

CEMETR-2015-06  
NOVEMBER-2015

# CEME

## Technical Report

The Center for Educational Measurement and Evaluation

Findings from the 2014 North Carolina  
Kindergarten Entry Formative Assessment Pilot

Angela M. Ferrara  
Richard G. Lambert

RICHARD LAMBERT  
CHUANG WANG  
MARK D'AMICO  
SERIES EDITORS

A PUBLICATION OF  
THE CENTER FOR  
EDUCATIONAL  
MEASUREMENT  
AND EVALUATION

**Findings from the 2014 North Carolina Kindergarten Entry Formative Assessment Pilot**

Angela M. Ferrara and Richard G. Lambert

Center for Educational Measurement and Evaluation

The University of North Carolina at Charlotte

November, 2015

In response to the requirements of the Race to the Top-Early Learning Challenge Grant and a mandate from the North Carolina (NC) State Legislature requiring the development of a kindergarten entry assessment, NC Superintendent June Atkinson organized a K-3 Assessment Think Tank in February of 2013. That group, which included K-3 teachers, administrators, parents, early childhood education scholars, and policymakers, were tasked with researching a plan for developing effective, developmentally appropriate, student-centered assessments to improve early childhood learning and instruction in NC public schools. Their final report (K-3 Assessment Think Tank, 2013) was heavily influenced by prior research on the effectiveness of formative assessment on student learning outcomes (Black & William, 1998; William & Thompson, 2007), and recommended the development of a K-3 formative assessment designed to help teachers individualize their instruction. The report further identified six domains of learning the assessment should cover: approaches to learning, cognitive development, socioemotional development, health and physical development, language development, and communication. Based on the recommendations of the K-3 Assessment Think Tank, an assessment design team consisting of current and former educators and early childhood education scholars developed a kindergarten entry formative assessment (KEA) which was piloted in 193 classrooms during the first 60 instructional days of the 2014-2015 academic year. NC intended the pilot to be a first step toward scaling-up to a full statewide K-3 formative assessment process.

The pilot assessment consisted of 10 individual construct progressions housed within the domains of learning identified by the NC K-3 Think Tank:

Table 1.

*Construct Progressions*

<b>Domain</b>	<b>Construct</b>
Approaches to Learning	Engagement in Self-selected Activities
Socioemotional Development	Emotional Literacy
Health and Physical Development	Grip and Manipulation Hand Dominance Crossing Midline
Cognitive Development	Object Counting
Language Development and Communication	Book Orientation Print Awareness Letter Naming Following Directions

Each construct outlined a fundamental skill children generally acquire around the age of kindergarten entry and included a developmentally sequenced list of behaviors describing a child's progression toward skill mastery. As an example, the emotional literacy skill was defined as the ability of students to "communicate about and use strategies to regulate responses to their own emotions" (NC Office of Early Learning, 2014). The progression began with the description of a child who would express a range of emotions, or have an outburst, when responding to an experience. It then incrementally progressed toward the child being able to express their emotions in

appropriate gestures or words, recognize emotions in others, and finally recognize that the same event can cause different emotions in different people. This construct included nine total developmental steps, each including example behaviors the teacher might observe during their normal instruction to identify a child's status along the progression.

Pilot teachers were tasked with documenting evidence of learning, through photographs, videos, audio recordings, student work samples, or anecdotal notes from student observations, in order to place each student's abilities along the construct progressions. Within the overall KEA formative assessment process, this placement on the progression was termed as a current learning status. Teachers uploaded their evidences of learning to an electronic platform developed by Teaching Strategies, LLC<sup>®</sup> which was customized to house the NC KEA pilot content. The online platform acted as a central location for teachers to review the student documentation they gathered and the learning status they assigned to each. Teachers then used that data to inform their instructional decisions in order to individualize instruction based on each student's unique learning needs.

