

# An Evaluation Design for a Study of **Florida's System for Measuring School Readiness**

**August 2001**

FJ King, Ph.D.  
Faranak Rohani, Ph.D.



# Table of Contents

<b>PURPOSE .....</b>	<b>3</b>
<b>BACKGROUND.....</b>	<b>3</b>
<b>MEASUREMENT SYSTEM.....</b>	<b>6</b>
<b>EVALUATION DESIGN .....</b>	<b>7</b>
<b>INSTRUMENTS.....</b>	<b>8</b>
OUTCOME INSTRUMENTS .....	8
UNIFORM SCREENING .....	11
<b>PROCEDURES.....</b>	<b>14</b>
SAMPLING PLAN .....	14
DATA MANAGEMENT .....	15
DATA COLLECTION.....	15
<i>Prekindergarten Data .....</i>	<i>16</i>
<i>Kindergarten through Third-Grade Data .....</i>	<i>18</i>
EVALUATION OBJECTIVES AND DATA ANALYSES .....	19
<i>Objectives.....</i>	<i>20</i>
<i>Reports to Be Prepared.....</i>	<i>23</i>
<b>REFERENCES .....</b>	<b>24</b>

## Tables

Table 1.....	8
Table 2.....	15
Table 3.....	17

## Figure

Figure 1.....	20
---------------	----

## **Purpose**

The purpose of this study is to develop an evaluation design for measuring school readiness in Florida as mandated by Title XXX, Section 411.01(4)(o), Florida Statutes (F.S.), Florida Partnership for School Readiness (called the School Readiness Act). The evaluation design is part of a measurement system mandated by this legislation. The design provides a description of the context for the evaluation activities, instruments, sampling plan, evaluation objectives and data collection and analysis, timelines, and types of reports to be prepared. The design covers all four years of a four-year longitudinal study in which program impact will be determined for a sample of children and three additional cohorts of students from the same school sites.

## **Background**

The School Readiness Act was enacted by the Florida Legislature in 1999 for two purposes: (1) to establish a statewide school readiness program and (2) to establish the Florida Partnership for School Readiness (Partnership), a board charged with adopting and maintaining certain standards and policies for all school readiness programs.

Among other tasks mandated by the 1999 legislation, the Partnership was charged with preparing and submitting to the State Board of Education “a system for measuring school readiness.” In response to this mandate, the Partnership created the Workgroup on School Readiness Assessment (Workgroup) and an Advisory Group to guide its efforts. The Workgroup was asked to recommend (1) a set of standards for school readiness, (2) an approach to assessment, and (3) specific assessment tools, parameters, and timelines. These were not easy recommendations to formulate because each of the three

involved consideration of many complex questions. It was obvious that tentative agreements on working definitions of certain terms and conditions had to be reached before answers could be found for some of them.

As a first step in formulating the recommendations, the Workgroup considered what had been done in the area of readiness by national, state, and district/municipal groups as well as by researchers and practicing professionals in the field. Published and unpublished studies and current practices were reviewed; advice was sought from recognized experts.

The Workgroup then addressed the specific requirements of relevant Florida statutes and related them to the specifics of the situation in the state. Of special concern were problems related to the great diversity among Florida's children.

One result of this work was a set of “Proposed Principles for Florida's Readiness Assessment System.” The system comprises four major categories:

- I. Principles Regarding Readiness—defines readiness and enumerates the domains of child development that influence it.
- II. Principles Regarding the Use of Data—sets forth specific ways in which assessment data may and may not be used.
- III. Principles Regarding the Assessment System—specifies the components of an assessment system and strategies for its implementation. It also calls for funding and time allocations sufficient for “training and the development of a management system that is coordinated with those existing in the state” (Florida Partnership for School Readiness, 2001, p. 11).

IV. Principles Regarding Assessment Instruments and Process—covers such things as the individuals who must be included in the assessment process, the necessary appropriateness of instruments for the children whose readiness they attempt to measure, the need to include assessment data from more than one source, and the need to make screening an easy procedure yielding easy-to-understand results.

