Improving Teacher-Child Interactions: Using the CLASS™ in Head Start Preschool Programs

SUMMER 2013
ACKNOWLEDGMENTS

The National Center on Quality Teaching and Learning adapted this guide from the Classroom Assessment Scoring System™ (CLASS™) Implementation Guide (Hamre, Goffin, & Kraft-Sayre, 2009).

TABLE OF CONTENTS

1. Investing in Effective Teacher–Child Interactions in Head Start Programs ................................................... 2

2. Using the CLASS to Inform Program Improvement .............................................................................................. 8
   Using the CLASS as a Program Support Tool
   Using the CLASS to Assess Individual Classrooms

3. Frequently Asked Questions on CLASS Implementation ....................................................................................... 13

4. Case Studies ............................................................................................................................................................ 17
   University Settlement Early Childhood Center:
   Sustaining In-Depth Program Support and Professional Development

   Bright from the Start: Georgia Department of Early Care and Learning (DECAL):
   Planning Statewide Implementation of the CLASS

Appendix: Resources and References .......................................................................................................................... 21
Introduction

The National Center on Quality Teaching and Learning (NCQTL) identifies, develops, and promotes evidence-based teaching and learning practices to help Head Start programs achieve the best possible outcomes for young children. NCQTL develops resources for teachers and others who work with young children to make evidence-based practices everyday practices.

Children enrolled in early childhood classrooms need engaging interactions and environments to support their learning. Quality teaching and learning occurs within the context of supportive relationships and intentional learning activities. And quality teacher–child interactions are essential for children's social and academic development and learning.

This guide, Improving Teacher-Child Interactions: Using the CLASS™ in Head Start Preschool Programs, is designed to help programs use an evidence-based tool to achieve such relationships and interactions. The tool is called the Classroom Assessment Scoring System (CLASS), and research shows that it can help programs build and measure the effectiveness of teacher–child interactions and environments.

The guide includes a description of how the CLASS relates to school readiness, directions on how to use the CLASS to collect and apply data, and case studies that show how programs have used the CLASS for support and improvement.

It does not address ways the Office of Head Start (OHS) uses the CLASS for monitoring purposes. For information on this topic, OHS has published a list of frequently asked questions and answers about the use of CLASS in monitoring reviews. See http://eclkc.ohs.acf.hhs.gov/hslc/sr/quality/class.

We hope programs find this resource helpful in using the CLASS to improve child outcomes.

The CLASS and School Readiness

NCQTL resources are organized around a Framework for Effective Everyday Practice: Supporting School Readiness for All Children.

This framework (shown in Figure 1) represents four integral components of quality teaching and learning: providing engaging interactions with children; choosing and implementing research-based curricula and teaching practices; using ongoing assessment of children's skills; and individualizing teaching and learning. These elements correspond, respectively, to a house foundation, two pillars, and a roof. When connected with one another, they form a single structure—the House Framework—that fosters children's learning and development.

Effective, engaging interactions and environments form the foundation for all learning in early childhood classrooms. These high-quality preschool practices include a well-organized and managed classroom, social and emotional support, and instructional interactions and materials that stimulate children's thinking and skills. Such interactions involve the back-and-forth exchanges among teachers and children that occur every moment of the day. While effective interactions are critical for children's school success, they are
only one piece of effective programs. The House Framework shows that to ensure positive outcomes for children, programs must work toward improvements in other areas as well—such as the use of research-based curricula and teaching practices and ongoing child assessment.

Improving the effectiveness of teaching practices in Head Start programs occurs in the context of broader program improvement efforts. The Office of Head Start (OHS; see Program Instruction 11-04) describes four steps to support school readiness in Head Start programs. The first step is to adopt and align established child goals. The second step is to create and implement a plan of action to achieve these goals. Steps three and four involve tracking progress and determining priorities for improvement.

As a part of the second step of creating and implementing a plan, Head Start preschool programs should attend to the effectiveness of teacher–child interactions in the classroom. For example, a school readiness goal to promote social and emotional development may be that children engage in and maintain positive adult–child relationships and interactions. To meet this goal, the action plan may include assessment of the daily interactions between teachers and children, and the CLASS is one tool to help support this work (Pianta, La Paro, & Hamre, 2008). Such an assessment supports programs in using observations to focus professional development (PD) plans. NCQTL resources, such as the in-service suites and crosswalk linking these to CLASS, can be included as part of the plan.

Why use the CLASS to assess classroom interactions?

In the 2007 reauthorization of Head Start (The Improving Head Start for School Readiness Act), Congress directed OHS to include as part of its program monitoring process a reliable and valid tool that assesses teacher–child interactions. The use of the CLASS fulfills this mandate by providing a reliable and valid assessment of three broad domains of effective interactions—Emotional Support, Classroom Organization, and Instructional Support—that characterize children’s experiences in early childhood education (ECE) classrooms. Research findings from more than 3,000 classrooms demonstrate that children in classrooms with higher CLASS ratings realize greater gains in social skills, language, early literacy, and math development.

Five overarching conclusions have emerged from over a decade of research on CLASS.

(See pages 6–7 for a more detailed discussion.)

1. Effective teacher–child interactions are an active and crucial ingredient for children’s social and academic development.
2. Children in ECE settings are not consistently exposed to effective teacher–child interactions.
3. Initial evidence suggests thresholds for effective teacher–child interactions, as measured by CLASS, in promoting children’s learning and development.
4. Quality improvement efforts that focus explicitly on teacher–child interactions maximize impacts for children.
5. Carefully designed and implemented professional development support can improve the quality of teacher–child interactions.

Although OHS is using the CLASS for monitoring purposes, programs are not required to use it for their own program improvement efforts. Nonetheless, many programs are interested in collecting their own CLASS data and developing PD plans that are aligned with the tool.