At the end of the 60-day pilot period, teachers were required to review each student's full documentation portfolio and assign a 'final' learning status to all construct progressions. While the intention of the NC Office of Early Learning is for teachers to use the KEA continually and formatively throughout the school year to inform instruction, the pilot period was set for a shortened timeframe so that pilot informed revisions to the assessment could be made prior to statewide implementation in the fall of 2015. This finalization requirement also ensured that pilot teachers would familiarize themselves with all construct progressions, and therefore could provide feedback regarding the content and procedure within all domains of the assessment.

### Study Purpose

Over the course of the NC KEA pilot, we conducted case studies of eight pilot schools selected for their demographic representativeness of the total NC kindergarten student population. These case studies involved classroom observations and semi-structured interviews with pilot teachers, administrators, and instructional coaches. Following the case studies, we conducted an electronic survey that was open to all pilot participants. The purpose of this qualitative study was to gather teacher and administrator perspectives of the NC KEA and determine whether:

1. changes were necessary to the professional development provided for the assessment.
2. the content of the learning progressions were developmentally appropriate.
3. the documentation process worked within teachers' routine classroom instruction.
4. teachers found the assessment's data useful in driving their instructional decisions.
5. the assessment aligned with the curriculum and assessment practices currently in place within their district and school.
6. there were characteristic patterns between schools and/or classrooms where the assessment was being implemented with fidelity and those where it was not.

Immediately following the pilot, we conducted a systematic, quantitative analysis of the electronic evidences teachers entered to the electronic platform. The purpose of this analysis was to determine:

1. the type of evidences teachers gathered and entered.
2. the number of evidences entered per student.
3. the number of evidences entered per construct.

4. whether the evidences entered were useful for driving instructional decisions.
5. whether there was a change in the number and/or usefulness of evidences entered over time.

### Qualitative Methods and Data Sources

#### Case Study School Selection

The NC Board of Education divides all public school districts into eight state board of education (SBE) regions. In order to allow for potential regional variations in training, curriculum, etc., we selected one school from each SBE region. Once we divided participating schools by SBE region, we gathered school and student level demographics for each from the National Center for Educational Statistics, including: school urbancentric locale, student-teacher ratio, percent English language learners (ELL), percent minority students, percent of students participating in the free or reduced lunch program (FRL), and percent of students with an individualized education plan (IEP). The same demographics were compiled for each SBE region as a whole and compared to the pilot schools. We selected the school that was most closely representative of their SBE region as the case study site. The following table illustrates the resulting sample demographics and their comparison to the NC kindergarten student population as a whole.

Table 2.  
*Demographics*

	State	Sample
<b>Urbanicity</b>	Diverse	Diverse
<b>STR</b>	15.06	15.48
<b>% ELL</b>	6.64	3.69
<b>% IEP</b>	12.57	13.70
<b>% FRL</b>	52.21	62.95
<b>% Minority</b>	50.74	49.33

#### Case Study Format

The eight case studies consisted of three parts: classroom observations of each participating pilot teacher, semi-structured interviews with a school and/or district administrator, and a semi-structured focus group with the pilot teachers. The purpose of the classroom observations was twofold. First, they provided the opportunity for us to see the assessment in use. Second, they provided a context for us to understand feedback from the teacher interviews that may have been unique to a particular classroom. We conducted the administrator interviews with the school principal and/or a district administrator familiar with the KEA formative assessment process. All pilot teachers in each school took part in the focus group interviews, and the school's instructional coach was invited to participate if they were familiar with the KEA. In total, the eight case studies included: 23 kindergarten teachers, 7 school principals, 4 district administrators, and 4 instructional coaches.

