#600

Joint Legislative Committee on Performance Evaluation and Expenditure Review (PEER)

Report to
the Mississippi Legislature

The Early Learning Collaborative Act of 2013: Evaluation of the Operations and Effectiveness of the Program

November 17, 2015

PEER: The Mississippi Legislature's Oversight Agency

The Mississippi Legislature created the Joint Legislative Committee on Performance Evaluation and Expenditure Review (PEER Committee) by statute in 1973. A joint committee, the PEER Committee is composed of seven members of the House of Representatives appointed by the Speaker and seven members of the Senate appointed by the Lieutenant Governor. Appointments are made for four-year terms, with one Senator and one Representative appointed from each of the U. S. Congressional Districts and three at-large members appointed from each house. Committee officers are elected by the membership, with officers alternating annually between the two houses. All Committee actions by statute require a majority vote of four Representatives and four Senators voting in the affirmative.

Mississippi's constitution gives the Legislature broad power to conduct examinations and investigations. PEER is authorized by law to review any public entity, including contractors supported in whole or in part by public funds, and to address any issues that may require legislative action. PEER has statutory access to all state and local records and has subpoena power to compel testimony or the production of documents.

PEER provides a variety of services to the Legislature, including program evaluations, economy and efficiency reviews, financial audits, limited scope evaluations, fiscal notes, special investigations, briefings to individual legislators, testimony, and other governmental research and assistance. The Committee identifies inefficiency or ineffectiveness or a failure to accomplish legislative objectives, and makes recommendations for redefinition, redirection, redistribution and/or restructuring of Mississippi government. As directed by and subject to the prior approval of the PEER Committee, the Committee's professional staff executes audit and evaluation projects obtaining information and developing options for consideration by the Committee. The PEER Committee releases reports to the Legislature, Governor, Lieutenant Governor, and the agency examined.

The Committee assigns top priority to written requests from individual legislators and legislative committees. The Committee also considers PEER staff proposals and written requests from state officials and others.

PEER Committee
Post Office Box 1204
Jackson, MS 39215-1204

(Tel.) 601-359-1226
(Fax) 601-359-1420
(Website) http://www.peer.state.ms.us

The Mississippi Legislature

Joint Committee on Performance Evaluation and Expenditure Review

PEER Committee

SENATORS
THOMAS GOLLOTT
Vice Chair
SAMPSON JACKSON II
Secretary
KELVIN E. BUTLER
VIDET CARMICHAEL
NANCY ADAMS COLLINS
GARY JACKSON
PERRY LEE

REPRESENTATIVES
BECKY CURRIE
Chair
RICHARD BENNETT
KIMBERLY L. CAMPBELL
STEVE HORNE
MARGARET ELLIS ROGERS
RAY ROGERS
PERCY W. WATSON

TELEPHONE:
(601) 359-1226

FAX:
(601) 359-1420

OFFICES:
Woolfolk Building, Suite 301-A
501 North West Street
Jackson, Mississippi 39201

Post Office Box 1204
Jackson, Mississippi 39215-1204

James A. Barber
Executive Director

www.peer.state.ms.us

November 17, 2015

Honorable Phil Bryant, Governor
Honorable Tate Reeves, Lieutenant Governor
Honorable Philip Gunn, Speaker of the House
Members of the Mississippi State Legislature

On November 17, 2015, the PEER Committee authorized release of the report entitled **The Early Learning Collaborative Act of 2013: Evaluation of the Operations and Effectiveness of the Program.**

Representative Becky Currie, Chair

This report does not recommend increased funding or additional staff.

Table of Contents

Table of Contents (continued)

List of Exhibits

The Early Learning Collaborative Act of 2013: Evaluation of the Operations and Effectiveness of the Program

Executive Summary

Introduction

Meta-analytic research on public preschool programs concludes that quality programs can have a positive impact on important outcomes that benefit society. Examples of these long-term outcomes include improved test scores and high school graduation rates, as well as decreases in K-12 grade repetition, special education, and crime.[*]

In 2013, the Mississippi Legislature passed the Early Learning Collaborative Act to help ensure that all of Mississippi's children have access to quality early childhood education and development services through a voluntary prekindergarten program for four-year-olds. State law requires that the Mississippi Department of Education (MDE) provide annual reports on the program to the Legislature and that the PEER Committee review those reports and conduct an independent evaluation of the program's operations and effectiveness.

Background

The Legislature contemplated funding the prekindergarten program on a phased-in basis, appropriating $9 million for fiscal years 2014 through 2016 for the first phase. Of this amount, MDE distributed approximately $8.4 million to eleven early learning collaboratives and retained $450,000 for administrative costs. The eleven collaboratives (including fifty-three providers) served approximately 1,580 students in the 2014-15 school year, the first full year of the program's operation.