Research shows that all children benefit from high-quality instruction and classroom interactions, regardless of language status, race/ethnicity, or special needs (August & Shanahan, 2006; Bowman, Donovan, & Burns, 2001). Programs that decide to use the CLASS as a resource for program improvement should note that the tool has been used to assess classroom quality across diverse populations, including dual language learners (DLLs), children from migrant families, tribal populations, and children with special needs and diverse cultural backgrounds (Downer et al., 2011). Additionally, while the CLASS was not designed for use in family child care settings because it describes general
adult–child interactions, the observation protocol can be modified for use in these kinds of Head Start settings. The CLASS should be used in conjunction with other tools and methods, such as aggregate child assessment results, and supports that are important to DLLs, inclusion classrooms, classrooms with children of diverse cultural backgrounds, and family child care settings. No single tool or system can assume sole responsibility for improving the quality of ECE programs or even the quality of one aspect of these programs, such as teacher–child interactions. The CLASS is an evidence-based tool that can be an important part of efforts to promote quality implementation of Head Start services. The goal of this document is to provide programs with guidance on these activities.

The rest of this chapter provides an overview of the CLASS and research findings. The next chapter focuses on how programs can self-assess using evidence-based tools that measure effective teacher–child interactions, how they can use CLASS data to inform program efforts, and how they can use CLASS data collected at the program level to support program and classroom professional development. The final two chapters offer answers to frequently asked questions about CLASS implementation and case studies.

Overview: What does CLASS measure?

The CLASS focuses on the quality of classroom interactional processes. This focus differs from other measurement tools that address the content of the physical environment, available materials, or a specific curriculum. For the CLASS, the physical environment (including materials) and curriculum are important considerations in the context of how all teachers and other staff in the classroom put them to use in their interactions with children.

The CLASS is organized to assess three broad domains of interactions among teachers and children: Emotional Support, Classroom Organization, and Instructional Support. As Table 1 shows, each domain includes specific dimensions. Collectively, these 10 dimensions assess the extent to which teachers are effectively supporting children’s development, both social and academic.

Table 1: CLASS Domains and Dimensions

<table>
<thead>
<tr>
<th>Emotional Support</th>
<th>Classroom Organization</th>
<th>Instructional Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Climate</td>
<td>Behavior Management</td>
<td>Concept Development</td>
</tr>
<tr>
<td>Negative Climate</td>
<td>Productivity</td>
<td>Quality of Feedback</td>
</tr>
<tr>
<td>Teacher Sensitivity</td>
<td>Instructional Learning Formats</td>
<td>Language Modeling</td>
</tr>
<tr>
<td>Regard for Student</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perspectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain</td>
<td>Dimension</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Emotional Support</td>
<td>Positive Climate</td>
<td>Reflects the overall emotional tone of the classroom and the connection between teachers and students. Considers the warmth and respect displayed in teachers’ and students’ interactions with one another as well as the degree to which they display enjoyment and enthusiasm during learning activities.</td>
</tr>
<tr>
<td></td>
<td>Negative Climate</td>
<td>Reflects the level of expressed negativity such as anger, hostility, or aggression demonstrated by teachers and/or children. Low scores represent fewer instances of expressed negativity in the classroom.</td>
</tr>
<tr>
<td></td>
<td>Teacher Sensitivity</td>
<td>Encompasses teachers’ responsivity to students’ needs and awareness of students’ level of academic and emotional functioning. The highly sensitive teacher helps students see adults as a resource and creates an environment in which students feel safe and free to explore and learn.</td>
</tr>
<tr>
<td></td>
<td>Regard for Student Perspectives</td>
<td>The degree to which the teachers’ interactions with students and classroom activities place an emphasis on students’ interests, motivations, and points of view, rather than being very teacher-driven. This may be demonstrated by teachers’ flexibility within activities and respect for students’ autonomy to participate in and initiate activities.</td>
</tr>
<tr>
<td>Classroom Organization</td>
<td>Behavior Management</td>
<td>Encompasses teachers’ ability to use effective methods to prevent and redirect misbehavior by presenting clear behavioral expectations and minimizing time spent on behavioral issues.</td>
</tr>
<tr>
<td></td>
<td>Productivity</td>
<td>Considers how well teachers manage instructional time and routines so that students have the maximum number of opportunities to learn. Not related to the quality of instruction but rather to teachers’ efficiency.</td>
</tr>
<tr>
<td></td>
<td>Instructional Learning Formats</td>
<td>The degree to which teachers maximize students’ engagement and ability to learn by providing interesting activities, instruction, centers, and materials. Considers the manner in which the teachers facilitate activities so that students have opportunities to experience, perceive, explore, and utilize materials.</td>
</tr>
<tr>
<td>Instructional Support</td>
<td>Concept Development</td>
<td>The degree to which instructional discussions and activities promote students’ higher-order thinking skills versus a focus on rote and fact-based learning.</td>
</tr>
<tr>
<td></td>
<td>Quality of Feedback</td>
<td>Considers teachers’ provision of feedback focused on expanding learning and understanding (formative evaluation) and not correctness or the end product (summative evaluation).</td>
</tr>
<tr>
<td></td>
<td>Language Modeling</td>
<td>The quality and amount of teachers’ use of language-stimulation and language-facilitation techniques during individual, small-group, and large-group interactions with children. Components of high-quality language modeling include self and parallel talk, open-ended questions, repetition, expansion/extension, and use of advanced language.</td>
</tr>
</tbody>
</table>
Research findings on the CLASS

Research on the CLASS provides evidence about the nature of teacher–child interactions in ECE settings and how these interactions promote children’s social and academic development. In this section of the guide, we share five overarching conclusions that have emerged from the research.