We visited six of the eight case study schools twice, once within the first 30 days of the pilot and once during the latter 30 days of the pilot. Due to scheduling conflicts, two case study schools

were only visited once, and they were visited during the latter 30 days of the pilot. During each school visit, the pilot teachers were observed for no less than 1 hour of instructional time. Observers first noted classroom characteristics, such as the total number of students present, whether a teaching assistant or co-teacher was present and their contributions to the classroom, whether the classroom was inclusive of exceptional children, instructional resources and technology available in the classroom and which were utilized during the observation period, etc. Below is a short ‘snapshot’ of the average case study classroom characteristics as documented during these observations:

Table 3.

*Classroom Characteristics*

Average Class Size	20 Students (Range = 13-25)
Average Number of Students of Color	12
Technology Available	Smart Board (86%) Desktop/Laptop Computer (92%) Tablet/Mobile Device (46%)

*Note:* 94% of observed classrooms had a trained teaching assistant present at the time of the observation

After noting classroom characteristics, observers recorded qualitative field notes capturing the teacher's instructional routine, the student's activities, and any use of the KEA formative assessment process. All observation notes were transcribed to document format for analysis.

All interviews were semi-structured. The interviewers utilized a prepared interview protocol consisting of open-ended questions for each group; however, they did at times ask probing or follow-up questions when warranted. We interviewed all participants during each school visit, meaning that if we visited the school twice, all the participants at that school were interviewed twice. This allowed us to note any changes in participants' experiences or attitudes toward the KEA over time. We audio recorded each interview, which amounted to approximately 27 hours of recorded dialogue in all. These recordings were then transcribed for analysis.

### Electronic Survey

Shortly after the 60-day pilot period, we conducted an electronic survey that was open to all 305 volunteer KEA pilot teachers and administrators. The survey provided an avenue for pilot participants outside of the case studies to give feedback regarding the KEA. It also allowed us to compare the data we gathered through our observations and interviews in the smaller case study school settings to the broader pilot population. The survey included 18 closed ended or Likert scale questions and 26 open-ended questions. A majority of the questions in the survey mirrored those from the teacher and administrator interview protocols. We received 72 total responses: 52 from teachers, 16 from administrators, and 4 from instructional coaches. The relatively low response rate was likely an issue of timing. The KEA pilot ended the week prior to Thanksgiving; therefore, we conducted our survey in early December. This coincided with the end of the second marking period of the school year and several benchmarking milestones for other district and state assessments.

## Data Analysis

We uploaded our observation field notes, interview transcriptions, and survey responses to the qualitative analysis program NVivo 10. We took a grounded approach when analyzing the data. In other words, we generated a codebook from the data itself rather than analyzing the data with a previously developed coding scheme or entering into the analysis with particular hypotheses. Three researchers coded the data and we compared their codes through the NVivo software to ensure inter-rater reliability. During the coding process, we created 193 unique codes and documented 3,952 individual references to those codes. Some of the codes were categorical (who, what, where, etc.), while others identified individual topics or themes (training, assessment content, application, etc.). Finally, we conducted frequency and cross-reference analyses on the coded data to identify the most frequent feedback threads and areas where codes often intersected.

## Qualitative Findings and Implications

### Professional Development

While participants walked away from training generally understanding the purpose and content of the KEA (32 interview references and 63% of survey responses), they felt they were unprepared to use the assessment's electronic platform to upload documentation and enter student learning statuses (22 interview references and 60% of survey responses). This indicates that teachers need additional hands-on training with the software in order to feel confident when using it in their classrooms.

Of those who did feel comfortable uploading evidences to the platform and marking student learning statuses, the majority felt they could not use that data in a meaningful way to inform their instruction (39 interview references and 57% of survey responses). One teacher summarized this perception quite succinctly: "I'm putting all of this information in, but I'm getting nothing out." This indicates that pilot teachers struggled to move from documenting student behaviors to the interpretation and application of that data to inform instruction, and highlights the need for greater focus in these areas during professional development for the assessment. We did not have sufficient data to determine what part(s) of the data analysis and application processes participants found most difficult. This is an area for future research.

