

Contract No.: 233-02-0086(03)
MPR Reference No.: 6094-500

MATHEMATICA
Policy Research, Inc.

**Implementation of the
National Reporting
System in Migrant/
Seasonal Head Start
Programs: Year Two
Update**

Final Report

November 2007

*Renée Nogales
Diane Paulsell
Louisa Tarullo*

Submitted to:

Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
370 L'Enfant Promenade
Washington, DC 20447

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Project Officer:

Jennifer Brooks

Project Director:

Louisa Tarullo

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

CONTENTS

Chapter		Page
I	INTRODUCTION	1
	KEY CHARACTERISTICS OF MIGRANT/SEASONAL HEAD START PROGRAMS	2
	SAMPLE SELECTION AND SITE VISITS	3
	ROAD MAP TO THE REPORT	4
II	ADMINISTERING THE SPANISH-LANGUAGE VERSION OF THE CHILD ASSESSMENT	7
	MEETING THE CERTIFICATION STANDARDS	8
	INTER-RATER RELIABILITY	9
	ERRORS IN PROCEDURES, ADMINISTRATION, AND SCORING	10
	Administering the Set-up and Warm-up Section	11
	Administering and Scoring Tío Simón (Pre-LAS Simon Says)	12
	Administering and Scoring Exposición de Arte (Pre-LAS Art Show).....	12
	Administering and Scoring the Test de Vocabulario en Imágenes Peabody (TVIP-Adapted).....	13
	Administering and Scoring Nombre de las Letras (Letter Naming)	14
	Administering and Scoring Básicos de Matemática (Early Math Skills)	15
	ISSUES UNIQUE TO ADMINISTRATION AND SCORING OF THE SPANISH-LANGUAGE ASSESSMENT	16

CONTENTS

Chapter	Page
II (<i>continued</i>)	
ASSESSORS' EXPERIENCES ADMINISTERING THE CHILD ASSESSMENT	18
Children's Responses to the Assessment Process.....	18
Strategies for Reacting to Children's Responses	19
Experiences of Assessors Who Were Also Children's Teachers.....	19
Experiences Assessing English-Language Learners	19
Experiences Assessing Children with Disabilities.....	20
III NRS TRAINING AND SUPPORT ACTIVITIES	23
LOCAL REFRESHER TRAINING ACTIVITIES	23
LOCAL QUALITY ASSURANCE PRACTICES	24
VIEWS OF LOCAL PROGRAM STAFF MEMBERS ON NRS TRAINING MATERIALS AND SUPPORT	24
NRS Web Casts.....	25
Format of Training Sessions	25
Training Materials	26
Technical Assistance.....	26
RECOMMENDATIONS FOR IMPROVING TRAINING AND SUPPORT	26
IV LOCAL APPROACHES TO IMPLEMENTING THE NRS CHILD ASSESSMENTS.....	29
COMMUNICATING WITH PARENTS AND POLICY COUNCILS	29
COORDINATING THE CHILD ASSESSMENTS	30
Staffing for Coordination and Oversight	30
Locations Where Assessments Were Conducted.....	30
Scheduling the Assessments.....	30
Tracking Progress in Completing Assessments.....	31

CONTENTS

Chapter	Page
IV (<i>continued</i>)	
APPROACHES TO ASSIGNING STAFF MEMBERS TO ADMINISTER THE CHILD ASSESSMENTS	31
COSTS OF CONDUCTING THE CHILD ASSESSMENTS	32
USING THE COMPUTER-BASED REPORTING SYSTEM.....	32
Experiences with the CBRS	33
V USING THE NRS FOR LOCAL PROGRAM IMPROVEMENT EFFORTS	35
REACTIONS OF LOCAL PROGRAM STAFF MEMBERS TO THE 2003-2004 GROWTH REPORT	36
HOW PROGRAMS USED THE GROWTH REPORT	36
FUTURE PLANS FOR USING NRS DATA.....	37
RECOMMENDATIONS FOR IMPROVING THE NRS REPORT.....	37
VI LOCAL PROGRAMS' VIEWS ON THE NRS	39
CONCERNS ABOUT THE NRS	39
Whether the NRS Fits into the Unique Structure of Migrant/Seasonal Programs	39
Whether NRS Results Accurately Reflect Program Effectiveness.....	40
Appropriateness of the NRS for Young Children.....	41
Time and Resources Dedicated to the NRS	41
Purpose of the NRS	41
SUGGESTIONS FOR IMPROVING THE NRS.....	42
Ways to Improve the Test Administration Process.....	42
Ways to Make the NRS More Suitable for Migrant/Seasonal Programs	42
Ways to Improve Communication between the Office of Head Start and Local Programs.....	42
Ways to Improve the Instrument.....	43

CONTENTS

Chapter	Page
VII	
SUMMARY OF FINDINGS AND IMPLICATIONS FOR SYSTEM IMPROVEMENT	45
OVERVIEW OF FINDINGS FROM THE SITE VISITS	45
IMPLICATIONS FOR SYSTEM IMPROVEMENT	46
REFERENCES	49

APPENDICES

**ADMINISTRATION AND SCORING ERRORS OF OBSERVED BASELINE AND GROWTH NRS
CHILD ASSESSMENTS IN MIGRANT/SEASONAL HEAD START PROGRAMS**

2005 END OF PROGRAM YEAR PROGRAM REPORT

T A B L E S

Table		Page
1	CHARACTERISTICS OF SAMPLE MIGRANT/SEASONAL HEAD START AGENCIES.....	5
2	DISTRIBUTION OF CERTIFICATION SCORES ACROSS OBSERVED BASELINE AND GROWTH SPANISH ASSESSMENTS	9

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

CHAPTER I

INTRODUCTION

In fall 2003, the Office of Head Start began implementing the National Reporting System (NRS), an ambitious initiative to assess systematically the early literacy, language, and numeracy skills of all 4- and 5-year-olds enrolled in Head Start. Required by a directive from the Office of the President as part of the administration's Good Start, Grow Smart initiative, the NRS aims to collect, in a consistent manner, information on a standard set of child outcomes from all Head Start programs. Analysis of these data will enable the Office of Head Start to determine how children progress on a limited set of outcomes during the year preceding kindergarten. The data will also provide the Office of Head Start with information it can use to enhance its current program-monitoring system and develop targeted training and technical assistance. The NRS includes a 15-minute child assessment battery, a system for training staff members from all Head Start grantees to administer the assessment, and a computer-based reporting system that programs use to enter the completion status of assessments and report information on the characteristics of participating Head Start programs, teachers, and children. The Office of Head Start provides to each program a summary report of average results for all children in the program who were assessed, as well as information broken down by the child's primary language. The report also provides information on national averages across all children assessed in the NRS. This report is available several months after each fall and spring administration (or baseline and follow-up administrations, respectively, for Migrant/Seasonal programs). Reference tables are also created to allow programs to compare their scores with national averages, with regional averages, and with scores of other programs similar along a number of characteristics, such as the percentage of children who are English-language learners.

In July 2003, the Administration on Children, Youth, and Families (ACYF) contracted with Mathematica Policy Research, Inc. (MPR) and its subcontractor, Juárez & Associates (J&A), to conduct the Head Start NRS Quality Assurance and System Development Project. The project had two components: (1) an implementation study to assess the quality and other aspects of the first year of NRS implementation (training, child assessments, data entry, and program perspectives), and (2) support for system development activities that could enhance the quality and usefulness of the NRS. A final report on the Year One

Quality Assurance Study was submitted to the Office of Head Start in December 2004 (Paulsell et al. 2004). ACYF decided to extend the Quality Assurance Study into a second year and in July 2004 again contracted with MPR and J&A to continue the study with the same two project components.¹ This phase yielded two interim reports on the fall 2004 and spring 2005 implementation periods (Paulsell et al. 2005 and 2006).

This report summarizes the experiences of a sample of Migrant/Seasonal Head Start programs during their second year of NRS implementation. We report findings from the second year of the NRS Quality Assurance Study, based on site visits conducted in April-June 2005 to observe baseline assessments and in May-August 2005 to observe growth (or follow-up) assessments.² Although our Migrant/Seasonal sample is small, and findings should be interpreted with caution, the information presented provides useful insights about implementing the NRS in Migrant/Seasonal programs. The following section describes the unique characteristics of Migrant/Seasonal Head Start programs, and our methodology for selecting the sample and conducting the visits.

KEY CHARACTERISTICS OF MIGRANT/SEASONAL HEAD START PROGRAMS

The Head Start program has served migrant farm worker families since 1969 and seasonal farm worker families since 1999. Currently, 24 migrant and seasonal grantees operate centers in 34 states and serve more than 32,000 migrant and 2,500 seasonal children.³ “Migrant” families are those engaged in agricultural work and who have changed residency within the past two years, while “seasonal” families are those who perform agricultural work but have not changed residency in the past two years. According to the Migrant and Seasonal Head Start Quality Improvement Center, approximately 97 percent of families are Latino, and 9 of 10 individuals speak Spanish as their dominant language.⁴

There are two types of migrant and seasonal grantees—“home based” and “upstream.” Home-based grantees are located in the southern United States and serve families in the geographic location that they consider to be their permanent home, generally from September through May, after families return from distant farm work. “Upstream” grantees

¹ In July 2005, oversight of the contracts for the Head Start NRS Quality Assurance and System Development Project was transferred from the Office of Head Start to the Office of Planning, Research, and Evaluation (OPRE), Administration for Children and Families, U.S. Department of Health and Human Services.

² We use the term “growth” because this is the word that appears on Migrant/Seasonal assessment forms. A growth assessment is, in effect, the second (follow-up) assessment administered to children within a given program year.

³ The National Migrant and Seasonal Head Start Association [www.nmshsa.org]. Accessed November 14, 2005.

⁴ Migrant and Seasonal Head Start Quality Improvement Center [www.mhsqic.org]. Accessed November 14, 2005.

provide services to families as they move north for short-term agricultural employment, typically during the summer months. (The United States is divided into the Western stream, the Midwestern stream, and the East Coast stream.) However, changes in the agricultural industry (for example, due to droughts) might prompt families to target different crops, move out of state unexpectedly, or relocate in atypical migration patterns, such as west to central instead of the traditional south to north.⁵ Consequently, families do not always access the same social service providers from season to season or year to year.

Migrant/Seasonal Head Start grantees and their delegate agencies have unique program characteristics that reflect their families' migratory lifestyles and participation in the agricultural industry. Children often need full-time care from 6 A.M. to 6 P.M., six days a week. Enrollment is open and continuous, as families arrive and leave a given geographic area to "follow the crops." Moreover, Head Start staff members frequently must provide top-priority, critical services—such as medical care, emergency food, clothing, and temporary housing—to families who are in the upstream programs and might be around for only two or three months. Families must also confront fundamental health issues and threats more often than the wider Head Start population, including higher infant mortality rates, chemical poisoning from pesticides, vitamin deficiencies, and substandard housing with poor sanitation.⁶ Ensuring that families meet basic health and safety needs sometimes takes precedence over academic components.

SAMPLE SELECTION AND SITE VISITS

To select a sample of Migrant/Seasonal Head Start programs, we used the most recent version of the 2003-2004 Head Start Program Information Report (PIR) as the basis for a sampling frame. We used the PIR reporting unit—grantees and delegate agencies—as the primary sampling unit. Because they serve a large proportion of children enrolled in Migrant/Seasonal Head Start programs, we selected the two largest grantees (sometimes referred to as "super grantees" because of their size) with certainty using explicit stratification. In addition, because the Office of Head Start required that migrant programs conduct NRS assessments with children who would be enrolled in the program for at least 16 weeks, we excluded from our sampling frame 16 of 61 Migrant/Seasonal Head Start programs on the basis of their program start and end dates in the PIR. We then applied the same procedures used for the main Quality Assurance Study to identify children to observe within programs (Paulsell et al. 2006).

We selected a sample of five Migrant/Seasonal Head Start programs to visit during the baseline assessments, and another five agencies to visit during the growth assessments. Table 1 presents characteristics of the sample of migrant programs visited. Two programs were selected for both the baseline and growth assessment visits; thus, we conducted a total

⁵ Ibid.

⁶ Ibid.

of 10 site visits to 8 programs. Because Migrant/Seasonal Head Start programs operate on a schedule that coincides with the growing season, baseline visits were conducted in spring 2005, and growth visits were conducted in summer 2005. As in the main Quality Assurance Study of regional programs, at each site we observed a sample of 10 child assessments, interviewed key program staff members, and conducted a focus group with NRS assessors to discuss their experiences with the assessment system.

ROAD MAP TO THE REPORT

We now turn to describing the results of our assessment of NRS implementation during the second year in a sample of Migrant/Seasonal Head Start programs. In Chapter II, we describe the quality of the child assessments we observed, the experiences of staff members in administering the child assessments, and unique challenges associated with the Spanish-language version of the assessment. Chapter III presents the approaches that programs took in training and certifying staff members to conduct the assessments. Chapter IV describes programs' approaches to implementing the NRS, including communication with parents and Policy Councils, coordination and staffing of the assessments, and associated costs. This chapter also describes programs' experiences using the Computer-Based Reporting System (CBRS). In Chapter V, we describe (1) how Migrant/Seasonal programs have used the NRS results for program improvement efforts, (2) their future plans for using the results, and (3) their reactions to the growth report on outcomes for their programs. In Chapter VI, we present what we learned from the on-site interviews and focus groups with program staff members regarding their concerns about the NRS, along with their suggestions for improving it. Based on findings from the spring and summer 2005 site visits, in Chapter VII we synthesize recommendations for improving the NRS system in Migrant/Seasonal programs.

