. Pianta, M. J. Cox, and K. L. Snow, eds., *School Readiness and the Transition to Kindergarten in the Era of Accountability*, Baltimore, Md.: Paul H. Brookes, 2007, pp. 11–30.
- Karoly, L. A., B. Ghosh-Dastidar, G. L. Zellman, M. Perlman, and L. Fernyhough, *Prepared to Learn: The Nature and Quality of Early Care and Education for Preschool-Age Children in California*, Santa Monica, Calif.: RAND Corporation, TR-539-PF/WKKF/PEW/NIEER/WCJVSF/LAUP, 2008. As of August 1, 2008: http://www.rand.org/pubs/technical_reports/TR539/
- Lamb, M. E., "Nonparental Child Care: Context, Quality and Correlates," in W. Damon, I. E. Siegel, and K.A. Renninger, eds., *Handbook of Child Psychology, Vol. 4, Child Psychology in Practice*, New York: Wiley, 1998.
- Le, V., M. Perlman, G. L. Zellman, and L. Hamilton, "Measuring Child-Staff Ratios in Child Care Centers: Balancing Effort and Representatives," *Early Childhood Research Quarterly*, Vol. 21, No. 3, 2006, pp. 267–279.
- Long, P., P. Wilson, P. Kutnick, and L. Telford, "Choice and Childcare: A Survey of Parental Perceptions and Views," *Early Child Development and Care*, Vol. 119, No. 1, 1996, pp. 51–63.
- Mani, M., *Quality Rating and Improvement Systems: Multi-State Conversations*, August 2007.
- Mitchell, Anne W., *Stair Steps to Quality: A Guide for States and Communities Developing Quality Rating Systems for Early Care and Education*, United Way, Success by 6, July 2005.
- Morris, J. R., and S. W. Helburn, "Child Care Center Quality Differences: The Role of Profit Status, Client Preferences, and Trust," *Nonprofit and Voluntary Sector Quarterly*, Vol. 29, No. 3, 2000, pp. 377–399.
- NAEYC—See National Association for the Education of Young Children.
- National Association of Child Care Resource and Referral Agencies (NACCRRRA), *Parents' Perceptions of Child Care in the United States: NACCRRRA's National Parent Poll*, Alexandria, Va.: NACCRRRA, May 2006.
- , "Threshold of Licensed Family Child Care (June 2007)," Web page, 2008. As of July 28, 2008: http://www.naccrra.org/randd/licensing_training_qr/fcc_threshold.php
- National Association for the Education of Young Children, "Teacher-Child Ratios Within Group Size," Web page, 2005. As of August 1, 2008: http://www.naeyc.org/academy/criteria/teacher_child_ratios.html
- National Early Childhood Accountability Task Force, *Taking Stock: Assessing and Improving Early Childhood Learning and Program Quality*, Washington, D.C.: Pew Charitable Trusts, October 2007. As of July 28, 2008: http://www.pewtrusts.org/uploadedFiles/wwwpewtrustsorg/Reports/Pre-k_education/task_force_report1.pdf
- National Institute of Child Health and Human Development Early Child Care Research Network, "The Relation of Child Care to Cognitive and Language Development," *Child Development*, Vol. 74, No. 4, 2000, pp. 960–980.
- , "Does Quality of Child Care Affect Child Outcomes at Age 4 1/2?" *Developmental Psychology*, Vol. 39, No. 3, 2003, pp. 451–469.
- Norris, Deborah, L. Dunn, and S. Dykstra, "Reaching for the Stars" *Center Validation Study, Executive Summary*, Norman, Okla.: Early Childhood Collaborative of Oklahoma, November 2003.
- North Carolina Division of Child Development, Web site, no date. As of August 1, 2008: <http://ncchildcare.dhhs.state.nc.us/general/home.asp>
- Peisner-Feinberg, E. S., and M. R. Burchinal, "Relations Between Preschool Children's Child-Care Experiences and Concurrent Development: The Cost, Quality, and Outcomes Study," *Merrill-Palmer Quarterly*, Vol. 43, No. 3, 1997, pp. 451–477.

