Supporting Quality Services:
An Examination of the Reliability and Validity of the
Early Learning Accomplishment Profile (Early LAP)

Paper presentation at the
American Educational Research Association Annual Meeting
April, 2003

Ellen S. Peisner-Feinberg
FPG Child Development Institute
University of North Carolina-Chapel Hill
Chapel Hill, NC

Belinda J. Hardin
Chapel Hill Training Outreach Project
Chapel Hill, NC

Contact Information: Belinda J. Hardin, Ph.D., Chapel Hill Training Outreach Project, 800 Eastowne Drive, Suite 105, Chapel Hill, NC 27514; E-mail: bhardin@chtop.org

Introduction

More than six million infants and toddlers in the United States receive some form of child care each day (Children's Defense Fund, 2001). However, many studies indicate that the average quality of care for infants and toddlers is inadequate, and tends to be even lower than the quality of services for preschool-age children. For example, the *Cost, Quality, and Outcomes Study* results indicated that 40% of the classrooms serving infants and toddlers had a rating of "poor quality" (Cost, Quality, and Outcomes Team, 1995). Yet, researchers have found that better quality care is associated with better cognitive and social outcomes for children (e.g., Peisner-Feinberg et al., 2001).

One indicator of high quality child care is individualized instruction that includes meaningful and stimulating experiences for children. In order to provide appropriate, individualized instruction, child care staff need to have a comprehensive understanding of the child, including gathering information from the family, reviewing official records (e.g., health records); and assessing the child's development. Currently, there are few assessment instruments for infants and toddlers that can be used by practitioners to create an in-depth profile of a child's development. Current examples of tools used by practitioners to understand infant and toddler development include screening tools (e.g., Brigance Screens, Denver Developmental Screening Test-II) or parent questionnaires (e.g., Ages and Stages Questionnaires).

A more comprehensive, criterion-referenced measure designed for use by practitioners to assess the development of infants and toddlers is the Early Learning Accomplishment Profile (Early LAP). This measure provides a systematic method for observing children functioning in the birth to 36-month age range across six domains of development: gross motor, fine motor, cognitive, language, self-help, and social emotional. The Early LAP is designed so that the results can be used by practitioners to facilitate the planning and implementation of individualized, developmentally appropriate activities in the classroom. However, there is little written information about the reliability and validity of the Early LAP, even though it has been widely used across the United States since the 1970s (Fleming, 2000). It is important that the tools used by practitioners meet the same measurement standards as other research tools if they are to offer an accurate picture of a child's functioning and provide an appropriate basis for individualized instruction.

Purpose of Study

The purpose of this research was to examine and document the reliability and validity of the Early LAP. Three types of studies were conducted:

- Criterion Validity--the extent to which individual scores on the Early LAP corresponded with scores on a similar test. Both the Early LAP and the Bayley Scales of Infant Development (BSID-II) were administered to the same child in the same or consecutive sessions.
- Test-Retest Reliability--the extent to which scores on the Early LAP were consistent from one period of time to the next. The Early LAP was administered and then re-administered to the same child on two different occasions by the same examiner one to three weeks apart.
- Interrater Reliability--the extent to which different examiners achieved the same results when

independently assessing the same child using the Early LAP. The Early LAP was administered and then re-administered to the same child by two different examiners one to three weeks apart.

In addition, the data were used to calculate other statistical properties of the Early LAP, including correlations between developmental age scores and chronological age, internal consistency, and construct validity.

Methods

Sample

The sample included 285 children ages 2 to 44 months old, including children with typical and atypical development. A stratified sampling procedure was used based on geographic region, age, race, gender, and type of setting.

Geographic Distribution of Project Sites

Four sites were selected to represent the major geographic regions of the United States with similar sample sizes in each region: Northeast (New York City and Westchester County, New York), South (Orange County, North Carolina), North Central (Greater Metropolitan Area of Kansas City and Lawrence, Kansas), and West (San Jose, California). Table 1 depicts the distribution of the sample across the four sites.