Table 1. Characteristics of Sample Migrant/Seasonal Head Start Agencies

Characteristics	Number of Agencies
ACF Region	
Region XII (Migrant/Seasonal Programs)	8 ^a
Head Start Grantee Status	
Grantee That Operates Programs Directly and Does Not Have Delegate Agencies	4
Grantee That Operates Programs Directly and Has Delegate Agencies	2
Delegate Agency	2
Head Start Program Option	
Center-Based Services Only	8
Type of Agency	
Community Action Agency	2
Private or Public Nonprofit	6
Number of Head Start Centers the Agency Operates	
1 to 5	2
6 to 10	1
11 to 20	1
21 to 35	3
More than 35	1
Number of Enrolled 4- and 5-Year-Olds	
1 to 50	2
51 to 100	0
101 to 150	0
151 to 200	0
201 to 300	1
301 to 400	2
401 to 500	1
501 to 600	0
601 to 700	0
More than 700	2

Source: Head Start Program Information Report, program year 2003-2004.

^a While we conducted a total of 10 site visits to Migrant/Seasonal programs for the second year of the Quality Assurance Study, our sample includes 8 unique agencies. As explained on pages 3-4, in the section describing our sampling plan, we included two programs in the first and second round of site visits, for a total of eight different agencies. Frequencies for Table 1 are based on a base of eight programs (n=8), not the total number of site visits (n=10).

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

CHAPTER II

ADMINISTERING THE SPANISH-LANGUAGE VERSION OF THE CHILD ASSESSMENT

As in our fall 2004 and spring 2005 site visits to 69 regional Head Start programs, evaluating the implementation of NRS child assessments was the central component of our 10 site visits to Migrant/Seasonal programs. Identifying patterns of errors can help the Office of Head Start understand the quality of assessments, the need for additional training and guidance, and whether certain items are unsuitable or too difficult and should be modified or eliminated.

Our approach to evaluating the quality of the NRS assessments remained the same from our evaluation of a nationally representative sample of regional programs (Paulsell et al. 2004, 2005, and 2006). We used an internally designed Assessment Observation Form, based on the assessor certification form used in the initial training. This observation form captured three components of test administration—scoring errors, administration errors, and procedural errors.⁷ We again conducted focus groups with local assessors about conducting assessments, and interviewed training staff members. Since most children served by these programs were Spanish speakers, we include only the Spanish observation data in this analysis.⁸

In this chapter, we describe the overall quality of assessment administration of the Spanish-language version in spring and summer 2005, and also make some comparisons to results from earlier rounds. Our analysis is based on the extent to which the assessments we observed met certification standards, on ratings of inter-rater reliability, and on the types and frequency of administration and scoring errors that assessors made. Next, we describe

⁷ Please see Paulsell et al. 2006, pages 18-19, for detailed explanations of the types of errors.

⁸ Of 50 observations conducted for the baseline assessments, 42 were in Spanish and 8 were in English; the growth assessments yielded 46 observations in Spanish and 3 in English.

assessors' experiences conducting the assessments, based on interviews and focus groups with assessors, lead NRS trainers, and Head Start directors. We also present some unique issues involved in administering and scoring the Spanish-language assessment.

MEETING THE CERTIFICATION STANDARDS

During our site visits, project team members observed Head Start assessors as if they were “certifying” them. The mean certification score for the Spanish-language baseline assessments was 93, which indicates that, on average, administration of the baseline assessments exceeded the certification standard by 8 points (see Table 2).⁹ Ninety percent of observed assessments achieved a certification score of at least 85, the level required for the initial certification following training. Seventy-six percent scored in the range of 92 to 100 percent. One assessor achieved a score of 64, the lowest we observed. These results closely resemble the baseline migrant certification scores derived from the first year of the Quality Assurance Study.

During the growth assessment (follow-up) observations, the mean certification score was 95, which indicates that, on average, administration of those assessments exceeded the certification standard by 10 points (see Table 2). Eight-nine percent of the observed assessors met certification standards with a minimum score of 85, and 80 percent scored in the range of 92 to 100. In contrast, only about 67 percent of migrant growth assessments during the first year of implementation had achieved at least an 85, and half scored in the 92-100 range. Overall, staff members administered the child assessments to a high degree of quality in Year Two, particularly given the unique operational challenges that Migrant/Seasonal programs face with their compressed calendars. Only one baseline assessment and two growth assessments missed the certification standards by more than one point.

Although it was not a criterion for certification, completing the assessment within a specified length of time relates both to operational and cost issues. The NRS is designed to be administered in 15 minutes. The average duration across all baseline observations in which the child passed the language screener and completed the entire assessment was 18 minutes; this increased to 19 minutes for the growth observations. Sixty-four percent of observed baseline assessments lasted between 12 and 19 minutes, and 59 percent of observed growth assessments lasted between 13 and 19 minutes. Eighteen percent (baseline) and 39 percent (growth) took 20 minutes or longer, but only 8 percent of baseline assessments and 9 percent of growth assessments lasted at least 25 minutes.¹⁰

⁹ We did not calculate certification scores for those observations in which the children completed only the Tío Simon and Exposición de Arte sections.

¹⁰ The length of the Spanish-language assessments for the regional Head Start programs visited in spring 2005 was similar, ranging from 11 to 28 minutes and averaging 19 minutes.

Table 2. Distribution of Certification Scores Across Observed Baseline and Growth Spanish Assessments

Certification Score	Number of Assessments	Percentage of Assessments	Cumulative Percentage
Observed Baseline Spanish Assessments^a			
64	1	2.4	2.4
84	3	7.3	9.8
88	6	14.6	24.4
92	12	29.3	53.7
96	10	24.4	78.1
100	9	22.0	100.0
Total	41	100	
Observed Growth Spanish Assessments^b			
80	2	4.4	4.4
84	3	6.7	11.1
88	4	8.9	20.0
92	9	20.0	40.0
96	9	20.0	60.0
100	18	40.0	100.0
Total	45	100	

^aObservations (N=41) of the Spanish-language version of NRS child assessments, April to June 2005.

^bObservations (N=45) of the Spanish-language version of NRS child assessments, May to August 2005.

INTER-RATER RELIABILITY

In addition to assessing the extent of errors on individual items (discussed in detail below), we looked at the reliability of total scores for each section of the NRS by calculating an intraclass correlation coefficient (ICC) for each scale (Understanding Spanish, Vocabulary, Letter Naming, Early Math Skills, and Counting). The ICC estimates the proportion of total score variance that is due to variance of the scores across children observed, rather than variance across the Head Start assessor and the MPR observer scores. High ICC scores indicate a small amount of variance that is attributable to assessment error.

Intraclass Correlation Coefficients for NRS Scales		
	Baseline ^a	Growth ^b
Understanding Spanish	0.98	0.93
Vocabulary	0.99	0.95
Letter Naming	0.98	0.94
Early Math Skills	0.99	0.97
Counting (item E20)	0.94	0.96
^a N=40 observations, N=42 for Understanding Spanish ^b N=45 observations, N=46 for Understanding Spanish		

Estimates of reliability using the ICC for both rounds of assessments indicate high inter-rater reliability across subscales. Errors on the individual items do not result in substantial differences in scores for the scales (see Box above).

ERRORS IN PROCEDURES, ADMINISTRATION, AND SCORING

During the assessment observations, site visitors coded scoring and administration errors for the assessment items, the introductory sections preceding each segment, and the practice items.¹¹ Assessors had difficulty administering and scoring some items, but their difficulties tended to cluster in specific sections of the assessment. The overall error rate was low (see Box below). Incidents of straying from the script and coaching increased on average from the first to the second year of NRS implementation. However, average scoring errors, using incorrect hand gestures, and cases of non-neutral encouragement declined between the two years. Samples are small, so any changes in scores should be interpreted with caution.

The most frequent types of errors observed were scoring, followed by incorrect hand gestures, straying from the script, coaching, and then non-neutral encouragement. The highest number of mean errors emerged in scoring for growth assessments (4.0). On

	Baseline ^a Year 1	Growth ^b Year 1	Baseline ^c Year 2	Growth ^d Year 2
Straying from the script	1.6	0.8	1.9	1.5
Coaching	0.3	0.1	0.9	0.7
Non-neutral encouragement	0.9	0.2	0.4	0.1
Incorrect hand gestures	1.7	3.5	2.6	0.6
Scoring errors	2.3	6.2	2.5	4.0
^a N=34 observations	^c N=42 observations			
^b N=21 observations	^d N=46 observations			

average, assessors scored about 97 percent of all responses per baseline assessment correctly, and about 96 percent of all responses per growth assessment correctly. During the first set of observations, the percentage of assessments that included at least one scoring error, by section, ranged from 29 percent (Exposición de Arte) to 43 percent (Test de Vocabulario en Imágenes Peabody, TVIP). For the second set of observations, the percentage of assessments that included at least one scoring error, by section, ranged from 28 percent (Exposición de Arte) to 49 percent (Básicos de Matemática). Most scoring errors clustered around Tío Simón items and the counting blocks/marbles activity (E20).

Assessors made an average of 3 incorrect hand gesture errors—or omitted them—per baseline assessment, and each growth assessment contained an average of 0.6 gesturing errors. Most gesturing errors emerged in Básicos de Matemática, which requires assessors to read the script carefully, make precise hand gestures, and record the child's answers accurately. These types of errors declined significantly by the next round, perhaps because assessors had familiarized themselves with the administration protocols.

¹¹ In this report, we present unweighted estimates of assessment errors.

Errors in straying from the script were similar over time (1.9 errors per baseline assessment versus 1.5 per growth assessment), though they were higher on average than in the previous year. Such errors usually involved either changing a word or two without significantly altering the meaning of the question, or else accidentally skipping an item or a scripted phrase of neutral encouragement. In some cases, straying from the script included pronunciation mistakes. (However, we did create a separate error code category for mispronunciation errors or inserting of articles into the TVIP section.) Coaching errors occurred in about one-third of baseline and growth assessments in the Básicos de Matemática section (usually repeating questions), and in about one-third of baseline assessments during Tío Simón (looking up for the second practice item); coaching errors in the other sections of the NRS were less frequent. Staff members made almost no errors in offering non-neutral encouragement.

For the rest of this section, we describe the frequency of administration errors, scoring errors, and some procedural errors not captured in the certification form (highlighted in the text boxes throughout this section) for specific items in baseline and growth assessments. It is important to note that overall error rates were low. However information about these errors is useful for making improvements to the training and materials provided to programs in the future; hence, we focus our discussion here on the errors that were observed.

Administering the Set-up and Warm-up Section

Most errors in this section involved failure to record background information properly on the cover sheet (see Box at right). Assessors occasionally did not have the required materials ready or relied on a noisy space, such as an office with traffic, for NRS testing. Assessors strayed from the script once or twice during the warm-up section in about 21 percent of baseline assessments and 33 percent of growth assessments. Before the warm-up, one assessor instructed a child to say, “I don’t know” if he could not give an answer in order to minimize the length of the assessment. Whenever the child answered this way throughout the sections, the assessor praised him.

Percentage of Assessments with Observed Setup Errors

Baseline^a

Did not fill out child’s ID number correctly	27
Did not fill out cover before starting assessment	26
Did not fill out date correctly	14
Did not have the paper ready for Simon Says	2

Growth^b

Did not fill out date correctly	9
Did not fill out cover before starting assessment	7
Assessment was not set up in quiet area	7
Area was not set up before child arrived	2

^aN=42 observations; ^bN=46 observations
Unweighted estimates

Administering and Scoring Tío Simón (Pre-LAS Simon Says)

In the Spanish version of the NRS assessment, a task called Tío Simón is part of the language screener. This is administered to determine if a child whose home language is recorded as Spanish has enough language proficiency to do the assessment in Spanish. Scoring and coaching were the most frequent types of errors (see appendix tables A and B).¹² Scoring errors clustered primarily around four items—“Tío Simón dice, junta los pies” (put your feet together), “Tío . . . abre una mano (open your hand), “Tío . . . toma el papel” (pick up the paper), and “Tío . . . pon una mano sobre la otra (put one hand on top of the other). Most incidents of coaching occurred when assessors modeled the second practice item by looking down, or hinted by looking under the table as well as looking or moving the piece of paper (see Box). These types of errors decreased substantially from baseline to growth assessments.

Percentage of Assessments with Coaching Errors on Tío Simón	
Baseline^a	
Modeled second practice by looking down	56
Hinted by looking at or moving paper	12
Hinted by looking under table too soon	12
Modeled first practice item incorrectly	2
Growth^b	
Did not give appropriate guidance to child on what to do with piece of paper	4
Hinted by looking under table too soon	2
^a N=42 observations; ^b N=46 observations Unweighted estimates	

Administering and Scoring Exposición de Arte (Pre-LAS Art Show)

On the Exposición de Arte task, the second part of the language screener, assessors made relatively few errors. There were more straying-from-the-script errors during the second round (10 assessments) than the first (3 assessments). Coaching and incidents of non-neutral encouragement were minimal (see appendix tables A and B). One assessment included eight gesturing errors because the assessor repeatedly neglected to point to each item when asking the child, “¿Qué es esto?” Scoring errors occurred during 29 percent of observed baseline assessments and 28 percent of growth assessments, nearly all of which contained one scoring error each.

At the conclusion of Exposición de Arte, assessors must sum the incorrect responses from Tío Simón and Exposición de Arte. If the total number of incorrect responses exceeds 14 and the child does not speak Spanish as the home language, the assessor is instructed to end the assessment; otherwise, it should continue. If English is the home language, the assessor can immediately assess the child in English or reschedule the assessment in English for another day. Children whose home language is neither Spanish nor English are not assessed in English if they do not pass the English screener. Staff members did not choose

¹² Frequencies and percentages cited in this section should be interpreted with caution, since the sample sizes for Migrant/Seasonal Head Start programs, as well as the observed assessments, are small.

the correct path for four baseline assessments (10 percent). One assessor continued with the next section even though the child made 15 errors. Another assessor stopped three assessments after the language screener because she assumed children needed to give 15 correct responses to complete the final three sections; she also entered onto the cover sheet the number of correct items instead of the number of incorrect responses. Moreover, for 19 percent of baseline observations and 15 percent of growth observations, assessors did not complete the language screener boxes before continuing on to the TVIP section.

Administering and Scoring the Test de Vocabulario en Imagenes Peabody (TVIP-Adapted)

Assessors made several administration errors throughout the vocabulary task during both sets of observations (see appendix tables A and B). At least one scoring error emerged in 45 percent of baseline and 38 percent of growth assessments. Observers typically recorded one error per assessment, though a small number included two or three scoring errors.

