

Contract No.: 233-02-0086 (03)
MPR Reference No.: 6094-710

MATHEMATICA
Policy Research, Inc.

**Implementation of the
Head Start National
Reporting System:
Spring 2006**

Executive Summary

February 2008

*Cheri A. Vogel
Renée Nogales
Nikki Aikens
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

Project Officer:

Jennifer Brooks

Submitted by:

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

Project Director:

Louisa Tarullo

ACKNOWLEDGMENTS

We would like to thank the many people who contributed to the successful completion of this study and report. Jennifer Brooks, Administration for Children and Families, provided invaluable leadership and vision to shape this effort. However, this report would have been impossible without the support and participation of the 35 Head Start program directors and staff who so generously spent their time to talk with us.

Numerous staff at MPR also played critical roles in the research effort. Susan Sprachman led the training of our talented group of observers/site visitors. The site visitors included Barbara Schiff Kennen, Brian Roff, David Eden, Betsy Santos, Charlotte Hansen, Emily Sama Martin, Nii Addy, Rob Buschmann, Margo Salem, and Mina Dadgar. Barbara Carlson, our sampling statistician, assisted by Yuhong Zheng, randomly selected programs, centers and children for site visits and assessment observation. Barbara Schiff Kennen and Jennifer McGovern contacted all programs to ensure their participation. Patricia del Grosso coded the site visit summaries, and managed the project finances. Megan Hague Angus assisted with data analysis. William Garrett patiently formatted the document, assisted by Alfreda Holmes. Barbara Geehan was our editor, with help from Daryl Hall. Diane Paulsell, who originally conceptualized this implementation study and led its previous reports, provided valuable insights during the review process. Sally Atkins-Burnett served as our quality assurance reviewer. We would also like to thank John Love, the original project director, for his ongoing support.

We also extend our thanks to staff from Juárez and Associates, Rita Rico and Jatzin Alvarado, who were exceptional site visitors.

The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Health and Human Services.

EXECUTIVE SUMMARY

In fall 2003, the Office of Head Start began implementing the Head Start 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.¹ 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 information on a standard set of child outcomes from all Head Start programs in a consistent manner. Head Start has a decade-long history of concentrating on child outcome measures within its performance expectations. Specific national program performance measures were developed in 1995 and, in 1996, the Head Start Program Performance Standards strengthened requirements for the ongoing screening and assessment of children throughout their Head Start participation. In 1998, after the Head Start reauthorization, all programs were required to include child outcomes in their self-assessment process by 2003. In addition to using child assessment to measure program performance, the Office of Head Start supports large scale projects that assess children's performance and experiences within Head Start using representative samples. These studies include the Head Start Impact Study and the Family and Child Experiences Survey (FACES).

Analysis of NRS data will enable the Office of Head Start to determine how children progress on a limited set of outcomes during the year preceding kindergarten. These data will provide the Office of Head Start with information that can be used to develop targeted technical assistance and to enhance its ability to report for federal accountability efforts. In addition, the Office of Head Start has suggested that the NRS can be useful for local program improvement efforts, providing an additional source of information to complement local surveys, needs assessments, and child assessments.

¹ The Improving Head Start Act, signed in December 2007, discontinued the NRS. This report focuses on NRS implementation in the spring of 2006 and was completed before the discontinuation of the NRS.

The NRS includes a 15-minute child assessment battery, a system for training staff 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 now provides each program with a summary report of average results for all children in the program who were assessed, available within a month after each fall and spring administration and in final form several months after the final completion date of administration.² Reference tables are also created to allow programs to compare their scores to national averages, to regional averages, and to programs similar to theirs based on a number of characteristics, such as the percentage of children who are English Language Learners or the program auspice.

In July 2003, the Administration on Children, Youth and Families (ACYF) contracted with Mathematica Policy Research, Inc. (MPR) and its subcontractor, Juárez and Associates (J&A), to conduct the Head Start NRS Quality Assurance and System Development Project. The project had two components—an implementation study to assess the quality and other aspects of the first year of NRS implementation (training, child assessment, data entry, and program perspectives) and support for system development activities that could enhance the quality and usefulness of the NRS. A final report on the Year 1 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.³

This report documents Head Start programs' experiences with the NRS during the third year of implementation. The report is based on information collected through visits to a nationally representative sample of 35 Head Start programs in spring 2006. During these visits, MPR and J&A observed a sample of child assessments and interviewed staff about their experiences implementing the NRS, particularly in spring 2006.

