EXECUTIVE SUMMARY

In 2012, the Office of Planning, Research, and Evaluation at the Administration for Children and Families (ACF) engaged Mathematica Policy Research and its partners to conduct a project titled "Assessing Early Childhood Teachers' Use of Child Progress Monitoring to Individualize Teaching Practices."[1] The purpose of the project is twofold: (1) to develop a conceptual framework of early childhood teachers' use of ongoing assessment to individualize instruction and (2) to create a measure to examine this process. This project can make a significant contribution to the early childhood field, both by defining the key aspects of ongoing assessment and individualization and by providing guidance on efficiently measuring how they are carried out. Ultimately, it may help teachers enhance their practices in the classroom which, in turn, will improve early childhood program performance.

Head Start recognizes the importance of using ongoing assessment to individualize instruction for each child. Over the past five years, the Office of Head Start (OHS) has elaborated on its vision for preschool child and family outcomes, strengthened its focus on monitoring program and classroom quality, and developed tools to support ongoing assessment in daily practice (U.S. Department of Health and Human Services 2010; Atkins-Burnett et al. 2009). Recently, the Advisory Committee on Head Start Research and Evaluation, convened by the secretary of the U.S. Department of Health and Human Services, advocated investing in evidence-based and data-informed practices across all domains of teaching and learning (Advisory Committee on Head Start Research and Evaluation 2012).

Despite the importance of using ongoing assessment data to guide instruction, and the Head Start program requirements to do so, information is sparse on how early education teachers actually collect and use these data to tailor their instruction. Policymakers, practitioners, and researchers continue to see an urgent need for research in this area in the quest for better educational outcomes (Bambrick-Santoyo 2010; Black et al. 2003, 2004; Fuchs and Fuchs 2006; Hamilton et al. 2009; Marsh et al. 2006).

A. Two Approaches to Ongoing Assessment

Ongoing assessment of children's progress is increasingly a priority in early childhood classrooms, yet teachers' use of these assessments has not been extensively researched. General outcomes measures (GOMs) and curriculum-embedded approaches are two common approaches to ongoing assessment used in these classrooms.

General outcomes measures. In the GOM approach, teachers use a brief measure with strong evidence of reliability and validity to conduct frequent, standard assessments of children's progress toward a long-term goal. Central to this approach is the repeated measurement of a few key skills that represent the entire skill set required to achieve a given goal, rather than measuring the full skill

---

[1] This project focuses on all forms of ongoing child assessment, of which child progress monitoring is a common form. The content of this report will be broader than the title of the project implies. For more information, see the section in Chapter 1 of this report entitled, "The Importance of Developing a Measure of Early Childhood Teachers' Use of Ongoing Assessment to Individualize Classroom Instruction."

set. A child's increasing proficiency on a GOM is indicated by improved performance on these same skills measured over time.

With GOMs, children's performance may be measured as infrequently as three times per year or as often as once per week (Jenkins et al. 2009). The probes to obtain these performance samples typically range from one to five minutes, depending on the outcome (that is, the knowledge, skill, or behavior) being measured. One common application of GOMs is Response to Intervention (RTI)—an approach to early intervention involving the regular screening of all children throughout the year. Children not progressing as expected receive intensive support as well as frequent assessments to test whether the support is helping (Hamilton et al. 2009; National Association for the Education of Young Children et al. 2012; Buysse and Peisner-Feinberg 2013). GOMs typically do not focus on the full set of child outcome domains. Most GOMs in preschool currently focus on language and literacy, and some focus on mathematics.

Curriculum-embedded approaches. The most commonly used systems for assessing the progress of children in early care and education are curriculum-embedded approaches. These assessments are used to examine children's progress relative to early learning standards and the skills and knowledge taught via a specific curriculum. Teachers using this approach often collect assessment information as they are teaching their normal curriculum. The assessment tasks are intended to be authentic in context; that is, they are "opportunities created for children that reflect typical experiences rather than discrete isolated tasks that are irrelevant to the child's daily life" (Pretti-Frontczack et al. forthcoming). Some curriculum-embedded approaches are developed by the curriculum developers to align closely with the material being taught ("curriculum-based assessments" such as the Teaching Strategies: GOLD assessment used with the Creative Curriculum), whereas other such assessments are derived from national standards and developmental expectations ("curriculum-embedded assessments" such as Galileo and the Work Sampling System).