1. **Effective teacher–child interactions are an active and crucial ingredient for children’s social and academic development.**

   Children in classrooms with higher CLASS ratings experience greater gains in academic achievement and social skill development during the school year (Howes et al., 2008; Mashburn et al., 2008). Classrooms in which teachers develop positive relationships with children and are sensitive to children’s needs (as measured by the CLASS domain Emotional Support) foster children’s social development. Classrooms in which teachers effectively manage behavior and take an active role in creating learning opportunities enhance children’s self-regulatory skills and help them get the most out of each day they spend in the classroom (as measured by the CLASS domain Classroom Organization). Children in classrooms in which teachers offer higher-quality feedback and more consistently support the development of thinking skills (as measured by the CLASS domain Instructional Support) show more academic progress in both pre-k and kindergarten than do their peers who receive lower levels of these supports. When ECE teachers provide effective emotional, organizational, and instructional supports, children are more successful as learners and demonstrate improved social and academic outcomes.

2. **Children in early childhood education settings are not consistently exposed to effective teacher–child interactions.**

   National data collection in state pre-k and Head Start programs indicates that Emotional Support and Classroom Organization typically are at moderate to high levels of quality. Instructional Support, however, is typically at a low level of quality. Similar findings have been replicated in several large national studies of ECE settings, including state pre-k, Head Start, and community-based child care centers (Maxwell et al., 2009). Recent data collected through Head Start monitoring efforts reveal similar trends.

   As discussed in detail in the next chapter, Head Start programs can use their CLASS data to help target their professional development (PD) efforts to improve the quality of these interactions.
3. **Initial evidence suggests thresholds for effective teacher–child interactions, as measured by CLASS, in promoting children’s learning and development.**

Recent research from national data of state pre-k programs, including many Head Start programs, suggests that classrooms need to have fairly high levels of Emotional Support and Classroom Organization, at or around a score of 5 on the CLASS, to promote positive social development and reduce problem behaviors. The threshold for quality in Instructional Support appears to be a bit lower, however. For example, when classroom interactions are characterized by CLASS Instructional Support scores of 3 or above, children demonstrate greater gains in early academic and language skills (Burchinal et al., 2010). More than a third of Head Start grantees monitored in FY 2011 were below this threshold, based on 2010–11 monitoring data. This also means that relatively small differences in the quality of teachers’ instructional interactions with children (promoting concept development, providing good feedback, stimulating language and conversations) may be especially important for helping children learn more. This is not to say that programs should strive just for a score of 3 on Instructional Support. Rather, programs should aim high to increase effective instructional interactions.

4. **Quality improvement efforts that focus explicitly on teacher–child interactions maximize impacts for children.**

While basic elements of program quality such as teacher education, class size, and classroom materials are important, their significance is measured in part by the extent to which they facilitate effective teacher–child interactions. Research shows that the classroom interactional components measured by the CLASS are more powerful predictors of children’s development and learning than are structural elements of program quality (Mashburn et al., 2008). Factors such as teacher qualifications and class size, though important, are not sufficient in and of themselves to ensure children’s positive development.

5. **Carefully designed and implemented professional development support can effectively improve the quality of teacher–child interactions.**

Research evidence from a variety of sources now shows overwhelmingly that high-quality and targeted PD programs can help teachers improve the quality of their interactions with children and that these improved interactions, in turn, foster greater social and academic development (Bierman et al., 2008; Domitrovich et al., 2009; Hamre et al., 2012; Pianta, Mashburn et al., 2008; Raver et al., 2008).

PD supports intended to improve the effectiveness of teachers’ interactions with children must be developed and chosen carefully to ensure success. The most typical form of PD experienced by early childhood educators continues to be brief workshops; there is little to no evidence that these efforts will lead to enduring changes in teachers’ interactions with children (Zaslow et al., 2010).

For Head Start leaders, the studies noted above provide compelling evidence that continuous improvement efforts must focus directly on the quality of teachers’ interactions with children to positively affect children’s learning gains at the individual teacher or program level. The remainder of this guide provides support for this work.
Head Start programs across the country want to know how to use CLASS data to inform program improvement. Some programs are also interested in using the CLASS to conduct their own assessments of the effectiveness of teacher–child interactions. The following chapters provide guidance regarding how to collect their own CLASS data in ways that can inform continuous program improvement efforts. The Office of Head Start (OHS) has published a list of frequently asked questions and answers regarding the use of CLASS in monitoring reviews, available at: http://eclkc.ohs.acf.hhs.gov/hslc/sr/quality/class.

Head Start programs may use CLASS observations in two ways: program planning and support as well as individual classroom support. At each level, a key distinction is whether the CLASS is used to evaluate a sample of classrooms served by a program versus collecting data in each classroom. Sampling allows program administrators to identify agency-wide strengths and develop targeted professional development plans at the program level. In contrast, individual classroom data can determine how to individualize PD for each teacher or teaching team based on their strengths and challenges.

CLASS data can also be used across time—at the program or classroom level—to assess change in the effectiveness of teacher–child interactions. For example, data can be collected each year before and after a sustained interactions-focused PD intervention to assess teacher progress on dimensions and to examine the effectiveness of the intervention.

USING THE CLASS AS A PROGRAM SUPPORT TOOL

Programs interested in using the CLASS as a program support tool should consider several factors. One of the first steps is to clearly articulate overall goals for CLASS data collection. For example, is your goal to improve teacher–child interactions program-wide or on a classroom-by-classroom basis? Your answer should guide your ultimate data collection plan and lead to several related decisions.

**General principles to consider:**

1. The more CLASS observation cycles you are able to obtain and aggregate, the more stable your estimates of typical classroom interactions will be.

2. In most cases, a two-hour observation (four CLASS cycles) provides a reliable estimate of the overall status of teacher–child interactions in a classroom.