Some assessors erroneously inserted articles (for example, “Pon el dedo sobre *el* perro” instead of “Pon el dedo sobre perro”) during the practice section. Some assessors also mispronounced certain vocabulary words; site visitors guessed that perhaps this was because they either were not native Spanish speakers or were not accustomed to speaking formal Spanish. For example, in some instances, assessors said *liquado* (not a Spanish word) instead of *liquido* (liquid), some said *hombre* (man) instead of *hombro* (shoulder), and some had difficulty pronouncing the word *aislamiento*. Straying from the script usually involved skipping a practice item. Some assessors made coaching errors by repeating questions during a small number of baseline assessments (10 percent) and growth assessments (4 percent). Using non-neutral encouragement and incorrect gestures rarely occurred on the TVIP section.

Additional administration errors were observed, nearly all at baseline. Some assessors either did not repeat at least one practice item so that the child would understand the activity, or else reread a practice item even if the child answered correctly the first time (31 percent); did not use encouragement or suggested probes as appropriate (17 percent); did not encourage a child to guess when appropriate (12 percent); did not help the child to master pointing and to choose one quadrant as a final answer (10 percent); and did not repeat the practice item until the child gave a correct response (5 percent).

Administering and Scoring *Nombre de las Letras* (Letter Naming)

Site visitors observed at least one scoring error in about one-third of all assessments for the *Nombre de las Letras* task (see appendix tables A and B). Assessors often made one or two scoring errors in an assessment, although during the second round, four assessments contained between 7 and 14 scoring errors each. Errors sometimes stemmed from assessors' failure to accept an English response as correct.

Coaching was observed during 9 percent of growth assessments. In one case, an assessor returned to *Nombre de las Letras* to give a child another chance to identify letters (originally the child was too

shy, and the assessor suspected—correctly—that he could name them). Straying from the script occurred in about 5 percent of baseline and growth assessments. Instances of non-neutral encouragement rarely occurred; one assessor said, “Muy bien” (“Very good”) and “Muy bien hecho!” (“Very well done!”) on two assessments. While only one child during the first round of assessments seemed to be aware that the assessor was marking responses on the bubble sheet, site visitors observed six children during the growth assessments who seemed to be aware that assessors were recording only correct answers.

Site visitors observed several other types of administration errors across all four plates (see Box above). Some assessors had difficulty managing this task, especially during the first round of observations. In about one-fourth of baseline assessments, they did not use suggested probes, such as prompting a child to say the name of a letter instead of a number. Some assessors did not appropriately encourage a child to say the names out loud instead of just pointing. In a handful of cases, assessors did not accept English responses, though such answers are permitted on the *Nombre de las Letras* task. During 13 percent of baseline assessments, assessors did not ask the child to continue with the task. According to one focus group, this practice may be a reaction to how time-consuming this activity can sometimes be.

Percentage of Assessments with Errors on <i>Nombre de las Letras</i> , Across All Four Plates	
Baseline^a	
Did not use suggested probes	23
Did not ask child to continue with task	13
Was unable to help child pace pointing and naming	10
Did not score correct letter if given once	3
Growth^b	
Did not score correct letter if given once	7
Did not use suggested probes	4
Was unable to help child pace pointing and naming	2
^a N=42 observations; ^b N=46 observations Unweighted estimates	

Administering and Scoring Básicos de Matemática (Early Math Skills)

During the Básicos de Matemática section, assessors had difficulty with several items that required a combination of carefully reading the script, gesturing appropriately, and scoring. Administration errors decreased over time (see appendix tables A and B). Gestures were the most common type of mistake assessors made in the math section at baseline (83 percent), although this frequency dropped to 28 percent during the second round of assessments. This section in particular requires assessors to use precise hand motions.

Sometimes assessors used a slightly different hand gesture to draw attention to the easel page. For example, they circled the eggs instead of sweeping them (E3), or circled the fish instead of pointing to them (E14). In other cases, they omitted gestures, such as not pointing to the digits when asking, “What is this?” (E6-E8). In other instances, assessors pointed to the wrong part of the page or pointed when they should not have. For example, assessors neglected to point to the orange for the size comparison question (E13), swept the ruler instead of circling the numbers (E17), and circled the crayons *and* the paintbrush instead of just the crayons (E12). Items E18 and E19 caused assessors the most difficulty. For E18, they pointed to the pictures of three animals or the squares on the graph instead of their names (dog, cat, rabbit), as instructed; for E19, they either pointed to the picture of the cat instead of the word *gato* or else did not point at all.

Baseline ^a	
Did not prompt for a number if child must say number out loud	21
Did not allow non-verbal response	12
Did not encourage child to keep counting or could not control counting	12
Growth ^b	
Did not encourage child to keep counting or could not control counting	7
Did not prompt for a number if child must say number out loud	7
Did not allow non-verbal response	2
^a N=42 observations; ^b N=46 observations Unweighted estimates	

Assessors altered the script in 45 percent of observed baseline assessments and 43 percent of growth assessments. In several instances, assessors changed a word or words from the script, but doing so did not change the meaning of the question. For example, one assessor asked, “Cuántos estrellas hay *aquí*?” (here) instead of *ahí* (there), another replaced “How many eggs are there altogether?” with “How many total eggs are in the nests?” Other script errors were more significant. One assessor omitted “She gave one piece to Sarah and one piece to Joey” (E12), a clause that provided information important for answering the question. For a comparison item, one assessor omitted the word *más* in the command, “Ahora señala la fruta *más* pequeña” (“Now point to the smallest fruit”) on two assessments, which changed the meaning to “Now point to the small fruit.” Several assessors interchanged the conditional (*quedarían*) and future (*quedarán*) verb tenses for items E15 and E16 (“would you have” versus “will you have”).

Assessors made coaching errors in 31 and 24 percent of baseline and growth assessments, respectively. The most common type of coaching was repeating a question when a child gave an incorrect response. In two cases, the assessor instructed the child to count just the dogs when he counted all the squares on the graph (E18). One assessor ran her finger along the graph squares above the picture of Billy when asking, “How many times did Billy use the computer?” instead of pointing to his name. During three assessments, assessors counted the first marble with the child (E20). Non-neutral encouragement errors were less common overall, occurring in 12 percent of baseline assessments but not at all in the growth assessments.

As in the assessments observed during the first year of NRS implementation, most scoring errors clustered around the final task in which children are asked to count blocks/marbles out loud while simultaneously pointing to each object (E20). Assessors often had trouble scoring the item accurately when children either skipped blocks/marbles or counted the same object twice. They scored this item incorrectly in 26 percent of observed baseline assessments and 24 percent of growth assessments. Sometimes assessors did not seem to understand exactly how to give a child credit for the counting task. For example, children get credit for this activity as long as they count the objects in numerical order (1, 2, 3 . . .) and maintain a one-to-one correspondence while pointing to the objects; it is irrelevant if the child’s finger “jumps around” the plate, as long as these two conditions are meant. Assessors must understand this rule, as well as be able to control the child’s counting as necessary.

In addition, some assessors had difficulty using prompts appropriately, failing (1) to ask a child to say the number after giving a letter name as an answer, (2) to encourage the child to keep counting after reaching the end of a row of blocks during the counting activity, or (3) to accept a non-verbal response (see Box above). These procedural errors may have contributed to the difficulty some assessors had in accurately filling out the answer sheets.

ISSUES UNIQUE TO ADMINISTRATION AND SCORING OF THE SPANISH-LANGUAGE ASSESSMENT

Observing the Spanish-language version of the NRS assessments at Migrant/Seasonal Head Start programs revealed some issues surrounding language proficiency and translation. The themes are similar to findings from our sample of migrant programs in the Year One Quality Assurance Study (Paulsell et al. 2004 and 2005)—as well as certain findings that emerged during our site visits to regional programs that serve English-language learners. We summarize these issues below:

- Some focus group participants reported difficulty keeping track of which items they could accept English answers for when administering the Spanish-language version of the assessment; this uncertainty was confirmed by site visitors’ observations.
- Assessors from one program were confused about why English responses were acceptable on the Spanish version of the NRS, but Spanish words were not

acceptable on the English version. They also questioned how the Office of Head Start would interpret the results if they cannot know whether the child answered in English or Spanish. They wondered how the results are affected if children can answer in one of two languages on one instrument but only one language on the other instrument.

- Assessors observed that sometimes children’s responses were not one of the words listed in the instrument as acceptable answers, yet they were still “correct.” Several programs faulted the test developers for not taking into account the wide range of Spanish dialects spoken by families, and for only including one or two allowable options for certain items.
- Mispronunciation among some assessors could make it more difficult for children to give a correct response. Site visitors observed occasional instances in which an assessor’s mispronunciation of a word changed its meaning. For example, “Señala *hombre*” (“Point to man”) instead of “Señala *hombro*” (shoulder), or replaced “Señala *líquido*” (liquid) with “Señala *líquado*” (a word that does not exist in Spanish). One assessor said, “Ahora señala la fruta pequeña” (“Point to the small fruit) instead of “Ahora señala la fruta más pequeña” (“Point to the smallest fruit”) on two assessments.
- Staff members pointed out that certain vocabulary items either have another meaning entirely or are conveyed using other terms in certain dialects (see Box). Likewise, children are more likely to recognize and understand the meaning of *apuntar* (to point) instead of *señalar*, which is used throughout the script for the command “Point to . . .”

One program encountered a few children who banged their heads on the table in response to the Tío Simón command “Golpea sobre la mesa” (knock on the table) because, according to staff members, some children associate the verb *golpear* with hitting

themselves or getting hit by something. Expanding the list of acceptable synonyms would address programs’ concerns about the narrow scope of the instrument. However, adding acceptable synonyms that are of similar complexity may be difficult. Sometimes words may be used because of more limited, or less formal vocabulary, while other times the replacement indicates a true difference in dialects.

Examples of Differences in Spanish Word Usage		
English	Accepted as Correct, Used in Directions	Used by Children
Fork	Tenedor, Cuchara	Trinche
Snake	Culebra	Víbora
Cage	Jaula	Some children know this word to mean “bird house”

ASSESSORS' EXPERIENCES ADMINISTERING THE CHILD ASSESSMENT¹³

During site visits, assessors shared their experiences administering the assessment, describing aspects that they felt worked well and those that they felt were difficult for them or the children. Examining assessors' input from the second year of NRS implementation can help pinpoint aspects of the child assessment for which additional training and support are needed, as well as sections or specific items that may require some revision.

Children's Responses to the Assessment Process

Staff members from nearly all Migrant/Seasonal programs in our sample noted that most children responded positively and were cooperative with assessors. During the second year, no child refused to take the assessment. Moreover, local staff members agreed that children were more comfortable and confident during the growth assessments. Not only did they learn classroom skills (for example, counting and alphabet knowledge) that enabled them to answer more questions, but they were familiar with the testing procedures and assessors. As one lead trainer noted, "It's like a review for them."

Nevertheless, half the programs observed that some children had difficulties, especially during the baseline assessments, when they were becoming acclimated to a new environment and frequently took the assessment within a few weeks of having begun Head Start. Some children became timid or bored; some "shut down" because they were unfamiliar with certain activities they were asked to do (for example, pointing to pictures) or were not familiar with the content of the test. One director characterized the children as having "tolerated" the NRS but not having enjoyed it; one reported seeing children cry during the test. She worried that their self-esteem could become deflated when they did not know the answers.

Concerning specific sections with which children struggle, a group of assessors said that children tended to grow restless during the vocabulary section because it is so long, often picking quadrants randomly without thinking through their answers. Similarly, some focus group participants observed that the letter-naming task can frustrate children because staff members are expected to ask repeatedly, "Look carefully at all of them. Do you know any others?" until the child says no. One site visitor noted that in some cases a child takes the instructions very literally, and simply nods when an assessor asks, "Do you know any of

¹³ For the rest of this report, we summarize the experiences of programs in our sample as they implemented the NRS in 2005. As previously highlighted in the sampling section (page 3), we included 10 Head Start agencies (5 for the baseline assessments and 5 for the growth assessments). However, we again selected 2 agencies in the second round of site visits, for a total of 8 different agencies. Importantly, to analyze the qualitative data from interviews and focus groups, we calculated frequencies based on $n=10$ as opposed to $n=8$. For example, if a program visited twice was the only program to select classroom teachers to conduct the child assessments, we counted this as 2 programs that used teacher-assessors and considered the percentage from a total of 10 agencies (2 divided by 10, or 20 percent), rather than a total of 8 agencies (1 divided by 8, or 13 percent). The denominator represents the number of visits for both rounds.

these?” When this occurs, the assessor directs the child to point to the letters that he or she knows, in which case the child may simply point without naming. Site visitors witnessed similar situations that prolonged child assessments at regional Head Start programs.

Strategies for Reacting to Children’s Responses

Local program staff members encountered a variety of responses and behaviors from children while administering the assessments. They shared a number of strategies they developed to respond to children who, for example, were withdrawn, distracted, tired, or afraid or wanted a break to get a glass of water or go to the bathroom. Several assessors (1) used redirection techniques suggested by the training manual, (2) took breaks, (3) rescheduled the assessments for a different day, or (4) repeated questions if a child was not paying attention. Staff members said that in a few instances they replaced one assessor with another to see whether a child responded better, as well as invited a child’s classroom teacher to sit in the same room during the assessment. For those assessors who were not the children’s teachers, one lead trainer encouraged her assessment team to visit the classroom and meet the child in advance of the scheduled testing date.

For the most part, assessors said they were comfortable responding to children’s behaviors, encouraging them when necessary and managing the assessments, especially since most of them had conducted three or four rounds of assessments. Still, some staff members described difficulties. For example, three programs noted challenges when children were shy and unresponsive, since the NRS protocols discourage them from providing too much encouragement. As one focus group participant explained, they could give neutral praise and redirect the child, but they also had to refrain from straying from the script too much, and they feel as if their “hands are tied.”

