



# High/Scope Program Quality Assessment

**PQA—PRESCHOOL VERSION**

## Administration Manual

# High/Scope Program Quality Assessment

**PQA—PRESCHOOL VERSION**

## Administration Manual

Published by **High/Scope® Press**  
A division of the  
**High/Scope Educational Research Foundation**  
600 North River Street  
Ypsilanti, Michigan 48198-2898  
Phone: 734/485-2000, FAX: 734/485-0704  
E-mail: [info@highscope.org](mailto:info@highscope.org)  
Web site: [www.highscope.org](http://www.highscope.org)

**Copyright© 1998 by High/Scope Educational Research Foundation,**  
600 N. River St., Ypsilanti, MI 48198. All rights reserved. Except as permitted under the Copyright Act of 1976, no part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage-and-retrieval system, without prior written permission from the publisher. High/Scope is a registered trademark and service mark of the High/Scope Educational Research Foundation.

ISBN 1-57379-043-5

## CONTENTS

Potential Uses of the PQA	2
Directions for Administering the PQA	4
Reliability and Validity of the PQA	6
References	9

## INTRODUCTION

### **Purpose of the PQA**

The High/Scope Program Quality Assessment (PQA): Preschool Version is a rating instrument for evaluating the quality of early childhood programs and identifying staff-training needs.<sup>1</sup> It was developed by the High/Scope Educational Research Foundation for use in all center-based preschool and child care settings, not just those using the High/Scope educational approach. As such, the PQA intentionally reflects “best practices” in the early childhood field as a whole. That is, the items consolidate what current theory, decades of practice, and ongoing research tell us about the environments and interpersonal relationships that best promote the development of young children, encourage the involvement of families, and create supportive working environments for staff.

The PQA is a comprehensive tool examining all aspects of program implementation, from the physical characteristics of the setting to the nature of adult-child interaction to program management. This comprehensiveness reflects a growing consensus that the assessment of program quality should not be based on a single type of data but requires a multidimensional approach (Fiene & Melnick, 1991). To be complete, a measure of program quality must include both the structural information used to determine compliance with licensing regulations and the dynamic information that captures children’s experiences in the physical and interpersonal learning environment. By employing a combination of observational and interview techniques, the PQA encompasses both these dimensions.

The PQA can be administered by trained independent raters conducting research and evaluation, or it can be used as a self-assessment by agencies interested in identifying program strengths and areas for improvement. Based on classroom observations and interviews with teaching and administrative staff, raters assess a broad array of program characteristics by using an objective five-point scale for each characteristic. To assist raters in making valid and reliable assessments, the endpoints and midpoint of each scale are clearly defined and

---

<sup>1</sup> The Program Quality Assessment (PQA) is a revision and expansion of the High/Scope Program Implementation Profile (PIP) (High/Scope Educational Research Foundation, 1989). The PQA reflects updated knowledge about the learning environment, daily routine, and adult-child interaction and greatly expands coverage of curriculum planning and assessment, work with families, staffing and training, and program management. In addition, the PQA is a generic instrument, while the PIP focused specifically on the implementation of High/Scope’s early childhood approach.

illustrated with examples. Unlike compliance measures that typically permit only “yes-no” responses to items, the PQA defines program quality along a continuum. The five-level scales given for each item allow raters to indicate with greater specificity where a program is and what it needs to do to achieve maximum quality. The PQA thus permits program quality to be measured both with breadth and precision.

### **Organization of the PQA**

The PQA has 72 items organized into seven sections that together cover all the areas essential to evaluating the quality of early childhood programs:

- I. Learning Environment (9 items)
- II. Daily Routine (12 items)
- III. Adult-Child Interaction (12 items)
- IV. Curriculum Planning and Assessment (5 items)
- V. Parent Involvement and Family Services (10 items)
- VI. Staff Qualifications and Staff Development (14 items)
- VII. Program Management (10 items)

Each item is scored using a five-point rating scale extending from a low (*1*) to a high (*5*) level of quality. Examples are provided for ratings of *1*, *3*, and *5*; the points in between (*2* and *4*) give the rater additional flexibility in assigning scores. Below each item, space is provided for recording supporting evidence and anecdotes about the settings, activities, individuals, and interactions assessed. The Learning Environment section provides a blank page for drawing a diagram of the room, while the Daily Routine section includes a page for listing both the posted schedule and the actual schedule observed. At the end of the PQA Assessment Form is a Summary Sheet for entering the scores for all 72 items and calculating the total and average PQA scores.