3. There typically is more variance in CLASS scores within a program than there is between programs. This means you have to assess a significant portion of classrooms within any one program to get a reliable estimate of that organization. Sampling a greater number of classrooms will provide a more accurate assessment of the program level.

4. Avoid conflict of interest between the observers and the classrooms they observe.

5. Even if all observers are CLASS certified, there will be small, systematic differences between their scoring. The best way to minimize any potential “observer effects” is to randomly assign observers to classrooms within an organization (program, school, grantee, etc.).

6. The CLASS observer should speak the language most common in the classroom.
Do we need to observe every classroom or a sample of classrooms?

The decision to observe in every classroom or in a sample of classrooms is based on the goals of data collection. When programs want to make PD decisions for individual classrooms or teachers, each classroom should be observed, as CLASS scores can differ greatly from classroom to classroom. When programs want to monitor their overall quality or set program-wide goals for improvement, observing in a sample of classrooms may be appropriate.

While observing in every classroom is ideal, a program’s infrastructure—specifically, staffing and budget constraints—may not provide enough capacity for this level of observation. In this case, observing in a sample of classrooms is a good option. Table 3 compares these two approaches.

<table>
<thead>
<tr>
<th>OBSERVING EVERY CLASSROOM</th>
<th>OBSERVING A SAMPLE OF CLASSROOMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>How:</strong> Collect CLASS data in every classroom in the program. For each classroom, observe a minimum of four cycles.</td>
<td><strong>How:</strong> Collect data in a sample of classrooms within the program. Sample should represent how teachers and children are distributed within the program. For example, if 50% of a program’s classrooms are urban, 25% suburban, and 25% rural, the percentage of classrooms sampled in each setting should reflect these percentages.</td>
</tr>
<tr>
<td>Purpose is to use data to make individual decisions about what PD options are most appropriate for each teacher or teaching team.</td>
<td>Purpose is to use data to make program-wide PD decisions. This is often a good first step for longer-term improvement projects.</td>
</tr>
<tr>
<td><strong>Examples of uses for PD:</strong></td>
<td><strong>Examples of uses for PD:</strong></td>
</tr>
<tr>
<td>• Share individual classroom results with teachers, describing strengths and areas for growth, and support teachers and teaching teams in setting goals for their interactions.</td>
<td>• Share aggregated CLASS score results with teachers, noting that results reflect the average of a sample of classrooms and not individual classrooms.</td>
</tr>
<tr>
<td>• Direct teachers and teaching teams to different PD options based on strengths and challenges. For example, direct the most intensive PD support to lower scoring teams.</td>
<td>• Base program-wide PD on sample results. For example, if Instructional Support scored lower, provide emphasis in this area during PD.</td>
</tr>
<tr>
<td>• Provide individualized coaching for teachers or teaching teams based on strengths and challenges.</td>
<td>• There will not be enough data to inform individualized PD at the classroom level.</td>
</tr>
<tr>
<td></td>
<td><strong>Additional recommendation:</strong></td>
</tr>
<tr>
<td></td>
<td>Move toward more intensive, individually targeted PD, based on data collection in every classroom when possible.</td>
</tr>
</tbody>
</table>
How do we set goals for quality using CLASS?

Programs may be interested in setting goals for the quality of their classrooms in terms of the effectiveness of teacher–child interactions. The CLASS manual establishes criteria for low, mid, and high ranges of interactions, but these may not be the most useful cut-points for program support purposes. Two important pieces of information help in determining goals for program quality:

• The levels of quality that are sufficient to promote positive child outcomes.
• How data on CLASS scores are distributed across a sample of classrooms in a program.

As discussed above, initial evidence suggests using CLASS scores of at least a 5 on Emotional Support and Classroom Organization and a 3 on Instructional Support. But it is important to remember that these points are based on an initial study, and more data are needed to provide definitive recommendations around cut-points.

Just as it is important to use research to help identify sufficient levels of teacher–child interactions, it is also important to examine patterns in program-level data and to individualize CLASS goals, as appropriate. If, for example, there are no classrooms in a program scoring above a 3 on Instructional Support, it will likely make sense to establish short-term goals for improvement at somewhat lower levels, while keeping in mind longer-term goals at higher levels. Short-term goals may include identifying very targeted strategies as areas of focus, such as increasing open-ended questions during book reading, centers, and meals, rather than multiple strategies across several dimensions. However, even when programs are meeting the quality thresholds, they should still be working toward improvement.

It is also necessary to understand that changing teacher–child interactions takes time and practice. For example, research demonstrates that after intensive, CLASS-focused PD lasting approximately 10 months, teachers increase their CLASS scores by an average of one-half to one point on the CLASS dimensions (Pianta, Mashburn et al., 2008). Goals for improvement on CLASS scores must be based on appropriate expectations and backed by ongoing, intensive, interactions-focused PD.

How should data be used to inform program improvement?

Data can be used to inform the allocation of resources, such as coaching and in-service trainings. Some programs may be in greater need than others. In addition, some domains of interactions as measured by the CLASS, such as Emotional Support, may be in need of particular attention.

Consider the following data from a Head Start program with six centers (Figure 2). Based on these data, the program director would conclude Center 3 is doing generally well, Center 1 is struggling primarily with Instructional Support, and Center 4 is struggling across all domains. Noting these and other patterns can help determine a plan of action. In this example, the education manager may decide to work closely with the director at Center 4 to conduct a more thorough needs assessment and Professional Development plan for each teacher at the center. For Center 6, the education manager might focus work on understanding why Emotional Support is particularly challenging and deploying coaching resources appropriately. The focus on Center 1 may be on Instructional Support. Center 3 may serve as a model center, with the center director leading a director group that explores ways to support teachers' classroom practices.

Recommendations for sharing data with program staff

• Make sure the program staff has enough information about the tool to understand results.
• Provide results within the context of national/regional averages to aid in interpretation.
• Data can inform program-wide areas of strength and areas with room for growth.
How should CLASS data be reported and shared at a program level?