---

[*]The Legislature of the State of Washington created the Washington State Institute for Public Policy (WSIPP) in 1983 to carry out practical, non-partisan research that answers relevant policy questions. WSIPP has developed a model to evaluate the costs and benefits of a wide range of public policies and programs, including Pre-K to 12 education, that affect many different outcomes. The institute follows a meta-analytic framework to assess systematically all relevant evaluations on a given topic (e. g., state and district early childhood education programs) meeting the institute's high-quality standards and computes an average effect on each measured outcome from all of the credible studies so identified. WSIPP's benefit-cost analyses, including estimated average effect sizes by outcome measured, are available on its website: http://www.wsipp.wa.gov/

MISS. CODE ANN. Section 37-21-51 (1972) gives MDE responsibility for administering the implementation, monitoring, and evaluation of the prekindergarten program, including the application process and awarding of funds. This section also states that effectiveness of the program is to be determined by the school readiness of participants. MDE determines "school readiness" by administering a Kindergarten Readiness Assessment near the end of the school year and setting a target score on that assessment.

Evaluation of Program Operations and Effectiveness

After the first full year of implementation of the Early Learning Collaborative Act, the average performance of students in non-collaborative publicly funded prekindergarten programs was better than the average performance of students in the collaboratives by a statistically significant amount. MDE should conduct the research and analysis necessary to identify those factors most associated with positive student outcomes and ensure that all students enrolled in the collaboratives receive the quality programming necessary for school readiness.

PEER found significant room for improvement in the implementation and effectiveness of the Early Learning Collaborative Act of 2013.

Regarding MDE's implementation of the act:

- Although required by state law to do so, MDE has not adopted a minimum rate of readiness as a benchmark for continued program funding. The department is researching the use of student growth as a future determinant for continued funding eligibility, possibly yielding a more lenient standard.
- MDE's annual report does not show the rates of school readiness for each collaborative and each provider, as is required by state law.
- MDE awarded funding to four collaboratives that utilized a prekindergarten curriculum found through rigorous research to have "no discernable effects" on student learning.
- MDE has not assessed students' progress on some content areas of the department's early learning standards.

Regarding program effectiveness, according to PEER's independent evaluation, after the first full year of implementation, prekindergarteners in the program's participating collaboratives achieved the end-of-the-year target score on the Kindergarten Readiness Assessment less often than children enrolled in other public prekindergartens, after adjusting for initial score differences. Also, there was an extremely wide range of end-of-school-year assessment pass rates by program provider, by collaborative, and by curriculum.

To improve the effectiveness of the program, MDE must conduct the research and analysis necessary to identify those

factors most associated with positive student outcomes and ensure that all students enrolled in the collaboratives receive the quality programming necessary for school readiness.

For More Information or Clarification, Contact:

PEER Committee
P.O. Box 1204
Jackson, MS 39215-1204
(601) 359-1226
http://www.peer.state.ms.us

Representative Becky Currie, Chair
Brookhaven, MS

Senator Thomas Gollott, Vice Chair
Biloxi, MS

Senator Sampson Jackson, Secretary
DeKalb, MS

The Early Learning Collaborative Act of 2013: Evaluation of the Operations and Effectiveness of the Program

Introduction

Authority

MISS. CODE ANN. Section 37-21-51 (1972) directs the Mississippi Department of Education (MDE) to implement a voluntary prekindergarten program in the state. According to CODE Section 37-21-51 (3) (g), MDE "shall make an annual report to the Legislature and the Governor regarding the effectiveness of the program." Further, this CODE Section directs the PEER Committee to:

> *…review those reports and other program data and submit an independent evaluation of program operation and effectiveness to the Legislature and the Governor on or before October 1 of the calendar year before the beginning of the next phased-in period of funding.*

The Committee acted in accordance with MISS. CODE ANN. Section 5-3-51 et seq. (1972).

Purpose and Scope

Because fiscal years 2014 through 2016 represented the first phase of funding for Mississippi's prekindergarten program under the Early Learning Collaborative Act of 2013, the Legislature will need information for the 2016 Legislative Session to determine whether to continue or expand the program for the next phase of funding beginning in FY 2017.

Therefore, the purpose of this report is to present an independent evaluation of the operation and effectiveness of Mississippi's prekindergarten program established under the Early Learning Collaborative Act that will be useful for decisionmakers when considering the program's next phase of funding.