These principles act to define and, to some extent, configure the proposed Statewide Readiness Uniform Screening System. They do not, however, provide an answer to the question that is fundamental to it—that is, the question of just how to determine whether an individual child is really ready to learn in the school program(s) available to him or her. This project is designed to further inform/respond to this question.

In 2001, the Florida Legislature passed several amendments in regard to the Partnership. As a result of this legislation, on July 1, 2001, the Partnership was transferred from the Executive Office of the Governor to the Agency for Workplace Innovation. The legislation further clarified that implementation of the uniform screening process for kindergarten would be the responsibility of the Department of Education (DOE). The amendment stated that the uniform screening process must include a pilot program during the 2001–02 school year to validate the system recommended by the Partnership as part of a comprehensive evaluation design.

According to the 2001 mandate, the evaluation design is to be fully in place by the 2002–03 school year and DOE shall require that all school districts administer the Statewide Readiness Uniform Screening System to each student in the district school system upon his or her entry into kindergarten. Children who enter public school for the first time in first grade must undergo a uniform screening upon entry. The collected data will then be incorporated into the K–20 data warehouse for longitudinal tracking to allow

policymakers and local school readiness coalitions to assess the progress of readiness programs toward goals and provide input for the continual improvement of local school readiness services and programs.

## **Measurement System**

The measurement system, prepared by the Workgroup and adopted by the Partnership, contains three components:

1. uniform screening system to provide objective information on each child's performance as he or she enters kindergarten—Section 411.01(4)(o), F.S.;
2. instructional improvement component designed to “enhance the cognitive, social, and physical development of children to achieve the performance standards and outcome measures specified by the Partnership”—Section 411.01(5)(c)2, F.S.—by providing teachers with the information they need to plan appropriate programs; and
3. an “evaluation of the effectiveness of the school readiness program”—Section 411.01(5)(g), F.S.—that follows children beginning in their preschool years through the elementary grades.

This document proposes a design to meet the third component of this list.

## **Evaluation Design**

The 2001 Florida Legislature (Section 411.01) specified two major purposes for the evaluation of the school readiness programs.

Adopt a system for measuring school readiness that provides objective data regarding the expectations for school readiness. . . . [It] must include the tracking of school readiness system information back to individual school readiness programs to assist in determining program effectiveness. (4)(j)8a

Adopt a system for evaluating the performance of students through the third grade to compare the performance of those who participated in school readiness programs with the performance of students who did not participate in school readiness programs in order to identify strategies for continued successful student performance. (4)(j)8b

To achieve these goals, the evaluation design must take into account information from individual school readiness programs, necessitating the collection of certain data elements by the Partnership and local school readiness coalitions. The characteristics of these data elements will be discussed under the data collection section.

This proposed evaluation study will extend over a four-year period beginning in 2002 and involve four cohorts of children who enter kindergarten and progress through grades one, two, and three. Each cohort will contain children who have attended a school readiness program and children who have not been through such programs. At the beginning of kindergarten, they will undergo the uniform screening process, and it is recommended that, at the end of kindergarten, students be assessed on relevant outcome

measures. Likewise, it is recommended that the students be assessed at the end of grades one through three in order to measure the academic and social/emotional growth of the children over that time period. Table 1 shows the years, grades, and student cohorts that will be involved.

**Table 1**  
**Years by Cohorts and by Grades for Program Evaluation**

<b>Cohorts</b>	<b>Year</b>			
	<b>2002–03</b>	<b>2003–04</b>	<b>2004–05</b>	<b>2005–06</b>
1	K	Grade 1	Grade 2	Grade 3
2		K	Grade 1	Grade 2
3			K	Grade 1
4				K

## Instruments

### Outcome Instruments

The measurement of academic growth from kindergarten through grade three is an integral part of the evaluation. Therefore, it is highly desirable that the instruments selected for use at the end of grades K through three offer vertically equated test forms that yield scores on a common scale over the time period involved. This means that developmentally linked appropriate instruments would be administered at the end of each school year from kindergarten through third grade. These instruments shall be referred to as outcome measures in the sections that follow. It would be desirable if all participating districts use the same outcome instruments.