DEVELOPMENT OF THE HEAD START NATIONAL REPORTING SYSTEM

Head Start has long emphasized continuous program improvement and outcomes-oriented accountability. With an extensive history of conducting research and program evaluations, Head Start began developing specific program performance measures in 1995, in part to be responsive to the Government Performance and Results Act (GPRA) requirements. In 1996, the Family and Child Experiences Survey (FACES) was launched to collect data on the performance indicators. The 1996 revisions of the Head Start Program

² Beginning with the spring 2006 data wave, programs received interim reports based on their local data in July, and full reports, with the final national averages included, in October. Note that our site visits occurred prior to this change in reporting policy.

³ 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.

Performance Standards instituted new requirements for the screening and assessment of children for the purpose of improving teaching and learning.

Following the 1998 reauthorization of Head Start, the Office of Head Start required all programs to include child outcomes in their self-assessment process by 2003. In August 2000, the Office of Head Start issued an information memorandum that laid out the steps programs must take to meet this requirement. These results-based standards and performance measures were presented in the context of a “Head Start Child Outcomes Framework,” comprising eight general developmental domains, including several—specifically, language development, literacy, and mathematics—that were targeted in the assessment requirements of the NRS (ACYF 2003). The child assessments required of all programs encompass 13 legislatively mandated indicators in language, literacy, and mathematics. In implementing these assessments, all Head Start grantees have been charged with (1) improving the objectivity of their assessments, (2) analyzing the data over time in order to understand the nature and patterns of children’s progress, and (3) incorporating the results into continuous program improvement efforts. To meet these requirements, Head Start programs were permitted to select their own assessment instruments, as long as their instruments measured progress in the required developmental domains. Programs currently use a wide range of assessment strategies and tools to measure children’s progress.

While a child outcomes approach is not new to Head Start, a national assessment system implemented consistently for all 4- and 5-year-olds is. The NRS, initiated in April 2002 when the Bush administration announced the *Good Start, Grow Smart* early childhood initiative, is a key element of the “Strengthening Head Start” component of this initiative. It creates a national assessment and reporting system out of the congressionally mandated “standards of learning,” thus carrying out the president’s directive to develop “a strategy to ensure that, for the first time, every Head Start center assesses the standards of learning in early literacy, language, and numeracy skills.” As the Assistant Secretary for Children and Families noted, “The President’s *Good Start, Grow Smart* initiative challenges us to improve the operational effectiveness of Head Start programs by developing a systematic, nationwide approach to assessing every child’s school readiness” (Horn 2003).

To aid in developing the NRS, the Office of Head Start contracted with Westat, Inc., and its subcontractor Xtria, LLC, in August 2002 and convened a Technical Work Group (TWG) of 16 experts in child development, child assessment, measurement, and program evaluation. Beginning in December 2002, the TWG met three times to advise the Office of Head Start on the design of the NRS and the selection of instruments to be included in the child assessment. In addition, the Office of Head Start convened several discussion sessions, focus groups, and workshops with Head Start program staff, early childhood researchers, and assessment experts, to discuss plans for the NRS.

The NRS Assessment

In April and May 2003, Westat field-tested an initial child assessment battery with 1,434 children in 36 programs. Based on the analysis of data from the field test, and

considering input from the TWG and others, ACYF finalized a 15-minute NRS assessment battery for fall 2003 that contained four components:

1. **English Language Screener.** This component screens children for comprehension of spoken English to identify those whose English is insufficient to participate in the full assessment. It is composed of two subtests from the Oral Language Development Scale (OLDS) of the Preschool Language Assessment Scale (PreLAS) 2000 (Duncan and DeAvila 1998). The first set of items uses the “Simon Says” game to request that children follow simple commands, such as “touch your ear” and “point to the door.” In the second set of items, children are asked to name or describe the function of objects in pictures. A Spanish-language version of the OLDS is also used with Spanish-speaking children.
2. **Vocabulary.** Adapted from the third edition of the Peabody Picture Vocabulary Test (PPVT-III; Dunn and Dunn 1997), this section includes 24 items assessing children’s receptive vocabulary that represent a range of difficulty.
3. **Letter Naming.** A test developed by Westat for the Head Start Quality Research Centers Consortium (Zill 2003a), this section presents all 26 pairs of upper- and lower-case letters of the alphabet in three groupings (with 30 letters in the Spanish version). Children are asked to identify the letters they know by name.
4. **Early Math Skills.** Adapted from the mathematics assessment used in the Early Childhood Longitudinal Study—Kindergarten cohort (ECLS-K), this section includes items on number understanding; shape recognition; relative size judgments and measures; and simple word problems involving reading graphs, counting, or basic addition and subtraction (Zill 2003b).

A Spanish-language version of the child assessment was also developed. During the first year of NRS implementation, all children whose home language was identified as Spanish were assessed in both English and Spanish, provided they passed the language screener for each version of the assessment. All children took the English-language assessment first. Spanish-speaking children who did not attain the threshold on the English language screener were assessed in Spanish only. During the second year of implementation, children whose home language was identified as Spanish were still assessed in both languages as long as they passed the language screener for each version. However, beginning in fall 2004, in response to feedback from the TWG, the Office of Head Start staff, and local Head Start staff who participated in the Year 1 Quality Assurance Study, children whose home language was identified as Spanish took the Spanish-language assessment first, followed by the English version.

Since first implementing the NRS in fall 2003, Westat has made minor changes to the NRS battery, based on recommendations from the TWG and findings from MPR’s reports on the Head Start NRS Year 1 Quality Assurance Study. These changes have included

shortening introductions, changing some words in the vocabulary section, simplifying the directions on the letter naming section, improving some pictures in the math section, simplifying question wording in the math section, and simplifying the hand gestures required for some items. In addition, Westat added clarifications to training materials to allow some minor deviations from the script; these changes were made to help children feel more comfortable during the assessment.

The Computer-Based Reporting System

The Office of Head Start implemented the Computer-Based Reporting System (CBRS) to collect background information on Head Start programs and children, to facilitate the identification of eligible children, and to track completed assessments. The CBRS is a web-based system where Head Start program staff members enter all relevant information. Included in the program-level data are contact information for the grantee, delegate agencies, centers, and program start and end dates. Classroom-level information includes the type of class (such as part-day or full-day), total enrollment, and number of classroom staff. Information on teacher qualifications and experience is also collected. For each eligible child, staff members enter the date of birth, classroom entry date, years in Head Start, disability status, language spoken at home, level of English proficiency, ethnicity, race, and assessment completion status. The CBRS is used to assign identification numbers and can print out class rosters for use in tracking assessments, as well as assessment completion reports.

Beginning in the second year of NRS implementation, Xtria, the subcontractor responsible for developing the CBRS, expanded its reporting capabilities and system functions. Local programs can now generate assessment completion reports at the program or center level, perform data searches, and view and operate the CBRS in Spanish. They can also take advantage of both a data copy feature—to reduce data entry duplication for certain fields that remain the same from the previous program year—and a data import feature—to import data from the program’s management information system. Centers can also enter or make changes to multiple children at one time. In addition, grantees can produce NRS data reports from the CBRS about their delegate agencies.

NRS QUALITY ASSURANCE AND SYSTEM DEVELOPMENT PROJECT

As noted, the NRS is the first implementation of a nationwide assessment of 4- and 5-year-olds enrolled in Head Start. Since the scale of this initiative is unprecedented, the Office of Head Start contracted with MPR and J&A to assess the extent to which locally trained Head Start program staff across the country could administer a standardized assessment with reasonable accuracy and fidelity to the assessment protocol. In addition, because this new policy of assessing all kindergarten-eligible children had generated some controversy within the Head Start community, the Office of Head Start sought to learn more about local program perspectives on the assessment process.