Table 1. Geographic Distribution of Sample

Geographic Area	Number of Children	Percentage of Sample
Northeast	59	20.7%
South	68	23.9%
North Central	81	28.4%
West	77	27.0%
Total	285	100%

Participant Characteristics

Age and Gender. Children were recruited for the study in the following seven age categories: birth to 6 months, 7-12 months, 13-18 months, 19-24 months, 25-30 months, 31-36 months, and 37-44 months. Children in the oldest category (37-44 months) were included to demonstrate the upper age limit for the Early LAP. An approximately equal number of males and females were selected for the sample. Table 2 shows the distribution of the Project Sample by age category and gender.

Table 2. Number of Participants by Age and Gender

Age Category	Mean Age (SD)	Females	Males	Total Number (percent)
Birth – 6 months	4.49 (1.38)	18 (51.4%)	17 (48.6%)	35 (12.3%)
7-12 months	9.77 (1.54)	30 (68.2%)	14 (31.8%)	44 (15.4%)
13-18 months	15.60 (1.72)	19 (42.2%)	26 (57.8%)	45 (15.8%)
19-24 months	22.00 (1.81)	19 (38.0%)	31 (62.0%)	50 (17.5%)
25-30 months	27.63 (1.74)	13 (31.7%)	28 (68.3%)	41 (14.4%)
31-36 months	33.70 (1.54)	16 (40.0%)	24 (60.0%)	40 (14.0%)
37-44 months	39.90 (2.22)	10 (33.3%)	20 (66.7%)	30 (10.5%)
Total	21.28 (11.29)	125 (43.9%)	160 (56.1%)	285 (100%)

Race/Ethnicity. To represent the variety of cultural and ethnic groups in the United States, children were proportionally selected for the sample to reflect the major racial/ethnic groups indicated in the U.S. Census Bureau population projections for the Year 2000 (U.S. Census Bureau, 1995). Table 3 depicts the racial/ethnic distribution by geographic region.

Table 3. Number of Participants by Race/Ethnicity and Geographic Region

Race or Ethnic Group	Northeast	South	North Central	West	Total**
African American	23	8	20	6	57 (20.0%)
American Indian, Eskimo, and Aleut	2	0	2	1	5 (1.8%)
Asian and Pacific Islander	1	1	1	5	8 (2.8%)
Hispanic origin	10	3	1	13	27 (9.5%)
White	15	52	50	52	169 (59.3%)
Other*	8	4	7	0	19 (6.6%)
Total	59	68	81	77	285 (100%)

* The majority of the children classified as Other were bi-racial children.

** The 1995 U.S. Census Bureau population projections for Year 2000 were: African American - 13%; American Indian, Eskimo, and Aleut - 1%; Asian and Pacific Islander - 4%; Hispanic origin - 11%; and White - 71%.

Program Types

Because the Early LAP is used for in-home services as well as group-based settings for infants and toddlers, children were recruited from three different settings: child care center programs (n=179, 62.8%); Early Head Start programs (n=37, 13.0%); and individual homes (n=69, 24.2%). A total of 29 child care programs participated in the study, reflecting a range of types of centers, including Early Head Start centers, developmental day centers (serving only children with disabilities), private child care centers, and other public centers.

Measures

Early LAP

The Early LAP is a criterion-referenced measure designed for children birth to 36 months of age. The Early LAP assesses skills in six developmental domains: gross motor, fine motor, cognitive, language, self-help, and social emotional. A hierarchy of developmental skills in each domain is

arranged according to chronological age categories. Each item contains specific material recommendations, procedures for administration, and scoring criteria.