Teachers typically assess children's performance in relation to criteria on rubrics provided by the assessment system. These rubrics specify different levels of performance based on end-of-year goals, but often provide no guidance regarding children's expected progress throughout the year. In addition, although the tasks being assessed are embedded within daily activities and aligned with curriculum goals, the tasks are not standardized and require teachers to collect assessment data from multiple sources. The assessments may use a variety of data collection methods, such as observation recording forms, worksheets, standardized assessments, and portfolios.

B.   A Focus on Curriculum-Embedded Approaches

This report presents a measure development plan for curriculum-embedded approaches. We chose to focus on these assessments because they are (1) more common in early childhood settings than GOMs; (2) more demanding for a teacher to implement (that is, they require greater teacher skills and knowledge); and (3) more comprehensive, as they traditionally cover several domains of development.

Figure ES.1 shows the conceptual framework for using curriculum-embedded approaches to monitor children's progress on an ongoing basis and individualize instruction. The model has four iterative stages: (1) selecting the assessment target and method; (2) implementing the assessment; (3) interpreting the assessment data, including hypothesis setting and selection of instructional

decisions; and (4) applying instructional decisions. We explain each stage and its quality indicators below.

Figure ES.1. Conceptual Framework for Curriculum-Embedded Approaches
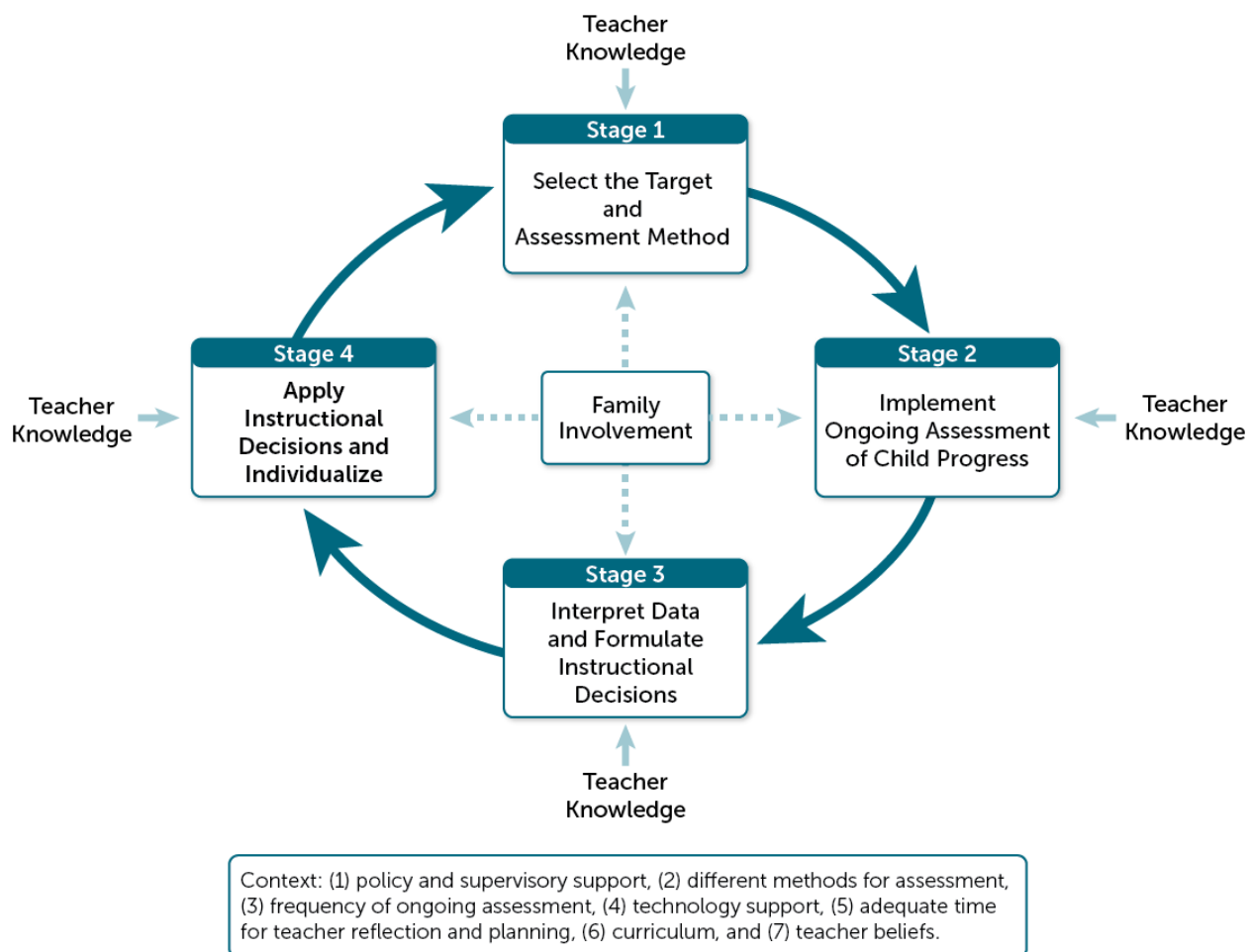
Stage 1: Selecting the target and assessment method. The assessment system is often selected by program staff rather than by teachers. However, teachers have some autonomy in selecting the assessment target (the skill, knowledge, or behavior to be assessed) and the assessment method (how that skill, knowledge, or behavior will be assessed, such as observations, structured tasks, or standard tests), although both are also influenced by the selected assessment system. In choosing assessment targets, teachers consider the desired end-of-year outcome and set targets that track progress toward that outcome. They should collect, interpret, and reflect upon the data and make instructional adjustments throughout the year, frequently enough to monitor and guide children's progress. In this way, teachers' decisions play a large role in the assessment process throughout the entire year. The supports available to the teacher within different assessment systems vary, as does the strength of the link to the curriculum.