Three options for instruments that meet the requirements for program evaluation are offered below.

Option 1. The first option would be to select an assessment instrument from existing nationally developed instruments in order to allow comparison with larger national studies. Indeed the Workgroup conducted an extensive review of existing instruments and initially suggested “using a comprehensive battery of accepted measures, such as those used in the national Family and Child Experience Survey (FACES) evaluation of Head Start . . . .” The FACES battery includes the following assessment instruments:

- *Peabody Picture Vocabulary Test (PPVT)*
- Letter-Word Identification, Applied Problems, and Dictation Tasks from the *Woodcock-Johnson Psycho-Educational Battery (Woodcock-Johnson)*
- *Story and Print Concepts*
- *McCarthy Scales of Children’s Abilities* subtest
- Phonemic Analysis subtest of the *Test of Language Development*
- Color Naming and Counting task (developed especially for FACES)
- Social Awareness items from the Comprehensive Assessment Program (CAP)
- *Early Childhood Diagnostic Instruments (Florida Partnership for School Readiness, April 30, 2001, p. 4)*

According to the Workgroup, this battery of tests requires about one hour of administration time for each child and would be administered at the end of kindergarten by persons specially trained in their administration.

It might be desirable to eliminate some of the aforementioned tests to reduce administration time. However, consideration should be given to the relationship between

length of administration time versus the depth of information gathered. In general, the shorter the administration time, the less information gathered.

Both the PPVT and Woodcock-Johnson provide national norms for ages 2 or 2 1/2 to over 90. Whether they provide common score scales for the grades relevant to this study is uncertain. In addition, their content may not be highly related to the school curriculum. A strength of these instruments is the inclusion of the measurement of social awareness.

Option 2. The second option is to obtain a measurement of growth by using standardized achievement tests such as the *Iowa Test of Basic Skills* (ITBS), *California Achievement Test* (CAT), and *California Test of Basic Skills* (CTBS). These are group-administered tests whose contents appear to be closely related to the school curriculum. The kindergarten and grade one battery of the ITBS, for example, includes tests of (1) vocabulary, (2) listening, (3) language, (4) mathematics, and (5) word analysis (optional). These tests are untimed and administered orally. The tests for grades two and three include (1) vocabulary, (2) reading, (3) listening, (4) language, (5) mathematics concepts, (6) mathematics problems, and (7) word analysis (optional).

The ITBS also provides vertically equated test forms and a common score scale for the relevant grades. Note, however, that it does NOT contain measures of social/emotional maturity. To accommodate the measures of social/emotional maturity, the social awareness items from the *CAP Early Childhood Diagnostic Instrument* could be used.

Option 3. The third option is development of specific instruments for the specified outcomes. A criterion-referenced instrument based on the Florida School Readiness

Performance Standards developed by DOE in conjunction with the Partnership could be constructed. It could have high content validity for measuring the growth of children as they pass through the various grade levels in terms of the increasing percentage of objectives they attained. It could contain both cognitive and social/emotional content.

It should be noted, however, that development of such instrument(s) might not be possible in a relatively short period of time and would require considerable financial investment. Also, such instruments would not have national norms.

## **Uniform Screening**

The Workgroup recommended that the uniform screening instruments be based on six domains of children's development: (1) physical health, (2) approaches toward learning, (3) communication and language development, (4) social/emotional development, (5) motor development, and (6) cognitive development and general knowledge.