## POTENTIAL USES OF THE PQA

The PQA has widespread applicability as a tool for training, monitoring, observation and feedback, and research and evaluation. The results can be shared with administrators, staff, parents, and funders associated with the program, as well as with early childhood practitioners and researchers in the field as a whole. Findings from the PQA can be used to define and illustrate best practices, to focus attention on program development issues in preservice and inservice training, to examine the relationship between program practices and children’s development, and to point to promising policy initiatives and investments for improving the quality of early childhood programs.

### **Training**

The PQA can be used in both preservice and inservice training. The instrument can be used in its entirety to provide students with a comprehensive picture of all the components of quality, or it can be used a section at a time to focus on specific elements of program implementation. The detailed examples in each item provide students with concrete examples of what best practices “look like” in actual programs. Users often comment that the PQA gives

definition to the term *developmentally appropriate practice* by translating an ideal into specific strategies for implementation. Even experienced teachers find that the depth and specificity of the instrument help them reconsider long-established practices from a new perspective.

### **Self-Assessment and Monitoring**

The PQA is a valuable tool for administrative and instructional staff to use in assessing their own practices and identifying areas for further development and training. It can also be used by agency supervisors, or by others responsible for quality control, to monitor program implementation at a single site or across multiple sites. Because the PQA is objective and quantitative, it can be used to set program goals in one or more areas and to provide a numerical and anecdotal record of progress toward these goals.

### **Observation and Feedback**

The PQA is an excellent resource for providing supervision through observation and feedback. A staff member and supervisor might, for example, agree to focus on one or more aspects of implementation covered in the PQA. The supervisor would then observe in the program setting, record detailed notes and make ratings, and discuss these with the staff member. Together, they could assess strengths and areas for improvement, with the PQA providing concrete examples of how implementation can progress along the rating continuum.

### **Research and Evaluation**

The PQA has been used extensively as a research tool to document program practices, compare quality in different program settings, and examine the relationship between program quality and children's development. (See page 6 for a discussion of the PQA's reliability and validity). This tool can also be used to evaluate the effectiveness of staff development by assessing program quality in one or more areas before and after the relevant inservice training activities.

### **Information and Dissemination**

Because of its straightforward language and detailed examples, the PQA can be used to explain the meaning of *best practices* or *developmentally appropriate practice* to a variety of individuals and agencies. Potential audiences include administrators and policymakers, particularly those new to the early childhood field, who need help identifying the elements of high-quality programs. Support staff can also benefit from becoming familiar with the PQA as a way of better understanding the practices and the requests of the instructional staff. The PQA is an effective tool for sharing program practices with parents and helping them understand comparable environments and interactions they can establish at home. Results of the PQA can be easily communicated to researchers, and the instrument's accessibility makes it possible for others to replicate and extend the lessons learned about effective program practices.

## DIRECTIONS FOR ADMINISTERING THE PQA

### Who Completes the PQA

The PQA can be completed by a trained, independent rater such a researcher, program evaluator, consultant, or agency administrator. It may also be completed as a self-assessment by site staff such as directors, early childhood specialists, curriculum coordinators, individual teachers or teaching teams, or parents. The PQA can be used by these raters to highlight program strengths and to identify problems requiring the revision of policies, program planning, or staff development.

### Time Period for Completing the PQA

Raters, whether from outside or inside the program, should conduct a comprehensive program observation and review. It is recommended that independent raters spend at least one full day reviewing a program before completing the ratings, allocating half the day to observing in the classroom (Sections I, II, and III) and half to conducting the interviews (Sections IV, V, VI, and VII). If more than one classroom in a center or agency is to be rated, the rater should observe in each classroom for one-half day (Sections I, II, and III) and interview the head teacher in each classroom (Sections IV and V). The interview with the director (Sections VI and VII) need only be conducted once if the same information applies to all classrooms in that center or agency.