This report will address:

- the creation and purpose of the prekindergarten program;
- the statutory requirements of the prekindergarten program;
- the funding appropriated to the prekindergarten program and how the funds have been disbursed;
- the statewide assessment for measuring the readiness of prekindergarten students for kindergarten;
- an evaluation of the program's operations and effectiveness; and,
- PEER's independent evaluation of program effectiveness.

Exhibit 1, page 3, contains the statutory definitions of terms used in this report in reference to the prekindergarten program.

Method

In conducting this review, PEER:

- reviewed relevant sections of state law;
- interviewed selected staff of MDE;
- reviewed MDE's self-reported information in an annual report[1] for the 2014-15 school year, as required by MISS. CODE ANN. Section 37-21-51 (3) (g) (1972), plus additional information requested by PEER; and,
- requested and analyzed the following test data from MDE:
  - pre-test and post-test scaled scores for each student who took the Kindergarten Readiness Assessment in School Year 2014-15 as part of the Early Learning Collaborative Act; and,
  - pre-test and post-test scaled scores for each student enrolled in a four-year-old prekindergarten program in Mississippi who took the Kindergarten Readiness Assessment but was not a part of the Early Learning Collaborative Act in School Year 2014-15.

[1]PEER did not independently verify all of the information provided in MDE's annual report for the 2014-15 school year.

Exhibit 1: Statutory Definitions of Terms Related to the Early Learning Collaborative Act of 2013

Term	Definition
Preschool or prekindergarten children	Children who have not entered kindergarten but will have obtained four years of age on or before September 1 of a school year.
Early Learning Collaborative	A district or countywide council that writes and submits an application to participate in the voluntary prekindergarten program. An early learning collaborative is comprised, at a minimum, of a public school district and/or a local Head Start affiliate if in existence, private or parochial schools, or one or more licensed child care centers. Agencies or other organizations that work with young children and their families may also participate in the collaborative to provide resources and coordination even if those agencies or organizations are not prekindergarten providers.
Prekindergarten Provider	A public, private, or parochial school, licensed child care center or Head Start center that serves prekindergarten children and participates in the voluntary prekindergarten program.
Lead Partner	A public school district or other nonprofit entity with the instructional expertise and operational capacity to manage the early learning collaborative's prekindergarten program as described in the collaborative's approved application for funds. The lead partner serves as the fiscal agent for the collaborative and shall disburse awarded funds in accordance with the collaborative's approved application. The lead partner must facilitate a professional learning community for the teachers in the prekindergarten program and lead the collaborative. The lead partner ensures that the collaborative adopts and implements curriculum and assessments that align with the comprehensive early learning standards. The public school district shall be the lead partner if no other qualifying lead partner is selected.
Comprehensive Early Learning Standards	Standards adopted by the State Board of Education that address the highest level of fundamental domains of early learning to include, but not be limited to, physical well-being and motor development, social/emotional development, approaches toward learning, language development and cognition and general knowledge. The comprehensive early learning standards shall also include standards for emergent literacy skills, including oral communication, knowledge of print and letters, phonological and phonemic awareness, and vocabulary and comprehension development.
Research-Based Curriculum	An age-appropriate curriculum that is based on the findings of current research and has been found to be effective in improving student learning.

SOURCE: MISS. CODE ANN. Section 37-21-51 (1) (1972).

Background

This chapter addresses:

- the creation and purpose of the prekindergarten program of the Early Learning Collaborative Act;
- the statutory requirements of the prekindergarten program;
- the funding appropriated to the prekindergarten program and how the funds have been disbursed; and,
- the statewide assessment for measuring prekindergarten students' readiness for kindergarten.

Creation and Purpose of the Prekindergarten Program

In its 2013 Regular Session, the Legislature passed the "Early Learning Collaborative Act of 2013," which directs the Mississippi Department of Education to implement a voluntary prekindergarten program in the state on a phased-in basis. The prekindergarten program's purpose is to help ensure that all children have access to quality early childhood education and development services.

MISS. CODE ANN. Section 37-21-51 (2) (1972) outlines the following findings of the Legislature with regard to ensuring that all children have access to quality early childhood education and development services:

> *(a) Parents have the primary duty to educate their young preschool children;*
>
> *(b) The State of Mississippi can assist and educate parents in their role as the primary caregivers and educators of young preschool children;*
>
> *(c) There is a need to explore innovative approaches and strategies for aiding parents and families in the education and development of young preschool children; and*
>
> *(d) There exists a patchwork of prekindergarten entities but no coordination of services and there needs to be a coordination of these services.*

Based upon these findings, the Legislature authorized and directed MDE to implement a prekindergarten program in the state on a phased-in basis. MISS. CODE ANN. Section 37-21-51 (3) (1972), also known as the Early Learning Collaborative Act of 2013, states:

> *Effective with the 2013-2014 school year, the Mississippi State Department of Education shall establish a voluntary prekindergarten program, which shall be a collaboration among the entities providing prekindergarten programs including Head Start, licensed child care facilities and licensed public, parochial and private school prekindergarten programs. This program shall be implemented no later than the 2014-2015 school year.*

In the 2013-14 school year, eleven early learning collaboratives began implementing a prekindergarten program. For the 2014-15 school year, the eleven collaboratives (encompassing fifty-three providers) served approximately 1,580 students.

