

**Preschool Curriculum Evaluation Research (PCER)
2000: Lessons Learned From Two Years
of Curriculum Implementation**

**Early Literacy and Learning Model (ELLM):
Instrument Reliability and Validity, Fidelity of
Implementation, and Impact on Children**

Montreal, Canada

April 11, 2005

ELLM Presenters

Dr. Madelaine M. Cosgrove

Associate Director for School Readiness, Florida Institute of Education

PCER, Co-Investigator

Dr. Cheryl A. Fountain

Executive Director, Florida Institute of Education

PCER, Principal Investigator

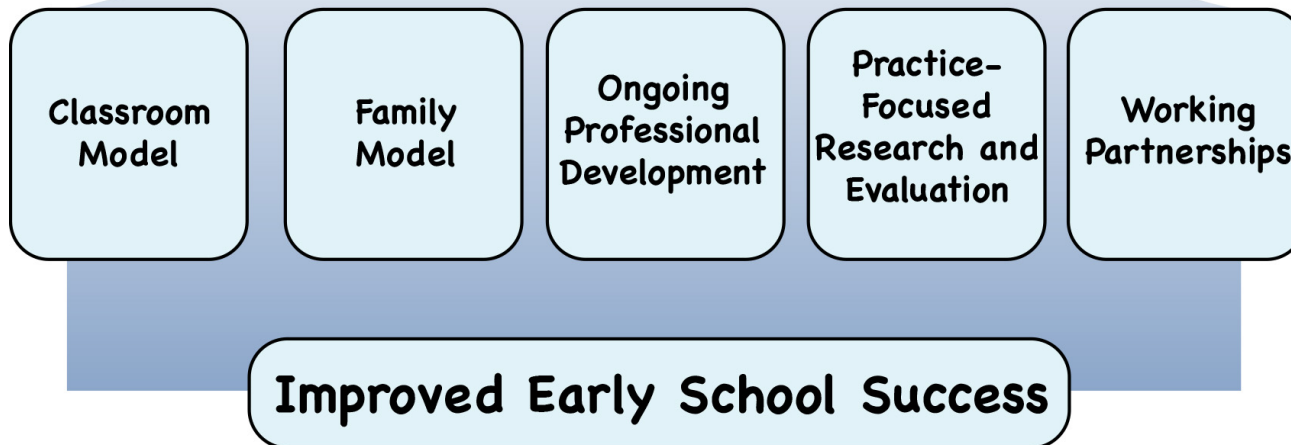
Dr. Janice A. Wood

Crawford Early Literacy Faculty Fellow, Florida Institute of Education

PCER, Co-Investigator

The Preschool Curriculum Evaluation Research (PCER) program funded by the Institute of Education Sciences (IES), U.S. Department of Education includes a national evaluation study conducted by RTI International and Mathematica Policy Research (MPR), and complementary research studies conducted by each grantee. The findings reported here are based on the complementary research activities carried out by the Florida Institute of Education at the University of North Florida under the PCER program. These findings may differ from the results reported for the PCER national evaluation study. The findings presented in the Paper at the Annual Meeting of the American Educational Research Association, Montreal, 2005 are based on a larger sample size of children, classroom and teachers and sought to answer complementary research questions including program effectiveness. The content of this presentation does not necessarily reflect the views or policies of the PCER Consortium including IES, RTI, and MPR, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Department of Education.

Early Literacy and Learning Model (ELLM) Early Literacy Curriculum and Instructional Support System



Introduction to the Study

This pilot study presents evidence of the reliability and validity of fidelity-of-implementation instrument designed to measure teachers' instructional practices using the critical components of emergent literacy:

- Read Aloud to Children/Print Concepts
- Oral Language
- Emergent Writing
- Letter/Sound
- Phonological Awareness

The instrument was designed to measure two aspects of each component, *Pedagogical Quality* and *Time-on-Task*.

Methodology

- To determine teachers' level of fidelity-of-implementation, teachers were videotaped for an entire day during March and again in May.

Participants

- Teachers participating in the PCER study were assigned to ELLM or wait-list control (W-L Control).
- ELLM teachers received intensive 2-day summer training and on-going professional development support focused on helping them to become more proficient in implementing instructional strategies and activities that address the 5 components.
- W-L Control teachers received no training in addition to that routinely provided by the curriculum developers and center directors.

Participants 2

Table of Participating Teachers

Time	Number of ELLM Teachers	Number of W-L Control Teachers
Time 1	19	10
Time 2	20	10
Both Times	17	5

Segmenting Tapes

Trained assessors segmented the tapes to:

1. identify the implementation of the literacy components,
2. determine the total length in minutes of the segments using each literacy component (*Time-on-Task*), and
3. rate the levels of *Pedagogical Quality* using specific literacy-component items on the instrument.