When examining teachers' selection of assessment targets and methods, there are several indicators of quality to consider. These include the assessment target's relation to meaningful and developmentally appropriate outcomes, key behaviors, knowledge, or skills; the ability to affect the assessment target through intervention and change; the alignment of the target with the curriculum

and method of data collection; the focus on observable and generalizable behaviors; the validity of the data collection method (including linguistic and cultural appropriateness); and the efficiency and feasibility of frequent data collection.

Stage 2: Implementing ongoing assessment. Typically, the teacher will weave assessment tasks into his or her instructional activities. Efficient assessment is therefore important to making the most of instructional time. The assessment should also be ecologically valid, using tasks that fit into a child's usual routine in class or at home. Teachers must document child progress objectively, accurately, and with relevant contextual information.

The indicators of quality to measure at this stage include the soundness of the assessment itself (whether it is ecologically valid, appropriate for the child, and fair) and the teacher's approach to documenting the findings (whether it is efficient, consistent, and objective, taking contextual factors into account).

Stage 3: Interpreting data and formulating instructional decisions. In an ongoing assessment system, teachers need to be able to interpret the data about each child's performance compared with performance expectations, which are usually based on developmental or curricular guidelines or the scores of typical same-age peers. Teachers may also need to combine the assessment data with other relevant data. The data are used to identify children's strengths, weaknesses, interests, and learning differences; based on the findings, teachers then select the best way to support each child's progress. This process may be conducted in teams with the support of other teachers, coaches, consultants, and family members.

When measuring quality at this stage, researchers should consider whether the assessment data are organized so as to (1) facilitate interpretation and easy communication with families; (2) impose a minimal burden on teachers; and (3) provide consistent, reliable data entry. Teachers' reflection on and interpretation of the data should also be evidence-based and consider alternative hypotheses.

Stage 4: Applying instructional decisions and individualizing. A requirement of the Head Start Performance Standards, individualization is important for maximizing child progress (*Federal Register* 2011). It involves planning and delivering high-quality, evidence-based instruction that is targeted to individual children and reflects the data collected about each child.

The indicators of quality to measure at this stage include the use of evidence-based strategies that are responsive to the data, implemented with fidelity, and evaluated in an ongoing manner. Teachers should also individualize their lessons using a variety of approaches, while building on children's strengths and interests.

Personnel, family, and contextual factors affecting most stages. The conceptual framework contains several factors that could potentially affect the implementation quality across stages. These factors include teachers' knowledge and beliefs about assessment, instruction, and children's development as well as families' involvement in the assessment process. In particular, each stage of the process calls for a specific kind of teacher knowledge. Stage 1 calls for teacher knowledge of assessment and child development, for example, whereas stage 3 calls for teacher knowledge of instruction, pedagogy, and child development. If a teacher has more knowledge in one area than in another, her quality of her implementation may vary across stages. The context in which the assessment occurs also affects the quality of implementation. Context includes key aspects of the program structure that help or hinder teachers' use of ongoing assessment, such as:

- The degree of policy and supervisory support for ongoing assessments

- The availability of adequate time for reflection and planning

- A culture of using data to inform instructional planning and opportunities for teachers to collaborate as they assess and interpret data

- Access to professional development opportunities and information about evidence-based or professionally recommended instructional strategies

C. Proposed Multi-Method Measure

The indicators of quality identified in the conceptual framework served as the foundation for the development of a plan to measure teachers' use of ongoing assessment to individualize instruction. This measurement plan draws on information from a number of sources, including a literature review, input from ACF and an expert consultant group, examples of teachers' assessment documentation, and reviews of manuals for ongoing assessments. Together, these sources helped us identify key constructs for measurement as well as data sources for measuring these constructs.