Statutory Requirements of the Prekindergarten Program

MISS. CODE ANN. Section 37-21-51 (1972) outlines MDE's responsibilities in administering the prekindergarten program, eligibility criteria for the awarding of funds to collaboratives, and funding requirements of the program.

MISS. CODE ANN. Section 37-21-51 (3) (b) (1972) states that the prekindergarten program "shall be a collaboration among the entities providing prekindergarten programs including Head Start, licensed child care facilities and licensed public, parochial and private school prekindergarten programs."

MDE's Responsibilities

State law designates MDE as the entity responsible for administering the implementation, monitoring, and evaluation of the voluntary prekindergarten program.

MISS. CODE ANN. Section 37-21-51 (3) (1972) designates MDE as the entity responsible for administering the implementation, monitoring, and evaluation of the voluntary prekindergarten program, including awards and the application process. Specifically, MDE's statutory responsibilities include:

- establishing a rigorous and transparent application process for the awarding of funds;
- establishing monitoring policies and procedures that, at minimum, include at least one site visit per year;
- providing technical assistance to collaboratives and their providers to improve the quality of prekindergarten programs;

- evaluating the effectiveness of each early childhood collaborative and each prekindergarten provider;
- ensuring that collaboratives provide each parent enrolling a child in the program with a profile of every prekindergarten provider participating in the collaborative's geographic catchment area;
- adopting a minimum rate of readiness that each prekindergarten provider must meet in order to remain eligible for program funds (MDE must recognize each child's unique pattern of development when adopting a rate); and,
- making an annual report to the Legislature and the Governor regarding the effectiveness of the program.

Eligibility Criteria for the Program

According to state law, prekindergarten program funds must be awarded to early childhood collaboratives whose proposed programs meet certain program criteria.

Per MISS. CODE ANN. Section 37-21-51 (3) (b) (1972), eligible entities may submit applications for program funds for the following purposes:

- to defray the cost of additional and/or more qualified teaching staff, appropriate educational materials and equipment and to improve the quality of educational experiences offered to four-year-old children in early care and education programs; and/or,
- to extend developmentally appropriate education services at such programs currently serving four-year-old children to include practices of high quality instruction; and,
- to administer, implement, monitor and evaluate the programs; and,
- to defray the cost of professional development and age-appropriate child assessment.

According to MISS. CODE ANN. Section 37-21-51 (3) (d) (1972), prekindergarten program funds must be awarded to early childhood collaboratives whose proposed programs meet certain program criteria (e. g., meet certain teacher-child ratios). (See Appendix A, page 33, for a complete list of program criteria.)

For the initial phase of implementation, MISS. CODE ANN. Section 37-21-51 (3) (h) (vi) (1972) states that MDE must award state funds based on a community's capacity, commitment, and need. To determine capacity, commitment, and need, MDE must require evidence of strong local collaborations of early education stakeholders. In determining community need, MDE must

consider low academic achievement within the public school districts participating in an applicant early learning collaborative and the number and percentage of children without quality prekindergarten options.

Requirements for Funding

State law requires that funding be provided to collaboratives on a basis of $2,150 per student in a full-day program and $1,075 per student in a half-day program. MDE cannot reserve more than 5% of the appropriation for administrative costs.

According to MISS. CODE ANN. Section 37-21-51 (3) (h) (iii) (1972), funding must be provided to collaboratives on a basis of $2,150 per student in a full-day program and $1,075 per student in a half-day program. MDE cannot reserve more than 5% of the appropriation for administrative costs. Funds may be carried over to the next year if they are not used.

Collaboratives must match state funds on a 1:1 basis. Local matching funds can include local tax dollars; federal dollars as allowed; parent tuition; philanthropic contributions; or in-kind donations of facilities, equipment, and services required as part of the program (such as food service or health screenings). The Early Learning Collaborative Act of 2013 provided for a state income tax credit for contributions paid to approved providers or collaboratives not to exceed $1 million by any individual, corporation, or other entity.