Assessor Training

- Assessors met weekly to practice segmenting the videotapes to identify literacy component episodes, clarify component boundaries, and establish videotaping guidelines.
- Group and individual practice continued until assessors were consistent segmenting tapes, establishing Time-on-Task, and rating Pedagogical Quality. At this point, assessors were certified to use the fidelity-of-use instrument.

Results

The pilot study of the fidelity-of-use instrument presents evidence of the reliability and validity of scores obtained by trained assessors using the instrument.

Results 2

Table of Means

Method	Trait	Time 1 W-L Control (n=10)	Time 1 ELLM (n=19)	Time 2 W-L Control (n=10)	Time 2 ELLM (n=20)
Pedagogical Quality	Print Concepts	14.7*	63.5*	52.4*	62.9*
Pedagogical Quality	Oral Language	75.3*	75.7*	76.3*	75.0*
Pedagogical Quality	Emergency Writing	25.0*	44.7*	32.5*	35.8*
Pedagogical Quality	Phonological Awareness	6.3*	50.7*	7.5*	29.4*
Pedagogical Quality	Letter/Sound	38.6*	51.1*	24.3*	55.0*
Time on Task	Print Concepts Time	4.4**	8.4**	8.9**	13.3**
Time on Task	Oral Language Time	12.1**	11.0**	27.1**	11.4**
Time on Task	Emergent Writing Time	4.6**	13.7**	6.5**	13.6**
Time on Task	Phonological Awareness Time	0.4**	5.6**	1.0**	1.9**
Time on Task	Letter/Sound Time	2.9**	11.3**	2.9**	6.4**

* Mean Percent of Items

** Mean Time on Task (in minutes)

Reliability

- Cronbach's alpha was calculated for each *Pedagogical Quality* literacy component for each videotaping time.
- The results were averaged to form overall measures of internal consistency.

Cronbach's Alpha Coefficient of Internal Consistency

Category	Number 1 (n=29)	Number 2 (n=30)	Average
Print Concepts	.9365	.9390	.9378
Oral Language	.7019	.6902	.6961
Emergent Writing	.8379	.8952	.8666
Phonological Awareness	.9006	.9209	.9108
Letter/Sound	.9105	.9353	.9229

- Inner-rater reliability for *Time-on-Task* and *Pedagogical Quality* is yet to be determined.

Validity

Evidence supporting the validity of the Pedagogical Quality component comes from three sources:

- External construct validity using a Multitrait-Multimethod
- Matrix to present evidence of trait validity.
- Construct validity evidenced by the convergence of the measures across groups. Because videotaped teachers' are either ELLM or W-L Control, their fidelity-of-use Pedagogical Quality scores should differ in predictable ways.
- Evidence of nomological validity showing that the theoretical basis of the Pedagogical Quality scores of teachers provide a link to their students' literacy-related outcomes.

Trait Validity: Multitrait-Multimethod Matrix

The coefficients in the bold (lower left) validity diagonal provides convergent evidence. The lack of significant correlations in the yellow triangles provides evidence that measures discriminate among traits.

Multi-Trait Multi-Method Matrix

Pedagogical Quality - Pedagogical Quality

Trait	PC	OL	EW	PA	L/S
Print Concepts (PC)	.94	-	-	-	-
Oral Language (OL)	.40	.70	-	-	-
Emergent Writing (EW)	.00	-.20	.87	-	-
Phonological Awareness (PA)	.41	-.24	-.04	.91	-
Letter/Sound (L/S)	.33	.23	.23	.28	.92

Pedagogical Quality - Time on Task

Trait	PC	OL	EW	PA	L/S
Print Concepts (PC)	-	-	-	-	-
Oral Language (OL)	-	-	-	-	-
Emergent Writing (EW)	-	-	-	-	-
Phonological Awareness (PA)	-	-	-	-	-
Letter/Sound (L/S)	-	-	-	-	-

Time on Task - Pedagogical Quality

Trait	PC	OL	EW	PA	L/S
Print Concepts (PC)	.79	.40	-.20	.07	.39
Oral Language (OL)	-.17	.51	-.40 ¹	-.46	-.05
Emergent Writing (EW)	.11	-.27	.69	.26	-.10
Phonological Awareness (PA)	.04	-.47	.09	.63	.26
Letter/Sound (L/S)	.44	.15	.35	.08	.70

Time on Task - Time on Task

Trait	PC	OL	EW	PA	L/S
Print Concepts (PC)	*	-	-	-	-
Oral Language (OL)	.01	*	-	-	-
Emergent Writing (EW)	-.15	-.29	*	-	-
Phonological Awareness (PA)	-.05	-.38 ¹	.26	*	-
Letter/Sound (L/S)	.52	.12	.21	.08	*

n=22 time1 and time2 averaged scores.