As a minimum for each child, the instruments must provide objective data on the following legislatively mandated expectations for school readiness—Section 411.01 (4)(o), F.S.:

1. immunizations and other health requirements as necessary, including appropriate vision and hearing screening and examinations conducted by qualified medical personnel
2. physical development
3. compliance with rules, limitations, and routines
4. ability to perform tasks
5. interaction with adults

6. interaction with peers
7. ability to cope with challenges
8. self-help skills
9. ability to express his or her needs
10. verbal communication skills
11. problem-solving skills
12. ability to follow verbal directions
13. demonstration of curiosity, persistence, and exploratory behavior
14. interest in books and other printed materials
15. attention to stories
16. participation in art and music activities
17. ability to identify colors, geometric shapes, letters of the alphabet, numbers, and spatial and temporal relationships

In the first and subsequent years of the study, students entering kindergarten will be administered the uniform screening instruments by their teachers within the first 45 days of school. The instruments will be selected by DOE through a competitive bidding process during Fall 2001. During the selection process, cooperation between DOE and the Partnership should result in the selection of instrument(s) that can be used at both the prekindergarten and kindergarten levels. Using the same uniform instrument(s) enables the evaluators to begin estimates of academic and social/emotional growth. In addition, it is recommended that a parent questionnaire be developed during the pilot year to provide additional information of relevance to the measurement system. It should include demographic information, items concerning the risk factors found by the National Center for Education Statistics (NCES) to be associated with inadequate knowledge and skills in

children entering kindergarten, and items concerning the child's background and developmental status. These risk factors are living in a single-parent household; having a mother with less than high school education; living in a family that receives food stamps or cash welfare payments; and having parents whose primary language is something other than English (U.S. Department of Education, 2001, p. 21).

It is recommended that, in the selection of the uniform screening instrument(s), DOE consider the following minimum criteria for publisher's materials and services:

- Age-appropriate and technically sound instrument(s) that can measure the legislatively mandated skills and abilities
- Ease of administration by classroom teachers or other district-level personnel
- Brief administration time (not to exceed 20 minutes)
- Training for teachers or other district-level personnel to administer the instrument
- Use of scannable answer sheet for ease of scoring and reporting
- Results that yield profile information to assist teachers in individualizing kindergarten instruction and inform parents of their child's strengths and weaknesses
- Capability to provide electronic data to DOE and the evaluators
- Capability to provide summary reports of results to schools and teachers to assist them to individualize the kindergarten instruction and to inform parents of child's strengths and weaknesses

Uniform screening should also include results of a physical examination and records of any immunizations that are required by state law.

## **Procedures**

It is assumed that DOE will hire a contractor who will be responsible for gathering, analyzing, and synthesizing the data and reporting the results of the evaluation study. The sampling plan and data analysis strategies are described below.

### **Sampling Plan**

Every child who enters kindergarten will be assessed with the uniform screening instruments; therefore, for the purpose of program evaluation, it will be sufficient to obtain data on a sample of students at the end of kindergarten and grades one, two, and three.

Data will be gathered through the use of stratified cluster sampling composed of clusters of kindergarten children from all schools within sampled districts. Districts will be cross-classified by five geographic areas and by their location in high or low population density districts. Two districts from each classification will be randomly selected and asked to participate in the study. If all schools in the 20 districts that contain kindergarten classes are used, it is anticipated that approximately 45,000 children will be involved. In 2002–03, it is estimated that about 200,000 children will attend kindergarten and about 30 percent of them will have attended a school readiness program (G. Wilson, Florida Partnership, personal communication, August 7, 2001). Thus, the sample will contain about 13,500 program children. If desired, a smaller sample of children can be selected using proportional sampling to select schools. Within the 20 districts, schools

with at least one class of 20 kindergarten children will be listed. Within each district, random samples of these schools will be selected.

## **Data Management**

The data management in this evaluation study is complex and requires cooperation of a number of different sources. Therefore, it is recommended that data management be conducted by the evaluation contractor. The contractor should obtain data from the Partnership coalitions, DOE, and school districts as appropriate.

## **Data Collection**

This evaluation design is based on data collection at two levels, prekindergarten and kindergarten through third grade. Table 2 provides the list of instruments and the responsibilities of the different agencies involved.