Results of the Year 1 Quality Assurance Study demonstrated that, while there was room for improvement, Head Start staff members were able to administer the NRS assessment

with a fairly good degree of fidelity to the protocol (Paulsell et al. 2004). Moreover, analyses of site visit interview data and assessment observations yielded a number of helpful suggestions for improving the assessment battery and process. The Office of Head Start decided to extend the study into a second year for several reasons. First, they wanted to determine whether the level of fidelity documented through the Year 1 study would be sustained over time. Second, they wanted to determine whether comparable levels of fidelity would be observed in different samples of programs (and consequently requested that MPR draw different samples in fall and spring of the second year, for a total of 70 programs). Third, the Office of Head Start wanted to collect ongoing information about programs' experiences and perspectives on NRS implementation, including their perspectives on changes made to the assessment process from Year 1 to Year 2. An extension of the study to collect data in spring 2006 was added later.

In addition to conducting the Quality Assurance Study, MPR and J&A were charged with recommending system development strategies for enhancing the quality and usefulness of the NRS. Below we describe these two primary components of the project:

1. **Quality Assurance Study.** To observe and collect information on various facets of the NRS implementation, MPR and J&A staff made site visits to a nationally representative sample of 35 Head Start programs in fall 2003, spring 2004, spring 2005, and spring 2006. During these visits, staff observed the assessments of a random sample of children and interviewed program staff members about their experiences implementing the NRS.
2. **System Development Support.** MPR and J&A supported the Office of Head Start in assessing all aspects of the ongoing implementation of the NRS with the ultimate goal of enhancing the methods of NRS child assessment, data management, and training; and developing ways to help programs understand how they can use NRS information to improve the quality and outcomes of Head Start services for all children. Information used to formulate recommendations for improvements comes from the NRS Quality Assurance Study, program practitioners' views, input from expert consultants, Office of Head Start staff, other support contractors, and other federal agencies that often partner with ACYF/ACF on issues related to children's programs and research.

The purpose of this report is to provide updated findings from the Quality Assurance Study—focusing primarily on NRS implementation in spring 2006—and to make recommendations for system improvement.

QUALITY OF ASSESSMENT ADMINISTRATION

As in earlier rounds of the NRS Quality Assurance Study, most of the English language assessments observed in spring 2006 met or exceeded the standard of quality used in certifying assessors. The average certification score was 92, and 83 percent of observed assessments exceeded the certification standard of 85 points. The inter-rater reliability of

assessment scale scores remained high, and the quality of the Spanish-language assessments observed (65 in 11 sites) was also high, on average (certification score of 97).

Errors due to coaching, inappropriate gestures, administering non-neutral encouragement, straying from the script, and scoring errors were higher in spring 2006 than in spring 2005; only errors for inserting articles such as “a” and “the” decreased. The increase in errors may reflect changes made to the assessment in spring 2005 (particularly those that simplified gesturing) and perhaps less attention to local refresher training that would underscore such changes in administration.

Although quality was high overall, some areas of the assessment remained difficult for staff. Modifications to the assessment easel or additional guidance and training might be helpful in these areas. The most difficult areas were also problematic in Years 1 and 2, including:

- Setting up the testing in a quiet area free of distractions
- Avoiding coaching, particularly in the Simon Says section
- Gesture errors on certain Art Show and Early Math items
- Scoring the counting item (E20) correctly in the Early Math section

Performance in these difficult areas was somewhat poorer than in spring 2005. Scoring errors on the counting item, however, remained essentially unchanged—it was scored incorrectly on roughly one in four assessments. Moreover, while the Letter Naming section yielded low error rates overall, with the exception of script errors in the introduction and transitions in between letter plates, it remains a section with which many programs continue to struggle. They worry that this part of the assessment is not valid, since children may very well know some or all of the alphabet letters, but the structure of this test administration might not enable children to demonstrate what they know. Allowing assessors to point to each letter and say, “What’s this?” or “What’s the name of this one?” could help ensure that children do not inadvertently omit a letter that they lost track of while scanning the plate.