Numbers on the main diagonal are reliabilities.

*Represents the non-available reliabilities of the Time on Task scores.

Bold-faced and italicized correlations are significantly different from zero, $\alpha=.05$.

¹ Indicates $\alpha=.10$

Lower diagonal bold-faced numbers represent monotrait-heteromethod correlations.

Numbers in the green triangles represent heterotrait-monomethod correlations.

Numbers in the yellow triangles represent heterotrait-heteromethod correlations.

Trait Validity: Multitrait-Multimethod Matrix 2

- The significant correlations between the scores of Oral Language and other components indicate Oral Language items do not distinguish the component from the other literacy-related components.
- This, together with the consistent *Pedagogical Quality* mean scores across ELLM and W-L Control teachers and the lowest of the reliabilities, led to the decision to remove the component from the pilot study.

Multi-Trait Multi-Method Matrix 2

Pedagogical Quality - Pedagogical Quality

Trait	PC	EW	PA	L/S
Print Concepts (PC)	.94	-	-	-
Emergent Writing (EW)	.00	.87	-	-
Phonological Awareness (PA)	.41	-.04	.91	-
Letter/Sound (L/S)	.33	.23	.28	.92

Pedagogical Quality - Time on Task

Trait	PC	EW	PA	L/S
Print Concepts (PC)	-	-	-	-
Emergent Writing (EW)	-	-	-	-
Phonological Awareness (PA)	-	-	-	-
Letter/Sound (L/S)	-	-	-	-

Time on Task - Pedagogical Quality

Trait	PC	EW	PA	L/S
Print Concepts (PC)	.79	-.20	.07	.39
Emergent Writing (EW)	.11	.69	.26	-.10
Phonological Awareness (PA)	.04	.09	.63	.26
Letter/Sound (L/S)	.44	.35	.08	.70

Time on Task - Time on Task

Trait	PC	EW	PA	L/S
Print Concepts (PC)	*	-	-	-
Emergent Writing (EW)	-.15	*	-	-
Phonological Awareness (PA)	-.05	.26	*	-
Letter/Sound (L/S)	.52	.21	.08	*

Construct Validity: Convergence Across Groups

ELLM teachers have experienced targeted professional development designed to improve critical literacy-related skills; therefore, higher *Pedagogical Quality* scores were expected.

ANOVA Table

Pedagogical Quality	Source of Variance	F Ratio Time 1 (n 29)	Probability Time 1	F Ratio Time 2 (n 30)	Probability Time 2
Print Concepts	Status*Education	0.02	.8857 [‡]	7.83	.0095 [‡]
Print Concepts	Education	0.12	.7304	-	-
Print Concepts	Status	29.24	<.0001	-	-
Emergent Writing	Status*Education	0.20	.6602 [‡]	0.64	.4317 [‡]
Emergent Writing	Education	0.01	.9045	0.58	.4522
Emergent Writing	Status	4.38	.0458	0.08	.7832
Phonological Awareness	Status*Education	0.55	.4664 [‡]	0.17	.6879 [‡]
Phonological Awareness	Education	0.01	.9045	0.47	.4970
Phonological Awareness	Status	15.78	.0458	3.12	.0880**
Letter/Sound	Status*Education	2.57	.1212 [‡]	0.00	.9576 [‡]
Letter/Sound	Education	0.60	.4447	1.71	.2020
Letter/Sound	Status	0.70	.4095	3.95	.0568**