- Peisner-Feinberg, E. S., M. R. Burchinal, R. M. Clifford, M. L. Culkin, C. Howes, S. L. Kagan, and N. Yazejian, "The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories Through Second Grade," *Child Development*, Vol. 72, No. 5, 2001, pp. 1534–1553.
- Peisner-Feinberg, E. S., M. R. Burchinal, R. M. Clifford, N. Yazejian, M. L. Culkin, J. Zelazo, C. Howes, P. Byler, S. L. Kagan, and J. Rustici, *The Children of the Cost, Quality, and Outcomes Study Go to School: Technical Report*, Chapel Hill, N.C.: University of North Carolina, Frank Porter Graham Child Development Center, 1999.
- Pennsylvania Early Learning, Keys to Quality, Web site, 2006. As of August 1, 2008:
<http://www.pakeys.org/docs.aspx>
- Perlman, M., G. Zellman, and V. Le, "Examining the Psychometric Properties of the Early Childhood Environment Rating Scale—Revised," *Early Childhood Research Quarterly*, Vol. 19, 2004, pp. 398–412.
- Public Law 107-110, The No Child Left Behind Act of 2001, January 8, 2002.
- Qualistar Early Learning, Web site, no date. As of July 28, 2008:
<http://www.QualistarColorado.org>
- Scarr, S., "American Child Care Today," *American Psychologist*, Vol. 53, No. 2, 1998, pp. 95–108.
- Scarr, S., M. Eisenberg, and K. Deater-Deckard, "Measurement of Quality in Child Care Centers," *Early Childhood Research Quarterly*, Vol. 9, No. 2, 1994, pp. 131–151.
- Stoney, L., *Financing Quality Rating Systems: Lessons Learned*, United Way, Success by 6, September 2004. As of July 28, 2008:
http://www.earlychildhoodfinance.org/handouts/Louise_Stoney_QRS_Financing_Paper.pdf
- U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, *The Role of Early Head Start Programs in Addressing the Child Care Needs of Low-Income Families with Infants and Toddlers: Influences on Child Care Use and Quality*, Washington, D.C.: Early Head Start Research and Evaluation Project, 2004. As of July 28, 2008:
<http://www.mathematica-mpr.com/publications/PDFs/roleofearly.pdf>
- Van Horn, M. L., S. L. Ramey, B. A. Mulvihill, and W. Y. Newell, "Reasons for Child Care Choice and Appraisal Among Low-Income Mothers," *Child and Youth Care Forum*, Vol. 30, No. 4, 2001, pp. 231–249.
- Vandell, D. L., and B. Wolfe, *Child Care Quality: Does It Matter and Does It Need to Be Improved?* Washington, D.C.: U.S. Department of Health and Human Services, 2000.
- Zellman, G. L., and S. M. Gates, *Examining the Cost of Military Child Care*, Santa Monica, Calif.: RAND Corporation, MR-1415-OSD, 2002. As of July 28, 2008:
http://www.rand.org/pubs/monograph_reports/MR1415/
- Zellman, G. L., and A. S. Johansen, "Investment or Overkill: Should Military Child Care Centers Be Accredited?" *Armed Forces and Society*, Vol. 23, 1996, pp. 249–268.
- Zellman, G. L., and M. Perlman, "Parent Involvement in Child Care Settings: Conceptual and Measurement Issues," *Early Child Development and Care*, Vol. 176, No. 5, 2006, pp. 521–538.
- Zellman, G. L., M. Perlman, V. Le, and C. M. Setodji, *Assessing the Validity of the Qualistar Early Learning Quality Rating and Improvement System as a Tool for Improving Child-Care Quality*, Santa Monica, Calif.: RAND Corporation, MG-650-QEL, 2008. As of July 28, 2008:
<http://www.rand.org/pubs/monographs/MG650/>
- Zellman, G. L., G. Ryan, L. Constant, H. Salem, G. Gonzalez, R. Karam, N. Orr, C. A. Goldman, H. Al-Thani, and K. Al Obaidli, *The Evolution of Qatar's K–12 Education Reform over Its First Three Years*, Santa Monica, Calif.: RAND Corporation, forthcoming.