About half of sample programs reported that most children reacted positively to the child assessment. Assessors noted that some children enjoyed the one-on-one time with the assessor and liked showing off what they knew, as well as the opportunity to “play some games.” Children’s behavior was much less of a concern in the spring than it had been in the fall, a finding similar to Years 1 and 2. The major behavioral issue was that children became bored or restless during the PPVT and/or the Letter Naming task and needed a lot of redirection. Nearly half of the programs thought that the NRS assessment is too long.

Most programs (24 out of 35, or 68 percent) administered the child assessment to English Language Learners. Spanish was the most common other language spoken, followed by Chinese and Arabic. Assessors reported that most children could pass the English language screener in the spring, even if they had not passed in the fall. Some children who are classified as having Spanish as a primary language frequently are more

fluent in English; similarly, some children lose their Spanish skills over the course of the school year, and perform better on the English version in the spring. Bilingual assessors, as has been the case in previous rounds of site visits, were critical about two overarching components of the NRS assessment. First, they felt that certain Spanish speakers are at a disadvantage due to the wide variety of dialects spoken by children and families who use colloquial words not considered to be correct on the NRS assessment. Second, these assessors asserted that children are tested on four additional alphabet letters (CH, LL, Ñ, and RR) that are no longer used in Spanish. However, further investigation has shown that only RR has been discontinued in current Spanish usage, and it is therefore being eliminated from the assessment.

Nearly all programs (33 out of 35 in the sample) assessed children with disabilities, using a wide range of accommodations for the child assessments. In some situations, assessors decided to administer the NRS assessment to a child who could reasonably have been exempted. Most staff members who administered assessments to children with disabilities felt comfortable with the process and with the accommodations made. However, several programs would like the NRS training materials to address explicitly the issues of conducting the NRS assessment to children with special needs and making appropriate accommodations for these testing situations. In each round of the QA study, a number of Head Start programs have expressed interest in these same issues.

LOCAL APPROACHES TO NRS TRAINING

To ensure consistent administration of the NRS assessment, Head Start programs need to follow standardized procedures when training local staff. Since the NRS was first implemented, the Office of Head Start has provided local programs with standardized training materials and guidelines to conduct training.

In summer 2005, few programs sent a representative to one of the “training-of-trainers” events offered for programs that have had a change in their NRS lead. Several programs, however, would like to see refresher trainings for NRS lead trainers, and to be able to send more than one staff member to a regional conference to build local capacity for NRS implementation. They would prefer training sessions that are available closer to their programs.

In spring 2006, all but three of the sample programs provided refresher training. However, only about one quarter adhered to the training guidelines. The other programs did not include all of the required training elements—viewing and discussing the training video, distributing the Assessor’s Guide, and role-playing the assessment—and thus most training sessions were shorter than the recommended four and a half hours. About half of the programs that needed to provide refresher training for Spanish assessors reported offering a formal group session; bilingual assessors at most other programs relied on self-study to prepare for the spring 2006 assessments. Some reasons why most programs did not adhere to the training protocol include late arrival of training and assessment materials, limited time to complete assessments by the end of the program year, and staff members’ perceptions that experienced NRS assessors did not need refresher training.

Eighteen percent of the sample programs trained new assessors in spring 2006. Although all of these programs provided separate training for new assessors and most programs certified them during a practice assessment, none of the programs followed the required training protocol for new assessors. Instead of providing a full-day training, most programs conducted the training in two to three hours. In rare cases, a small number of new assessors did not get formally certified on the NRS after attending a training session.

As in past rounds of site visits, the responses of local Head Start staff to the NRS materials and information provided by the Office of Head Start were mixed. At least one staff member at most programs watched one of the NRS broadcasts or webcasts, but programs' views of their usefulness were uneven. Problems continued with timely delivery of training and assessment materials, with nearly half of programs reporting a delay in receipt, and few programs took advantage of technical assistance offered through the help line or other means. While staff from two-thirds of programs found the video to be the best or one of the best components of local NRS training, other staff found the video too long and repetitive and preferred the role-playing exercises or the Assessor's Guide to help prepare for the spring assessment. Lack of thorough refresher training in many programs, coupled with a tendency not to seek technical assistance, may have contributed to some of the errors in administration and scoring described in Chapter II. This point is important because analysis of error rates between programs that did and did not provide refresher training showed that assessors in programs without refresher training committed significantly more errors in administration than assessors from programs that did hold them.