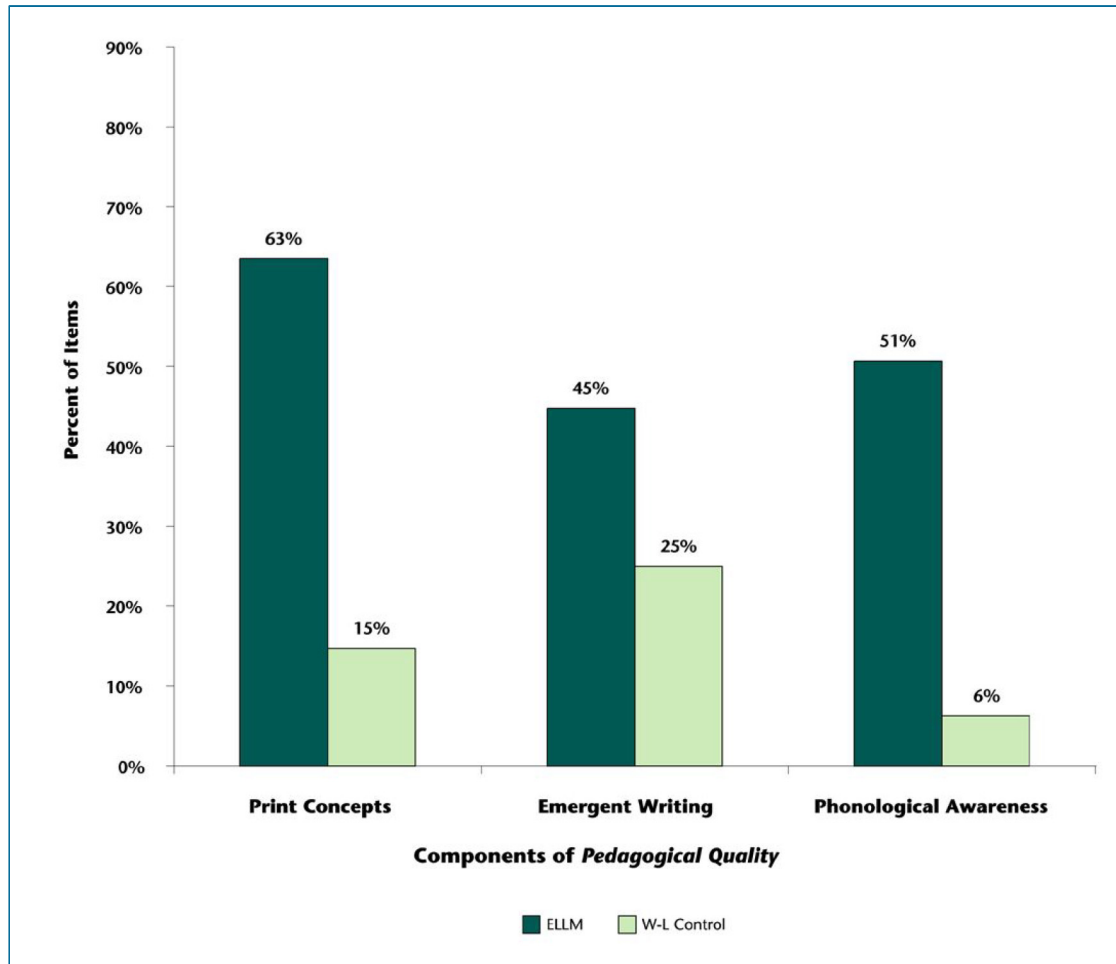
Note: [‡] Indicates an education level and status interaction that was not significant and was removed from the analysis.

* Indicates a significant effect at $\alpha = .05$. ** Indicates a significant effect at $\alpha = .10$.

Time 1: March 2003

ELLM teachers, regardless of educational level, have higher *Pedagogical Quality* scores at Time 1 for Print Concepts, Emergent Writing, and Phonological Awareness than W-L Control teachers.