Bayley Scales of Infant Development—Second Edition (BSID-II)

The *Bayley Scales of Infant Development—Second Edition (BSID-II)* was used as the criterion validity instrument for the study. The BSID-II is an individually administered and norm-referenced instrument that assesses the developmental functioning of children ages 1-42 months old. The BSID-II consists of three scales: the Mental Scale, Motor Scale, and Behavior Rating Scale. Only the Mental and Motor Scales were used in the study. The BSID-II is a well-known and widely-used measure with good reliability and validity characteristics. For example, the test-retest reliability coefficients in the test development sample were .83 for the Mental Scale and .77 for the Motor Scale and the interrater reliability coefficients were .96 for the Mental Scale and .75 for the Motor Scale. The individual Scales also were reported to have good internal consistency. The alpha coefficient for the Mental Scale was .88 and for the Motor Scale was .84.

Parent Questionnaire

A parent questionnaire was distributed with the permission letters. The parent questionnaire contained demographic information, child background information (e.g., primary language, family income, parents' education and age), and five items on the Early LAP that data collectors could not observe or caregivers might not be able to report (e.g., bathing practices). Completion of the demographic information was required for study participation. However, the other information was optional.

Procedures

A team of nine data collectors trained and supervised by the project co-directors collected the data. Each data collector had a master's or doctoral degree in education, early intervention services, or another related field. The data collectors participated in a two-and-a-half day training session on the Early LAP, BSID-II, and procedures for data collection, and collected the data over an 8-month period.

Within each site, children were recruited from child care centers and individual families. An effort was made to include settings representing children from a range of socioeconomic groups. Each center director, or parent in the case of home settings, was contacted by phone and invited to participate in the study. Copies of the Early LAP manual and letters describing the study and requesting consent to participate were shared and discussed during a subsequent meeting. In the case of child care and Early Head Start programs, each center director agreed to distribute and collect permission forms for parents interested in participating in the study.

Each child in the study was administered both the Early LAP and the BSID-II Mental and Motor Scales. These assessments were administered during the same testing session or in two sessions (30 minutes to one hour each) in close proximity.

A second Early LAP was given to 141 children for the test-retest (n=92) or interrater reliability (n=49) studies. The children participating in these two studies reflected a similar distribution in geographic region, age, gender, and race/ethnicity to the overall sample. For the test-retest reliability study, the same examiner administered the Early LAP on two separate occasions one to three weeks apart. For the interrater reliability study, two different examiners administered the Early LAP on two separate occasions one to three weeks apart. Because the Early LAP measures a continuum of developmental skills, the test-retest and interrater reliability were measured over a short period of time so that any differences between administrations were more likely to reflect reliability rather than individual development.

Results

The project sample, consisting of 285 children from 2 to 44 months old, included 242 typically developing children from 2 to 36 months old (referred to as the core sample), 26 children with typical development from 37 to 44 months old, and 17 children with professionally diagnosed disabilities from 6 to 43 months old. While the Early LAP is designed for children in the birth to 36 month age range, a sample of older children was included to verify the upper age limit of the measure. On average, 81% of these children completed the items in a domain before reaching a ceiling, confirming that the Early LAP is not an appropriate instrument for children with typical development above 36 months of age. The sample of atypically developing children was included to examine whether the Early LAP could be used appropriately with children with disabilities. These children represented 6% of the total project sample, which reflects the national proportion of children under age 18 with disabilities (U.S. Census Bureau, 1995). Of these children, 10 had developmental delays, five had motor disabilities, and two had speech and language disabilities. As expected, the developmental age scores for each domain (means=14.94 to 19.42) were substantially lower than the children's chronological ages (mean=25.47), and the correlations between developmental age scores and chronological age (.69 to .74) were substantially lower than the correlations for children with typical development.

Table 4 presents the means and standard deviations for the Early LAP developmental age domain scores and each of the age categories for the core sample.