### Determining What PQA Rating to Give

To complete the PQA Assessment Form, **circle only one numerical rating for each of the items**. Ratings are clearly defined at levels *1*, *3*, and *5*. Ratings of *2* and *4* should be used for gradations between these points, that is, when program quality falls between the defined levels. The following general rules apply to the ratings:

- Use a rating of *1* for behavior or situations characterized by the words *never* or *rarely*, *none* or *few*, *little*, *infrequently*, and so on.
- Use a rating of *3* for behavior or situations characterized by the words *half the time*, *half the requirements*, *sometimes*, *some*, and so on.
- Use a rating of *5* for behavior or situations characterized by the words *always* or *almost always*, *all* or *most*, *regularly*, *consistently*, *frequently*, and so on.

Following each item is space to write down **supporting evidence and anecdotes**; use this documentation to explain the rating and, if appropriate for monitoring or training, to indicate any action plans or follow-up activity.

Try to complete *all the items*. Only if the relevant information cannot be obtained through all available methods should the rater check “**not observed or reported.**”

The rater should compile all supporting evidence and anecdotes before completing the PQA ratings. The ratings should be completed as soon as possible after the observation and interview period. If the PQA is being conducted as an independent assessment, the rater should complete the ratings after leaving the site, even though the results may later be shared

with site staff. Completing the ratings on the spot can promote anxiety and a fixation on “numbers” among the staff being observed and interviewed. It can also interfere with the rater’s ability to objectively review the supporting evidence and anecdotes and to deliberate about the appropriate rating to assign.

### **Completing the PQA as an Independent Assessment**

**Observation items.** To complete the observation items (Sections I, II, and III), the rater observes in the classroom, writes supporting evidence and anecdotes, and circles one rating for each item. Where necessary, the rater may ask program staff (for example, the teacher) for additional information to supplement what can be observed.

**Interview items.** To complete the interview items (Sections IV, V, VI, and VII), the rater begins by saying the following to the appropriate staff member:

Interviewer [At the Beginning of Each Section]: *“I’m going to read you a series of statements about (curriculum planning and assessment/parent involvement/etc.). Please tell me to what extent and in what ways each statement is true or not true of your program. Give me examples.”*

The rater then reads the main statement at the top of the item (not the descriptors under 1, 3, and 5) and writes down the respondent’s answers in the “Supporting Evidence/Anecdotes” section. (If the main statement is long or detailed, it is acceptable for the rater to repeat it.) When the respondent’s answer is unclear, the rater uses probe(s) as needed:

Interviewer: *“Can you tell me more about \_\_\_\_\_?”* [Ask respondent to specify the missing or unclear piece of information, for example, *“How many years did you do that?”* or *“What is the name of the instrument you use?”*]

If the person being interviewed cannot answer one or more questions but knows another individual in the program who can, it is acceptable to use information from the other person to complete the rating. For example, the teacher may refer the rater to the parent coordinator to answer questions related to parent involvement. Or the director may refer the rater to the curriculum specialist to answer questions on staff development.

### **Completing the PQA as a Self-Assessment**

**Observation items.** To complete the observation items (Sections I, II, and III), the rater writes supporting evidence and anecdotes describing the program, refers to the examples provided, and circles one rating for each item. Where necessary to supplement what he or she knows about the program, the rater may ask other staff members (for example, the curriculum specialist) for additional information.

**Interview items.** To complete the interview items (Sections IV, V, VI, and VII), the rater writes supporting evidence and anecdotes describing the program, refers to the examples provided, and circles one rating for each item. Where necessary to supplement what he or she knows about the program, the rater may ask other staff members (for example, the social services coordinator) for additional information.

## Completing the PQA Summary Sheet

At the end of the PQA Assessment Form is a Summary Sheet. To complete it, the rater transfers the numerical rating (1, 2, 3, 4, or 5) from each item on the assessment form to the corresponding item on the summary sheet. If an item was not rated, the rater enters "NR." At the bottom of the sheet, the rater indicates the number of items not rated and subtracts these from 72 (the full number of items) to compute the number of items rated. The total PQA score is calculated by adding the scores for all the items. A separate score may also be calculated for each section. To compute the average PQA score, the rater divides the total score by the number of items rated.