Experiences of Assessors Who Were Also Children’s Teachers

Site visitors spoke with a handful of individuals from four programs who administered the NRS assessment to children from their own classrooms. These teacher-assessors reported that children’s performance sometimes aligned with their expectations and perceptions of skill levels, and sometimes it did not. One assessor concluded that the NRS does not provide any unique insight into children’s abilities. But another said she likes conducting the assessments because they enable her to observe how children perform and respond to certain tasks and activities that she does not observe in the classroom, such as letter naming and counting. Finally, a group of assessors reported that it is routine for Head Start teachers to give children positive encouragement and to answer the question, “How did I do?” after assessments and screenings. Because assessors can give only neutral encouragement on the NRS, one of these assessors felt that the children get “nervous” when the feedback they receive is not more enthusiastic.

Experiences Assessing English-Language Learners

All programs in our sample assessed children whose home language was not English. Spanish was the most common home language, followed by indigenous Mexican languages

(Misteco and Triki). Some children spoke English, and a very small percentage spoke Haitian Creole or Arabic.

Staff members in the sample programs reported that most children were able to complete the growth assessments in both English and Spanish. As a result, they were often more comfortable with the NRS toward the end of the program year because they had learned more English in the classroom and better comprehended the assessor's questions and commands. In contrast, Spanish speakers who passed the English assessment at baseline often grew frustrated or quiet because they could not understand what the assessor was saying. Some assessors observed that children whose home language was neither Spanish nor English rarely passed the screener at baseline; they simply parroted what the assessor asked them to do (Tío Simón) or name (Exposición de Arte).

Some respondents noted that although some families reported Spanish as the home language at enrollment, their children actually spoke and understood English much better, making the first assessment (Spanish version) more difficult. In fact, staff members said that bilingual children often performed better on the English-language assessment. One program would like the option of assessing bilingual children in English first, depending on their fluency.

In addition, many programs took issue with the language used on the Spanish instrument. Forty percent of programs observed either that the language was more formal than what the children used at home (for example, *aislamiento*) or that it used vocabulary words different from those that are common where their families originate. Some of the children's words are technically correct but are not acceptable answers because they do not appear in the easel, which staff felt unfairly places children from certain ethnic subgroups at a disadvantage.

Experiences Assessing Children with Disabilities

Ninety percent of programs in our sample assessed children with disabilities. Speech and language delays were the most common type of disability encountered by assessors, followed by cognitive and developmental delays, and autism. Other less common disabilities included cerebral palsy, attention deficit disorder (ADD), and visual impairment.

Local program staff members identified several accommodations they used while assessing children with special needs. In many cases, assessors administered the NRS and "proceeded the best way they could." They most often mentioned taking breaks, using redirection, and allowing longer response times (for example, for a speech impediment) as ways to facilitate the assessment process. One program permitted a therapist who worked with an autistic child to sit with him during the assessment. Three programs did not need to make special accommodations, since children's disabilities had not affected test administration.

Four programs relayed difficulties in conducting or completing assessments. Staff members from two programs stopped during Tío Simón or Exposición de Arte when children were unresponsive as the result of a developmental delay or cerebral palsy. An

assessor from another program had difficulty comprehending a child with severe speech impairment during Nombre de las Letras, and another assessor noted that testing children with speech disabilities was difficult when it was not possible to find a quiet space. One focus group reported that they had not received adequate guidance on administering the NRS to special-needs children and would like more support from the Office of Head Start.

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

CHAPTER III

NRS TRAINING AND SUPPORT ACTIVITIES

To prepare for the second year of NRS implementation in Migrant/Seasonal Head Start programs, the Office of Head Start did not sponsor another “train the trainers” conference for all lead NRS trainers from the migrant programs. Instead, lead NRS trainers who had been certified at one of the training conferences in 2004 conducted refresher and new-assessor training for their local program staff members. A smaller summer training was held for only those programs that experienced a transition in their lead NRS trainer.

Since most programs in our sample did not need to certify new assessors in 2005, this chapter concentrates on the approaches agencies took to implement refresher training.¹⁴ We also describe the perspectives of local program staff members on the NRS training materials and the support they received from the Office of Head Start in 2005, as well as their recommendations for improvement.

LOCAL REFRESHER TRAINING ACTIVITIES

Most local Migrant/Seasonal Head Start programs in our sample offered refresher trainings prior to the baseline and growth NRS assessment periods. All “returning” bilingual and English-only assessors attended refresher training led by a trainer who had attended a train-the-trainers conference previously—either one for regional programs or one for Migrant/Seasonal programs. Depending on a program’s size, local training was conducted by a single individual or a team who had been certified by the lead trainer. Some programs elected to provide one group training, while others led multiple training sessions to minimize

¹⁴ Because the grantee had expanded, one agency certified additional assessors for the baseline assessments. During a day-long bilingual training session, participants reviewed the easels, watched the videos, reviewed the *Assessor’s Guide*, and role-played with each other. All staff members were certified on their first attempt; the lead trainer reported observing occasional instances of non-neutral encouragement.

travel time for assessors and to accommodate centers within an agency that operated on different schedules during the year. Because of logistical challenges, one grantee conducted one-on-one refresher trainings with some assessors who were not able to travel for a scheduled session at the central office, and another agency hosted teleconference calls for the same reason.

According to lead trainers and participants, trainers generally followed the model agenda included in the national training materials. Common activities included a review of the assessment easels, a viewing of the training video, role-playing exercises, and a question-and-answer session. English and Spanish trainings took place on the same day and lasted, on average, about 3.5 hours for each language. One program reported that it did not follow the Office of Head Start's suggested training plan because it was too time consuming. Instead, assessors met as a group with the lead trainer to discuss new items or changes to the easel, and some watched the video individually or in pairs. Only one delegate agency did not offer formal refresher training, since the lead trainer was the only person who conducted assessments during 2005.

LOCAL QUALITY ASSURANCE PRACTICES

Lead trainers at three programs visited said they had already implemented or planned to implement quality assurance procedures to monitor the administration of child assessments using parallel scoring and certification forms provided at regional train-the-trainers conferences. At one program, the trainer observed each assessor at least once; at another, the trainer observed about half the assessment team. At a third program, the lead trainer said she observed assessors (without using the certification forms) and gave feedback if an assessor requested it. Two other programs monitored assessment administration in 2004 but opted not to do so in 2005, primarily because their team had more experience. Two additional programs neither have performed nor plan to perform quality assurance activities.

Local trainers noted the following types of errors during their quality assurance observations: assessors who (1) did not accept a non-verbal response when permitted, (2) displayed confusion about when they could accept an answer in English on the Spanish assessment, (3) made minor gesturing errors, (4) deviated slightly from the script, (5) made minor scoring errors, and (6) occasionally mispronounced words in Spanish. Those staff members who performed parallel scoring met with assessors individually soon after they completed the certification forms to give feedback. One program used results from these observation sessions to guide which procedures to highlight during the next NRS refresher training.

VIEWS OF LOCAL PROGRAM STAFF MEMBERS ON NRS TRAINING MATERIALS AND SUPPORT

During site visits, we asked local program staff members involved in NRS implementation about the helpfulness of national training materials, technical assistance, and other support the Office of Head Start provided for implementation. Assessors reported that the refresher training prepared them adequately to conduct child assessments during the

second year of NRS implementation. Lead trainers said that their assessment teams were well prepared to conduct assessments, since the vast majority was experienced assessors who have been involved since the inception of the NRS. Although trainers and participants commended certain aspects of the refresher training, they identified several problems with the training and technical assistance process. Next, we describe their views on the NRS Web casts, the refresher training, training and assessment materials, and technical assistance.

NRS Web Casts

At least one representative from each Migrant/Seasonal program watched one or more of the three most recent broadcasts sponsored by the Office of Head Start. While some programs found them to be helpful and informative, many did not. Some directors and lead trainers enjoyed hearing about the experiences of other programs, learning how their program compared with the national averages, and receiving updated information about NRS implementation and impending revisions to the instruments. Some staff members found the explanation of the growth report from the first year to be helpful and appreciated the opportunity to call in and ask questions. One lead trainer noted that the broadcasts helped “motivate staff to share information about their program,” and “made us feel good about our work.”

On the other hand, local staff members frequently noted that the Web casts targeted regional Head Start programs and did not address specific issues facing programs that serve migrant and seasonal families. Some respondents described them as repetitive, “vague” (two programs), and “politically motivated” (one program).

Format of Training Sessions

All programs in our sample found some elements of the refresher training to be helpful. The most useful component cited by assessors was the opportunity to review the new versions of the easels and discuss changes with lead trainers. A few programs appreciated reading through the scripts verbatim as a group, particularly the Spanish assessment, which enabled them to practice pronouncing unfamiliar terms. Additional training benefits included performing role-plays and receiving clarification on how much neutral encouragement was permitted.

Five programs identified challenges with scheduling the refresher training; they arranged for either regional or one-on-one sessions for those who were unable to attend a group training, which added another layer of complexity to the local training timetable. One program director from a super grantee noted that time and resource constraints limited each regional training session to five hours (both languages), which reduced time available for role-play practice. Moreover, a few programs found the refresher trainings to be repetitive and too lengthy. Staff members said that requiring experienced assessors to meet each fall and spring wasted time and resources that could be spent on direct services for children and families.

Training Materials

Half the programs in our sample reported problems in obtaining NRS training and assessment materials from the Office of Head Start. Several programs reported delays in receiving the materials or mistakes with the orders (for example, not enough answer sheets, or missing easel pages). In one instance, children had already left the program before materials for the growth assessments arrived. Another program never received a training manual or Assessor's Guides—only the videos. Consequently, the lead trainer copied materials from its affiliated regional Head Start program so that the refresher training could take place as scheduled. A few lead trainers described their frustration in getting information from the Information and Publication Center on their order status, often needing to make several phone calls. As one noted, “The center does not seem to know anything about the migrant programs.”

Two programs identified the training video as an effective tool because it (1) clarified which answers assessors could accept in complex cases, and (2) modeled different testing scenarios. More often than not, however, local staff members often criticized the training videos for not including Hispanic children. Others observed that while the video has improved over time, the scenarios are too “perfect” and “easy,” even with added “staged” events (for example, a child who becomes unfocused and needs to be redirected). In addition, one director reported that the Assessor's Guide is too long and that staff members find it difficult to use as a quick reference tool.

Technical Assistance

Most programs did not need to contact the technical assistance helpline. However, one lead trainer called to inquire about such topics as the correct order of the English and Spanish assessments, and cutoff dates for completing baseline and growth assessments. As someone who was involved with NRS training for both regional and migrant programs, she felt that it was more difficult to get answers to questions about procedures for Migrant/Seasonal programs than for regional programs, because the helpline staff members were less familiar with Migrant/Seasonal programs' unique structure.

RECOMMENDATIONS FOR IMPROVING TRAINING AND SUPPORT

Directors, lead trainers, and assessors made several suggestions for improving the refresher training and technical assistance provided to local programs. The most common suggestion across programs was to represent Migrant/Seasonal Head Start programs on the Web casts since, as one lead trainer noted, “They feel as if they are treated [by the Office of Head Start] like they did not exist.” For example, it would be helpful to give an overview of how they performed, not just the regional programs. Other recommendations included:

- If Migrant/Seasonal programs cannot be highlighted on the Web casts, then appoint a Head Start official from the Region XII ACF office who can address issues and procedural questions from these programs either when they contact

the office directly through the regular contact number or else through the NRS helpline, almost like a Migrant/Seasonal NRS “expert”

- Compile and distribute quarterly “Frequently Asked Questions” memos that document inquiries from the Web casts and helpline so that programs can learn from each other
- Ensure that the Office of Head Start’s Information and Publication Center is aware of Migrant/Seasonal programs’ unique operational schedules, and that it ships materials in a timely fashion
- Design a short document that summarizes for easy reference all changes made on the most recent version of the easels
- Consider shortening the length of refresher training, or permit experienced assessors to prepare for the next round of child assessments through self-study
- Include scenarios on the training video that demonstrate dealing with special-needs children, as well as scenarios that are more challenging and “realistic”

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

CHAPTER IV

LOCAL APPROACHES TO IMPLEMENTING THE NRS CHILD ASSESSMENTS

Several dimensions of NRS implementation were subject to the decisions of local programs. While the Office of Head Start provided guidance on administering the assessments, the latitude programs had in implementing the NRS resulted in different local approaches. This chapter describes the Migrant/Seasonal programs' approaches to communicating with parents and Policy Councils, coordinating the assessments, and assigning staff members to be assessors, as well as the costs related to the NRS. We also discuss programs' experiences using the CBRS.

COMMUNICATING WITH PARENTS AND POLICY COUNCILS

All eight Migrant/Seasonal Head Start programs we visited shared information about the NRS with parents, typically by sending letters (in English and Spanish) and by informing families about the NRS at enrollment and orientation. One program sent reminder letters to parents before growth assessments took place. Teachers or program administrators at some agencies spoke about the NRS at parent meetings. Nearly all programs shared NRS information with their Policy Councils, sometimes in a formal presentation with PowerPoint slides. One lead trainer wrote a report about its purpose and distributed copies to council members. Six of 10 programs required that parents sign a consent form authorizing their children to take the NRS assessment, usually as part of a generic form that families sign at enrollment for any assessment or screening administered during the year.

Staff members reported that many parents had questions and concerns about the NRS. They asked for clarification about its purpose, wanted to know whether its results would be used to cut Head Start funding, wanted to know whether it was age appropriate, and requested children's individual scores. A very small percentage of parents at two programs opted not to allow their child to be assessed; at another program, some parents requested that their children be assessed only in English. Policy Council members raised few questions or concerns and were mostly interested in the purpose of the NRS.

COORDINATING THE CHILD ASSESSMENTS

To implement the NRS effectively, program directors and other senior staff members needed to decide how to coordinate the assessment process. This section focuses on approaches to four coordination issues: (1) who would be responsible for the NRS, (2) where to conduct assessments, (3) when to schedule assessments, and (4) how to track progress. Site visit interviews revealed that these procedures seemed to be working well this year, although staff members at several programs noted difficulties associated with finishing all assessments within NRS time frames.