6

Table 4. Means and Standard Deviations of Early LAP Developmental Age Domain Scores by Chronological Age Clusters for Core Sample

Domain/ Chronological Age Clusters	Mean	Standard Deviation
Gross Motor		
Birth -6 months	4.45	1.66
7-12 months	10.95	2.90
13-18 months	17.88	5.41
19-24 months	25.05	5.24
25-30 months	29.51	6.02
31-36 months	33.56	4.48
Fine Motor		
Birth -6 months	4.78	2.08
7-12 months	10.73	1.94
13-18 months	16.16	2.96
19-24 months	22.91	4.46
25-30 months	28.60	5.09
31-36 months	33.35	3.60
Cognitive		
Birth -6 months	5.03	1.69
7-12 months	10.48	2.14
13-18 months	15.77	3.51
19-24 months	22.69	4.59
25-30 months	28.81	5.55
31-36 months	34.37	2.94
Language		
Birth -6 months	5.03	1.27
7-12 months	10.28	3.13
13-18 months	16.60	4.29
19-24 months	23.41	5.60
25-30 months	29.19	6.07
31-36 months	34.20	3.93
Self Help		
Birth -6 months	6.78	.67
7-12 months	12.76	2.53
13-18 months	16.80	2.99
19-24 months	23.61	4.96
25-30 months	28.65	5.25
31-36 months	33.77	5.06
Social Emotional		
Birth -6 months	4.61	1.61
7-12 months	10.63	2.92
13-18 months	17.62	7.26
19-24 months	28.67	6.91
25-30 months	32.50	5.02
31-36 months	34.96	2.88

Reliability

The reliability of an assessment instrument refers to its accuracy and consistency over time. For example, an assessment instrument should produce roughly the same results when the same individuals are tested under similar conditions within a short period of time. Analyses of the reliability for each domain of the Early LAP, including correlations with age, internal consistency, test-retest reliability, and interrater reliability were conducted. Although children were recruited for the sample in 6-month increments of age, for purposes of analysis the data were collapsed into 12-month increments to ensure sufficient sample sizes in each cell. Every

effort was made to gather complete data for each child; however, in some cases there were missing items that prevented calculation of a domain score. In most cases, the missing data were caused by the inability to observe particular behaviors due to the unavailability of large materials (e.g., furniture, stairway) or a restricted number of test items in a developmental range (e.g., self-help begins at six months and social emotional has a limited number of items for some age ranges).

Correlations Between Chronological Age and Developmental Age Scores

The correlations between the Early LAP developmental age scores and chronological age were computed for the core sample (242 children with typical development in the birth to 36 month age range) using Pearson product-moment correlation coefficients (*r*). Table 5 presents the correlation coefficients by domain and age group. These results indicate strong correlations (.90 to .95) between chronological age and developmental age in each domain for the overall sample. Within age groups, the correlations for the 2-12 month old age range (.80 to .91) and the 13-24 month old age range (.63 to .73) are fairly high, while the correlations for the 25-36 month old age range (.31 to .61) are moderate. These findings suggest that the developmental age scores on the Early LAP are reliably associated with chronological age for younger children, but that the association decreases somewhat as children become older and begin to age out of certain items and/or domains on the test.

Table 5. Correlations Between Chronological Age and Developmental Age Scores by Domain and Age Group

DOMAINS	2 months to 12 months[a]	13 months to 24 months[b]	25 months to 36 months[c]	Total[d]
Gross Motor	.88	.63	.45	.91
Fine Motor	.91	.72	.51	.94
Cognitive	.92	.73	.61	.95
Language	.82	.70	.43	.92
Self-Help	.80	.73	.60	.91
Social Emotional	.85	.68	.31	.90

Note: For all correlations, p <. 01

N: a (GM=73, FM=73, C=73, L=68, SH=47, SE=69)   b (GM=85, FM=86, C=88, L=87, SH=85, SE=75)
    c (GM=71, FM=74, C=72, L=72, SH=68, SE=57)   d (GM=229, FM=233, C=233, L=227, SH=200, SE=201)

Internal Consistency

The internal consistency of the Early LAP was examined to determine how well the items in each domain relate to one another. The internal consistency coefficient indicates how effectively the individual domain scores on the Early LAP are measuring defined constructs (e.g., gross motor, fine motor, cognitive skills). Cronbach's coefficient alpha was used to calculate the internal consistency of each domain for the total core sample (n=242) by age groups.