CLASS data collected by individual programs can be shared with administrators and other stakeholders such as governing bodies, Policy Council, or Tribal Council to provide an overview of the quality of interactions in the program. For these broad purposes, sharing data at the domain level (e.g., Emotional Support) is likely sufficient. It is often helpful to compare these program-level data to national or regional averages. Data collected by a program may also be used to identify individual centers in need of additional support. In these cases, it may be helpful to share dimension-level data (e.g., Teacher Sensitivity) because these dimensions provide more specific information about the types of interactions that may be in need of improvement.

USING THE CLASS TO ASSESS INDIVIDUAL CLASSROOMS

Programs that are interested in using the CLASS to help develop individual PD plans for teachers should consider several pertinent factors.

How long and when should each classroom be observed?

Each classroom should be observed for at least two hours (four CLASS cycles). CLASS observations should last 20 minutes, and then the observer should take no more than 10 minutes to code. In rare cases, a cycle may terminate after 10 minutes of observation. In the majority of these cases, termination after 10 minutes occurs when the children transition from free play to recess. If the observation has lasted more than 10 minutes before recess, the observer may score that cycle. The number of cycles the observer should complete depends on the goal for the program and how the program intends to use the CLASS data. If you are interested in change over time, observe in the fall and spring of the same school year or at the same time the next year. It is best to avoid the first and last weeks of the school year.
How should CLASS data be shared at the teacher level?

Remember that the CLASS assesses classroom interactions, not a specific teacher. When sharing observations at the classroom level, include all teachers who work in that classroom.

It is extremely important that teachers have sufficient knowledge about the CLASS prior to receiving feedback from an observation. Feedback on a teacher’s interactions with children related to Instructional Learning Formats, for example, will have greater meaning when the teacher has a clear understanding of what specific behaviors are noted in this dimension. The CLASS Manual and the CLASS Dimensions Guide provide information that can be helpful to teachers in understanding the specific behaviors they can use to improve their practice.

We generally recommend sharing results with individual teachers at the dimension level and describing patterns and examples of teacher–child interactions rather than specific scores. As much as possible, include specific, behavioral notes from the actual observation so that the teachers can really understand what the CLASS assessed in their classrooms. For example, it may not be helpful to tell a teacher that when observed, he or she received a score of 3 on Concept Development. The teacher might immediately focus on whether a 3 is good or bad rather than identifying behaviors that can help “move up” in that dimension regardless of the specific score obtained. Rather than sharing specific scores, a coach or consultant may identify a specific dimension of focus with the teacher. Then, based on observed interactions, engage the teacher in a discussion about specific strategies the teacher used, such as open-ended questions to promote analysis and reasoning, how children responded, as well as how the strategies could be extended to further understanding. Sharing scores with teachers presents the risk of becoming bogged-down in a discussion of the number of behaviors rather than on the specific behaviors that are critical targets for change. For these reasons, we recommend discussing behaviors instead of sharing scores with the teachers.

There may be times, however, when sharing scores with teachers is required or desired. In these instances, it is important to provide a good description about what was observed as well as ways to understand their scores. You may want to use the CLASS Dimensions Guide to help teachers understand their scores and why they are important. To promote more careful listening and openness, consider using individual meetings with teachers to share information about their strengths and areas of challenge.

Should those who are sharing results with teachers talk to observers? If so, what should be the focus of these interactions?

When sharing results with teachers, include the more behavior-specific observations obtained by observers during ratings. If someone other than the observer will be sharing results—such as an education manager, director, or coach—the observer must first share information with these individuals in a one-way communication process that happens shortly after the observation is completed. To avoid potential conflicts of interest and reduced objectivity, these individuals should not talk to observers about their knowledge of the teacher prior to the observation. The communication from the observer may take the form of written notes that provide context for the results and are not intended for sharing with the program or teacher—e.g., the observer may note that there was a marked difference between the lead teacher and assistant teacher in terms of the Emotional Support in the classroom.
**Who should conduct CLASS observations?**

The answer to this question depends in part on how the data will be used. For program support purposes, it is essential to have as little bias as possible in the results. Thus, if feasible, observers should be free of conflict of interest with the classrooms they will be observing. It is difficult, for example, for a program director to collect program support data for the classrooms he or she supervises. Ideally, programs may ask an outside group (e.g., observers from a nearby Head Start center) to conduct classroom observations for program support or hire independent, part-time staff specifically for this purpose. The ECE specialists are a resource to help explore options, such as other managers, coaches, and center directors.

**How does a program get trained to implement CLASS?**

ECE specialists in Head Start's technical assistance state-based system are certified CLASS affiliate trainers and are able to train a group of observers for your program. Observers participate in a two-day intensive overview of the CLASS and learn how to use the tool reliably. This training is followed by an online reliability test for certification as a certified CLASS observer. Consult with your ECE specialists and/or NCQTL to learn about available CLASS training resources.

**How many observers do we need?**

Programs can answer this question after considering factors such as the length of the observation period, the number of classrooms within each program, the times of day being observed, and other logistical decisions. For larger grantees a good rule of thumb is to have about 10% to 15% more observers than will be needed and to allow for staff turnover and the fact that not everyone will pass the reliability test.

For example, if a large grantee wished to observe 16 classrooms at a standardized time of day (e.g., 8:30–10:30 a.m.) and specific time of year (e.g., the month of October), it would need to staff observers to complete approximately four observations per week for four weeks. Assuming 20 working days, it would be reasonable to have a staff of about two observers who would be available to observe two days per week each. Many programs using the CLASS assume that an observer will complete one classroom observation per day, which includes travel time and writing up detailed reports.