Staffing for Coordination and Oversight

Some Migrant/Seasonal programs are delegate agencies under the auspices of a grantee; some are part of a super grantee that oversees programs in multiple states. As a result, in our sample, responsibility for coordinating and overseeing NRS implementation often fell to someone other than the lead trainer. Other local staff members, such as a site supervisor or head NRS assessor, assumed responsibility for managing the NRS within a particular center or cluster of centers. They gave regular progress reports to someone from a central administrative office—usually a lead trainer—who then shared information with multiple program directors to keep them up to date on NRS activities and developments. According to program staff members, allowing center staff members to take the lead on monitoring progress and scheduling assessments usually worked well.

Locations Where Assessments Were Conducted

Assessors frequently used classrooms or offices for the assessments, although program staff members also used hallways, libraries, conference rooms, a staff lunchroom, and a gym. While most programs in our sample did not face significant difficulties in securing a quiet space for conducting child assessments, one director said it can be challenging at small centers with little room.

Scheduling the Assessments

As in 2004, all sampled programs scheduled assessments during regular program hours, taking children out of their classrooms. Local staff members reported that assessors and teachers used discretion in deciding when it would be most appropriate to assess children. Nearly all sites favored the mornings, when children were most alert. Programs used a range of strategies for scheduling the English and Spanish assessments. Some conducted both assessments on the same day with short breaks in between; others preferred to conduct the Spanish assessment first and then schedule the English-language version on a different day to give the child a break.

Since Migrant/Seasonal programs operate for shorter periods during a calendar year, several assessment teams were instructed—usually by a lead trainer or other administrator—that they needed to complete the assessments within a defined time frame. For example, programs often required that children take the baseline assessment within two weeks of enrolling in Head Start, and complete the growth assessments within two weeks of leaving to begin kindergarten. However, one director created a predetermined schedule indicating

when each child would be tested since assessors needed to cover a wide geographic area, and it would not be logistically feasible to travel back to certain locations multiple times.

While nearly all programs said they would be able to complete assessments within required time frames, respondents from five programs described the challenges associated with scheduling child assessments in Migrant/Seasonal programs. Centers in some sites were spread out over a large geographic area. In addition, because of the transient nature of the migrant communities, some children leave the program before taking the follow-up assessment.

Tracking Progress in Completing Assessments

Almost all programs (90 percent) in our sample reported using the CBRS classroom rosters to track when assessments were completed. In several cases, lead trainers checked in with center-level staff members on a regular basis (daily, weekly, or monthly) to ensure that they were “on pace” to complete the child assessments on time, and in turn gave updates to program directors and other central office staff members. So that it could monitor progress, one grantee designed a tracking spreadsheet for its delegate agency to complete and submit periodically.

APPROACHES TO ASSIGNING STAFF MEMBERS TO ADMINISTER THE CHILD ASSESSMENTS

Programs used one of three approaches to assign staff members to conduct the NRS assessment: (1) teachers only (one program); (2) no teachers (five programs); or (3) a combination of teachers, specialists, and other staff members (four programs). No program relied on outside contractors. Since these programs serve migrant families and have Spanish-speaking staff members, it was not difficult to identify bilingual assessors. About two-thirds of programs trained all or nearly all assessors in both languages, with the percentage of bilingual assessors for any given agency ranging from 50 to 100 percent. For the most part, these staffing strategies have worked well, and programs rarely reported instances of turnover among assessors since the NRS began. Nearly all agencies anticipate using the same type of staff members in the future.

The number of assessors certified within each agency varied according to its enrollment of kindergarten-eligible children and the number of centers. Some programs aimed to certify at least a few assessors at each center—usually classroom teachers, head teachers (much like a mentor teacher), site directors, or family service workers. Other programs assigned groups of assessors to different regions of their service areas to conduct assessments at clusters of centers or in a particular county. A few programs trained administrative staff members as backup assessors on an as-needed basis. Numbers ranged from 1 assessor (3 centers) to 50 assessors (61 centers in multiple states).

Selecting appropriate staff members to administer the NRS child assessments often depended on the perceived advantages and disadvantages of using classroom teachers for this task. Programs that opted not to use teachers did not want (1) to cause them to lose valuable classroom time with children, or (2) to add to their full workloads. One director

who did not use teachers cited her desire to avoid (1) possible bias in test administration, and (2) the added costs of substitutes. However, to provide a degree of familiarity and comfort, these programs tried to identify people who had some contact with children, including head teachers (who are not assigned to a permanent classroom), education coordinators, social workers, family service workers, paraprofessionals, and bus drivers.

Programs that used teachers exclusively or in combination with other staff members acknowledged some drawbacks, namely the possibility of coaching and lost classroom time. But they pointed to several other benefits. First, administering NRS assessments helped teachers gain insight into the children's learning abilities. Second, teachers were well equipped to manage behavioral issues and respond to shy children. Third, teachers were experienced assessors and skilled at putting children at ease, which may enable children to perform better and yield a more accurate picture of their skills. Providing this degree of comfort is particularly important among migrant children, many of whom are either transitioning to a formal learning environment for the first time, or are routinely enrolled in different Head Start programs as their families follow the crops.

COSTS OF CONDUCTING THE CHILD ASSESSMENTS

Based on four waves of child assessments, programs offered estimates of incurred costs over this period and perceived levels of burden that implementing the NRS had placed on local agencies. Three programs had not estimated costs, because they saw the NRS as being folded into staff members' normal responsibilities. However, four programs shared ballpark cost figures. For example, one director guessed that her program incurred about \$10,000 per assessment cycle to implement the NRS in six centers (which included training costs, assessor time, and travel expenses), while a smaller program with four centers spent about \$1,500 during the second year, not including training costs.

Local staff members perceived varying degrees of burden that the NRS has placed on them. During the second year of NRS implementation, some programs did not think that the Office of Head Start provided enough additional grant dollars to offset the costs of the NRS. To make up the difference, some programs tapped additional resources, such as a general fund, an education budget, or a grant from a local foundation. For example, one program used grant money to help cover expenses related to travel, time that the program director and lead trainer devoted to NRS-related activities, time staff members needed to conduct assessments, and data entry tasks. However, staff members from four programs felt strongly that the Office of Head Start needs to provide more funding to support implementation at the local level, such as for training, transportation, substitutes, and hiring of consultants to do assessments. Moreover, these four programs expressed that they do not see any benefits for staff members or families in return for the expenses and burden on staff members' time.

USING THE COMPUTER-BASED REPORTING SYSTEM

The CBRS was created to collect background information on Head Start programs and children, to facilitate the identification of children eligible for the NRS assessment, and to track the completion status of NRS assessments. The CBRS is managed by the contractor

Xtria, LLC. This next section briefly describes the overall experiences of the 10 Migrant/Seasonal programs using the CBRS. Most information that we gathered is consistent with findings from the sample of regional programs that implemented the CBRS during the first Quality Assurance Study (Paulsell et al. 2004). We then highlight unique obstacles that some programs encountered.

Experiences with the CBRS

Programs used a range of approaches to delegate CBRS data entry responsibilities, with most programs (80 percent) consolidating most or all data entry tasks under one person. Five agencies opted for people who already played a role in NRS implementation, such as lead trainers or assessors, either to handle all data entry and oversight or to share CBRS responsibilities with an administrative or information technology staff member. The other five programs assigned CBRS duties exclusively to administrative or information technology staff members, such as administrative assistants or data clerks.

All programs had enough computers and Internet access to fulfill data entry tasks, and all local program staff members felt comfortable with the system's security. No program incurred additional hardware or software costs, and no staff members reported any difficulty in gathering demographic and other background information on children and classroom teachers. People either had direct access to the data through their internal information system or contacted relevant staff members to collect information and update the records.

Nearly all Migrant/Seasonal programs had established internal quality assurance procedures. In many cases, staff members used CBRS rosters to verify when assessments had been completed and whether the information entered was accurate. For example, some data managers generated rosters on a regular basis for a supervisor or assessment teams to review and check for accuracy. Three programs also double-checked CBRS data against other data sources, such as the Head Start Family Information System (HSFIS).

Overall, sites were pleased with the system's reliability and straightforward procedures, and no one reported problems with passwords. Furthermore, most programs did not request technical assistance from the CBRS helpline. Staff members either did not encounter any problems or else consulted a supervisor familiar with the system.

However, staff members from 40 percent of programs called the helpline to report their inability to access the CBRS at certain times during the program year. One director was particularly frustrated, noting that the grantee's first-year growth report did not include data from one region because the data manager could not upload information into the system. Eventually, a temporary Web site was created for Migrant/Seasonal programs, but the director said it took three months to resolve. Representatives from Xtria, the CBRS contractor, explained to callers that this "lock down" was necessary to prepare the system for the next round of assessments at regional programs.

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

CHAPTER V

USING THE NRS FOR LOCAL PROGRAM IMPROVEMENT EFFORTS

A fundamental goal of the NRS has been to support local Head Start programs in their program improvement efforts. After each round of assessments, the Office of Head Start provides grantees and delegate agencies with summary reports that present program-level results of how children performed in all four skill areas covered by the NRS: (1) Understanding Spoken English/Understanding Spoken Spanish (the language screener), (2) Vocabulary, (3) Letter Recognition, and (4) Early Math Skills. Programs can then compare their average scores with the national averages for all Head Start programs, as well as with programs that are similar to their own (for example, other programs in Region 12). These average scores can be used in concert with other data sources to help guide programs in their quality improvement efforts.

In 2004, the Office of Head Start distributed the first NRS growth reports, covering the 2003-2004 program year, to Migrant/Seasonal programs. Unlike other NRS reports that document how children performed on the baseline assessments, the 2003-2004 growth report documents (1) progress made during the Head Start year among children who were assessed both in the beginning and toward the end of the program year, before children left for kindergarten, and (2) the growth point-in-time data, which include children who enrolled in Head Start after baseline assessments were completed. Accompanying materials offered guidelines and suggestions for reviewing the reports and using them for local program improvement efforts. Our spring and summer 2005 site visits afforded us the opportunity to discuss with local Migrant/Seasonal Head Start staff members their views on the usefulness of the report, how they have used the NRS results so far, and how they plan to use these reports in the future.

In this chapter, we discuss how local program staff members reacted to the 2003-2004 program-level growth reports sent out by Westat, the implementation contractor, including their views on how useful the report was for their program. Next, we describe how Migrant/Seasonal programs have used the NRS growth reports for program planning, what

changes they made to classroom practices, and how they intend to use data in the future. Finally, we describe their suggestions on how to improve the NRS reports.

REACTIONS OF LOCAL PROGRAM STAFF MEMBERS TO THE 2003-2004 GROWTH REPORT

Nearly all Migrant/Seasonal programs in our sample shared the NRS results with staff members and/or other stakeholders. In most cases, results were shared with senior staff members (for example, center directors and education coordinators), the Policy Council, and, when applicable, the agency's board of directors. Three programs shared the data with teachers at pre-service or staff meetings, and two programs also discussed NRS results at parent meetings.

Four programs found the growth reports to be in an easy-to-read format they could understand without difficulty; one lead trainer noted that the graphs helped her share results with the Policy Council. On the other hand, staff members from six programs had mixed reactions to the summary data, and some found the results difficult to interpret. One director contacted Westat because she did not understand how average Spanish comprehension scores dropped from baseline to growth. Some respondents acknowledged that while they were initially confused by the data, the supplemental materials and their coworkers helped them to decipher the results.

Overall, program staff members offered mixed views on how closely the NRS results mirrored their expectations of how much children should be learning over time. Staff members from three programs were pleased with the growth in scores, which on average reflected their expectations. Three different programs said scores were lower than expected, and speculated that this was due to the formal testing format that was unfamiliar to children. Four programs that compared the NRS report with local assessment outcomes found them to be comparable. However, two of them did not place much weight on the comparisons, since they felt that the methodologies used for local assessments—observations and teacher reports—are too different from the NRS.

HOW PROGRAMS USED THE GROWTH REPORT

The Office of Head Start provided some general guidelines and suggestions for how programs could apply and incorporate NRS results into local program planning in effective, appropriate ways. While 40 percent of programs indicated that their local assessments provide more relevant and useful information—in part because many children from migrant families were not eligible to take the NRS, having been enrolled less than 16 weeks—some staff members identified ways in which they have used the growth reports.

Three programs used the outcomes data to inform staff development. One group of education coordinators reviewed the scores to determine the skill areas in which children did not perform well. They then discussed how they could guide and support teachers to modify classroom practices and better prepare children for kindergarten. In addition, two directors met with senior staff members to discuss how to improve performance levels in

the classroom using the Head Start outcomes as a guide. After another agency saw lower math scores on the English assessment as compared with the Spanish one, it concluded that this was because teachers conducted math activities in Spanish. As such, the management team decided that teachers should also conduct math activities in English so that children can learn basic mathematical terms in preparation for taking the NRS.

FUTURE PLANS FOR USING NRS DATA

Sixty percent of Migrant/Seasonal programs in our sample indicated they were considering using or intended to use the NRS report in the future. Staff members tended to describe fairly general plans, such as using outcomes data for informing program planning decisions, setting long-term goals, identifying training needs, and determining which skill areas required more attention in the classroom. One program was considering using the data to inform annual staff performance reviews. In contrast, one director said that she has no intention of ever using the data, since she felt that the local assessment data they gather is more informative.

RECOMMENDATIONS FOR IMPROVING THE NRS REPORT

Several respondents offered recommendations to improve the NRS outcomes report. Of the seven programs that gave preferences on the level of data, five concluded that classroom- and center- level data would be most useful. Not only could outcomes data help teachers shape lesson plans, they would enable education supervisors to “tap into the expertise of better-performing centers and classrooms” and draw upon their practices to inform program-wide training and curriculum planning. Staff members from one program had mixed opinions: administrators said program-level data would help with overarching programmatic changes, but the lead trainer preferred center-level data, since the agency’s service area is quite large and managers oversee different regions. Because of their small enrollment, two delegate agencies said program-level data were sufficient.