Program Funding

Under the Early Learning Collaborative Act of 2013, Mississippi's prekindergarten program received legislative appropriations of $9 million for fiscal years 2014 through 2016. This funding assisted eleven early learning collaboratives in implementing prekindergarten programs that served approximately 1,580 students in the 2014-15 school year.

Initial Appropriations and Program Phases

With passage of the Early Learning Collaborative Act of 2013, the Legislature contemplated funding the prekindergarten program on a phased-in basis, appropriating $9 million for fiscal years 2014 through 2016. Of this amount, MDE distributed approximately $8.4 million to eleven early learning collaboratives and retained $450,000 for administrative costs.

According to CODE Section 37-21-51 (3) (h) (i) (1972), funding for the prekindergarten program would be appropriated on a phased-in basis, as follows:

- First phase: "shall be based on an annual state appropriation of not more than $8 million and shall serve approximately 3,500 children through 5-8 early learning collaboratives and their pre-K providers;"
- Second phase: "shall be based on an annual state appropriation of not more than $16 million and shall serve approximately 7,000 children through 10-15 collaboratives and their providers;"
- Third phase: "shall be based on an annual appropriation of not more than $33.95 million and shall serve approximately 15,000 children through 20-25 collaboratives and their providers."

According to CODE Section 37-21-51 (3) (h) (ii) (1972), future phases must be based on interest in the program and the effectiveness of the program as determined by the school readiness of participants. Each phase lasts for at least three years but no more than five years, as determined by MDE.

For fiscal years 2014 through 2016, the Legislature appropriated $3 million per year ($9 million total) for the prekindergarten program. MDE distributed approximately $8.4 million to eleven early learning collaboratives and retained $450,000 for administrative costs from FY 2014 through FY 2016. (See Exhibit 2, below.)

Exhibit 2: Prekindergarten Appropriation Amounts and Distribution of Funds, FY 2014-FY 2016

Year	Legislative Appropriation Amount	Funds Distributed to Collaboratives	MDE's Administrative Costs (5% of Appropriation)	
FY 2014	$3 million	$2,072,023	$150,000	
FY 2015	3 million	3,182,320	150,000	
FY 2016	3 million	3,183,481	150,000	
Total	$9 million	$8,437,824	$450,000	$112,176 not spent*

*As provided in MISS. CODE ANN. Section 37-21-51 (3) (h) (v) (1972), funds remaining after awards to early learning collaboratives and administrative costs may be carried over in the following year. In its 2015 report, MDE states that the $112,176 not spent will be used to fund a collaborative during the next round of grants for FY 2017, pending legislative approval.

SOURCE: MDE.

Selection Process and Collaboratives that Received Funding for FY 2014-FY 2016

In December 2013, the Board of Education approved funding for eleven collaboratives that were selected using a competitive process. These eleven collaboratives represent 111 classrooms that can serve 1,713 children when

at capacity. In the spring of 2015, 1,580 students enrolled in the collaboratives took the statewide assessment for measuring kindergarten readiness.

As noted previously, according to MISS. CODE ANN. Section 37-21-51 (3) (c) (1972), MDE is responsible for administering the implementation, monitoring, and evaluation of the prekindergarten program, including awards and the application process.

According to MDE, on October 20, 2013, MDE posted the Early Learning Collaborative Act Request for Proposals (RFP) on its website. MDE shared the RFP with school districts and early childhood educators throughout the state through its listserv email.

Thirty collaboratives applied for funds by submitting proposals to MDE. MDE created an evaluation team of six individuals, each of whom scored proposals based on a set of criteria outlined in MISS. CODE ANN. Section 37-21-51 (3) (d) (1972) (see Appendix A, page 33, for a complete list of criteria). This was MDE's Level 1 evaluation (i. e., evaluation of the proposal in response to the RFP).

In its Level 2 evaluation, MDE conducted and scored interviews with collaboratives that had the highest scores on their proposals. Based on the combined scores from the proposal and interviews from all evaluation team members, MDE made recommendations to the Board of Education for approval.

Level 1 evaluation criteria could result in a maximum of 170 points per team member (1,020 points total). Level 2 evaluation criteria could result in a maximum of 30 points per team member (180 points total). Therefore, the maximum total points any one collaborative could receive was 200 points per team member (1,200 points total).

Evaluation Criteria for RFPs

In its Request for Proposals, dated October 10, 2013, MDE included in its participation criteria the following: "All participating Pre-K programs must use the *Early Learning Standards for Classrooms Serving Four-Year-Old Children* adopted by the Mississippi Department of Education." (See page 23 for more information on these standards.)