**Table 2**  
**Data Collection and Agency Responsible**

<b>Data to Be Collected</b>	<b>Agency Responsible</b>
<b>Pretest</b> School Readiness Program	<b>Partnership</b>
<b>Posttest</b> School Readiness Program	
<b>Uniform Screening</b> Kindergarten Readiness	<b>DOE/Evaluation Contractor</b>
<b>Outcome Instruments</b> Kindergarten	
<b>Outcome Instruments</b> Grades 1–3	

The Partnership, through its coalitions, will be responsible for the collection of data and its transfer to the evaluation contractor for the prekindergarten level. DOE will be responsible for transfer of certain data elements for kindergarten through third grade to the evaluation contractor. The following list includes data elements required from the Partnership and DOE, as well as data the evaluation contractor should obtain from school districts.

### **Prekindergarten Data**

#### Program Data to Be Collected for Each Child

- Type of school readiness programs attended and length of attendance time in each program (e.g., Head Start, Pre-K Early Intervention [EI], etc.)
- Types of administrative organization (school-based, licensed, registered, religious-exempt) of school readiness program attended
- Curriculum used in school readiness program attended (e.g., High Scope, Programming Early Education for Children with Handicaps [PEECH], etc.)
- Percentage of instructional staff with early childhood certification and percentage who have additional qualifications

#### Student Data

- Pre- and postmeasures administered by coalitions prior to the child's entry into kindergarten
- Health records
- Personal and family data, including the age of the child at time of entry into school readiness program; gender; ethnicity; primary language of parents;



mother's age, education, and marital status; and socioeconomic data (e.g., welfare, food stamps)

- Number of days child attended school readiness program
- Number of years child has been in a school readiness program
- Parent questionnaire or interview that contains information on the child's background and development status

As an example of the variety of pre-/postmeasures used, Table 3 shows the instruments that are being used by several programs in the Clay/Nassau Coalition, together with the organizations providing the assessment.

**Table 3**  
**Clay/Nassau Coalition Programs and Data Instruments Used**

<b>Program</b>	<b>Organization</b>	<b>Instrument</b>
Head Start	Episcopal Children's Services	LAP-D Screening and Assessment, DIAL-R
Pre-K EI	Nassau County School District	Brigance
Pre-K EI Disabilities	Clay County School District	Dial-R, Vineland, several others
Subsidized Child Care	Episcopal Children's Services	Brigance
CCRR*	Episcopal Children's Services	Ages and Stages

\*Child Care Resource and Referral

Data transmitted to the evaluation contractor should include the name of the program in which the child was enrolled, the organization providing the test names, and test scores.

## **Kindergarten through Third-Grade Data**

### Program Data to Be Collected for Each Child

It should be noted that existing data elements available through the DOE database have a data element number next to each of them. The evaluation contractor will be responsible for obtaining the remaining data elements in this list from school districts or other appropriate entities.

- Participation in supplemental programs such as Head Start during Pre-K (#161825)
- Participation in special education programs during Pre-K (#160637)

### Student Data

- Data from uniform screening tests administered at the beginning of kindergarten (#143025)
- Grade promotion status from kindergarten and grades one through three (#126425)
- Scores from outcome instruments administered at the end of kindergarten and grades one, two, and three
- Third-grade Florida Comprehensive Achievement Test scores (this data can be obtained from the DOE)
- Personal and family data, including the child's birth date (#104025), country of birth (#108800), gender (#173625), ethnicity (#168025), and free or reduced-cost lunch status (#146025); the primary language of the parents (#160450); limited English proficiency (#144025); the mother's age, education, and marital

status; the number of days the child attended school (#113225); and the number of days the child was absent from school (#112025)

- Child's primary exceptionality (#118575)
- Participation in Even Start family literacy program (#116925)
- Participation from birth to three years of age in First Start program (#122363)
- Parent questionnaire or interview that contains information on the child's background and developmental status
- Health data