When asking participants whether additional, non-assessment specific professional development would be helpful for implementing the KEA, an interesting pattern emerged. There were 49 total references describing a need for additional early childhood education training that both outlined and modeled best practice in early childhood instruction. We cross-referenced those statements with the characteristic codes to determine who made these suggestions and whom they identified as needing that additional training. All 49 references came from pilot teachers who identified either their school principal (32 cross-references) or a district administrator (17 cross-references) as needing this additional training. Each reference stated that the administrator did not have early childhood training or teaching experience and, therefore, could neither see the value of the KEA assessment process nor adequately support the teachers during the implementation of a developmentally focused assessment.

## Gathering Evidence during Daily Instruction

Observers noted three instances of KEA documentation gathering. In all three instances, the teachers used an electronic tablet to take a picture of a student as they worked and uploaded it to the child's electronic portfolio along with a note describing the action they were documenting. These instances were not artificially created, but occurred in the course of the teacher's routine daily instruction. This illustrates that some of the pilot teachers understood the documentation process and were able to incorporate the process seamlessly into their normal instructional routine. Many teachers, however, could not identify documentation opportunities during their instruction that illustrated the content of the KEA construct progressions (42 total references). There were several requests for an alignment guide between the KEA progressions and teachers' current curriculum and state standards to assist with this process in the future (39 total references). Teachers who struggled to identify documentation gathering opportunities often resorted to creating new direct assessment tasks to gather the needed documentation, in essence approaching the formative process as a summative one (21 total references).

## Interpreting and Applying KEA Data

Only six of the 23 teachers interviewed and 43% of survey respondents felt that the KEA formative data was meaningful for driving their instruction. While all case study and survey participants saw value in the assessment's whole-child approach with its inclusion of socioemotional and physical development domains, 57% could not identify ways to use such data for instructional planning. These findings mirror those of other recent scholarship on the formative assessment process that indicates that teachers struggle to move from gathering evidences of learning to using such data to inform their instruction (see review of the literature by Akers, Del Grosso, Atkins-Burnett, Monahan, Boller, Carta, & Wasik, 2015). It is unclear, however, what step in the process of moving from gathering data, to analyzing the data, to applying the data that teachers struggle with. This is another area for future research.

## Developmental Appropriateness

Participants overwhelmingly felt that the construct progression content was developmentally appropriate (39 interview references and 73% of survey responses). Furthermore, they felt the socioemotional and physical domains validated kindergarten practices often overlooked by practitioners who do not have training in early childhood education (23 total references). As one teacher stated:

“This really validates what we do and deal with everyday...there's so much that needs to happen before you see a lot of academic changes. These young children are going to grow socially tremendously [in the beginning of the year] and administrators need to understand we have all this other stuff to get in place before they can start moving academically.”

Teachers in exceptional children (EC) inclusive classrooms especially valued the whole-child approach of the KEA as it allowed them to document skills and progress in areas outside of the strictly 'academic' (10 total references). As one EC teacher explains, “I like the KEA assessment because it can be used with any child in kindergarten. I have an autistic, but very bright student. He doesn't respond to standardized tests, but with the KEA I had the time and ability to observe, identify, and document many strong skills from this child.” This indicates that the KEA process is appropriate for use in both EC inclusive and self-contained classrooms and may align more closely with student IEP goals than other strictly academic assessments.



## Alignment with Current Curriculum and Assessments

More than 60% of all pilot feedback concerned this area, indicating it was at the fore front of participants' minds. The participants felt the KEA modeled the best-practices for the assessment of young children, yet it was too misaligned with current state and district mandates for them to perform it as intended (39 cross references of developmentally appropriate/misalignment with curriculum and assessments). Teachers stated that the KEA was more developmentally appropriate than many current state and district mandated assessments, but due to tightening accountability guidelines in literacy and mathematics they did not have time to focus on their students' development in other areas (39 total references.) Some teachers also voiced concerns that the inclusion of activities to foster socioemotional and physical development would be interpreted as 'playing' by their administrators which would negatively affect their performance reviews (10 total references). While this speaks volumes to the overall instruction and assessment climate in early childhood education currently, it highlights the need for research into the demands on kindergarten students and teachers, the developmental appropriateness of current early childhood assessments, and the basis for the policies governing early childhood assessments.