Level 1 Evaluation

Level 1 evaluation criteria and maximum points that could be awarded by each team member were:

- Community partnership and lead partner capabilities – 30 points;
- Leveraging and sharing resources – 20 points;

- Commitment – 20 points;
- Evidence of need – 20 points;
- Parent involvement and ease of family access – 10 points;
- Qualified staff – 10 points;
- Curriculum and assessments aligned to standards – 25 points;
- Daily schedule – 5 points;
- Class size and staff-child ratio – 5 points;
- Professional development – 15 points; and,
- Budget – 10 points.

One of the above criteria, which requires that programs use curricula and assessments aligned to the early learning standards, is imperative in analyses of prekindergarten program effectiveness. First, MDE required that collaboratives use a research-based curriculum that is designed to prepare students to be ready for kindergarten. (For the next phase of funding beginning in FY 2017 [pending legislative approval], MDE indicated to PEER that it plans to require collaboratives to use evidence-based curricula, which is a higher standard than research-based curricula. See pages 20-21 for evidence-based curricula standards.) Whether the particular curriculum used has any effect on student learning is an important research question (see pages 20-22 and 27-28 for further discussion.) Secondly, MDE established the *Early Learning Standards for Classrooms Serving Four-Year-Old Children*, which indicates what four-year-olds should know and be able to do. Analyses of program effectiveness should include an analysis of whether students are ready for kindergarten based on the early learning standards. Therefore, any statewide assessment used to measure kindergarten readiness should be aligned to the early learning standards. (See page 23 for further discussion.)

Level 2 Evaluation

MDE conducted interviews with the highest-scoring applicants. Level 2 (interview) evaluation criteria and maximum points that could be awarded by each team member were:

- Understanding of how the collaborative will leverage resources – 10 points;
- Understanding the role of each participating provider – 10 points; and,
- Understanding of how the prekindergarten providers will use curriculum, instruction, and assessments aligned to standards – 10 points.

Awarding of Funds to Collaboratives

Based on the scores from the application and interviews, MDE made recommendations to the Board of Education for approval. On December 19, 2013, the State Board of Education approved a total of $8,437,824 in funding over a three-year period to these eleven early learning collaboratives. According to MDE, these collaboratives include 111 classrooms and can serve 1,713 children when at capacity. (See Exhibit 3, page 12.) In the spring of 2015, 1,580 students enrolled in the collaboratives took the statewide assessment for measuring kindergarten readiness.

After the approval of funds in December 2013, for the remainder of School Year 2013-14, these collaboratives began implementing their programs, which included activities such as enrolling students, conducting screenings of students, hiring teachers, and purchasing classroom materials. Collaboratives completed their first full school year in 2014-2015.

MDE's annual report for 2015 highlighted several actions taken by MDE and the collaboratives that are worth noting. For example, in the area of professional development, MDE notes that collaborative staff were trained on the administration of the Classroom Assessment Scoring System (CLASS), which is an assessment tool designed to document classroom quality. Also, the report mentions monitoring visits, which are conducted by MDE to help gauge the implementation status of the program in each collaborative.

Exhibit 3: FY 2014-16 Funding, Number of Classrooms and Capacity, by Collaborative

Collaborative	FY 2014	FY 2015	FY 2016	Total Funding	Number of Classrooms	Number Students (capacity)	Number Tested (Spring 2015)
Clarke County Early Learning Partnership	$172,000	$172,000	$172,000	$516,000	6	111	111
Coahoma County Pre-K Collaborative Initiative	327,554	294,846	331,459	953,859	18	330	330
Corinth-Alcorn-Prentiss Early Learning Collaborative	433,225	519,225	519,225	1,471,675	15	237	224
DeSoto County Early Learning Collaborative	54,348	124,700	124,700	303,748	2	40	39
Gilmore Early Learning Initiative Collaborative (in Monroe County)	174,210	501,247	466,236	1,141,693	14	223	213
Lamar County Early Learning Collaborative	105,320	236,500	236,500	578,320	5	90	87
McComb Community Collaborative for Early Learning Success	462,250	462,250	462,250	1,386,750	22	261	202
Petal Early Learning Collaborative/Petal Excel by 5	107,500	107,500	107,500	322,500	4	77	69
Picayune School District	43,000	43,000	43,000	129,000	1	20	19
Sunflower County Early Learning Collaborative	82,616	316,050	316,050	714,716	9	147	138
Tallahatchie Early Learning Alliance	110,000	405,002	404,561	919,563	15	177	148
TOTAL	**$2,072,023**	**$3,182,320**	**$3,183,481**	**$8,437,824**	**111**	**1,713**	**1,580**

SOURCE: MDE.