## **Evaluation Objectives and Data Analyses**

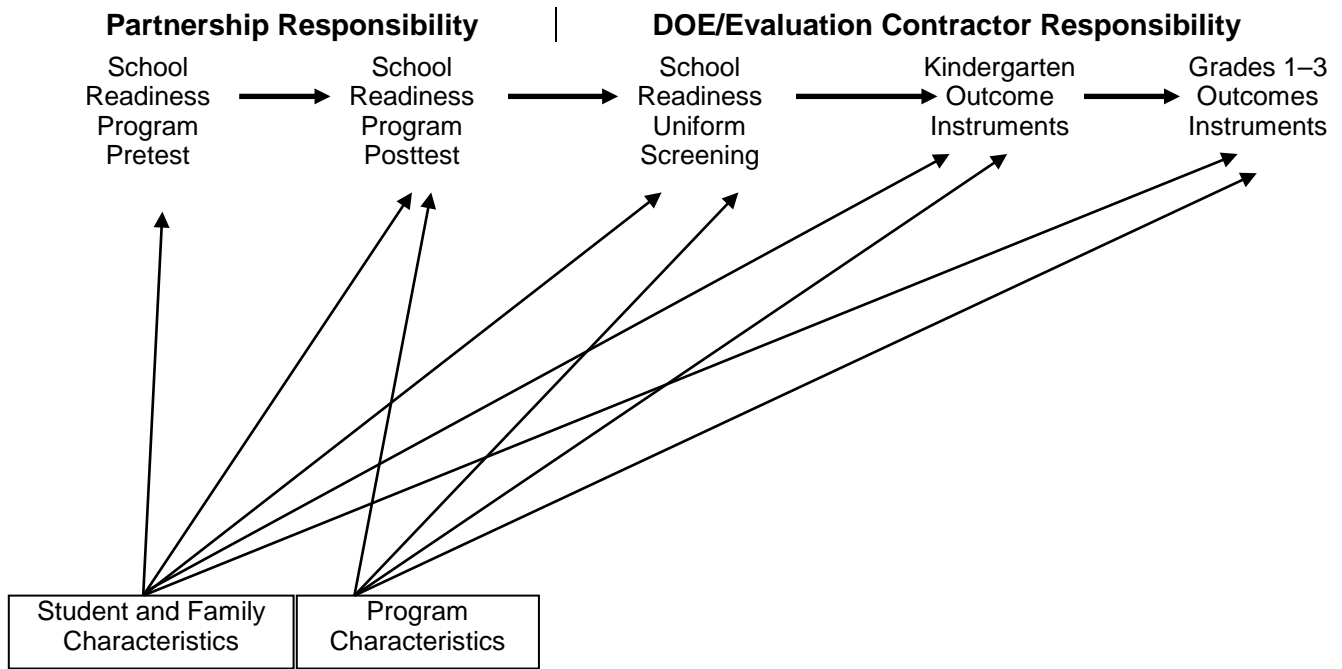
Nine sets of variables will be involved in the evaluation of the effectiveness of the school readiness programs.

- Student and family background characteristics
- School readiness program pretest scores
- School readiness program characteristics
- School readiness program posttest scores
- Kindergarten uniform screening (readiness) scores
- Kindergarten outcome measures
- First-grade outcome measures
- Second-grade outcome measures
- Third-grade outcome measures

Figure 1 is a model of the evaluation design, which illustrates the variables involved and the responsible agencies.

**Figure 1**

**Evaluation Design for Measuring School Readiness**



The school readiness programs employ a variety of instruments as pre- and posttests. Evaluation objectives that involve these instruments will require separate statistical analyses for each instrument. The results of these separate analyses can be combined through the use of statistical procedures to aid in the interpretation of the overall results for the objective under consideration.

**Objectives**

Objective 1

The first evaluation objective will be to determine the effects of student and family background characteristics on school readiness pretest scores. Total and subtest scores will be regressed separately on the set of student and family characteristics. A subset of these characteristics will be identified for use as control variables in subsequent analyses.

## Objective 2

The second evaluation objective is to determine the extent to which school readiness posttest scores are dependent on program characteristics when the effects of school readiness pretest scores and student and family background characteristics are controlled. Analysis of variance (General Linear Model) will be used to determine whether there are statistically significant differences among the school readiness posttest means associated with each set of program characteristics. While this analysis will be helpful in determining which kinds of programs are most effective, it will not allow conclusions to be drawn about the benefits of program versus nonprogram membership.