LOCAL APPROACHES TO IMPLEMENTATION

Overall, the 35 Head Start programs we visited in spring 2006 took an approach to NRS implementation similar to that of programs visited in previous rounds of site visits. Program directors assigned a lead NRS trainer the responsibility of overseeing implementation, including assigning, training, and certifying assessors; scheduling and tracking the completion of assessments; overseeing quality assurance activities; and submitting score sheets by the deadline set by the Office of Head Start. In more than a third of programs, the lead trainer was also responsible for maintaining the CBRS.

Most programs maintained the same basic staffing structure for the NRS in the spring that they had instituted in the fall. When programs did make changes, most were due to staff turnover within the program rather than a rethinking of the program's approach to the NRS. On average, the programs trained 17 assessors. As in the sample of programs we visited in spring 2005, most programs did not rely exclusively on teaching staff to conduct the assessments. Program staff saw many advantages to having teachers administer the assessment, including that the children were comfortable and familiar with teachers; children responded better to teachers; and teachers had better knowledge of children's behavior, child development, and child assessments. Learning more about children's knowledge in specific skill areas could inform teachers' classroom practice. However, many programs expressed concerns about teacher burden, reduction in instructional time, potential for coaching or bias in administration, and the cost of hiring substitutes to cover for teachers while they were assessing children.

Nearly all programs communicated with parents and Policy Councils about the NRS. While the means of communication varied, programs made an effort to inform parents of the assessment at the start of the program year. Three-quarters of programs sought to obtain parents' written consent for the NRS assessment, far more than in previous rounds of site visits. Most programs had few or no parent refusals. Concerns among parent and Policy Council members were similar to those expressed in previous rounds of visits. They included parent requests to see their child's individual results; concerns about bias of specific items; and questions about the purpose of the NRS, how the results would be used, and how the results would affect the Head Start program.

Few Head Start directors estimated the monetary and in-kind costs of NRS implementation. However, a number of directors identified significant costs, such as staff time, travel costs, costs for paying substitute teachers, and overtime for staff working on the NRS.

USING THE COMPUTER-BASED REPORTING SYSTEM

As in past rounds of site visits, programs reported that the CBRS was easy to learn and use. Programs had adequate numbers of computers and Internet connections for accessing the system. Many programs used and liked the enhanced features of the CBRS introduced in Year 2, especially the data copy, data import, and expanded report features. Some programs reported accessing the CBRS help line at least once and most of these reported receiving prompt and helpful support from the help line staff. Data managers made some suggestions for further enhancements to the CBRS, including reducing the cost of the data import feature, providing more support for using the data copy feature, and adding other user-friendly features.

USING THE NRS FOR LOCAL PROGRAM IMPROVEMENT EFFORTS

The majority of Head Start staff who participated in the site visits found the 2004–2005 Growth Report easy to understand, although some said it was somewhat confusing, particularly for non-technical audiences. Staff made several recommendations to improve the report's content (for example, provide information on how subgroups of children performed) and format (for example, distribute a version that yields better photocopies). Staff in about two-thirds of the sample programs agreed that center or classroom-level reporting would be more useful in targeting practices and resources; about a quarter of programs requested individual child-level data. A majority of the sample programs shared NRS results with staff and key stakeholders, such as managers, specialists, and teachers; Policy Councils; boards of directors; and, to a lesser degree, parents. Some programs shared details on how they have used NRS data to modify classroom practices, such as spending more time on alphabet knowledge, literacy development, and, to a lesser degree, counting and other early math skills. Because the NRS reports did not provide center-, classroom- or child-level information, 11 percent of the programs reported tracking the item responses of individual children using the raw data on the Scantron answer sheets to make decisions on what areas needed improvement to better prepare children for kindergarten.

