Technical Appendix

This technical appendix reports data that were collected from 1997 to 2002 as part of the development of the ELLCO Toolkit, Research Edition, as well as additional data that were collected from 2002 to 2007 using the Research Edition. As described in Chapter 1, on the basis of this data, along with feedback from the field, we made numerous changes that serve to make the ELLCO Pre-K easier to use and score. The most significant changes are the integration of the Literacy Environment Checklist and the Literacy Activities Rating Scale into the architecture of the observation, and the inclusion of detailed descriptive indicators for each of the five scale points. Specific psychometric analyses on the current ELLCO Pre-K will be reported as the tool is used. For reasons that we outline at the end of this appendix, however, we believe that the ELLCO Pre-K will prove to be as reliable, if not more reliable, than the Research Edition.

The ELLCO Toolkit, Research Edition, was pilot tested and used in several research studies from its development, with the minor revisions that typically occur through practical use and feedback, through its original publication in 2002; this included research conducted in more than 150 preschool classrooms for the Head Start–funded New England Quality Research Center (NEQRC) (funder: U.S. Department of Health and Human Services, Administration for Children and Families; 1995–2000) and the Literacy Environment Enrichment Project (LEEP) (funder: U.S. Department of Education, Office of Educational Research and Improvement; 2000–2003), both based in the Center for Children & Families (CC&F) at Education Development Center, Inc., in Newton, Massachusetts. Since its initial publication, researchers at CC&F have used the ELLCO Toolkit, Research Edition, in more than 250 classrooms as part of six different projects:

- The New England Quality Research Center: The Next Generation (funder: U.S. Department of Health and Human Services, Administration for Children and Families; 2001–2006)

- Examining the Efficacy of Two Models of Preschool Professional Development in Language and Literacy (funder: U.S. Department of Education, Institute of Education Sciences; 2005–2007)
- Child Care Quality: Does Partnership Make a Difference—an Extension of the Partnership Impact Project (funder: U.S. Department of Education, 2004–2007)
- Evaluation of the Newport Early Reading First Collaborative (funder: U.S. Department of Education, 2003–2006)
- Evaluation of the Springfield Early Reading First Initiative (funder: U.S. Department of Education, 2003–2007)
- Connecticut Is Reading First (funder: U.S. Department of Education, 2005 to present)

All of these projects are concerned with the language and literacy development of children from lower income families and communities. Because of this, data reported here are based on centers and classrooms in lower income communities.

The data reported in the main body of the appendix come from 30 classrooms from the NEQRC study and a total of 117 classrooms for the LEEP study. Each of the NEQRC classrooms were observed on one occasion, whereas most of the LEEP classrooms were observed on two occasions (fall and spring), and a few LEEP classrooms were visited a total of three times. In the data used to calculate means, correlations, and Cronbach's alpha, each visit to a classroom is counted as a separate observation. In the data used to report stability and change (LEEP classrooms only), each LEEP classroom is counted only once, and fall and spring scores are treated as distinct variables.

An addendum to the original research at the end of this technical appendix describes Cronbach's alpha analyses performed using a larger sample of data collected between 2001 and 2007. These data are from a total of 259 classrooms from the following projects:

- The New England Quality Research Center: The Next Generation ($n = 57$)
- Examining the Efficacy of Two Models of Preschool Professional Development in Language and Literacy ($n = 67$)
- Child Care Quality: Does Partnership Make a Difference—an Extension of the Partnership Impact Project ($n = 66$)
- Evaluation of the Newport Early Reading First Collaborative ($n = 26$)
- Evaluation of the Springfield Early Reading First Initiative ($n = 22$)
- Connecticut Is Reading First ($n = 21$)

As with the data used for the Cronbach's alpha analyses initially reported in the technical appendix of the *User's Guide to the Early Language & Literacy Classroom Observation Toolkit, Research Edition* (Smith & Dickinson, 2002), some classrooms were visited on multiple occasions, and each classroom visit was counted as a separate observation.

PSYCHOMETRIC PROPERTIES OF THE LITERACY ENVIRONMENT CHECKLIST

The psychometric properties presented for the Literacy Environment Checklist (ELLCO Toolkit, Research Edition) are based on data from Year 4 of the NEQRC project combined with data from Years 1–3 of the LEEP project. Data from the NEQRC project were collected during the winter of 1998–1999 (n = 29). The data from Year 1 of the LEEP project were collected in the fall of 1998 (n = 26) and the spring of 1999 (n = 26). Data from Year 2 of the LEEP project were collected in the fall of 1999 (n = 42) and spring of 2000 (n = 38). Data from Year 3 of the LEEP project were collected in the fall of 2000 (n = 47) and spring of 2001 (n = 47). Together, the projects resulted in a total sample size of 255, although the actual subsample sizes vary depending on the analyses conducted. Many of the classrooms included were in Head Start programs. Unlike the Classroom Observation, the Literacy Environment Checklist and the Literacy Activities Rating Scale have been used for research only in preschool classrooms and were designed specifically to help identify the impact of our literacy intervention in those classrooms. They have not been used to predict children's growth; rather, they have been used in conjunction with the Classroom Observation to pinpoint the specific effects of a literacy intervention. The items from the Literacy Environment Checklist of the ELLCO Toolkit, Research Edition, have been incorporated into the main body of the ELLCO Pre-K. The Literacy Activities Rating Scale is not included in the ELLCO Pre-K. (See Table A.15 later in this appendix for item-level correspondences between the ELLCO Pre-K and the ELLCO Toolkit, Research Edition.)