Statewide Assessment for Measuring Student Readiness

MDE adopted a statewide kindergarten readiness assessment to measure the readiness of each prekindergarten student for kindergarten.

In March 2014, MDE adopted a statewide assessment to assess the readiness of each prekindergarten student for kindergarten—the Kindergarten Readiness Assessment.

The Kindergarten Readiness Assessment is one component of a system called the Mississippi K-3 Assessment Support System (MKAS[2]). The MKAS[2] includes the following three components:

- a universal screener[2] assessment for kindergarten through third grade,
- the Kindergarten Readiness Assessment, administered to all public prekindergarten and kindergarten students; and,
- the third grade reading summative assessment.

The Kindergarten Readiness Assessment is a computer-based adaptive assessment[3] that classifies student performance into one of four levels: Early Emergent Reader, Late Emergent Reader, Transitional Reader, and Probable Reader. (See Appendix B, page 34, for definitions of these literacy classification levels.)

In August 2014, early childhood stakeholders in Mississippi, including public prekindergarten teachers, child care directors, Head Start leaders, school district administrators, principals, and collaborative directors, met with MDE and Renaissance Learning (i. e., the Kindergarten Readiness Assessment vendor). Renaissance Learning provided general information to the stakeholders about the assessment and data from nationwide research regarding the trajectories of students at various scaled scores. Renaissance Learning and MDE also provided stakeholders with the results from the spring 2014 pilot administration of the assessment in Mississippi.

The stakeholder group then set two target scores (i. e., benchmark scores) based on this information. According to MDE, the following scores indicate a seventy percent mastery of early literacy skills:

---

[2]The *universal screener assessment* is given at least three times during each school year. This assessment is intended to provide data on individual students quickly in order to inform instruction and intervention decisions.

[3]A *computer-based adaptive assessment* is one in which the computer program adjusts the difficulty of questions throughout the assessment based on student responses. For example, a student who answers a question correctly will receive a more challenging item, while an incorrect answer generates an easier question.

- a score of 498 as the end-of-the-year target score for exiting prekindergarten four-year-old students; and,
- a score of 530 as the beginning-of-the-year target score for entering kindergarten students.[4]

Thus, if a prekindergarten student met or exceeded a score of 498 on the post-test of the Kindergarten Readiness Assessment, that student would be considered "on track" to reach or exceed the kindergarten readiness score of 530 upon entering kindergarten. According to MDE, based on a four-year longitudinal study, 84 percent of students entering kindergarten with a scaled score of 530 or above on STAR Early Literacy (i. e., the Kindergarten Readiness Assessment) met or exceeded the criterion for proficient reading at the end of third grade that was set by a national standard-setting committee.

MDE required all prekindergarten students enrolled in the collaboratives to participate in the assessment during the fall of 2014 (pre-test) and the spring of 2015 (post-test). Also, MDE indicated in its 2014 annual report that the Kindergarten Readiness Assessment would be required for all public prekindergarten programs beginning in School Year 2014-15.

[4]MDE did not identify research that supported the basis of the target scores (i. e., 498 end-of-year score for exiting prekindergarten students and 530 beginning-of-year score for entering kindergarten students) that the department believes indicates seventy percent mastery of early literacy skills. It is also not clear why the department believes 70% is an appropriate measurement of mastery of early literacy skills.

Evaluation of Operations and Program Effectiveness

MDE and the collaboratives should be recognized for their accomplishment in implementing a state-funded prekindergarten program that served approximately 1,580 students in its first full year of operation and for the success of the 58.86% of students who achieved the end-of-school-year target score or higher on the Kindergarten Readiness Assessment. However, PEER's independent analysis of student test data shows that there is significant room for improvement in the program's effectiveness. For example, PEER determined that the average performance of students in non-collaborative publicly funded prekindergarten programs was better than the average performance of students in the collaboratives by a statistically significant amount. PEER also determined that there was an extremely wide range of end-of-school-year assessment pass rates by program provider, by collaborative, and by curriculum. To improve the effectiveness of the program, MDE must conduct the research and analysis necessary to identify those factors most associated with positive student outcomes and ensure that all students enrolled in the collaboratives receive the quality programming necessary for school readiness.

According to MISS. CODE ANN. Section 37-21-51 (3) (g) (1972), MDE "shall make an annual report to the Legislature and the Governor regarding the effectiveness of the [prekindergarten] program." MDE's annual report submitted to PEER in July 2014 did not include information on students for a complete school year, as collaboratives were utilizing their funding to expand the number of children served by creating new sites or new classrooms or by increasing the number of slots available at existing sites and classrooms. The July 2015 report included information on the eleven collaboratives' first complete school year. Therefore, PEER analyzed MDE's 2015 annual report with regard to whether it provides adequate evidence to demonstrate program effectiveness.