Staff members from seven programs offered other suggestions for enhancing the way in which data are shared. Four programs expressed interest in receiving additional guidance on interpreting and using the data. Two agencies suggested that the Office of Head Start host a Web cast that focuses exclusively on the growth reports. One director added that without a demonstration on how to use the data, they are at a loss on how to proceed: “The NRS could be used as a tool for curriculum improvement . . . but the big question is, How [can programs] turn the testing into effective lesson plans?” Other suggestions included receiving support on how to share data with legislators; narrative descriptions to help programs present results to stakeholders; and a scripted letter, created by the Office of Head Start, that they could share with parents. Two programs wanted to receive the NRS report sooner, which would enable staff members to make timely, appropriate adjustments in the classroom to better prepare children for kindergarten.

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

CHAPTER VI

LOCAL PROGRAMS' VIEWS ON THE NRS

During site visit interviews, program staff members shared both their concerns about the NRS and their suggestions for improving it. Such feedback can facilitate the development of reporting formats and materials to support local program improvement efforts, help the Office of Head Start communicate the purpose of the NRS and the way results will be used at the national level, and inform ongoing system development activities. The information can also be reviewed and considered to determine whether the NRS is really suitable for these types of unique programs. This chapter describes these findings.

CONCERNS ABOUT THE NRS

For the most part, concerns clustered around five main topics: (1) implementation of the system in Migrant/Seasonal programs, (2) whether the data accurately reflect children's abilities and programs' performance, (3) the developmental appropriateness of the child assessment, (4) the amount of time and resources dedicated to the NRS, and (5) its purpose.

Whether the NRS Fits into the Unique Structure of Migrant/Seasonal Programs

Program staff members expressed serious reservations about whether the NRS is an appropriate tool for Migrant/Seasonal programs. At least five programs in our sample reported that the system is designed for regional programs that operate year-round or concurrently with the local public school system. Common concerns raised across programs included:

- Some respondents, including those that run both kinds of programs, felt that the Office of Head Start “ignores” Migrant/Seasonal programs and considers them an “afterthought.” As with the first Quality Assurance Study, programs questioned why Migrant/Seasonal programs have never been addressed during the Web casts.

- Operational schedules made NRS implementation challenging. Within a given “season,” centers were open anywhere from 4½ weeks to 7 months. For those programs that remained open for at least 16 weeks and could administer both baseline and growth assessments, staff members were frequently overwhelmed with completing two rounds of NRS assessments, other local assessments and screenings, classroom activities, and other basic services that Head Start provides to migrant families. They reported how difficult—if not impossible—it can be to schedule centralized trainings and assessments across centers within the same agency when those centers operate on different calendars or are hundreds of miles from each other. It was not uncommon for one region to conduct baseline assessments while another conducted growth assessments.
- Because they operate on compressed schedules, many programs conduct assessments within the first week to qualify as many children as possible for both assessments using the 16-week rule. Respondents from one program questioned how appropriate it is to administer a standardized test to 4- and 5-year olds who are still transitioning to a new environment; for some it is the first formal classroom or program setting they have ever been in. Ideally, children would be in Head Start for at least three or four weeks before taking the NRS, but the migrant calendar makes this very difficult.
- Several programs worried that migrant children may be “double-tested” during the baseline or growth assessments within a given calendar year. Double-testing is likely to be more prevalent for this population, whose families often travel multiple times during a growing season. As one director explained, it is common for families to leave suddenly because unexpected weather patterns forced them to move to a different harvest region. Some families move away and then return within a few months. A child could, for example, take the baseline assessment at one program, move to a different area and enroll in a program that opens later in the year, and take the baseline version again.

Apart from logistical and operational issues, a few programs questioned whether the NRS child assessment is appropriate for Migrant/Seasonal programs. Among these transient subgroups, staff members are frequently confronted by families with basic, immediate needs such as housing, health care, and nutrition. Addressing these priorities before families move on to another region often supersedes working on academic activities, such as teaching early math skills. One director remarked, “This approach to testing does not take into consideration what Head Start does for migrant communities.”

Whether NRS Results Accurately Reflect Program Effectiveness

Four programs questioned the validity of the NRS as a mechanism for assessing program performance. Some staff members questioned the validity of NRS outcomes since schedules are so compressed, and doubt that children could possibly demonstrate any noticeable growth after only a few months. Moreover, centers constantly enroll new children whose families move into the community, affecting what is covered in the classroom. One group of teachers described how they must return to speaking Spanish until

the new child learns enough English to begin “catching up” with the others. These interruptions slow the rate of English acquisition for the group and require that teachers repeat previously covered material. Staff members worry that a revolving enrollment might hinder growth in the NRS results, and that the Office of Head Start will not take this into consideration when evaluating the effectiveness of the Migrant/Seasonal programs.

Appropriateness of the NRS for Young Children

Six programs questioned the suitability of administering a standardized test to 4- and 5-year-olds. Several respondents said the tool is too formal or too long, or is presented in a “non-nurturing” format (i.e., following a rigid script). One director described the assessment as “demeaning, not respectful, and flat-out harmful.” Three programs worried that there is too much testing of young children who are already “over assessed,” that the NRS will evolve into an extension of No Child Left Behind, and that Head Start teachers will become more likely to “teach to the test.” One assessor noted, “Children are not getting enough time to play and are losing some of the socialization skills that organized play teaches.” Five programs reported that some math items, especially the coin, graph, and pie items, are too difficult for preschoolers. One lead trainer observed teachers incorporating fractions into lesson plans, which she considered to be a topic more appropriate for elementary school.

Time and Resources Dedicated to the NRS

Five programs faulted the amount of time and effort that the NRS takes away from classroom instruction and other activities, noting that it interferes with Head Start’s core services. The staff members felt that time spent with migrant children is especially important, since families may not be enrolled in a particular agency for very long.

Five programs also mentioned the cost burden associated with the assessment process. One agency director noted only two children qualified to take both the baseline and the growth assessments in one state and questioned whether the minimal amount of data justify the funding allocated for this initiative. Another program described the NRS as “a big waste” of resources that has been implemented “largely for political reasons” and is “being forced on preschool programs because of No Child Left Behind.” Respondents commonly referred to the NRS as an “unfunded mandate.”

Purpose of the NRS

After about two years of implementation, staff members from half the Migrant/Seasonal programs in our sample did not have a clear sense of the purpose of the NRS. They expressed skepticism and uncertainty of its “real” intent, especially since, in their opinion, local assessments are more useful for program planning. Two programs reported a fear that the federal government will use scores to justify funding cuts for Head Start, and three programs worry that the NRS is shifting the emphasis toward a more academic program while deemphasizing the socialization skills and support services provided to children and families. Two other programs expressed hope that the assessment would not become the primary means of shaping the Head Start curriculum. They were concerned that

this could encourage teachers to focus on raising test scores through drills that mirror the assessment's content.

SUGGESTIONS FOR IMPROVING THE NRS

Based on their experiences, local staff members from the 10 programs made recommendations about the NRS implementation process, as well as the content of the child assessment. Suggestions remained fairly consistent across both rounds of site visits.

Ways to Improve the Test Administration Process

Programs proposed several ways to make the test administration process easier for local staff members. Three programs would like the Scantron score sheets to reflect the easel's layout so as to facilitate marking children's answers (for example, create four alphabet grids on which assessors can record which letters were correctly named). One director would like assessors to receive clearer guidance on how they can explain to children taking the Spanish-language assessment that they can answer in English during *Nombre de las Letras*. One program suggested that the answer sheet include a means of recording when a special-needs child is unresponsive during *Tío Simón/Simon Says* so that assessors can stop after this section. Otherwise, it is a very uncomfortable environment for the child and assessor.

Ways to Make the NRS More Suitable for Migrant/Seasonal Programs

Two programs would like the Office of Head Start to take their operational timelines into consideration. For example, receiving materials as early as possible and granting programs the flexibility to administer assessments earlier—especially since some centers within an agency open much earlier than others—would enable them to plan accordingly and be less rushed to meet deadlines. In addition, the CBRS should be a seamless system that can identify which children have already taken the assessment so that they are not double-tested when the family moves and enrolls in a different program within the same year.

Ways to Improve Communication between the Office of Head Start and Local Programs

Three programs wanted to receive specific information on the purpose of the NRS, how its data will be used, and its concrete benefits. One director wanted specific guidance on how to develop more effective lesson plans while taking the NRS results into consideration, and a lead trainer from another program suggested that the Office of Head Start assign technical assistance advisors who would work exclusively with Migrant/Seasonal programs on the NRS. Staff members would also like to receive information on the early research conducted to determine the feasibility of the NRS in Migrant and Seasonal Head Start programs.

Ways to Improve the Instrument

Four programs thought that the NRS takes too long to administer. Staff members in these programs felt that children often become bored, uncooperative, restless, or unfocused. Shortening the assessment, in particular the vocabulary section, would reduce the likelihood that children will start to point randomly to quadrants just to finish the task. Some local staff members offered recommendations to shorten the scripted instructions. Four programs said that the *Nombre de las Letras* section would unfold more smoothly and quickly if assessors could point to each letter and ask, “What letter is this?” In its current format, children often nod when asked, “Do you know any more?” and sit silently until assessors prompt them to name the next letter that they know, or else they point to letters without naming them.

During the site visits, staff members considered whether additional domains should be incorporated into the NRS (respondents from a given program did not always have the same opinion). Respondents from seven programs identified additional domains to be included, but only if the child assessment would not exceed 15 to 20 minutes. Specific domains suggested were social and emotional development, approaches to learning, creativity, art, and science. Staff members from three programs recommended adding domains without being concerned about lengthening the assessment, including physical development, health, and creative arts. One focus group wanted the vocabulary section expanded to gauge emotional development by asking children to point to pictures that represent happiness, sadness, and other emotions. On the other hand, staff members from 60 percent of programs did not think that any domains should be added, because doing so would lengthen the assessment and duplicate efforts of collecting mandated outcomes using local instruments. One focus group asserted that the NRS should examine only social and emotional developmental outcomes.

Staff members also made some section-specific recommendations:

- Make *Básicos de Matemática*/Early Math Skills sections more age-appropriate by eliminating items that are too advanced, including the pie question, the money item, and the graphs (six programs)
- Add color to the TVIP/PPVT to help keep children’s attention and interest (three programs)
- Change the nickels to pennies for the money item because it is an easier addition task (one program)

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

CHAPTER VII

SUMMARY OF FINDINGS AND IMPLICATIONS FOR SYSTEM IMPROVEMENT

If the NRS is to be successful, local Migrant/Seasonal Head Start programs must be able to implement it in a way that produces reliable and valid information. It is also important for local program improvement efforts that programs accept the NRS as a valid and reliable assessment, and understand how it can lead to program improvement. Based on findings from site visits to Migrant/Seasonal programs during the first two years of NRS implementation, we suggest some implications for system improvement for the NRS. While these findings were not derived from a nationally representative sample of programs (unlike the sample of regional Head Start programs participating in the Quality Assurance Study), the information may be useful to the Office of Head Start as it considers ways to build upon the improvements that have been made to the NRS since its inception in fall 2003. This chapter summarizes our finding from the 2005 site visits, including a discussion of issues that are unique to Migrant/Seasonal programs, and then presents implications for system improvement.

OVERVIEW OF FINDINGS FROM THE SITE VISITS

The Migrant/Seasonal Head Start programs in our sample from 2005 had some experiences in NRS implementation similar to those of the regional programs in the second year of the Quality Assurance Study, as well as to those of the migrant programs from the first Quality Assurance Study. Staff members in migrant and regional programs identified common strengths and weaknesses in the training and assessment materials and generally found the CBRS to be a user-friendly system. In addition, staff members in both presented some similar suggestions (for example, allowing more flexibility on word usage to account for variation according to ethnicity and geography) and concerns (for example, whether the NRS accurately portrays program performance).

Assessors from the Migrant/Seasonal programs also achieved high levels of quality as established by the certification standards and demonstrated similar administration and scoring errors. Reliability of the assessment scale scores was high, which indicates that scoring errors on individual items do not substantially affect the total score for each scale.

However, we also identified some sections that were difficult for staff members to administer or score, and some sections in which they could use more guidance and training. Primary challenges included:

- Omitting or misusing hand gestures for various items
- Not understanding when non-verbal or English responses can and cannot be accepted as correct answers on the Spanish-language version of the NRS
- Appropriately probing for answers, especially on the *Nombre de las Letras* section
- Knowing when it is acceptable to encourage and prompt a child to guess appropriately, without giving non-neutral encouragement or coaching

Two implementation themes that emerged during the first year of the NRS (Paulsell et al. 2004) were observed again during the second year of site visits to Migrant/Seasonal programs. First, program staff members offered insights regarding the Spanish-language version of the assessment, including a preference (1) to accept colloquial vocabulary to accommodate the diversity of the Spanish language; and (2) to modify the instructions as necessary if children are unfamiliar with the word usage on the instrument. Second, the operational schedules of Migrant/Seasonal programs bring an important set of challenges. Abbreviated calendars, ongoing enrollment, and the transitory patterns of families increase the likelihood that many children will not be present for both baseline and growth assessments. Program staff members worried how these factors could affect a program's NRS outcomes. Moreover, the compressed timeline makes it especially burdensome to implement two rounds of local training and assessments. Staff members also noted that they routinely address the critical needs of their transient families—nutrition, housing, and health care—which sometimes supersede other non-emergency Head Start services. The schedules of Migrant/Seasonal programs and the characteristics of the families they serve require careful consideration from federal officials and experts about how best to implement the NRS in these environments.