Interrater Reliability

Research use of the Literacy Environment Checklist was predicated on the appropriate training of observers. We have required that prospective observers be familiar with theories of early literacy development and have an understanding of the range of instructional methods that are typically used in classrooms.

Prospective observers received a daylong training session on using the ELLCO Toolkit, Research Edition, which included background information on language and literacy development, explanation of the toolkit, and videotape examples; then training session participants received a second day of supervised practice in using the toolkit. When observers were trained and supervised appropriately, we achieved an average interrater reliability of 88% with relative ease. (This interrater reliability rate is for agreements within 1 point of each other on the rating scale.)

General Statistics

On the basis of our theoretical beliefs and preliminary analysis of the data, we created three summary variables for the Literacy Environment Checklist: the *Books* subtotal, the *Writing* subtotal, and the *Total* score. The *Books* subtotal includes all items from the Book Area, Book Selection, and Book Use sections of the checklist. The *Writing* subtotal includes all items from the Writing Materials and Writing Around the Room sections. Table A.1 reports descriptive statistics for Literacy Environment Checklist data gathered as part of the NEQRC and LEEP studies ($n = 255$).

Reliability Analysis

Reliability analysis was conducted to examine the internal consistency of the Literacy Environment Checklist. Table A.2 shows the alphas obtained for the *Total* score as well as for the two subtotals. Cronbach's alpha of .84 for the *Total* score shows good internal consistency. All item–total correlations were moderate to high ($r = .54$ to $r = .55$).

Cronbach's alpha of .73 for the *Books* subtotal shows good internal consistency for this composite. All item–total correlations were moderate ($r = .21$ to $r = .54$) with the exception of Item 1 in the Book Area section ("Is an area set aside just for book reading?"), which exhibited a correlation of .16.

Table A.1. Descriptive statistics for subscale and total score data for the Literacy Environment Checklist in the ELLCO Toolkit, Research Edition ($n = 255$)

Composite variable	Mean	Standard deviation	Minimum	Maximum
Books subscale	11.13	3.90	2.00	20.00
Writing subscale	10.44	4.22	1.00	20.00
Literacy Environment Checklist Total score	21.57	7.37	5.00	40.00

Table A.2. Cronbach's alpha for data for the Literacy Environment Checklist in the ELLCO Toolkit, Research Edition ($n = 255$)

Composite variable	Alpha
Books subtotal	.73
Writing subtotal	.75
Literacy Environment Checklist Total score	.84

Cronbach's alpha for the *Writing* subtotal was .75, also indicating somewhat low but still acceptable internal consistency. Item–total correlations ranged from a low of .21 for Item 15 in the Writing Materials section ("Are there templates or tools to help children form letters?") to a high of .59 for Item 21 in the Writing Around the Room section ("How many varieties of children's writing are on display in the classroom?").

Measuring Stability and Change

Using the data collected from the LEEP classrooms, we reported preliminary findings on the ability of the Literacy Environment Checklist to measure both stability and change over time (see Table A.3). When one looks at mean scores across the 3 years of the LEEP project, the fall scores of the intervention are

Table A.3. Stability and change in Literacy Environment Checklist scores (ELLCO Toolkit, Research Edition), fall and spring means, for Years 1–3 of the Literacy Environment Enrichment Project (LEEP)

	Fall		Spring	
Composite variable	Comparison group ($n = 38$)	LEEP intervention ($n = 40$)	Comparison group ($n = 38$)	LEEP intervention ($n = 40$)
Books subtotal	9.25	10.53	10.45 ($t = 3.27$, $p < .01$)	14.84 ($t = 7.18$, $p < .001$)
Writing subtotal	8.77	11.21	9.12 ($t = 1.14$, $p =$ n.s.)	14.26 ($t = 5.72$, $p < .001$)
Literacy Environment Checklist Total score	18.12	20.86	19.52 ($t = 2.87$, $p < .01$)	29.03 ($t = 7.82$, $p < .001$)

n.s., not significant.

slightly higher on the three dimensions of the Literacy Environment Checklist than the comparison group. (For the fall scores in the LEEP study, differences between the intervention group and comparison group on the Literacy Environment Checklist were statistically significant for the *Writing* subtotal only (t = –2.62, p .05). In the spring, the comparison group showed significant change on the *Total* score as well as on the *Books* subtotal yet remained stable on the *Writing* subtotal. As hoped, the intervention group scores changed significantly from fall to spring in all categories. These changes resulted in intervention group scores that were statistically significantly different from the comparison group scores in every category *and* statistically significantly different from the intervention group fall scores in every category.

PSYCHOMETRIC PROPERTIES OF THE CLASSROOM OBSERVATION

Like the other parts of the ELLCO Toolkit, Research Edition, the Classroom Observation has been used for research for the NEQRC and LEEP. The Classroom Observation also has been used as a part of a school improvement project in the Philadelphia public school system in classrooms that range from kindergarten through grade 5. It has also been introduced to school systems in Connecticut and Maine. In these settings, it is being used both to collect data on and to provide a basis for discussions about classroom quality.