Table 6 presents the results of the internal consistency analyses. The alpha coefficients for the total core sample (.96 to .99) indicate strong internal consistency for each domain. The alpha coefficients for the individual age groups are also quite high (.84-.98). These results indicate that the Early LAP items show strong internal consistency within each domain.

Table 6. *Internal Consistency of Early LAP Developmental Age Domain Scores by Age Group*

DOMAINS	2 months to 12 months[a]	13 months to 24 months[b]	25 months to 36 months[c]	Total[d]
Gross Motor	.98	.97	.84	.99
Fine Motor	.96	.94	.90	.98
Cognitive	.97	.96	.96	.99
Language	.91	.96	.95	˙.98
Self-Help	.97	.96	.93	.98
Social Emotional	.91	.91	.87	.96

Note: For all correlations, p <.01

N: a (GM=75, FM=75, C=75, L=75, SH=72, SE=74)   b (GM=80, FM=88, C=88, L=88, SH=87, SE=84)

   c (GM=62, FM=75, C=72, L=73, SH=72, SE=72)   d (GM=217, FM=238, C=235, L=236, SH=231, SE=230)

Test-Retest Reliability

Test-retest reliability indicates the extent to which scores on an assessment instrument are consistent from one time period to the next. Test-retest data were collected for 92 children, where the same examiner administered the Early LAP on two separate occasions. Test-retest reliability was determined by calculating the correlations between scores from the first and the second test administrations for each domain using Pearson's *r*. Table 7 presents the means and standard deviations for the first and second test scores and the test-retest correlation coefficients for each domain. The resulting correlations (.96 to .99) indicate a high degree of stability in individual test scores over short intervals of time.

Table 7. Means, Standard Deviations, and Correlations of Early LAP Developmental Age Domain Scores for Test-Retest Reliability Sample

	First Testing		Second Testing		
DOMAINS	Mean	SD	Mean	SD	*r*
Gross Motor	19.16	11.33	19.15	11.65	.98
Fine Motor	19.03	10.13	19.81	10.72	.97
Cognitive	18.59	10.43	19.26	10.52	.99
Language	18.93	10.44	18.57	10.52	.99
Self-Help	20.91	9.29	20.94	9.41	.96
Social Emotional	19.85	11.66	20.20	11.90	.99

Note: For all correlations, p <.01

N: GM=79, FM=85, C=84, L=81, SH =72, SE =69

Interrater Reliability

The results of an assessment should reflect the developmental skills of the child independent of the particular person administering the test, assuming proper procedures have been followed. Interrater reliability measures the extent to which the same results are achieved by different examiners when independently assessing the same child. Interrater reliability data were gathered for 49 children, where the Early LAP was administered to the same child by two different examiners on two separate occasions. Interrater reliability was determined by computing the correlations between the scores from the two test administrations for each domain using Pearson's *r*. Table 8 presents the means and standard deviations for both test administrations and the interrater reliability correlation coefficients for each domain. The resulting correlations indicate a high degree of reliability (.96 to .99) when two different examiners administer the Early LAP.

Table 8. *Means, Standard Deviations, and Correlations of Early LAP Developmental Domain Age Scores for Interrater Reliability Sample*

	First Testing		Second Testing		
DOMAINS	Mean	SD	Mean	SD	*r*
Gross Motor	22.63	10.41	22.02	10.02	.97
Fine Motor	21.46	10.06	21.74	9.89	.99
Cognitive	21.42	10.37	21.98	10.32	.99
Language	22.63	10.37	22.89	10.28	.96
Self-Help	24.17	10.78	23.27	9.03	.96
Social Emotional	23.45	12.06	24.07	11.60	.98

Note: For all correlations, p < .01
N: GM=44, FM=47, C=47, L=46, SH=39, SE=43

Validity

The foremost authoritative reference on validity and other test matters, the *1999 Standards for Educational and Psychological Testing*, defines validity as, "The degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses of a test." (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999, p.184). This definition emphasizes that inferences derived from test scores give meaning to them beyond simply reporting numbers. Two types of validity analyses were conducted as part of this study: construct validity and criterion validity.