IMPLICATIONS FOR SYSTEM IMPROVEMENT

Based on our analysis of information collected during the 10 site visits (and reinforced with data gathered during the first year's visits), we suggest the following list of implications for system improvement if the Office of Head Start plans to continue to implement the NRS in Migrant/Seasonal programs, some of which resonate with these programs in particular:

-
- ***Providing local programs more precise information on the purpose of the NRS and how the data will be used by the Office of Head Start.*** Some programs remain skeptical about how the Office of Head Start will use the results. They wonder whether the system is ultimately intended to become an extension of No Child Left Behind that would shift focus away from critical services for children and families toward a “high-stakes” assessment that links test outcomes to program performance. Federal officials may be able to address these concerns by providing concrete information on how the data will be interpreted and used.
 - ***Facilitating communication with families.*** To help programs communicate more easily and consistently with families about the NRS, create a scripted, easy-to-understand letter or document that staff members could read to parents and that would address common questions, such as “What is this assessment for?” and “What is done with the data?”
 - ***Providing more training for local staff members to enhance the instrument’s reliability.*** Some site visitors observed certain assessment errors that were not present during site visits to the regional programs. Specifically, a few assessors did not understand when it was appropriate to stop the assessment or continue with the TVIP, depending on the number of correct responses given in the first two sections. Similarly, several assessors forgot to complete the language screener before continuing with the assessment. Furthermore, mispronunciation altered the meaning of items in isolated cases, which could affect a child’s ability to give a correct response. Some assessors did not know when it was appropriate to accept a non-verbal answer or when they could accept an English response on the Spanish assessment, and some could not prompt appropriately on certain test items.
 - ***Consider granting assessors some latitude in accepting different answers in Spanish, aside from those included in the Assessor’s Guide.*** While it is difficult to design one universal instrument for all Spanish-speaking children, there are many linguistic differences attributable to ethnicity and region. Programs suggested that giving assessors more latitude would allow them to be able to accept synonyms that they judge to be accurate (for example, *trinche* instead of *tenedor* for “knife”). However, the NRS does not allow this latitude for assessors in the English assessments: although some additional answers have been allowed, they must be listed on the assessor’s form to be acceptable. It is important to ensure that the addition of acceptable answers should be the same for all children taking the assessment, and not left to the judgment of individual assessors. In order to help resolve this issue, a further study of the test administration with Spanish-speaking children could be carried out.

- ***Expanding access to the CBRS.*** To allow program staff members to complete data entry tasks in a timely manner, consider modifying the CBRS to create a separate system that would enable Migrant/Seasonal programs to enter the system even when regional programs are barred access to the main system. In addition, to avoid double-testing a child coming from another Head Start program during the program “year,” programs should be able to verify whether the child has already taken the baseline or growth assessment.
- ***Providing programs with more guidance on interpreting and applying NRS data.*** Programs have requested more information on how the Office of Head Start plans to use the NRS results nationally. In addition, some programs want more information on the instrument’s validity and reliability. For example, programs would like to know how scores of children with disabilities affect the program’s overall scores (and how staff members should interpret and incorporate these two sets of averages into program planning for two groups of children). Programs also mentioned wanting more information about how scores are affected by the fact that children can “get credit” for answering in English or Spanish for certain items on the Spanish version of the child assessment. Again, analyses of NRS child cognitive assessment data from Migrant/Seasonal programs nationally might be helpful in responding to these questions.
- ***Highlighting communications with Migrant/Seasonal programs.*** During site visits for the first and second year of the Quality Assurance Study, several respondents expressed feeling “ignored” by the Office of Head Start and sensed that the NRS is geared toward regional Head Start programs. Ensuring that Migrant/Seasonal programs receive their materials in a timely manner, and including them in the Web casts (or perhaps having a dedicated helpline for migrant programs) would help demonstrate that the Office of Head Start recognizes the unique challenges these programs face.

REFERENCES

- Migrant and Seasonal Head Start Quality Improvement Center. Available at [www.mhsqic.org]. Accessed November 14, 2005.
- National Migrant and Seasonal Head Start Association. Available at [www.nmshsa.org]. Accessed November 14, 2005.
- Paulsell, Diane, Anne Gordon, Renée Nogales, Patricia Del Grosso, Susan Sprachman, and Louisa Tarullo. "Implementation of the Head Start National Reporting System: Spring 2005 Update." Princeton, NJ: Mathematica Policy Research, Inc., January 2006.
- Paulsell, Diane, Anne Gordon, Renée Nogales, Susan Sprachman, Cheri Vogel, Barbara Schiff, and Louisa Tarullo. "Implementation of the Head Start National Reporting System" Fall 2004 Update." Princeton, NJ: Mathematica Policy Research, Inc., January 2005.
- Paulsell, Diane, Linda Rosenberg, Renée Nogales, Charles Nagatoshi, Susan Sprachman, Louisa Tarullo, and John Love. "Meeting the Challenge: How Head Start Programs Implemented the National Reporting System." Princeton, NJ: Mathematica Policy Research, Inc., December 2004.

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

APPENDIX

**ADMINISTRATION AND SCORING ERRORS
OF OBSERVED BASELINE AND GROWTH
NRS CHILD ASSESSMENTS (SPANISH-
LANGUAGE VERSION) IN
MIGRANT/SEASONAL HEAD START
PROGRAMS**

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

Table A: Percentage of Observed Baseline Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”	
Warm Up							
Introducción: Ahora, tengo algunos dibujos para mostrarte y algunas cosas...	14	NA	NA	NA	NA	NA	NA
Vamos a usar esta hoja de papel en un minuto.	17	NA	NA	NA	NA	NA	NA
Tío Simón							
Introducción a Tío Simón: Vamos a jugar Tío Simón Dice. Tío Simón es...	0	NA	NA	NA	NA	NA	NA
Practice A: Tío Simón dice mira para arriba.	2	5	0	NA	NA	NA	NA
Practice B: Tío Simón dice mira para abajo.	2	31	2	NA	NA	NA	NA
A1: Tío Simón dice tócate una oreja.	0	0	0	NA	NA	NA	0
A2: Tío Simón dice señale la puerta.	2	2	5	NA	NA	NA	0
A3: Tío Simón dice levanta un pie.	2	5	2	NA	NA	NA	5
A4: Tío Simón dice abre una mano.	5	0	0	NA	NA	NA	7
A5: Tío Simón dice coge el papel.	2	12	2	NA	NA	NA	5
A6: Tío Simón dice da vuelta el papel.	0	2	2	NA	NA	NA	7
A7: Tío Simón dice pon una mano sobre la otra.	2	0	0	NA	NA	NA	2
A8: Tío Simón dice golpea sobre la mesa.	5	0	0	NA	NA	NA	5
A9: Tío Simón dice señale el medio del papel.	2	0	2	NA	NA	NA	10
A10: Tío Simón dice junta los pies.	0	0	0	NA	NA	NA	10
Transition to next section.	NA	NA	2	NA	NA	NA	NA
Exposición de Arte							
Practice A: A ver si me puedes decir cómo se llaman algunas cosas. ¿Qué es esto? (perro)	0	0	0	0	0	NA	NA
IF INCORRECT OR NO RESPONSE REPEAT: “Este es un perro.”	2	0	0	0	0	NA	NA
Practice B: ¿Qué es esto? (mesa)	0	0	0	0	2	NA	NA
IF INCORRECT OR NO RESPONSE REPEAT: “Esta es una mesa.”	2	0	0	0	2	NA	NA
B1: ¿Qué es esto? (globo, bomba)	0	0	0	0	2	NA	2
B2: ¿Qué es esto? (gato)	2	0	0	0	2	NA	2
B3: ¿Qué es esto? (vaso)	0	0	0	0	2	NA	0
B4: ¿Qué es esto? (avión, aeroplano)	2	0	0	0	2	NA	2
B5: ¿Qué es esto? (peine, peinilla)	0	0	0	0	2	NA	7
B6: ¿Qué es esto? (pelota, bola, balón)	2	0	0	0	2	NA	2
B7: ¿Qué es esto? (falda, vestido, pollera)	0	0	0	0	2	NA	7
B8: ¿Qué es esto? (tenedor, cubierto)	2	0	0	0	2	NA	2
B9: ¿Qué es esto? (escoba)	0	0	0	0	2	NA	0

Table A: Percentage of Observed Baseline Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”	
B10: ¿Qué es esto? (taza)	0	0	0	0	0	NA	2
Transition to next section.	NA	NA	0	NA	NA		NA
TVIP (Adapted)							
Introducción: Ahora, quiero que mires unos dibujos conmigo...	2	0	0	0	NA	5	NA
Practice A: ¿Ves todos los dibujos en esta página?... Pon el dedo sobre “tenedor.”	0	2	0	0	2	10	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “tenedor.”	2	0	0	0	5	0	NA
Practice B: ¡Bien! Ahora hagamos otro. Pono el dedo sobre “perro.”	2	5	0	0	NA	12	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “perro.”	0	0	0	0	NA	0	NA
Practice C: Ahora mira todos los dibujos en esta página... Indica “hombre.”	2	2	0	0	5	2	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “hombre.”	10	0	0	2	NA	0	NA
Practice D: ¡Bien! Ahora hagamos otro. Indica “peine.”	2	0	0	2	NA	2	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “peine.” ¡Bien!	5	0	0	0	NA	0	NA
Transition to next section.	NA	NA	0	NA	NA	NA	NA
C1 Señala barco/bote.	0	0	0	2	NA	0	7
C2 Señala leer.	0	0	0	0	NA	0	0
C3 Señala accidente/choque	0	0	0	0	NA	0	2
C4 Señala roto.	5	5	0	0	NA	0	0
C5 Señala vaca.	0	0	0	0	NA	0	0
C6 Señala lámpara.	0	0	0	0	NA	0	5
C7 Señala cuello.	0	0	0	0	NA	0	5
C8 Señala abeja.	0	0	0	0	NA	0	0
C9 Señala pintor.	0	0	0	0	NA	0	0
C10 ¡Estás señalando muy bien! Señala grupo.	0	2	0	0	NA	0	0
C11 Señala flecha.	0	0	0	0	NA	0	5
C12 Señala vela/candela.	2	0	0	0	NA	0	2
C13 Señala trompeta.	0	0	0	0	NA	0	2
C14 Señala codo.	2	0	2	0	NA	0	2
C15 Señala dentista.	0	0	0	0	NA	0	2
C16 Señala joyas.	0	0	2	0	NA	0	0
C17 Señala bebida.	0	0	0	0	NA	0	2

Table A: Percentage of Observed Baseline Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”	
C18 Señala pelar.	0	0	0	0	NA	0	2
C19 Señala águila.	2	0	0	0	NA	0	5
C20 Señala ceremonia.	0	0	0	0	NA	0	2
C21 Señala aislamiento.	7	0	0	0	NA	2	0
C22 Señala premiar.	0	0	0	0	NA	0	5
C23 Señala medir.	0	0	0	0	NA	0	5
C24 Señala tronco.	0	0	0	0	NA	0	5
Transition to next section.	NA	NA	0	NA	NA	NA	NA

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gestures	Gestures Omitted	Not Slowing Child Down	
Nombre de las Letras							
Introducción Panel 1: Aquí hay algunas letras del alfabeto. Señale a todas...	0	NA	0	NA	NA	NA	NA
Identificación de letras.	5	0	2	NA	NA	0	NA
A	NA	NA	NA	NA	NA	NA	0
O	NA	NA	NA	NA	NA	NA	0
B	NA	NA	NA	NA	NA	NA	0
L	NA	NA	NA	NA	NA	NA	0
E	NA	NA	NA	NA	NA	NA	0
U	NA	NA	NA	NA	NA	NA	0
X	NA	NA	NA	NA	NA	NA	3
I	NA	NA	NA	NA	NA	NA	0
Introducción Panel 2: Aquí hay algunas letras más del alfabeto. ¿Conoces algunas de éstas?	0	NA	0	NA	NA	NA	NA
Identificación de letras.	0	0	0	NA	NA	0	NA
Ñ	NA	NA	NA	NA	NA	NA	0
P	NA	NA	NA	NA	NA	NA	5
J	NA	NA	NA	NA	NA	NA	3
CH	NA	NA	NA	NA	NA	NA	0
S	NA	NA	NA	NA	NA	NA	0
M	NA	NA	NA	NA	NA	NA	0
F	NA	NA	NA	NA	NA	NA	0

Table A: Percentage of Observed Baseline Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gestures	Gestures Omitted	Not Slowing Child Down	Scoring Errors
Introducción Panel 3: Aquí hay algunas letras más del alfabeto. ¿Conoces...	0	NA	0	NA	NA	NA	NA
Identificación de letras.	0	0	0	NA	NA	0	NA

Item	Administration Errors						
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gestures	Gestures Omitted	Not Slowing Child Down	Scoring Errors
D	NA	NA	NA	NA	NA	NA	5
C	NA	NA	NA	NA	NA	NA	11
Z	NA	NA	NA	NA	NA	NA	5
LL	NA	NA	NA	NA	NA	NA	3
T	NA	NA	NA	NA	NA	NA	0
H	NA	NA	NA	NA	NA	NA	3
R	NA	NA	NA	NA	NA	NA	0
Y	NA	NA	NA	NA	NA	NA	3
Introducción Panel 4: Aquí hay algunas letras más del alfabeto. ¿Conoces...	0	NA	0	NA	NA	NA	NA
Identificación de letras.	0	0	0	NA	NA	0	NA
G	NA	NA	NA	NA	NA	NA	3
RR	NA	NA	NA	NA	NA	NA	11
V	NA	NA	NA	NA	NA	NA	3
K	NA	NA	NA	NA	NA	NA	5
Q	NA	NA	NA	NA	NA	NA	0
N	NA	NA	NA	NA	NA	NA	0
W	NA	NA	NA	NA	NA	NA	0
Transition to next section.	NA	NA	0	NA	NA	NA	NA

Item	Administration Errors						
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”	Scoring Errors
Básicos de Matemática							
Introduction: ¡Tú eres un buen ayudante! Ahora te voy a hacer algunas...	0	NA	0	NA	NA	NA	NA
E1: ¿Cuántos lápices hay en esta página?	0	0	0	10	NA	NA	0