The psychometric properties presented in the sections that follow come from various analyses of data from Year 4 of the NEQRC research project combined with data collected from Years 1–3 of the LEEP project. Data from the NEQRC project were collected during the winter of 1998–1999 (n = 29). The data from Year 1 of the LEEP project were collected in the fall of 1998 (n = 27) and the spring of 1999 (n = 27). Data from Year 2 of the LEEP project were collected in the fall of 1999 (n = 42) and spring of 2000 (n = 38). Data from Year 3 of the LEEP project were collected in the fall of 2000 and spring of 2001 in New England (fall: n = 34; spring: n = 37) and North Carolina (fall: n = 37; spring: n = 37). Together, the projects resulted in a total sample size of 308 classrooms, though the actual subscale size varies depending on the analyses conducted. As with the other parts of the ELLCO Toolkit, the data reported here for the Classroom Observation come from centers and classrooms in lower income communities.

Interrater Reliability

Research use of the Classroom Observation was predicated on appropriate training of observers, as explained in the section of this appendix on the Literacy Environment Checklist. Novice observers' initial observations were con-

ducted with an experienced observer in order to ensure appropriate calibration to the rubrics in the Classroom Observation. When observers were trained and supervised appropriately, we consistently achieved interrater reliabilities of 90% and better for this part of the ELLCO Toolkit.

General Statistics

On the basis of our theoretical beliefs and preliminary analyses of the data, we chose to create three summary variables for the Classroom Observation: the *General Classroom Environment* subtotal, the *Language, Literacy, and Curriculum* subtotal, and the *Total* score. One item (Item 3), Presence and Use of Technology, was problematic[1] and was excluded from all summaries and analyses. Items included in the two subtotals in the ELLCO Toolkit, Research Edition, are as follows.

General Classroom Environment subtotal:

1. Organization of the Classroom
2. Contents of the Classroom
4. Opportunities for Child Choice and Initiative
5. Classroom Management Strategies
6. Classroom Climate

Language, Literacy, and Curriculum subtotal:

7. Oral Language Facilitation
8. Presence of Books

9P. Approaches to Book Reading (Prekindergarten and Kindergarten Version)

10P. Approaches to Children's Writing (Prekindergarten and Kindergarten Version)

11. Approaches to Curriculum Integration
12. Recognizing Diversity in the Classroom
13. Facilitating Home Support for Literacy
14. Approaches to Assessment

[1]By *problematic*, we mean that scores for Presence and Use of Technology did not cluster with scores for the other items, suggesting that effective use of technology reflects capabilities somewhat distinct from those captured by the other items in the Classroom Observation. In addition the scores for Presence and Use of Technology did not relate clearly to the *General Classroom Environment* subtotal or to the *Language, Literacy, and Curriculum* subtotal; therefore, it was not included in the reported averages or calculations of Cronbach's alpha for the total tool. For this item in 308 classrooms, the mean was 2.45, with a standard deviation of 1.09 and a minimum of 1.0 and a maximum of 5.0.

Table A.4. Descriptive statistics for data for the Classroom Observation in the ELLCO Toolkit, Research Edition ($n = 308$)

Composite variable	Mean	Standard deviation	Minimum	Maximum
General Classroom Environment subtotal	3.44	0.79	1.20	5.00
Language, Literacy, and Curriculum subtotal	3.02	0.75	1.13	5.00
Classroom Observation Total score	3.15	0.71	1.29	5.00

Using these subtotals, we obtained data from classrooms throughout New England that provide some indication of observed levels of performance in classrooms that serve low-income children. As with the Literacy Environment Checklist, many of the classrooms included were in Head Start programs. Tables A.4 and A.5 report descriptive statistics for the Classroom Observation data gathered as part of the NEQRC and LEEP studies ($n = 308$).

Reliability Analysis

Reliability analysis was conducted to examine the internal consistency of the Classroom Observation using data from 308 classrooms. Table A.6 shows the Cronbach's alphas obtained for the two composites, *General Classroom Environment* and *Language, Literacy, and Curriculum,* and for the *Total* score of all the items on the Classroom Observation that were included in these analyses.

Cronbach's alpha of .83 for the *General Classroom Environment* shows good internal consistency for this composite. All of the item–total correlations

Table A.5. Frequencies of classrooms ($n = 308$) with Classroom Observation (ELLCO Toolkit, Research Edition) scores in each of the following categories: *high-quality support* (scores ranging from 3.51 to 5), *basic support* (scores ranging from 2.51 to 3.5), and *low-quality support* (scores less than or equal to 2.5)

Composite variable	High-quality support		Basic support		Low-quality support	
General Classroom Environment subtotal	47.4%	(146)	42.2%	(130)	10.4%	(32)
Language, Literacy, and Curriculum subtotal	24.0%	(74)	45.8%	(141)	30.2%	(93)
Classroom Observation Total score	27.9%	(86)	52.6%	(162)	19.5%	(60)

Table A.6. Cronbach's alpha for data for the Classroom Observation in the ELLCO Toolkit, Research Edition ($n = 308$)

Composite variable	Alpha
General Classroom Environment subtotal	.83
Language, Literacy, and Curriculum subtotal	.86
Classroom Observation Total score	.90

were high—with correlation coefficients ranging from .60 for Item 1, Organization of the Classroom, to .75 for Item 6, Classroom Climate—with the exception of Item 2, Contents of the Classroom. This item had the lowest item–total correlation, which was nonetheless a moderate correlation ($r = .53$).