Construct Validity

Evidence of construct validity can be inferred by examining the intercorrelations among different areas of an assessment instrument. Thus, to examine the extent to which the different domains of

the Early LAP measure different skills, the intercorrelations among domains were calculated. High correlations among areas would suggest that they are measuring similar underlying constructs, while low correlations would suggest that they are measuring different underlying constructs. Domains that are more strongly related conceptually and that have more items in common would be expected to have relatively stronger intercorrelations. Zero-order correlations using Pearson's r were calculated between developmental age domain scores for the core sample (n=242) as shown below the diagonal in Table 9. While these high positive correlations (.87 to .97) potentially indicate a single underlying construct, because these zero order correlations were calculated across age groups, they also indicate differences in skill performance as a result of age.

To separate these two elements, partial correlations controlling for age were calculated between developmental age domain scores as depicted above the diagonal in Table 9. The magnitudes of the partial correlation coefficients are substantially smaller than the zero-order correlations (.27 to .76), in the modest to moderate range. The relatively higher correlations among the fine motor, cognitive, and language domains are likely a result of a number of shared items, while the less conceptually related domains evidence lower correlations. These results suggest that while the different domains of the Early LAP are somewhat related, they are also measuring somewhat independent aspects of development.

Table 9. *Zero-order Correlations and Partial Correlations Controlling for Age Among Early LAP Domains*

DOMAINS	Gross Motor	Fine Motor	Cognitive	Language	Self-Help	Social Emotional
Gross Motor		.50	.53	.51	.31	.58
Fine Motor	.93		.76	.57	.33	.39
Cognitive	.93	.97		.71	.48	.34
Language	.92	.94	.96		.27	.44
Self-Help	.88	.91	.93	.88		.27
Social Emotional	.93	.97	.90	.91	.87	

Note: For all correlations, p <. 01
N: GM=229, FM=233, C=233, L=227, SH=200, SE=201

Criterion Validity

Criterion validity (also known as concurrent validity) is the extent to which individual scores on one test correspond to scores on an established test of similar constructs. These two tests must be administered consecutively, so as to minimize differences due to development or other variations in test conditions. The established test is the criterion used to validate the new test (Gall, Borg, & Gall,1996). In this study, the correspondence between the Early LAP and the Mental and Motor

Scales of the BSID-II was examined to investigate the criterion validity of the Early LAP. The core sample (n=242) was administered both the Early LAP and the BSID-II Mental and Motor Scales during the same testing session or in two sessions in close proximity. Criterion validity was determined by examining the correlations, using Pearson's r, between the Early LAP domain scores and the BSID-II Mental and Motor Scale scores for conceptually related areas.

Table 10 presents the correlations between the developmental age scores for the Early LAP domains and the BSID-II Mental and Motor Scales by age group. The results indicate a strong correlation (.90 to .97) between the Early LAP and BSID-II scores in each domain for the overall sample. Fairly high correlations were found within the 2-12 month old age range (.83 to .95) and the 13-24 month old age range (.72 to .88). The correlations for the 25-36 month old age range (.47 to .83) were somewhat lower, particularly in the self-help and social emotional domains. These somewhat lower correlations for the oldest age group may reflect some aging out of the Early LAP assessment for these children, similar to the analyses discussed earlier.