In addition, determination of the number of observers must also take into consideration the languages spoken in the classroom. Observers should speak the language(s) that the teacher uses for instruction. Careful planning is needed to ensure valid administration and use of the CLASS.

**What data should be collected?**

Data collection includes a recording of CLASS scores and behavioral notes obtained for each classroom observed. In addition, CLASS scoring sheets include places to record some information about the context for the observations (e.g., number of teachers). In many cases, CLASS observers will also want to record some notes about each classroom at the CLASS dimension level and to share this information later with teachers and administrators. Some programs have observers write brief summary statements about what they observed for each dimension across the observation period. When results are shared with teachers, they include the observers’ statements.
How do we assign observers to classrooms?
The best way to minimize any potential “observer effects” is to randomly assign observers to classrooms within any organization (program, school, grantee, etc.) so that observers are selected by chance. Even if all observers are CLASS-certified, there will be small, systematic differences in their scoring. Some observers may tend to give slightly higher scores, while others may tend to be slightly more critical. Although slight differences fall within the threshold for “reliability,” collectively, they can produce inaccurate results. To ensure more reliable results, it is preferable to send a team of observers to centers rather than assigning one observer to each center. In addition, be sure the observer is proficient in the primary language(s) of the classroom.

Do we need to send more than one observer to each classroom?
One of the best ways to improve the reliability of CLASS scores is to have at least two observers make ratings of the same classroom. Although the associated expense of “double coding” may be prohibitive, at least a portion (from 5% to 15%) of classroom observations should be double-coded to assess reliability. These data will help you communicate to stakeholders about the fairness of the tool in practice.

How do factors such as time of day and year affect CLASS scores?
Evidence suggests that observations completed during the first 30 minutes of the day may yield lower ratings on some aspects of teaching, such as instructional practices, than observations conducted during the rest of the day. This is perhaps because the beginning of the day is typically used to complete transition activities such as eating breakfast and unpacking bags. There is also some evidence that more social aspects of the classroom environment, such as classroom climate, may decrease slightly over the course of the day. This may reflect teachers and children getting tired as the end of the day approaches. These variations tend to be quite small, however.

Other aspects of teaching practice, such as instruction, seem to be more consistent after the first 30 minutes. There may be good reasons to observe during the beginning of the day, for example, to observe the way a teacher handles transition routines. If scores on observations will be used to compare teacher–child interactions across classrooms in a program, we recommend standardizing the observational protocol to either include or exclude these first 30 minutes.

Findings from observations throughout the year indicate that, by and large, there is consistency in classroom interactions across the year. There are some indications that scores are lower at the very beginning of the year, around the winter holidays, and at the very end of the year. Overall, there is a tendency for scores to decline from fall to spring. For these reasons, if possible, avoid the first and last weeks of program’s calendar year as well as days leading up to the winter holidays if your objective is to obtain scores that accurately represent typical practice. Data are not yet available on year-round programs to know if fluctuations in CLASS scores appear over the summer.

How often should we conduct CLASS observations if we are interested in detecting change over time as a result of professional development?
When you work with teachers to improve the effectiveness of their interactions, the CLASS tool can be used to determine whether the intervention was successful. Because of the tendency for CLASS scores to decline from fall to spring in classrooms without intervention (e.g., PD), it is very important to have some comparison data such as from a group that did not participate in PD, or CLASS data for that classroom from the previous year in which no PD was offered. This information will help you interpret results from fall to spring observations intended to show improvements in teachers’ practice.

For example, a director may look at CLASS data from fall to spring and be disappointed to see no improvement, despite substantial efforts to improve the quality of teacher–child interactions through coaching and workshops. Without knowing what would have happened without these improvement efforts, it is impossible to know whether the interventions didn’t work or whether they represent an improvement over the typical situation in that program.
It may be that without the additional PD, classrooms would have declined in quality from fall to spring. The only way to clearly interpret these kinds of data is to have a comparison group within the same organization that did not receive the additional supports. If this is not possible, it may also be feasible to use data collected from a previous year to demonstrate the impact of PD.

Research using the CLASS tool has shown that targeted, intensive PD focused on improving teacher-child interactions can show effects in as little as six months. But programs should think carefully about their expectations for change and take into consideration the intensity and intended duration of the intervention.

Is the CLASS a valid measure of classroom interactions for dual language learners (DLLs)?

The National Center for Early Development and Learning (NCEDL) has conducted studies to examine whether the CLASS is a valid measure of classroom interactions for DLLs. NCEDL collected data from more than 700 state-funded pre-k and 700 kindergarten classrooms across 11 states. The classrooms in these studies were linguistically diverse: 15% of the pre-k classrooms had 50% or more children who were identified by their school programs as having "limited English proficiency" at the beginning of the year; 50% of pre-k classrooms had children who spoke Spanish; and 21% of classrooms had children who spoke a language other than English or Spanish.

In a recent study, NCEDL separated data on pre-k classrooms into “no DLL” classrooms, “mid DLL” classrooms (proportion of DLL children in classroom was 0% to 50%), and “high DLL” classrooms (proportion of DLL children in the classroom was greater than 50%; Downer et al., 2011). Its analysis reveals that the CLASS functions very similarly across these settings and that it validly assesses the quality of teacher-child interactions, regardless of the proportion of DLLs in the classroom. Furthermore, CLASS mean scores were not significantly different based on the percentage of DLLs in the classroom. Observers were a combination of English-speaking and bilingual in Spanish/English.

This study also examined whether the CLASS is predictive of child outcomes in classrooms with DLLs. Results showed that children in classrooms with higher CLASS scores made greater academic and social progress, regardless of the child’s individual language abilities or the proportion of DLL children in the classroom. Furthermore, the strength of the association between teacher-child interaction quality and children’s outcomes did not differ based on children’s language status (Downer et al., 2011).