The internal consistency of the *Language, Literacy, and Curriculum* composite is very good, with an alpha of .86. All of the item–total correlations were moderate to high, ranging from .55 for Item 8, Presence of Books, to .65 for Item 13, Facilitating Home Support for Literacy.

Cronbach's alpha of .90 also shows very good internal consistency for all items combined on the Classroom Observation. All of the item–total correlations for the *Classroom Observation Total* were moderate to high ($r = .39$ to $r = .68$).

Measuring Stability and Change

Again, in the LEEP project, classrooms were observed in the fall and in the spring of Years 1–3, yielding the ability to measure change over time using the Classroom Observation. Some of the teachers were taking a yearlong course that focused on early language and literacy (our intervention group); the remaining teachers were not (our comparison group). Using data from the comparison group classrooms, we have data on the ability of the Classroom Observation to measure both stability and change over time (see Table A.7).

The two groups began the fall with similar scores on the three dimensions of the Classroom Observation, with the comparison group scores being slightly lower overall, though not statistically significantly lower, than the intervention group scores. In the spring, the comparison group scores remained stable, though slightly higher overall, with no statistically significant changes from fall to spring. As hoped, the LEEP intervention group scores changed significantly from fall to spring in all categories. These changes resulted in intervention group scores that were statically significantly different from the comparison group scores in every category *and* statistically significantly different from in-

Table A.7. Stability and change in Classroom Observation scores (ELLCO Toolkit, Research Edition), fall and spring means, for Years 1–3 of the Literacy Environment Enrichment Program (LEEP)

	Fall		Spring	
Composite variable	Comparison group ($n = 65$)	LEEP intervention ($n = 42$)	Comparison group ($n = 65$)	LEEP intervention ($n = 42$)
General Classroom Environment subtotal	3.26	3.61	3.42 ($t = 1.96$, $p =$ n.s.)	3.91 ($t = 2.26$, $p < .05$)
Language, Literacy, and Curriculum subtotal	2.85	3.01	2.93 ($t = 1.13$, $p =$ n.s.)	3.75 ($t = 5.50$, $p < .0001$)
Classroom Observation Total score	2.97	3.19	3.08 ($t = 1.53$, $p =$ n.s.)	3.74 ($t = 4.88$, $p < .0001$)

n.s., not significant.

tervention group fall scores in every category.

From our comparison group data, we were able to conclude that the Classroom Observation is able to capture *stability* in classroom quality. This is a good indicator of the Classroom Observation's test–retest reliability. Our data also show that the Classroom Observation is able to capture *changes* in classroom quality associated with a literacy-focused intervention. These findings come from two sources: evidence of fall-to-spring growth and differences between the intervention and comparison groups. These data provide evidence of the *instructional sensitivity* of this tool. The concept of instructional sensitivity is an important factor in determining the quality of research instruments. Our data suggest that the Classroom Observation is both stable and sensitive to interventions that target literacy in ways that are consistent with its assumptions about what constitutes appropriate early literacy practices.

Correlation with Another Widely Used Measure

As part of the NEQRC project, the Classroom Observation has been used in conjunction with the Classroom Profile (Abbott-Shim & Sibley, 1998), a widely used tool for assessing the overall quality of early childhood classrooms. One

reason that the Classroom Observation was initially developed was that existing observation tools did not adequately or systematically address early language and literacy experiences or classroom features that are known to support literacy development (Dickinson & Tabors, 2001). Thus, it was our belief that the Classroom Observation would exhibit divergent validity when used in conjunction with these other tools, indicating that it is measuring something qualitatively different. To examine this hypothesis we correlated the *General Classroom Environment* subtotal, the *Language, Literacy, and Curriculum* subtotal, and the *Classroom Observation Total* score with the raw scores from two subscales from the Classroom Profile that we employed, *Learning Environment* and *Scheduling.* We found moderate correlations for all three Classroom Observation variables with scores on the Classroom Profile's *Learning Environment* subscale (r = .41, .31, and .44, respectively) but not the profile's *Scheduling* subscale (r = .12, .09, and .07, respectively). We take the finding of the modest positive relationship to the Classroom Profile's *Learning Environment* subscale as providing convergent validity for the Classroom Observation. The absence of relationship with the profiles *Scheduling* subscale provides divergent validity because the Classroom Observation was developed to tap a construct that is distinct from that examined by the *Scheduling* subscale.