Table 10. *Correlations Between the Early LAP Domains and the BSID-II Mental and Motor Scales for Developmental Age Scores by Age Group*

	2 months to 12 months[a]		13 months to 24 months[b]		25 months to 36 months[c]		Total[d]	
Early LAP DOMAINS	BSID-II Motor	BSID-II Mental	BSID-II Motor	BSID-II Mental	BSID-II Motor	BSID-II Mental	BSID-II Motor	BSID-II Mental
Gross Motor	.95		.72		.61		.92	
Fine Motor	.90		.72		.66		.94	
Cognitive		.93		.88		.83		.97
Language		.87		.88		.75		.96
Self-Help		.83		.83		.58		.91
Social Emotional		.85		.72		.47		.90

Note: For all correlations, p <.01

N: a (GM=73, FM=73, C=73, L=68, SH=47, SE=69)   b (GM=85, FM=86, C=88, L=87, SH=85, SE=75)

  c (GM=71, FM=74, C=72, L=72, SH=68, SE=57)   d (GM=229, FM=233, C=233, L=227, SH=200, SE=201)

Educational and Scientific Importance

Overall, this research found the Early LAP to be reliable and valid in assessing the development of young children. The Early LAP was found to have relatively high correlations between developmental age domain scores and chronological age, especially for children in the birth to two-year-old range, while older children aged out on some items and/or domains. The Early LAP also evidenced good internal consistency for each domain. Good test-retest reliability and interrater reliability were found for all domains of the Early LAP. Evidence of adequate construct validity was also shown. The Early LAP was found to have very good criterion validity, based on comparisons with the BSID-II. These analyses suggest that the Early LAP is an appropriate tool for use in assessing young children's developmental functioning. However,

we suggest that both the Self-Help and Social/Emotional domains be re-examined to see if revisions would improve their effectiveness in assessing two- to three-year-olds.

Early childhood educators need valid and reliable tools that will assist them in providing high quality services to infants and toddlers. This research provides information on the Early LAP that can assist practitioners and clinicians in making informed decisions about the types of tools that will help them in their efforts to provide effective, individualized services for the children in their care. Further research could explore the effectiveness of the Early LAP as a tool for enhancing the quality of services for infants and toddlers. For example, teachers could be trained on the Early LAP and intervention strategies using Early LAP results to see if such individualized developmental instruction improves the quality of program practices and results in better outcomes for children.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *1999 Standards for Educational and Psychological.* Washington, DC: American Educational Research Association.

Bayley, N. (1993). *Bayley Scales of Infant Development.* San Antonio, TX: The Psychological Corporation.

Children's Defense Fund. (2001). *Child Care Basics.* Available:http://www.childrensdefense.org /cc_facts.htm [April 9, 2003].

Cost, Quality & Child Outcomes Study Team (1995). *Cost, quality and child outcomes in child care centers.* Denver, CO: Economics Department, University of Colorado at Denver.

Fleming, J. A. (2000). An examination of interrater reliability of the early learning accomplishment profile. Doctoral dissertation. Baltimore, MD: Johns Hopkins University.

Gall, Meredith D., Borg, Walter R. Gall, Joyce P. (1996). *Educational Research.* White Plains, NY: Longman Publishers USA.

Glascoe, F. P. (1998). *Technical Report for the Brigance Screens.* North Billerica, MA: Curriculum Associates, Inc.

Glover, M. E., J. L. Preminger, et al. (1995). *The Early Learning Accomplishment Profile for Developmentally Young Children.* Chapel Hill, NC: Chapel Hill Training Outreach Project, Inc.

Kaplan Press. (n.d.). *A Workshop for the E-LAP.* Lewisville, NC: Kaplan Press.

Nehring, A. D. and E. F. Nehring (1992). *LAP-D: Learning Accomplishment Profile-Diagnostic Standardized Assessment.* Lewisville, NC: Kaplan Press.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M.L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool quality to children's cognitive and social developmental trajectories through second grade. *Child Development, 72(5),* 1534-1553.

U.S. Census Bureau. (1995). *Population Profile of the United States: 1995.* Washington, DC: U.S. Government Printing Office.