## KEA Implementing vs. Non-Implementing Classrooms

The KEA formative assessment process is a cycle consisting of five steps: selecting a learning target, developing criteria for success, eliciting evidence of learning, interpreting that evidence to assign a learning status, and adapting instruction to respond to the child's learning needs. Out of the 17 classrooms observed, we noted three teachers implementing the entire process. The other 14 teachers were unable to move from eliciting evidence of learning to interpreting and applying that evidence to their instruction. We compared demographic and behavioral characteristics of the classrooms and schools to determine any differences between implementation groups.

### Implementing Classrooms:

- Small class sizes: 14 student average
- Students easily transitioned from one activity to another independently.
- School/District had a strong background in the use of formative assessment.
- Teachers used self-created implementation resources to assist KEA documentation.
- Teachers worked collaboratively
- Schools had strong PLCs with a continual focus on data driven instruction.

### Minimally and Non-Implementing Classrooms:

- Large class sizes: 22 student average
- Students struggled to transition independently between classroom activities.
- Teachers often preoccupied with behavioral interventions.
- School/District did not have a strong background in the use of formative assessment.
- Teachers approached the KEA as a summative assessment, using new/additional activities to "test" each child's ability rather than using current instructional or assessment data.

## Quantitative Methods and Data Sources

Initially, 253 teachers volunteered to participate in the KEA pilot. Of those volunteers, 193 teachers attended the training and used the pilot assessment in their classrooms. Only 177 of those pilot teachers populated classes in the electronic platform. It is unclear why 16 teachers did not access the platform; however, it is possible that technical issues prevented them from accessing the online system. During our case studies, we noted two teachers documenting KEA data in rich paper portfolios, but were unable to access the system until the last two weeks of the pilot due to their district's internet security protocols.

Over the 60-day pilot, teachers uploaded 12,554 pieces of documentation to the online system. We reviewed each piece of evidence to determine its type, its associated construct, the number of children associated with it, and whether it contained specific information that a teacher could assign a learning status along a construct progression. An example of specific information would be "Mary can cross the midline while dancing the Macarena," while non-specific information would be "Mary danced today." We ran descriptive statistical analyses in order to gain a picture of how teachers used the online system during the pilot.

## Quantitative Findings and Implications

On average teachers uploaded 64.96 evidences (standard deviation=76.39), with an average of 3.67 (standard deviation=4.05) evidences per child. Since there were 10 constructs used during the pilot, an average of 3.67 pieces of evidence per child indicates that the majority of teachers did not upload evidences to all constructs for each of their students. Since we do not have data regarding paper portfolios teachers may have kept over the course of the pilot, we do not have enough data to conclude that the pilot teachers did not collect evidences in all areas, only that they did not upload them to the electronic platform.

The majority of the evidences were anecdotal notes (59%), followed by photographs (31.3%), and video recordings (9.7%). Most of the evidences were assigned to the book and print awareness, grip and manipulation, letter naming, and number counting constructs. Teachers uploaded the least number of evidences in the emotional literacy and follows directions constructs. Considering the current curricular focus on mathematics and literacy, it is likely that teachers recognized more opportunities to gather evidences in the more academic areas of the KEA during their routine classroom instruction.

Of the evidences entered, 81.6% were assigned to a single child while 18.4% were assigned to multiple children. Those assigned to multiple children were generally videos or photographs taken during a whole class or small group activity. It is important to note that assigning a piece of evidence to multiple children did not exclude that documentation from being categorized as containing child specific information as long as a note accompanied the entry that explained each child's abilities associated with the construct.