Predicting Child Outcomes

Possibly the most important test for a tool that purports to evaluate the quality of support provided for children's literacy development is the capacity of the tool to predict children's literacy development. The Classroom Observation has been used in correlational research and employed in hierarchical linear modeling designed to determine the contributions of classroom quality to children's receptive vocabulary (Peabody Picture Vocabulary Test–Third Edition; Dunn & Dunn, 1997) and early literacy scores (Profile of Early Literacy Development; Dickinson & Chaney, 1998). This sophisticated analytic approach allows identification of different sources of variation in children's scores, distinguishing variation between classrooms that is associated with children's backgrounds (e.g., income, gender) from variation associated with their classroom experiences. Level 1 models examining between-group variability took into account variables such as home language (English, Spanish, or other), gender, and age. The variance in scores that was not accounted for by the background factors (15% for vocabulary, 20% for literacy) was attributed to classroom factors. Our models examining sources of classroom-related variance found that scores on the Classroom Observation accounted for 80% of the between-classroom variance in vocabulary and 67% of the between-classroom variance in early literacy (Dickinson et al., 2000). Although revealing the power of the Classroom Observation to predict child outcomes, these analyses also provide evidence

that the quality of preschool classrooms attended by children from low-income families can play an important role in supporting their vocabulary growth and early literacy development.

PSYCHOMETRIC PROPERTIES OF THE LITERACY ACTIVITIES RATING SCALE

Like the Classroom Observation and the Literacy Environment Checklist, the Literacy Activities Rating Scale of the ELLCO Toolkit, Research Edition, has been used to conduct research as part of the NEQRC and LEEP, and the data presented here are from centers and classrooms in lower income communities. The psychometric properties presented for the rating scale are based on data from Year 4 of the NEQRC project, combined with data from Years 1–3 of the LEEP project. Data from the NEQRC project were collected during the winter of 1998–1999 (n = 30). The data from Year 1 of the LEEP project were collected in the fall of 1998 (n = 28) and the spring of 1999 (n =28). Data from Year 2 of the LEEP project were collected in the fall of 1999 (n = 42) and spring of 2000 (n = 40). Data from Year 3 of the LEEP project were collected in the fall of 2000 (n = 47) and spring of 2001 (n = 47). Together, the projects resulted in a total sample size of 262, although actual subsample size varies depending on the analyses conducted. As with the Literacy Environment Checklist, the rating scale has been used for research only in preschool classrooms, many of which are in Head Start programs. Some items from the Literacy Activities Rating Scale have been incorporated into the ELLCO Pre-K (see Table A.15 later in this appendix), but there is no Literacy Activities Rating Scale in the ELLCO Pre-K.

Interrater Reliability

Observers underwent a training process explained in the section of this appendix that describes interrater reliability for the Literacy Environment Checklist and the Classroom Observation. As with those parts of the ELLCO Toolkit, Research Edition, novice observers' initial observations are conducted with an experienced observer. Because the Literacy Activities Rating Scale was used simply to describe activities observed during the classroom visit, we did not maintain formal records of interrater reliability. Observers who visited classrooms together, however, had little difficulty arriving at the same ratings for the classrooms they visited. When observers were trained and supervised appropriately, we achieved an average interrater reliability of 81% with relative ease.

Table A.8. Descriptive statistics for data for the Literacy Activities Rating Scale in the ELLCO Toolkit, Research Edition ($n = 262$)

Composite variable	Mean	Standard deviation	Minimum	Maximum
Full-Group Book Reading subtotal	2.86	1.95	0	6.00
Writing subtotal	2.10	1.39	0	5.00
Literacy Activities Rating Scale Total score	5.80	2.63	0	13.00

General Statistics

On the basis of our theoretical beliefs and preliminary analysis of the data, we created three summary variables for the Literacy Activities Rating Scale: the *Full-Group Book Reading* subtotal, the *Writing* subtotal, and the *Total* score. The *Total* score includes all but two items, which were problematic: Item 4 ("Did you observe an adult engaged in one-to-one book reading or small-group book reading?") and Item 5 ("Is time set aside for children to look at books alone or with a friend?"). These two items were excluded from all analyses. The *Full-Group Book Reading* subtotal includes Items 1–3, which address the number of book reading sessions observed, the length of time spent on full-group book reading, and the total number of books read. The *Writing* subtotal includes Items 6–9, which catalog any observations of children writing as well as any instances of adults assisting children with or modeling writing. Table A.8 reports descriptive statistics for Literacy Activities Rating Scale data gathered as part of the NEQRC and LEEP studies ($n = 262$).

Reliability Analysis

Reliability analysis was conducted to examine the internal consistency of the Literacy Activities Rating Scale. Table A.9 shows the alphas obtained for the *Total* score (excluding the two problematic items mentioned previously), as well as the two subtotals. Cronbach's alpha of .66 for the *Total* score shows somewhat low but acceptable internal consistency for this measure. Item–total correlations ranged from a low of .17 for Item 9 ("Did an adult model writing?") to a high of .49 for Item 1 ("How many full-group book-reading sessions did you observe?").

Cronbach's alpha of .92 for the *Full-Group Book Reading* subtotal shows excellent internal consistency for this composite. All item–total correlations were high ($r = .79$ to $r = .88$). The Cronbach's alpha for the *Writing* subtotal was .73, indicating good internal consistency. Item–total correlations were moderate to high, ranging from a low of .37 for Item 9 ("Did an adult model

Table A.9. Cronbach's alpha for data for the Literacy Activities Rating Scale in the ELLCO Toolkit, Research Edition ($n = 262$)

Composite variable	Alpha
Full-Group Book Reading subtotal	.92
Writing subtotal	.73
Literacy Activities Rating Scale Total score	.66

writing?") to a high of .64 for Item 7 ("Did you see children attempting to write letters or words?"). Given the stronger psychometric properties of the two subscales, we advise using the scores on the distinct subscales of the Literacy Activities Rating Scale instead of the total score when analyzing data from this part of the ELLCO Toolkit, Research Edition.