Taken together, these findings suggest that the CLASS functions well as an assessment of the quality of teacher-child interactions in classrooms with language diversity and that the CLASS predicts gains in the school readiness skills of DLL children.

Should I modify the CLASS for use in classrooms with DLLs or multilingual children?

Many people have suggested that it would be helpful to make changes to the CLASS instrument to accommodate classrooms with many non-English speaking children. People have expressed the most concern about the Instructional Support dimensions of the CLASS. This is because so many of the indicators and behavioral markers of the Instructional Support dimension are language-based.

The CLASS should be used in the same way across classroom settings. The dimensions should not be tailored in terms of coding to try to accommodate differences across settings, including the language or special needs of children. The CLASS is a standardized tool, and this is one of its advantages. If modifications were made, it would be difficult to make comparisons within and across programs.

It is important to remember that the CLASS measures interactions among teachers and children that promote development—not simply teacher behaviors. For example, some people have expressed concern that the indicator of Advanced Language under the Language Modeling dimension may not work in classrooms with a large percentage of DLLs. CLASS training takes these variables into consideration and emphasizes that decisions about what is advanced are based on the best information available in the classroom and on the observation of the interaction. By placing an emphasis on the context and the interaction, the CLASS does attend to these relatively nuanced issues in classrooms.
How should we decide which language a CLASS observer needs to speak?

If instruction occurs in more than one language, then the observer must be bilingual and speak the languages that the teacher uses for instruction. This will help observers to accurately assess the quality of interactions in DLLs’ primary languages. For example, if observers coding in DLL classrooms are bilingual in Spanish, they will be able to pick up on back-and-forth exchanges in instances of Spanish or English for children who are DLLs.

What is interaction-focused and CLASS-focused professional development?

Interaction-focused PD is highly targeted to help teachers understand and practice effective interactions. Rather than focus on generalities about teachers’ practices, interaction-focused PD drills down to specific teacher behaviors that support children’s learning. For example, how does the teacher facilitate language and foster children’s thinking skills? Other types of PD offered to teachers that address the structure of the classroom environment, curricula, and child assessment may not be explicitly and primarily focused on teacher–child interactions. CLASS-focused PD is a subset of interaction-focused PD, and it is specifically designed to provide PD supports to teachers using the lens of the CLASS observational tool.

How can NCQTL resources be used to improve interactions in PD?

Head Start programs can receive additional assistance with implementing their program improvement from NCQTL and ECE specialists. We are creating a series of in-service suites that describe specific ways to promote engaging, social, organizational, and instructional interactions to improve practice. Each in-service suite includes a PowerPoint presentation, video examples, learning activities, and a Supervisor Guide, Tips for Teachers, and Helpful Resources. These in-service suites are available on our portal on the Head Start Early Childhood Learning and Knowledge Center (ECLKC) website at: http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/center/practice/ISS, or through your Head Start ECE specialists.

Additional NCQTL resources include Practice-Based Coaching, Teachers Learning and Collaborating, and higher education courses. See http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/.
University Settlement Early Childhood Center: Sustaining In-Depth Program Support and Professional Development

Improving classroom quality and outcomes for children involves a strategic and coordinated process of ongoing program support, professional development, and child assessment. The University Settlement’s Early Childhood Center has worked to create an infrastructure for this effort using the CLASS to focus on effective interactions. The center is a Head Start and child care collaboration program located in New York City that serves 158 children in eight classrooms and is funded through the New York City Administration for Children’s Services.

Building Infrastructure

In 2006, the center began the process by hiring a program research consultant to assess structural and process aspects of quality in the center’s classrooms using the Early Childhood Environment Rating Scale-Revised (ECERS-R) and the Arnett Caregiver Interaction Scale (ARNETT). Based on findings, the center partnered with a local university to provide PD focused on emotional support for children. In order to meet its quality improvement goals, the center expanded infrastructure in 2008 by conducting research to:

- Examine curriculum implementation in all its preschool classrooms, implement related PD, and evaluate child outcomes and classroom quality.
- Investigate the relationships among teachers’ professional growth, classroom quality, and child outcomes.
- Improve each of these critical areas of overall program quality.

The program research consultant role became a full-time Director of Early Childhood Programs Evaluation who, together with the Director of Early Childhood Programs and two new research consultants, performed child assessments, measured classroom quality using the ECERS-R, and found links between higher ECERS Interaction scores and children’s receptive vocabulary outcomes. Based on these findings, the center intensified its focus on teacher–child interactions and began implementing the CLASS in fall 2009, due to the instrument’s specific focus on interactions.

Implementing CLASS

In 2009–10, the center conducted fall and spring CLASS assessments with CLASS-focused PD in between, including:

- A one-day CLASS overview to help teachers understand their assessment results, focusing on the lower-scoring Instructional Support dimensions.
- A meeting between each teaching team and the Director of Early Childhood Programs Evaluation to discuss CLASS results, including classroom domain scores, scores in the context of national averages, strengths and challenges, and individualized strategies to improve interactions.

The following two case studies describe how programs have used the CLASS in their program improvement efforts.
The CLASS assessment was performed again in the spring of 2010, and average scores for all classrooms improved, demonstrating the effectiveness of their CLASS-focused PD. The evaluation team analyzed links between CLASS data and child outcome data and found that children in classrooms with higher Positive Climate and Teacher Sensitivity scores tended to have higher emergent literacy scores. Classrooms that were strong on aspects of Instructional Support tended to have children with better literacy skills.

The center then reinvigorated its focus on effective teacher–child interactions by performing CLASS assessments again in the fall of 2010. While scores across domains dropped slightly from the spring, they remained higher than the fall 2009 scores. The center attributed this slight drop to the summer gap in PD, as well as personnel changes.