## RELIABILITY AND VALIDITY OF THE PQA

The PQA has been used extensively as a research tool by trained, independent raters and also as a self-assessment tool by early childhood practitioners. An early version of the PQA, the High/Scope Program Implementation Profile (PIP), was used in the *Training for Quality* study, a national evaluation of 366 classrooms and 201 children in diverse early childhood settings (Epstein, 1993). The PQA was expanded to its current comprehensive form in the High/Scope *Michigan School Readiness Program Evaluation* (High/Scope Educational Research Foundation, 1997). As part of this statewide evaluation of programs for at-risk 4-year-olds, trained independent observers completed the PQA in 49 preschool classrooms, 642 teachers and program administrators used the PQA as a self-assessment, and 229 program children were assessed using a variety of observational and rating instruments at kindergarten. The Michigan Department of Education adopted the PQA as the official monitoring tool of its early childhood consultants and requires that all directors and teachers in the statewide at-risk preschool program complete the PQA as an annual self-assessment. In addition, the PQA was adapted and used by the High/Scope Head Start Quality Research Center in its *Head Start Staff Development Study* (Schweinhart, Epstein, Okoloko, & Oden, 1998) that includes observational assessments of 44 classrooms and 155 children. All these research and evaluation efforts have provided much data on the PQA's reliability and validity in relation to other indicators of program quality and young children's developmental progress. The significant findings<sup>2</sup> throughout these studies attest to the PQA's strength as a reliable and valid assessment instrument.

### Reliability

Training to acceptable levels of interrater reliability on the PQA takes three days. The first two days are used to review and practice with the PQA; the third day is used to determine interrater reliability. Trainees are first given an overview of the instrument and its uses. During training, they view videotapes and observe early childhood programs, practice writing supporting evidence and anecdotes, and discuss how they would score the PQA items. Trainees then score practice forms based on pre-written evidence and anecdotes and compare their

---

<sup>2</sup> Significance throughout this manual is defined as  $p \leq .05$ .

ratings to the “standard” scores assigned by trained raters. Differences are discussed until trainees understand the criteria for scoring all the items. At the end of the training, trainees observe and interview staff in actual program settings and complete all the PQA items. Their ratings are then compared to those of the staff trainer(s) to compute interrater reliability.

Interrater reliability, defined as the percentage of agreement between the trainer and trainees, is computed in two ways: as exact agreement (the trainer and trainee give the same numerical rating) and close agreement (the trainer and trainee give ratings within one point of each other). In all the studies where the PQA or its predecessor, the PIP, have been used, interrater reliability has been very high. In *Training for Quality*, the 25 observers reached an average exact agreement of 89.5% and an average close agreement of 98.6% on the PIP. Among the 18 PQA observers in the Michigan School Readiness Program Evaluation and the 8 PQA observers in the Head Start Staff Development Study, exact agreement averaged 79.4% and close agreement averaged 96.7%. Table 1 contains the average interrater reliability figures for the 26 observers in the latter two studies, using items comparable to the current version of the PQA. These figures indicate that the PQA can be used with confidence by trained observers to examine the quality of early childhood settings.

TABLE 1. PQA INTERRATER RELIABILITY

PQA Section	% Agreement (N=26 trained observers)	
	Exact	Close
Learning Environment	80.8	96.2
Daily Routine	91.2	100.0
Adult-Child Interaction	75.6	96.1
Curriculum Planning and Assessment	79.4	99.2
Parent Involvement and Family Services	75.0	96.6
Staff Qualifications and Staff Development	76.1	97.0
Program Management	83.8	93.9
OBSERVATION ITEMS	80.8	96.9
INTERVIEW ITEMS	78.5	96.5
PQA TOTAL	79.4	96.7

The reliability of the PQA was also examined in two other ways in the Michigan School Readiness Program Evaluation. Cronbach’s alpha coefficient on the PQA items was .952 for the 49 independent observations and .956 for the 642 teacher self-assessments. Correlations between independent observer ratings and self-assessments averaged .492 for the PQA as a whole. Self-ratings were generally higher than independent ratings, particularly when ratings fell in the mid-range. At the extremes, however, both groups identified many similar strengths and areas for improvement. It is not surprising that self-ratings would be inflated, especially if staff felt the scores would be used to monitor their performance relative to state standards. Nevertheless, higher correlations between independent ratings and self-ratings are desirable and would certainly make the PQA more reliable and valid as a self-assessment. Toward this

end, High/Scope is conducting statewide PQA training for all administrative and instructional staff in Michigan's at-risk preschool program. The training emphasizes the PQA's use as a formative rather than a purely summative evaluation tool. Similar training in other programs where the PQA is used as a self-assessment tool can likewise add to its value.