Measuring Stability and Change

Given the data collected from the LEEP classrooms, we have reported preliminary findings on the ability of the Literacy Activities Rating Scale to measure both stability and change over time (see Table A.10). To determine the stabil-

Table A.10. Stability and change in Literacy Activities Rating Scale scores (ELLCO Toolkit, Research Edition), fall and spring means, for Years 1 and 2 of the Literacy Environment Enrichment Project (LEEP)

	Fall		Spring	
Composite variable	Comparison group ($n = 38$)	LEEP intervention ($n = 53$)	Comparison group ($n = 38$)	LEEP intervention ($n = 53$)
Full-Group Book Reading subtotal	2.13	2.79	1.47 ($t = -2.07$, $p < .05$)	2.89 ($t = 0.28$, $p =$ n.s.)
Writing subtotal	1.57	2.17	2.16 ($t = 2.81$, $p < .01$)	2.68 ($t = 2.18$, $p < .05$)
Literacy Activities Rating Scale score	4.70	5.73	4.70 ($t = 0$, $p =$ n.s.)	6.68 ($t = 1.94$, $p =$ n.s.)

n.s., not significant.

ity of the Literacy Activities Rating Scale, we examined the fall and spring scores of the comparison group and LEEP intervention classrooms. We noted that on the *Total* score and the *Full-Group Book Reading* subtotal, the intervention group showed no significant change but that it did show significant change on the *Writing* subtotal. In contrast, the comparison group showed significant changes on both subtotals but not on the total score. We concluded that the Book Reading portion of the Literacy Activities Rating Scale and the overall scale are reasonably stable but that the Writing portion may be relatively more labile, possibly reflecting the developmental changes that occur as children gain literacy skill over the course of the year. Evidence of the *instructional sensitivity* of the Literacy Activities Rating Scale comes from data for the LEEP intervention approach, which we noted reflected significant fall-to-spring change on all dimensions.

CORRELATIONS AMONG THE ELLCO TOOLKIT (RESEARCH EDITION) MEASURES

In Table A.11, we report correlations among the three measures that make up the ELLCO Toolkit, Research Edition ($n = 248$). The variables included in these analyses are as follows:

- The *Books* subtotal, the *Writing* subtotal, the *Literacy Environment Checklist Total*
- The *Language, Literacy, and Curriculum* subtotal; the *General Classroom Environment* subtotal; and the *Classroom Observation Total*
- The *Full-Group Book Reading* subtotal, the *Writing* subtotal, and the *Literacy Activities Rating Scale Total*

We found that the *Language, Literacy, and Curriculum* subtotal and the *General Classroom Environment* subtotal are highly correlated with the *Classroom Observation Total* ($r = .95$ and .87 respectively) though not as highly with each other ($r = .69$). This modest correlation between the two subscales of the Classroom Observation provides support for the fact that the two subscales should be examined separately.

In addition, there are moderate-to-strong correlations for all three *Classroom Observation* variables with both the *Books* subtotal ($r = .65$, .47, and .62, respectively) and the *Writing* subtotal ($r = .64$, .51, and .63, respectively) of the Literacy Environment Checklist. The *Literacy Environment Checklist Total* exhibits an even stronger relationship with the *Classroom Observation* scores ($r = .67$, .69, and .53, respectively). The *Books* and *Writing* subtotals for

Table A.11. Correlations for data from New England Quality Research Center (NEQRC) Year 4 and Literacy Environment Enrichment Project (LEEP) Years 1 and 2 ($n = 92$)

	Literacy Environment Checklist			Classroom Observation			Literacy Activities Rating Scale	
Composite variable	1	2	3	4	5	6	7	8
1. *Literacy Environment Checklist: Books*	–							
2. *Literacy Environment Checklist: Writing*	.62***	–						
3. *Literacy Environment Checklist Total*	.89***	.90***	–					
4. *Classroom Observation: General Classroom Environment*	.47***	.51***	.53***	–				
5. *Classroom Observation: Language, Literacy, and Curriculum*	.65***	.64***	.69***	.69***	–			
6. *Classroom Observation Total*	.62***	.63***	.67***	.87***	.95***	–		
7. *Literacy Activities Rating Scale: Full-Book Reading*	.10	.11	.11	.06	.14*	.11	–	
8. *Literacy Activities Rating Scale: Writing*	.36***	.43***	.43***	.37***	.47***	.46***	.04	–
9. *Literacy Activities Rating Scale Total*	.33***	.37***	.38***	.31***	.44***	.41***	.75***	.63***

*$p < .05$ **$p < .01$ ***$p < .001$

the checklist are highly correlated with the *Literacy Environment Checklist Total* (r = .89 and .90, respectively), but not as highly correlated with each other (r = .62).