With regard to using the NRS results in the future, more than half of the sample programs said they planned to use the reports to some degree. In many cases, staff did not provide much detail on these plans, but they intend to continue incorporating NRS outcomes as one source of information for program-planning efforts. About 20 percent of programs either remained undecided about whether they would use the results or did not plan to use them at all. However, 11 percent of programs planned to observe trends over time, now that multiple years of data are available.

PERSPECTIVES OF LOCAL HEAD START STAFF ON THE NATIONAL REPORTING SYSTEM

When asked to list the contributions the NRS had made to their programs, staff members in sample programs were able to attribute a number of positive effects of the NRS. Primarily, the contributions that staff listed were that it (1) helped to improve classroom practice and inform teacher training, (2) proved Head Start is effective, (3) raised program accountability by providing national comparisons, and (4) validated local assessment results. In a few programs, a contribution of the NRS was to raise staff morale. In two programs, staff indicated that a contribution of the NRS was that it helps to get children accustomed to a testing environment.

Local program staff raised several concerns about the NRS and its implications for future directions Head Start may take; most of these concerns also had been raised in previous rounds of site visits. The concern expressed most often by local staff members was about how the NRS results would be used at the national level. Many staff said that they still did not have a clear understanding of the purpose of the NRS and its implications for local programs where children did not perform well on the assessment. Local staff also expressed concern about whether the NRS results accurately reflect program performance, the amount of staff time and financial resources dedicated to the NRS, and whether it was valid to compare fall and spring assessments.

To improve NRS implementation, programs suggested that the Office of Head Start share more information about how the results would be used, send training materials and outcome reports to programs sooner, provide more information about the assessment's development and validity, and consult more with the Head Start community about future changes to the assessment. Many programs requested written materials about the NRS for parents.

Regarding the assessment battery, programs were split on whether new domains should be added. Many staff thought that at least one new domain—particularly social-emotional development—should be added, but at the same time they did not want the assessment to become longer. (The social-emotional component of the NRS was implemented for the first time in fall 2006, after our site visits.) Program staff continued to recommend that the NRS be combined with the local assessments. They also suggested changes to improve the Spanish-language version and to modify specific assessment procedures and items.

IMPLICATIONS OF FINDINGS FROM THE QUALITY ASSURANCE STUDY

The value of the NRS for local program purposes is dependent upon how well local Head Start programs implement it, whether they accept it as a valid and reliable assessment, and whether they can use the information effectively for program improvement. Accuracy of administration, validity and reliability are also important to the other OHS goals for the NRS: to develop targeted technical assistance and to enhance the ability to report for federal accountability efforts. On the basis of findings from site visits to nationally representative samples of programs over the first three years of NRS implementation and our understanding of concerns raised by the Office of Head Start, expert consultants including Head Start program directors, and others in the field, we make the following suggestions for system improvement of the NRS. These suggestions are of several types: (1) ideas for increasing communication with local Head Start programs to improve knowledge and understanding of the purposes of the system, (2) ways to improve access to and usefulness of NRS results for local programs, (3) options for supporting programs in administering the assessment (including guidance on assessing children in Spanish and children with disabilities), and (4) ideas for changes to the assessment battery. Over the course of five rounds of visits to representative samples of Head Start programs, we have seen evidence that feedback on problems has led to improvements in the system. However, we have also found evidence that some problems persist and deserve further attention.

Communication

Several concerns raised by local program staff during three years of the Quality Assurance Study appear to be fueled by uncertainty and lack of information. For example, not knowing the specifics of how the NRS results will be used at the national level has led to speculation about whether it will become a “high-stakes” test, helping to make decisions on funding or employment at the program, center, or classroom level. Increased communication about the Office of Head Start’s plans could address some of these concerns. Suggestions for improving communication follow:

- Provide more information about how the Office of Head Start will use the NRS assessment results.
- Provide a technical report or background materials about how the NRS assessment battery was developed and the validity and reliability of each task, including predictive validity for later school performance. Analyses of NRS test data could be done to investigate this issue, leading either to changes in items or reassurance that children’s skills are not being unfairly measured by test items.
- Explore options for increasing consultation with Head Start practitioners on future changes to the NRS.
- Provide written materials about the NRS for parents. During site visits, some staff reported feeling unsure about how to report NRS results to parents and Policy Councils or respond to their questions about the NRS.