**Lessons Learned and Next Steps**

- Creating infrastructure with a trained evaluation team is an important first step to improving program quality.
- Obtaining periodic observational assessments is critical to informing the direction of PD. This led to a targeting of Instructional Support interactions as an area of focus based on observation results.
- The center will intensify interaction-focused PD to support teacher growth, including partnering with a local university to provide coaching that targets interactions that support children’s language and literacy development.
- The center’s PD team will continue to provide specific strategies to use quality of feedback, concept development, and language modeling (dimensions of the CLASS Pre-K Instructional Support Domain) in linguistically diverse classrooms, and more seamlessly integrate Instructional Support interactions into curriculum implementation.
- The center will maintain and expand infrastructure to continue this work, as well as continue to examine links between classroom quality—as measured by CLASS, teachers’ professional growth, and child outcomes—and use findings to guide ongoing work.

**Acknowledgments and Contact:** Nina Piros, Director of Early Childhood Programs, and Tonia N. Cristofaro, Ph.D., Director of Early Childhood Programs Evaluation, University Settlement Society of New York.
Georgia’s pre-k program demonstrates that careful, detailed planning results in effective implementation of the CLASS. After working with the Frank Porter Graham Child Development Institute (FPG) at the University of North Carolina-Chapel Hill on a statewide child care study, Bright from the Start: Georgia DECAL made the decision to implement CLASS observations and data collection throughout their 4,200-plus pre-k classrooms during the 2010–11 academic year. Many of Georgia’s pre-k classrooms blend resources with Head Start funding as part of a braided model.

**Preparation for Large Scale Evaluation**

Georgia’s pre-k management and consultants worked consistently to observe every pre-k classroom using the CLASS and to analyze the data and report it in a timely manner. They built infrastructure for evaluation by training their pre-k consultants to assure reliability on the CLASS and by obtaining a system for gathering, entering, and analyzing evaluation data. To ensure efficacy, DECAL carefully planned several quality assurance activities to occur before and during evaluation.

Approximately 40 DECAL pre-k consultants completed their initial training in October 2009, followed by practice CLASS observations during spring 2010. Many of the consultants were able to practice further during the summer of 2010 as DECAL received stimulus funding to conduct a summer pre-k program. For the summer practice sessions, expert CLASS observers double coded with the pre-k consultants to ensure inter-rater reliability and a high level of fidelity to the tool.

Official implementation of evaluation began in the fall of 2010, and the pre-k consultants participated in a two-day enhanced observation skills training to further refine their knowledge. During their observations, the pre-k consultants were often paired to double code, again to ensure the quality and reliability of CLASS observations across 4,200 classrooms. Additionally, the consultants were charged with quarterly calibration or reliability tests. Upon completion of the calibration exercises, a CLASS expert hosted mandatory webinars to further sharpen their skills. Throughout the process, the focus was on ensuring the integrity of observations.

As observations were completed, data were entered into an electronic database for purposes of analysis and comparison to national and statewide averages. At the end of the observations, electronic reports were generated and disseminated to all sites. DECAL is currently training pre-k consultants in working with program directors to share evaluation results with teachers. It is important to note that the report does not cite specific scores but rather indicates low, mid, or high range scores for each domain and dimension. Scores are discussed in the context of statewide averages. Teachers are encouraged to focus on their strengths and challenges and will be provided with specific supports to learn more about the important interactions the CLASS measures.

**Planning Phased Professional Development**

Now that the baseline observations have been completed and compiled into data sets, Georgia DECAL plans to phase in PD over the next three years. The goal of this PD is to continue to raise teachers’ awareness of the importance of their interactions with children and promote an understanding of the behaviors that constitute high-quality interactions, while continuing to support them in improving such interactions—and hence, positively affect student outcomes.
Lessons Learned and Next Steps

• Careful, detailed planning of this statewide evaluation included how to train and support evaluators while also entering and analyzing the data. Overall, the implementation was successful as 92% of the approximately 4,225 classrooms were observed.

• Regular quality assurance activities with CLASS evaluators support fidelity to the tool and accurate collection of CLASS data across classrooms. In the future, DECAL evaluators will perform more of their own double coding to check for and maintain inter-rater reliability.

• DECAL also plans to provide program reports closer to the observation. This was one piece of the implementation that proved more challenging than originally anticipated.

• Observer fatigue was a concern. Therefore, in subsequent years, DECAL will not observe every pre-k classroom but will sample about a third of the state each year.

• DECAL will randomly assign pre-k teachers into one of four professional development programs. FPG will conduct pre- and post- observations to determine the effectiveness of each of the PD programs.

Acknowledgments and Contact: Monica Warren, Director of Georgia’s Pre-K, and Bentley Ponder, Director of Research and Evaluation, Bright from the Start: Georgia Department of Early Care and Learning.
APPENDIX: RESOURCES AND REFERENCES

Organizations focused on supporting the use of the CLASS:

National Center on Quality Teaching and Learning (NCQTL)
Website:  http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching
Email:  ncqtl@uw.edu
Phone:  877-731-0764

Use of the Classroom Assessment Scoring System (CLASS) in Head Start
Website:  http://eclkc.ohs.acf.hhs.gov/hslc/sr/quality/class

University of Virginia, Center for Advanced Study of Teaching and Learning (UVA/CASTL)
Address:  350 Old Ivy Way, Suite 100, Charlottesville, VA 22903
Website:  http://curry.virginia.edu/research/centers/castl
Email:  castl@virginia.edu

Teachstone, LLC.
Address:  105 Monticello Avenue, Suite 101, Charlottesville, VA 22902
Website:  http://www.teachstone.org
Phone:  866-998-8352

University of Minnesota, Center for Early Education and Development (CEED)
Address:  1954 Buford Ave, Suite 425, St. Paul, MN 55108
Website:  http://www.cehd.umn.edu/ceed/
Email:  ceed@umn.edu
Phone:  612-625-3058
References