The *Literacy Activities Rating Scale Total* score and the *Writing* subtotal for the rating scale are moderately related to the three *Classroom Observation* scores (r = .44, .31, and .41 respectively, and r =.47, .37, and .46, respectively). Although the *Full-Group Book Reading* subtotal does not show a statistically significant relationship to the *Classroom Observation Total* score (r = .11) or the *General Classroom Environment* subtotal (r = .06), there is a statistically significant correlation between the *Full-Group Book Reading* subtotal and the *Language, Literacy, and Curriculum* subtotal. The *Literacy Activities Rating Scale Total* score and the *Writing* subtotal are moderately correlated with all three Literacy Environment Checklist scores (r = .38, .33, and .37, respectively, and r = .43, .36, and .43, respectively). The rating scale's *Full-Group Book Reading* and *Writing* subtotals are both highly correlated with the *Literacy Activities Rating Scale Total* score (r = .75 and r = .63, respectively) yet are not statistically significantly correlated with one another, indicating that the two subscales are measuring different constructs.

ADDENDUM

Cronbach's alpha analyses parallel to those described thus far were performed using a larger sample (n = 634) of data collected with the ELLCO Toolkit, Research Edition, between 2002 and 2007. These data come from the following sources:

- The New England Quality Research Center: The Next Generation (2001–2006, n = 182)
- Examining the Efficacy of Two Models of Preschool Professional Development in Language and Literacy (2005–2007, n = 213)
- Child Care Quality: Does Partnership Make a Difference: An Extension of the Partnership Impact Project (2004–2007, n = 66)
- Evaluation of the Newport Early Reading First Collaborative (2003–2006, n = 66)
- Evaluation of the Springfield Early Reading First Initiative (2003–2007, n = 66)
- Connecticut Is Reading First (2005–2007, n = 53)

Table A.12. Cronbach's alpha for larger sample data (2001–2007) for the Literacy Environment Checklist in the ELLCO Toolkit, Research Edition ($n = 616$)

Composite variable	Alpha
Books subtotal	.76
Writing subtotal	.75
Literacy Environment Checklist Total score	.84

Results of the Cronbach's alpha analyses, described in detail next for each section of the ELLCO Toolkit, Research Edition, corroborate the other findings reported in this appendix and thereby strengthen confidence in the internal reliability of the tool.

Alpha coefficients for the *Literacy Environment Checklist Total* score as well as the *Books* and *Writing* subtotals all show good internal consistency (see Table A.12). Item–total correlations for the *Books* subtotal were moderate, ranging from .23 to .59. Item–total correlations for the *Writing* subtotal also were moderate ($r = .23$ to $r = .53$), with the exception of Item 13 ("Is an alphabet visible?") and Item 24 ("Are there puzzles with words available for children's use?") ($r = .17$ for both). Those two items also exhibited low item–total correlation with the *Literacy Environment Checklist Total* score ($r = .13$ and .16, respectively), whereas the remaining items displayed moderate ($r = .21$ to $r = .57$) item–total correlation with the total score.

The Cronbach's alpha coefficients for the *General Classroom Environment* subtotal, the *Language, Literacy, and Curriculum* subtotal and the *Classroom Observation Total* score all show good to excellent internal consistency (Table A.13). Furthermore, both of the Classroom Observation subtotals as well as the *Total* score exhibited moderate to high item–total correlations, ranging from .57 to .73 for the *General Classroom Environment* subscale, .60 to .73 for the *Language, Literacy, and Curriculum* subscale, and .56 to .77 for the *Total* score.

Table A.13. Cronbach's alpha for larger sample data (2001–2007) for the Classroom Observation in the ELLCO Toolkit, Research Edition ($n = 634$)

Composite variable	Alpha
General Classroom Environment subtotal	.84
Language, Literacy, and Curriculum subtotal	.89
Classroom Observation Total score	.93

Table A.14. Cronbach's alpha for larger sample data (2001–2007) for the Literacy Activities Rating Scale in the ELLCO Toolkit, Research Edition ($n = 547$)

Composite variable	Alpha
Full-Group Book Reading subtotal	.90
Writing subtotal	.74
Literacy Activities Rating Scale total score	.72

In Table A.14, the *Full-Group Book Reading* alpha of .90 shows very good internal consistency for this subtotal. The *Writing* subtotal and *Literacy Activity Rating Scale Total* score both show good internal consistency with alphas of .74 and .72, respectively. Item–total correlations for the *Full-Group Book Reading* subscale were high, ranging from .75 to .85. The items that compose the *Writing* subtotal showed moderate to high item–total correlations of .47 to .64. The *Literacy Activities Rating Scale Total* score also had moderate to high item–total correlations that ranged from .30 to .56.

On the basis of on the psychometric properties of the ELLCO Toolkit, Research Edition, as well as the theoretical and practical considerations outlined in Chapter 1, we revised the ELLCO to include more specificity, in the form of detailed descriptive indicators for each scale point, as well as a broader range of measures of quality in early literacy, such as phonological awareness, efforts to build vocabulary, opportunities for extended conversations, and environmental print. Although the ELLCO Pre-K is more thorough and expansive than the Research Edition, it does include the same content covered by the earlier version, with the exception of Research Edition Item 3: Presence and Use of Technology. Cronbach's alpha analyses on the ELLCO Toolkit, Research Edition, described earlier in this technical appendix, indicated that this construct was not statistically related to the other items; therefore, it was not included in the ELLCO Pre-K.