## **Validity**

The PQA (or its predecessor, the PIP) has been significantly correlated with other measures of program quality and with indicators of children's development.

***PQA and other program quality indicators.*** In the Training for Quality study, the early version of the PQA was significantly correlated overall (.86) with the Early Childhood Environment Rating Scale (ECERS; Harms & Clifford, 1980) and was also significantly correlated overall (.48) with the Arnett Global Rating Scale (Arnett, 1989). The associations were strongest in those subscales most similar in the instruments, for example, PQA Learning Environment and ECERS Furnishings were correlated at .73 and PQA Adult-Child Interaction and Arnett Sensitivity were correlated at .77.

Teachers' qualifications are also differentially related to scores on the PQA. In the early version used in Training for Quality, classroom observation ratings were significantly correlated with teachers' years of formal education (.44), amount of inservice training (.33) and years of early childhood program experience (.32). Similarly, in the Head Start Staff Development Study, PQA scores were significantly correlated with teachers' years of formal education ( $r=.33$ ), number of hours of training in the program's curriculum ( $r=.47$ ), and years of experience working with young children ( $r=.47$ ). In both of these studies, then, program quality was positively and significantly related to teachers' education, experience, and curriculum-specific inservice training.

***PQA and children's development.*** Studies repeatedly show that levels of program quality as measured on the PQA are positively and significantly associated with indicators of children's development. In Training for Quality, all the subscales and the total score on this program assessment were positively and significantly correlated with the language scale of the DIAL-R (Mardell-Czudnowski & Goldenberg, 1990). Concurrent correlations between program quality and child development in preschool ranged from .35 to .56. Program quality was also significantly related to several subscales on the High/Scope Child Observation Record (COR) (High/Scope Educational Research Foundation, 1992). The strongest associations were between the program's daily routine and adult-child interactions, on the one hand, and children's initiative and creative representation on the other. These significant correlations between program quality and child development ranged from .33 to .52.

Similar relationships between the PQA at preschool and children's development in kindergarten appeared in the Michigan School Readiness Program Evaluation. Nearly one-quarter (18 out of 80) of the correlations computed between the PQA and the COR proved significant. Significant correlations ranged from .33 to .65. Key program quality factors were the extent to which programs had a consistent and articulated philosophy and curriculum model, the use of funding to support staff development, and the consistency of administrative support and supervision for staff. These PQA variables were in turn related to children's initiative, social relations, language and literacy, logic and mathematics, and total COR scores. There were also

significant correlations between these same dimensions of program quality and children's development at kindergarten as rated by their teachers. Correlations between the PQA and teachers' ratings of children ranged from .40 to .48.

## REFERENCES

- Arnett, J. (1989). Caregivers in day care centers: Does training matter? *Journal of Applied Developmental Psychology, 10*, 541–552.
- Epstein, A. S. (1993). *Training for quality: Improving early childhood programs through systematic inservice training*. Ypsilanti, MI: High/Scope Press.
- Fiene, R., & Melnick, S. (1991, April). *Quality assessment in early childhood programs: A multi-dimensional approach*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Harms, T., & Clifford, R. M. (1980). *The Early Childhood Environment Rating Scale*. New York: Teachers College Press.
- High/Scope Educational Research Foundation. (1989). *High/Scope Program Implementation Profile (PIP)*. Ypsilanti, MI: High/Scope Press.
- High/Scope Educational Research Foundation. (1992). *High/Scope Child Observation Record (COR) for Ages 2½–6*. Ypsilanti, MI: High/Scope Press.
- High/Scope Educational Research Foundation. (1997). *Early returns: First-year report of the Michigan School Readiness Program Evaluation*. Ypsilanti, MI: High/Scope Research Division.
- Mardell-Czudnowski, C., & Goldenberg, D. S. (1990). *Developmental Indicators for the Assessment of Learning—Revised*. Circle Pines, MN: American Guidance Services, Inc.
- Schweinhart, L., Epstein, A., Okoloko, V., & Oden, S. (1998, April). *Assessing the quality of Head Start programs: The Quality Research Consortium*. Paper presented at the American Educational Research Association, San Diego, CA.

# NOTES