Nearly half (49.8%) of the evidences entered during the pilot did not contain child specific information that could assist the teacher in assigning a learning status on its associated construct progression. The other 50.2% of evidences contained specific information that was used to assign a learning status appropriately to the child. To determine whether the number of specific evidences changed over time, we performed crosstab analysis between the data and child specific information variables. As the chart below illustrates, there was not a significant pattern between groups:

Table 4.  
*Crosstabulations*

		Child Specific Info		Total	
		Yes	No		
Date	August	Count	453	905	1358
		% within Date	33.4%	66.6%	100.0%
		% within ContainsChildSpecificInfo	7.2%	14.5%	10.8%
		% of Total	3.6%	7.2%	10.8%
	September	Count	2857	2666	5523
		% within Date	51.7%	48.3%	100.0%
		% within ContainsChildSpecificInfo	45.4%	42.7%	44.1%
		% of Total	22.8%	21.3%	44.1%
	October	Count	1867	1798	3665
		% within Date	50.9%	49.1%	100.0%
		% within ContainsChildSpecificInfo	29.7%	28.8%	29.2%
		% of Total	14.9%	14.3%	29.2%
	November	Count	1114	878	1992
		% within Date	55.9%	44.1%	100.0%
		% within ContainsChildSpecificInfo	17.7%	14.1%	15.9%
		% of Total	8.9%	7.0%	15.9%
	Total	Count	6291	6247	12538
		% within Date	50.2%	49.8%	100.0%
		% within ContainsChildSpecificInfo	100.0%	100.0%	100.0%
		% of Total	50.2%	49.8%	100.0%

Of the evidences that did not contain specific information, several of them appeared to be uploaded for the express purpose of the teacher letting any potential reviewer know that they were attempting to complete the process. For instance, one note stated simply, “I hope I’m doing this right, I’m trying!” Notes such as this imply that the teachers believe ‘someone’ is analyzing their portfolios to see if they are completing the process or analyzing the quality of the evidence they are entering. During our school visits, similar concerns arose with teachers asking questions such as “Who’s looking at all of this? Is DPI reviewing it?” and “What if an administrator looks at my portfolios and thinks I rated a student incorrectly?” Comments such as this illustrate that a portion of the pilot participants misunderstood the true purpose of the KEA process: to inform their own instruction. It also highlights a broader anxiety among educators regarding new initiatives given the current high-stakes assessment climate.

## References

- Akers, L., Del Grosso, P., Atkins-Burnett, S., Monahan, S., Boller, K., Carta, J., & Wasik, B.A., (2015). *Issue Brief—Tailored teaching: The need for stronger evidence about early childhood teachers' use of ongoing assessment to individualize instruction*. Washington, DC: Mathematica Policy Research. Retrieved from [http://www.mathematica-mpr.com/~media/publications/pdfs/earlychildhood/tailored\\_teaching\\_child\\_assessment.pdf](http://www.mathematica-mpr.com/~media/publications/pdfs/earlychildhood/tailored_teaching_child_assessment.pdf).
- Black, P. J., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139-48.
- North Carolina K-3 Assessment Think Tank (2013). *Assessment for learning and development in K-3: A report by the state assessment think tank*. Raleigh, NC: Retrieved from <http://www.ncpublicschools.org/docs/newsroom/news/2013-14/20131009/think-tank.pdf>.
- North Carolina Office of Early Learning (2014). *NC KEA progression posters*. Retrieved from [http://keapilot.ncdpi.wikispaces.net/file/view/MIA-posters\\_9.5.14.pdf/521677878/MIA-posters\\_9.5.14.pdf](http://keapilot.ncdpi.wikispaces.net/file/view/MIA-posters_9.5.14.pdf/521677878/MIA-posters_9.5.14.pdf).
- William, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-84). Mahwah, NJ: Lawrence Erlbaum Associates.