Users of the ELLCO Pre-K who are familiar with the Research Edition will notice that there is now less emphasis on teacher responses to the Teacher Interview, which, in the Research Edition, informed scoring decisions for several items (e.g., Item 13: Facilitation of Home Support for Literacy, Item 14: Approaches to Assessment). Aspects of these and other items from the Research Edition have now been integrated with items in the Pre-K version. All items are now based predominantly on observable classroom indicators to increase validity and reliability. Table A.15 outlines the relationship between the items in the ELLCO Pre-K and the ELLCO Toolkit, Research Edition.

Given that the ELLCO Pre-K includes the content covered in the Research Edition while providing more specificity and a broader range of items, we have every reason to believe that the ELLCO Pre-K will exhibit similar, if not

Table A.15. Item–level correspondences between ELLCO Pre-K and ELLCO Toolkit, Research Edition

ELLCO Pre-K item	Related items from ELLCO Toolkit, Research Edition
1. Organization of the Classroom	Classroom Observation 1. Organization of the Classroom
2. Contents of the Classroom	Classroom Observation 2. Contents of the Classroom
3. Classroom Management	Classroom Observation 5. Classroom Management Strategies
4. Personnel	N/A (item new to ELLCO Pre-K)
5. Approaches to Curriculum	Classroom Observation 11. Approaches to Curriculum Integration
6. Opportunities for Child Choice and Initiative	Classroom Observation 4. Opportunities for Child Choice and Initiative
7. Recognizing Diversity in the Classroom	Classroom Observation 12. Recognizing Diversity in the Classroom Aspects of Classroom Observation 13. Facilitating Home Support for Literacy
8. Discourse Climate	Aspects of Classroom Observation 6. Classroom Climate
9. Opportunities for Extended Conversations	Aspects of Classroom Observation 7. Oral Language Facilitation Aspects of Classroom Observation 14. Approaches to Assessment
10. Efforts to Build Vocabulary	Aspects of Classroom Observation 7. Oral Language Facilitation
11. Phonological Awareness	N/A (item new to ELLCO Pre-K)
12. Organization of Book Area	Aspects of Classroom Observation 8. Presence of Books Literacy Environment Checklist Items 1, 2
13. Characteristics of Books	Aspects of Classroom Observation 8. Presence of Books Literacy Environment Checklist Items 4, 6
14. Books for Learning	N/A (item new to ELLCO Pre-K)
15. Approaches to Book Reading	Aspects of Classroom Observation 9P. Approaches to Book Reading Literacy Activities Rating Scale Items 1, 4, 5
16. Quality of Book Reading	Aspects of Classroom Observation 9P. Approaches to Book Reading
17. Early Writing Environment	Aspects of Classroom Observation 10P. Approaches to Children's Writing Literacy Environment Checklist Items 13, 14, 15, 18, 19, 20, 22a, 22b
18. Support for Children's Writing	Aspect of Classroom Observation 10P. Approaches to Children's Writing Literacy Activities Rating Scale Items 6, 7, 8, 9
19. Environmental Print	N/A (item new to ELLCO Pre-K)

stronger, psychometric properties than the Research Edition. Because both instruments share the same general structure, it is appropriate to compare the ELLCO Pre-K and the Classroom Observation from the Research Edition. The ELLCO Pre-K includes 19 items, whereas the Classroom Observation contains 14 items. If all else remains constant, the mere increase in the number of items allows for more variance, which should lead to increased reliability of the overall scale. Given the greater level of detail provided by having descriptive indicators for each scale point, we would also anticipate improved levels of interrater reliability. Clearly, this hypothesis will need to be verified empirically; therefore, we are currently collecting data on the ELLCO Pre-K in order to perform psychometric analyses, the findings of which will be provided online at http://www.brookespublishing.com/ellco in the near future.

REFERENCES

Abbott-Shim, M., & Sibley, A. (1998). *Assessment Profile for Early Childhood Programs.* Atlanta, GA: Quality Assist.

Dickinson, D.K., & Chaney, C. (1998). *Profile of Early Literacy Development.* Newton, MA: Education Development Center, Inc.

Dickinson, D.K., Sprague, K., Sayer, A., Miller, C., Clark, N., & Wolf, A. (2000). Classroom factors that foster literacy and social development of children from different language backgrounds. In M. Hopman (Chair), *Dimensions of program quality that foster child development: Reports from 5 years of the Head Start Quality Research Centers.* Poster session presented at the biannual National Head Start Research Conference, Washington, DC.

Dickinson, D.K., & Tabors, P.O. (Eds.). (2001). *Beginning literacy with language: Young children learning at home and school.* Baltimore: Paul H. Brookes Publishing Co.

Dunn, L.M., & Dunn, L.M. (1997). *Peabody Picture Vocabulary Test–Third Edition.* Circle Pines, MN: American Guidance Service.

Smith, M.W., & Dickinson, D.K. (with Sangeorge, A., & Anastasopoulos, L.). (2002). *User's Guide to the Early Language & Literacy Classroom Observation Toolkit* (Research ed.). Baltimore: Paul H. Brookes Publishing Co.