We propose using a multi-method measure—the Tool for Tailored Teaching (T3)—to assess teachers' use of ongoing assessment data. The T3 will draw on three data sources: a document review, video-based observations, and a one-hour teacher interview with a reflective "think-aloud" protocol.[2] Collecting data from these sources, which will provide both overlapping and distinct information, will be critical to understanding all aspects of the assessment process. Teachers will video-record certain assessment and instructional activities over a two- to three-week period, followed by a one-day visit from researchers to conduct the document review, rate the videos, and interview the teachers. We will develop scoring systems for each data source. The T3 is designed to capture the constructs in each stage of the conceptual framework while balancing the competing considerations of (1) the measure's reliability and validity; (2) the burden placed on researchers, teachers, and classrooms; and (3) budgetary limits.

Document review and ratings. For the document review, researchers will gather ongoing assessment data (such as a portfolio) for two children, one who is performing well and another who is struggling, to see how teachers are actually using the data to individualize their instruction for each child.[3] The researchers will also review the teachers' lesson plans for evidence of individualization. Rubrics, checklists, and ratings will be used to evaluate each document.

Video-based observations. The teacher will video-record a combination of assessments and small-group instruction that includes one or both of the target children. The researchers will watch the video after rating the documents and analyze it using rubrics, checklists, and ratings.

Teacher interview with reflective think-aloud protocol. During the one-hour teacher interviews, the researchers will ask for additional details on the documents and videos as well as

---

[2] Teachers will be asked to reflect and "think aloud" about how they made decisions as they conducted assessments and used the data to inform their teaching. Throughout this volume, we use the term "think-aloud" to refer to this reflective process.

[3] By "performing well," we refer to children meeting or exceeding developmental expectations for their age.

teachers' planning for and use of adaptations, modifications, and individualized teaching strategies. Teachers will describe how they use the data (for example, to determine whether children are making adequate progress, to make instructional decisions, or to involve families). Researchers will rate this information using a rating or rubric, coding the teachers' responses about their interpretations of the data, any alternatives considered, their decisions about how to individualize, and the success of their efforts.

D. Pre-testing

Pre-testing will encompass the earliest stages of developing the T3 and testing its feasibility. The pre-test will consist of three rounds of data collection, with visits to five centers and 10 classrooms that use ongoing assessment systems. During this iterative process, we will try two different approaches to the focus and frequency of the video recording. After each round of data collection, the full team will debrief and consider changes to the items, protocols, and procedures. ACF and the expert consultants will have the opportunity to review the recommended changes, and we will incorporate their feedback.

For the pre-test, we propose focusing on two domains: (1) language and literacy and (2) social and emotional development. Per the expert panel's recommendation, we suggest limiting our focus to two domains to ensure the feasibility of measure development within the scope of this project. We chose these domains due to their prevalence in early childhood curricula and the likelihood that we might observe variability in the use of assessment data to tailor instruction, which would enable us to see if our measure can capture differences in the quality of teachers' implementation. Limiting the number of domains will also help narrow our focus, allowing greater opportunity to refine the measurement of one or two domains and stay within our budget. However, please note that we will also include items in the T3 that examine whether teachers are drawing on information across domains to interpret and use the assessment data.

E. Potential Uses

By exploring the role of ongoing assessment in delivering high-quality, individualized instruction, this project can significantly strengthen the knowledge base for early childhood education, especially for Head Start. The T3 in particular could provide valuable information about how teachers use curriculum-embedded assessment data to understand children's development and to plan their instruction.

The final, validated version of the T3 could be used by researchers, sponsoring agencies, administrators, teachers, mentors, education coordinators, and coaches from individual programs or by networks of programs for an array of purposes. Researchers will be the most frequent T3 users at first, as they help staff at Head Start and in the early childhood field understand how ongoing assessments is used in classrooms. The experts we spoke with repeatedly noted that we lack even basic information about whether and how early childhood teachers use ongoing assessment to individualize instruction. Consequently, the T3 offers an unprecedented opportunity to inform the early childhood field's basic understanding of a process that is valued and even mandated but previously has not been measured. The T3 could also inform teachers, mentors, and coaches about the strengths and weaknesses of their programs' use of ongoing assessment. This could lead to teaching that is better adapted for each child and, ultimately, to stronger outcomes for all children.