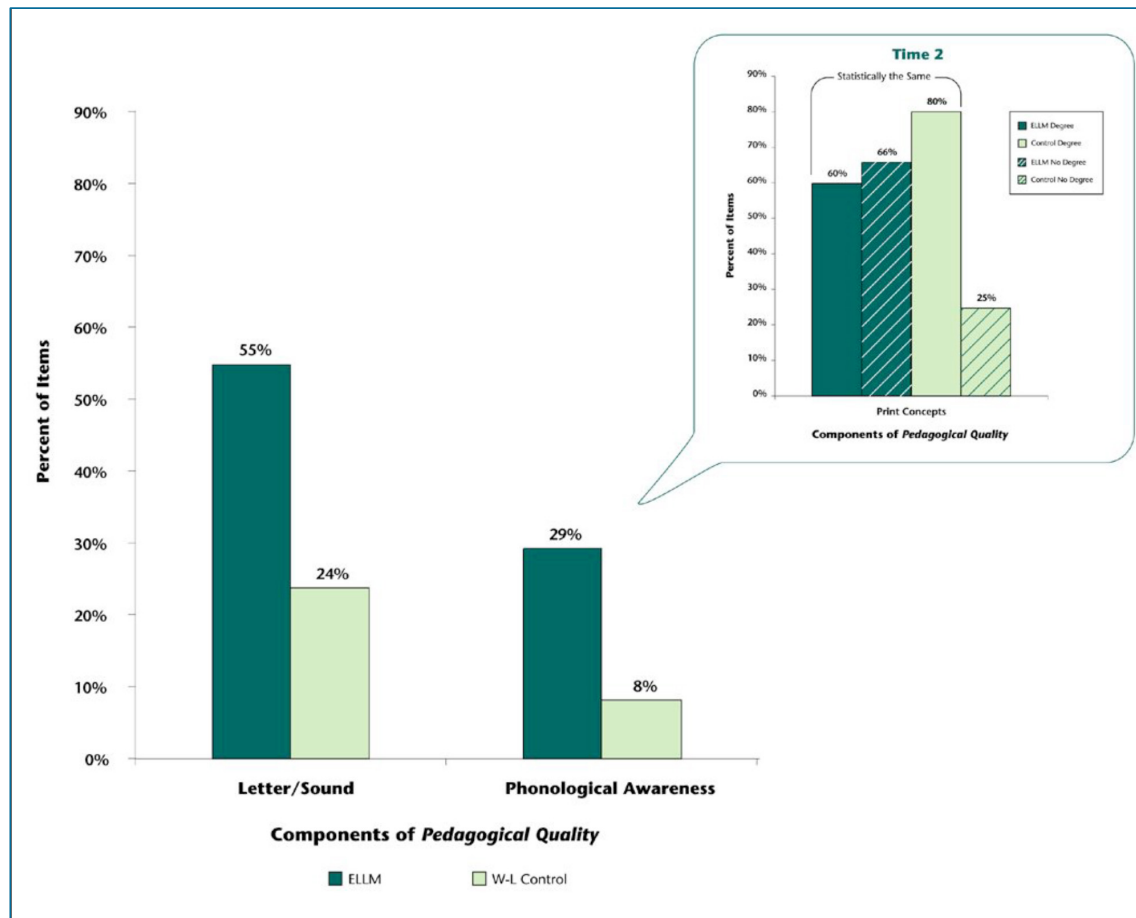
Time 1: March 2003 2



Time 2: May 2003

- At Time 2, ELLM Teachers, regardless of educational level have higher *Pedagogical Quality* scores for Phonological Awareness and Letter/Sound.
- Additionally, at Time 2, ELLM teachers had higher *Pedagogical Quality* scores for Print Concepts than W-L Control non-degreed teachers.

Time 2: May 2003 2



Nomological Validity: Links to Student Outcomes

- Teachers' scores on the *Pedagogical Quality* measure of the components were expected to correlate with students' literacy related outcomes, as measured by the Test of Early Reading Ability, Third Edition (TERA-3) and the Alphabet Letter Recognition Inventory (ALRI).
- The ALRI measures a child's ability to recognize all upper- and lowercase letters of the alphabet arranged in non-alphabetical order. With the exception of the correlations involving the TERA-3 Meaning subtest, the correlations converged as expected.

Print Concepts

The *Pedagogical Quality* of Print Concepts is measured during segments in which the teacher is reading aloud to children; therefore, it is not surprising the children also gain understanding that print conveys meaning during these activities.

First Order Correlations

Test	Posttest	Pedagogical Quality Component Correlations
TERA-3	Reading Quotient	Print Concepts (.1787)
TERA-3	Conventions of Print	Print Concepts (.1544), Letter/Sound (.1184*)
TERA-3	Meaning	Print Concepts (.2440), Phonological Awareness (.1688)
ALRI	-	Emergent Writing (.1993)

Note: Numbers in parenthesis are correlations.

Conclusions

- Evidence of the reliability of *Pedagogical Quality* was found using the Cronbach's coefficient alpha measure of internal consistency.
- Evidences of three types of validity were also presented for *Pedagogical Quality*.
 - External Construct Validity
 - Construct Validity
 - Nomological Validity

Conclusions 2

- Analyses of the videotapes indicate that ELLM teachers' *Pedagogical Quality* scores on the literacy components were generally higher than the scores of the W-L Control teachers.
- These differences in *Pedagogical Quality* scores were found regardless of education level, with the exception of one comparison.
- *Pedagogical Quality* scores on the literacy components are positively correlated with students' posttest scores on the TERA-3 and ALRI.

Next Steps

- Provide evidence of inner-rater reliability for both the *Time-on-Task* and *Pedagogical Quality* measures of teachers' use of the critical literacy components in their instruction.
- Restructure the Oral Language component items to clearly establish distinguishable boundaries for the construct.

Next Steps 2

- Improve the Letter/Sound measure. The teachers' scores should correlate with students' test scores of alphabet letter knowledge.
- Provide evidence of the validity of the *Time-on-Task* measures.
- Use the videotape fidelity-of-use instrument to develop a classroom observation scale to measure teachers' use of the critical components of emergent literacy.

*For more information visit us at:
www.unf.edu/dept/fie/ellm*