## Objective 3

The third evaluation objective is to determine whether beginning kindergarten students who participated in school readiness programs have higher mean scores on the uniform screening instruments than students who did not participate in school readiness programs when student and family characteristics are controlled. Analysis of variance (General Linear Model) will be used to determine whether there are statistically significant differences between the school readiness means of program and nonprogram students.

If educationally significant differences are found among school readiness programs in investigation of Objective 2, then individual program characteristics should be included in the analysis of this objective as well as the remaining objectives.

## Objectives 4 through 7

The fourth through seventh evaluation objectives are to determine whether the performance of kindergarten through third-grade students who participated in school

readiness programs show higher levels of performance on appropriate outcome measures (at the end of kindergarten and grades one, two, and three) than students who did not participate in school readiness programs. Analysis of variance (General Linear Model) will be used to determine whether there are statistically significant mean differences between the program and nonprogram students when student and family background characteristics variables are controlled.

#### Objectives 8 through 11

The eighth through the eleventh evaluation objectives are to determine whether kindergarten through third-grade students who participated in school readiness programs show lower proportions of retentions in grade than students who did not participate in school readiness programs. Logistic regression analysis will be used to determine whether there are statistically significant differences in proportions of retentions between program and nonprogram students when student and family background variables are controlled.

#### Objective 12

The twelfth objective is to determine whether performance on appropriate outcome measures increases over time as readiness teachers and other school personnel gain more experience with the program. For example, program and nonprogram kindergarten students in 2002–03, 2003–04, 2004–05, and 2005–06 will be compared in order to determine whether differences between the two groups increase over the four years or remain about the same. The same comparisons will be made for first-grade students during years two through four of the study and for second-grade students during the last two years. Analysis of variance (General Linear Model) will be used to determine

whether there are statistically significant mean differences between the program and nonprogram students when student and family characteristic variables are controlled.

### Objective 13

The thirteenth evaluation objective is to determine whether the academic and social/emotional growth of students over the four-year period is the same for program and nonprogram students. For example, the same program and nonprogram students who entered the program in 2002–03 will be measured at the end of each of the four years to determine whether differences on growth indices increase over time. Repeated measures analysis of variance or the Hierarchical Linear Model (HLM) will be used to determine whether there are statistically significant differences on growth indicators between the program and nonprogram students when student and family background variables are controlled.

### **Reports to Be Prepared**

The evaluation contractor is responsible for two types of reports: management status reports and program reports. The project management reports should be submitted to DOE at six-month intervals. These reports will contain progress status, list of obstacles in accomplishing objectives, and request for assistance needed from DOE.

At yearly intervals, the contractor should provide program reports concerning relevant evaluation objectives to both DOE and the Partnership. A final report covering the entire project should be submitted to both groups by December 2006. These reports should address the 13 evaluation objectives individually (listed in previous section) at appropriate times. In addition, the reports should compare and integrate study findings with results from research literature.

## References

Florida Partnership for School Readiness. (2001). *School readiness in Florida: Strategies for defining, measuring, and advancing children's success*. Retrieved August 2, 2001, from [http://www.myflorida.com/myflorida/government/learn/schoolreadiness/pdf/florida\\_assessment\\_final\\_report.pdf](http://www.myflorida.com/myflorida/government/learn/schoolreadiness/pdf/florida_assessment_final_report.pdf)

Florida Partnership for School Readiness. (2001). *Recommendations to the state board of education: A system for measuring school readiness*. Tallahassee: Author.

Florida Partnership for School Readiness Act, Fl. Stat. §§ 411-01 (1999 & Supp. 2001). Retrieved August 2, 2001, from <http://www.leg.state.fl.us/statutes/index.cfm>

U.S. Department of Education, National Center for Education Statistics. (2001). *Entering kindergarten: A portrait of American children when they begin school: Findings from the Condition of Education 2000*, Nicholas Zill and Jerry West (NCES Publication No. 2001-035). Washington, DC: U.S. Government Printing Office.