Table A: Percentage of Observed Baseline Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”	
E2: ¿Cuántas estrellas hay en esta página?	0	0	0	0	NA	NA	3
E3: Aquí hay algunos nidos con huevos. Señala el nido con 2 huevos.	5	0	2	17	NA	NA	3
E4: Ahora señala el nido con 3 huevos.	2	0	0	0	NA	NA	0
E5: ¿Cuántas huevos en total hay aquí?	10	5	0	2	5	NA	3
E6: ¿Qué es esto? (8)	7	2	2	17	NA	NA	3
E7: ¿Qué es esto? (5)	5	0	0	14	NA	NA	3
E8: ¿Qué es esto? (6)	5	2	0	12	NA	NA	0
E9: Mira estas formas. Una de ellas es un círculo. Señala el círculo.	5	0	0	2	0	NA	3
E10: Mira estas formas. Una de ellas es un cuadrado. Señala el cuadrado.	0	0	2	2	0	NA	0
E11: Mira estas formas. Una de ellas es un triángulo. Señala el triángulo.	10	0	0	0	0	NA	3
E12: ¡Eres bueno(a) para señalar! ¿Cuál de estos crayones es más corto que el pincel?	24	2	0	2	0	NA	0
E13: ¿Cuál de estas frutas es más grande que la naranja?	0	0	0	2	2	NA	0
E14: Bobby tiene tres pecesitos. Su amigo le acaba de dar uno más. ¿Cuántos pecesitos tiene Bobby ahora?	2	5	7	5	0	NA	0
E15: Imagínate que estos son tus libros. ¿Si tú les das a un amigo dos de estos libros, cuántos libros te quedarían?	0	2	0	17	2	NA	3
E16: ¿Cuando tres de estas burbujas revienten, cuántas burbujas quedarían?	2	0	2	2	0	NA	0
E17: Señala el número que marca cuántas pulgadas de alto mide el osito.	2	0	0	21	2	NA	3
E17a: ¡Estás trabajando mucho, y te lo agradezco! ¡Gracias!	0	0	0	NA	NA	NA	NA
E18: Esta gráfica muestra cuántos niños de...¿Cuántos niños tienen perros?	2	12	0	48	12	NA	3
E19: ¿Cuántos tienen gatos?	0	0	0	19	29	NA	8
E20: Cuando yo diga “empieza,” quiero que empieces a contrar....	2	14	0	5	2	NA	26

SOURCE: Observations of 2005 NRS baseline child assessments; N=42.

^aAssessment errors are based on unweighted estimates.

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

Table B: Percentage of Observed Growth Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”	
Warm-up							
Introducción: Ahora, tengo algunos dibujos para mostrarte y algunas cosas...	13	NA	NA	NA	NA	NA	NA
Vamos a usar esta hoja de papel en un minuto.	24	NA	NA	NA	NA	NA	NA
Tío Simón							
Introducción a Tío Simón: Vamos a jugar Tío Simón Dice. Tío Simón es...	2	NA	NA	NA	NA	NA	NA
Practice A: Tío Simón dice mira para arriba.	0	2	0	NA	NA	NA	NA
Practice B: Tío Simón dice mira para abajo.	0	0	0	NA	NA	NA	NA
A1: Tío Simón dice tócate una oreja.	0	2	0	NA	NA	NA	4
A2: Tío Simón dice señale la puerta.	9	9	0	NA	NA	NA	4
A3: Tío Simón dice levanta un pie.	2	0	0	NA	NA	NA	4
A4: Tío Simón dice abre una mano.	0	0	0	NA	NA	NA	13
A5: Tío Simón dice coge el papel.	0	9	0	NA	NA	NA	13
A6: Tío Simón dice da vuelta el papel.	11	7	0	NA	NA	NA	7
A7: Tío Simón dice pon una mano sobre la otra.	0	0	0	NA	NA	NA	11
A8: Tío Simón dice golpea sobre la mesa.	0	0	0	NA	NA	NA	7
A9: Tío Simón dice señale el medio del papel.	4	9	0	NA	NA	NA	7
A10: Tío Simón dice junta los pies.	0	0	0	NA	NA	NA	16
Transition to next section.	NA	NA	0	NA	NA	NA	NA
Exposición de Arte							
Practice A: A ver si me puedes decir cómo se llaman algunas cosas. ¿Qué es esto? (perro, chucho)	17	0	0	0	0	NA	NA
IF INCORRECT OR NO RESPONSE REPEAT: “Este es un perro.”	0	0	0	0	0	NA	NA
Practice B: ¿Qué es esto? (mesa)	0	0	2	0	0	NA	NA
IF INCORRECT OR NO RESPONSE REPEAT: “Esta es una mesa.”	0	0	0	0	0	NA	NA
B1: ¿Qué es esto? (globo, bomba)	7	2	0	0	0	NA	7
B2: ¿Qué es esto? (gato)	0	0	0	0	0	NA	4
B3: ¿Qué es esto? (vaso)	0	0	0	0	0	NA	2
B4: ¿Qué es esto? (avión, aeroplano)	0	0	0	0	0	NA	2
B5: ¿Qué es esto? (peine, peinilla)	0	0	0	0	0	NA	7
B6: ¿Qué es esto? (pelota, bola, balón)	0	0	0	0	0	NA	0
B7: ¿Qué es esto? (falda, vestido, pollera)	0	0	0	0	0	NA	2
B8: ¿Qué es esto? (tenedor, cubierto)	0	0	0	0	0	NA	7
B9: ¿Qué es esto? (escoba)	0	0	0	0	0	NA	2

Table B: Percentage of Observed Growth Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”	
B10: ¿Qué es esto? (taza)	2	2	4	0	0	NA	2
Transition to next section.	NA	NA	0	NA	NA		NA
TVIP (Adapted)							
Introducción: Ahora, quiero que mires unos dibujos conmigo...	0	0	0	0	NA	0	NA
Practice A: ¿Ves todos los dibujos en esta página?... Pon el dedo sobre “tenedor.”	0	0	0	0	0	0	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “tenedor.”	0	0	0	0	0	0	NA
Practice B: ¡Bien! Ahora hagamos otro. Pono el dedo sobre “perro.”	0	0	0	0	NA	0	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “perro.”	0	0	0	0	NA	0	NA
Practice C: Ahora mira todos los dibujos en esta página... Indica “hombre.”	0	0	0	0	0	0	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “hombre.”	0	0	0	0	NA	0	NA
Practice D: ¡Bien! Ahora hagamos otro. Indica “peine.”	0	0	0	0	NA	0	NA
IF POINTED INCORRECTLY... Tú tratastes, pero... Pon el dedo sobre “peine.” ¡Bien!	0	0	0	0	NA	0	NA
Transition to next section.	NA	NA		NA	NA	NA	NA
C1 Señala lámpara.	0	0	0	0	NA	0	2
C2 Señala vaca.	0	0	0	0	NA	0	0
C3 Señala ballena.	0	0	0	0	NA	0	4
C4 Señala hora.	0	0	0	0	NA	0	2
C5 Señala vela/candela.	0	0	0	0	NA	0	4
C6 Señala flecha.	0	2	0	0	NA	0	2
C7 Señala jaula.	2	0	0	0	NA	0	2
C8 Señala mueble.	0	0	0	0	NA	0	0
C9 Señala sobre.	0	0	0	2	NA	0	2
C10 ¡Estás señalando muy bien! Señala romper.	0	0	0	2	NA	0	7
C11 Señala canguro.	0	0	0	0	NA	0	4
C12 Señala vacío.	0	0	0	2	NA	0	2
C13 Señala culebra.	0	0	0	2	NA	0	4
C14 Señala trompeta.	0	0	0	2	NA	0	2
C15 Señala líquido.	0	0	0	0	NA	0	2
C16 Señala dirigir.	2	0	0	0	NA	0	2
C17 Señala ambulancia.	0	0	0	0	NA	0	0

Table B: Percentage of Observed Growth Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors							Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Pronunciation or Using “a” or “the”		
C18 Señala acariciar.	0	0	0	0	NA	0	0	
C19 Señala mecánico.	0	2	0	2	NA	0	0	
C20 Señala río.	0	0	0	0	NA	0	2	
C21 Señala barco/bote.	0	0	0	0	NA	0	4	
C22 Señala humano.	0	0	0	2	NA	0	2	
C23 Señala medico.	0	0	0	0	NA	0	2	
C24 Señala hombro.	4	0	0	2	NA	0	0	
Transition to next section.	NA	NA	0	NA	NA	NA	NA	

Item	Administration Errors							Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gestures	Gestures Omitted	Not Slowing Child Down		
Nombre de las Letras								
Introducción Panel 1: Aquí hay algunas letras del alfabeto. Señale a todas...	0	NA	0	NA	NA	NA	NA	
Identificación de letras.	2	9	0	NA	NA	0	NA	
A	NA	NA	NA	NA	NA	NA	7	
O	NA	NA	NA	NA	NA	NA	2	
B	NA	NA	NA	NA	NA	NA	7	
L	NA	NA	NA	NA	NA	NA	2	
E	NA	NA	NA	NA	NA	NA	7	
U	NA	NA	NA	NA	NA	NA	2	
X	NA	NA	NA	NA	NA	NA	7	
I	NA	NA	NA	NA	NA	NA	5	
Introducción Panel 2: Aquí hay algunas letras más del alfabeto. ¿Conoces algunas de éstas?	0	NA	0	NA	NA	NA	NA	
Identificación de letras.	0	2	0	NA	NA	0	NA	
Ñ	NA	NA	NA	NA	NA	NA	5	
P	NA	NA	NA	NA	NA	NA	11	
J	NA	NA	NA	NA	NA	NA	2	
CH	NA	NA	NA	NA	NA	NA	12	
S	NA	NA	NA	NA	NA	NA	0	
M	NA	NA	NA	NA	NA	NA	12	
F	NA	NA	NA	NA	NA	NA	9	

Table B: Percentage of Observed Growth Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

	Administration Errors						
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gestures	Gestures Omitted	Not Slowing Child Down	Scoring Errors
Introducción Panel 3: Aquí hay algunas letras más del alfabeto. ¿Conoces...	2	NA	0	NA	NA	NA	NA
Identificación de letras.	0	2	2	NA	NA	0	NA
	Administration Errors						
Item	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gestures	Gestures Omitted	Not Slowing Child Down	Scoring Errors
D	NA	NA	NA	NA	NA	NA	5
C	NA	NA	NA	NA	NA	NA	2
Z	NA	NA	NA	NA	NA	NA	2
LL	NA	NA	NA	NA	NA	NA	5
T	NA	NA	NA	NA	NA	NA	2
H	NA	NA	NA	NA	NA	NA	2
R	NA	NA	NA	NA	NA	NA	5
Y	NA	NA	NA	NA	NA	NA	7
Introducción Panel 4: Aquí hay algunas letras más del alfabeto. ¿Conoces...	0	NA	0	NA	NA	NA	NA
Identificación de letras.	0	2	0	NA	NA	0	NA
G	NA	NA	NA	NA	NA	NA	2
RR	NA	NA	NA	NA	NA	NA	2
V	NA	NA	NA	NA	NA	NA	14
K	NA	NA	NA	NA	NA	NA	0
Q	NA	NA	NA	NA	NA	NA	2
N	NA	NA	NA	NA	NA	NA	7
W	NA	NA	NA	NA	NA	NA	9
Transition to next section.	NA	NA	0	NA	NA	NA	NA

Table B: Percentage of Observed Growth Assessments with Errors in Administration and Scoring, by Assessment Item (Migrant and Seasonal Programs)^a

Item	Administration Errors						Scoring Errors
	Straying from Script	Coaching	Non-neutral Encouragement	Incorrect Gesture	Gesture Omitted	Did not Allow Non-verbal Response	
Básicos de Matemática							
Introduction: ¡Tú eres un buen ayudante! Ahora te voy a hacer algunas...	0	NA	0	NA	NA	NA	NA
E1: ¿Cuántos carritos hay aquí?	0	0	0	0	NA	0	2
E2: ¿Cuántas estrellas hay ahí?	7	0	0	0	NA	0	2
E3: Aquí hay algunos dibujos de uvas. Señala el dibujo con 3 uvas.	2	4	0	2	NA	NA	13
E4: Ahora señala el dibujo con 2 uvas.	0	2	0	0	NA	NA	2
E5: ¿Cuántas uvas en total hay ahí?	7	11	0	2	0	0	2
E6: ¿Qué es esto? (4)	0	2	0	0	NA	0	7
E7: ¿Qué es esto? (7)	0	2	0	0	NA	0	4
E8: ¿Qué es esto? (9)	0	0	0	0	NA	NA	2
E9: Mira estas formas. Una de ellas es un cuadrado. Señala el cuadrado.	2	0	0	0	0	NA	2
E10: Mira estas formas. Una de ellas es un círculo. Señala el círculo.	7	0	0	0	0	NA	2
E11: Mira estas formas. Una de ellas es un triángulo. Señala el triángulo.	4	0	0	0	0	NA	4
E12: La mamá de José cortó un pastel de manzanas fresco en ocho partes. Ellas le dió una parte a Sara., y una parte a José. ¿Cuál de estos dibujos...	9	4	0	4	9	NA	2
E13: Señala la fruta más grande.	0	0	0	0	0	0	0
E14: Ahora señala la fruta más pequeña.	4	0	0	4	0	0	0
E15: Imagínate que estos son tus bloques. ¿Si tú les das a un amigo dos de estos bloques, cuántos bloques te quedarían?	7	0	0	0	0	0	2
E16: Si tres de estas ranas se alejaron saltando, ¿cuántas ranas quedarían?	13	0	0	2	0	0	0
E17: Señala el número que marca cuántas pulgadas de alto mide el osito.	0	2	0	0	0	NA	0
E17a: ¡Estás trabajando mucho, y te lo agradezco! ¡Gracias!	0	0	0	NA	NA	NA	NA
E18: Esta gráfica muestra...¿Señala quién usó la computadora más veces?	2	0	0	13	0	0	2
E19: ¿Cuántos veces usó la computadora Billy?	2	13	0	2	2	0	0
E20: Cuando yo diga “empieza,” quiero que empieces a contrar....	7	4	0	0	0	NA	24

SOURCE: Observations of 2005 NRS growth child assessments; N=46.

^aAssessment errors are based on unweighted estimates.