Access to and Usefulness of Assessment Results for Local Programs

Each round of data collection reflects more programs that are attempting to make use of the NRS results to improve their classroom practice, but they continue to manifest a need for guidance in how to accomplish this goal. As in spring 2005, program staff suggested several ways to make the reports more useful to them and requested help in interpreting the results and implementing appropriate changes to their programs:

- Send reports sooner after each round of data collection, preferably in time to use them in planning for the next program year.
- Consider providing results disaggregated to the center or classroom level.
- Provide programs with more guidance on how to use the NRS results for program improvement, including building more linkages with local assessment results.
- Provide access to resources to support local program improvement efforts undertaken in response to assessment results.

Support for Administering the Assessment

Timely shipment of training and assessment materials and modest changes in the NRS training protocols have the potential to help staff feel more confident and conduct the assessments more smoothly and accurately. Specifically, recommended changes include the following:

- Ensure that programs receive materials and scheduling information on time.
- Ensure that adequate refresher training is carried out to reduce errors in administration. We conducted analyses that demonstrated a link between offering assessor trainings and assessor performance on the assessment.
- Consider requiring recertification of assessors during refresher training.
- Provide more guidance on how to interact appropriately with children during the assessment.

Guidance on Assessing Children in Spanish

Assessors and trainers continued to have some concerns about administering the Spanish version of the NRS and expectations for growth in Spanish-language skills. Some possible approaches to addressing these concerns include the following:

- Allow for more regional differences in Spanish terms used in the assessment.

-
- Clarify expectations for growth in Spanish skills.

Guidance on Assessing Children with Disabilities

All but two programs in the sample assessed some children with identified disabilities during spring 2006. Although the majority of programs felt comfortable with the accommodations they made, some requested more guidance on how to assess children with disabilities. More guidance would be helpful in the following areas:

- Clarify when a child should not be assessed because of a disability.
- Provide more examples of appropriate accommodations.

Improving the Assessment Instrument

Programs acknowledge that improvements have been made to the assessment instrument over time, but continue to raise several concerns:

- A large majority of sample programs (86 percent) reported some concerns with the receptive vocabulary (PPVT) section of the test that they fear may affect overall validity, or may differentially affect validity for some subgroups, like urban versus rural children.
- The primary concerns in the Letter Naming section were due to the layout of the plates and the method of administering them.
- The form or content of the Early Math section was the source of concern for close to half of programs; for example, many programs felt that the concepts such as subtraction and reading graphs were too advanced for this age group.

In summary, based on our visits to a representative sample of Head Start programs, we suggest implications for system improvement in two major areas: (1) helping program staff better understand the purposes, interpretation, and potential uses of the NRS through improved communication and guidance (including a special focus on children with disabilities and children assessed in Spanish) and (2) making improvements in the training, administration guidelines, format, and content of the battery itself. Through the Quality Assurance Study, MPR will continue to work with the Office of Head Start, expert consultants including Head Start program staff, and the implementation contractors to foster improvements to the entire assessment system.

REFERENCES

- Administration on Children, Youth and Families. “The Head Start Leader’s Guide to Positive Child Outcomes.” Washington, DC: U.S. Department of Health and Human Services, 2003.
- Dunn, L.M., and L.M. Dunn. *Peabody Picture Vocabulary Test—Third Edition*. Circles Pines, MN: American Guidance Services, 1997.
- Duncan, S., and E. DeAvila. *PreLAS 2000 Technical Report*. Monterey, CA: CTB/McGraw Hill, 1998.
- Horn, Wade F. “Improving Head Start: A Common Cause.” *Head Start Bulletin*, no. 76, 2003. Available at [www.headstartinfo.org/publications/hsbulletin76/hsb76_02.htm].
- 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.
- Zill, Nicholas. *Letter Naming Task*. Rockville, MD: Westat, Inc. 2003a.
- Zill, Nicholas. *Early Math Skills Test*. Rockville, MD: Westat, Inc., 2003b.