MISS. CODE ANN. Section 37-21-51 (h) (ii) (1972) requires evidence of program effectiveness for continued funding and defines program effectiveness as "the school readiness of participants." MDE's annual report, required by MISS. CODE ANN. Section 37-21-51 (3) (g) (1972), does not provide adequate evidence to demonstrate program effectiveness for the following reasons:

- Although required by state law to do so, MDE has not adopted a minimum rate of readiness as a benchmark for continued program funding. The department is researching the use of student growth as a future determinant for continued funding eligibility, possibly yielding a more lenient standard;
- MDE's annual report does not show the rates of school readiness for each collaborative and each provider;

- MDE awarded funding to four collaboratives that utilized a prekindergarten curriculum found through rigorous research to have "no discernable effects" on student learning; and,
- MDE has not assessed students' progress on some content areas of the department's early learning standards.

Also, as required by MISS. CODE ANN. Section 37-21-51 (3) (g) (1972), PEER conducted an independent evaluation of the program's effectiveness. Pages 24 through 28 contain the results of this evaluation.

Issues Regarding the Benchmark for Continued Program Funding

MDE has not adopted a minimum rate of readiness that providers must meet to remain eligible for funding, as required by state law. Because MDE is using the results of the 2014-15 Kindergarten Readiness Assessment to establish its minimum rate of readiness (rather than adopting a rate in the earlier phases of program planning), there is an increased opportunity for MDE to choose its rate based on a desired level of funding rather than a desired level of student performance. Also, MDE is considering changing the criteria by which a minimum rate of readiness is calculated, a change that could possibly yield a more lenient standard.

No Minimum Rate of Readiness Set for Continued Funding

MISS. CODE ANN. Section 37-21-51 (3) (c) (iv) (1972) requires that MDE adopt a minimum rate of readiness on the Kindergarten Readiness Assessment that each prekindergarten provider must meet in order to remain eligible for program funds. As of July 24, 2015, MDE had not adopted a minimum rate of readiness.

According to MISS. CODE ANN. Section 37-21-51 (3) (c) (iv) (1972):

> *The department will evaluate the effectiveness of each early childhood collaborative and each prekindergarten provider. If the State Department of Education adopts a statewide kindergarten screening that assesses the readiness of each student for kindergarten, the State Department of Education shall adopt a minimum rate of readiness that each prekindergarten provider must meet in order to remain eligible for prekindergarten program funds.*

According to MDE's 2014 annual report on the prekindergarten program, the minimum rate of readiness is the percentage of exiting prekindergarten students within a collaborative that meet the target score of 498 on the Kindergarten Readiness Assessment (post-test). Thus

the minimum rate of readiness is to serve as the benchmark for whether providers will continue to receive funding through the Early Learning Collaborative Act.

As noted previously, MDE has adopted the Kindergarten Readiness Assessment as its statewide kindergarten screening instrument; therefore, as required by state law, MDE must adopt a minimum rate of readiness that each prekindergarten provider must meet in order to remain eligible for funds. The minimum rate of readiness should be set prior to implementation of the prekindergarten program because when the choice of assessment methods is made after the data to be assessed are collected, it would be possible to use the results of an assessment as a factor in choosing the assessment method. To do so would be to choose a benchmark based on a desired level of funding rather than on a desired level of performance.

As of July 24, 2015, MDE had not set the minimum rate of readiness. According to MDE staff, the department will establish the rate during the fall of 2015 based on assessment results from the 2014-15 school year and this information will be included in the FY 2016 annual report on the prekindergarten program. This report would not be available until July 2016, after the legislative session in which the Legislature could consider continued funding for the prekindergarten program.

Establishing standards for a program post hoc (in this case, two years after initial implementation) is procedurally inappropriate for the reason mentioned above--it opens the possibility of letting funding determine evaluative methods, rather than evaluative methods determining funding.

Possible Change in Criteria for Benchmark for Continued Funding

MDE is researching the use of student growth as a future determinant for continued funding eligibility. Thus, the department is considering changing the criteria for the benchmark for continued funding for collaboratives, possibly yielding a more lenient standard.

Although the Early Learning Collaborative Act of 2013 does not specify a minimum rate of readiness (such as a statutorily established percentage) or criteria for MDE to use in establishing a minimum rate of readiness, MDE indicated in its 2014 annual report to PEER that the minimum rate of readiness would be a percentage of students who meet a certain score on the Kindergarten Readiness Assessment.

Based on information provided in its annual reports, MDE is considering changing the criteria by which minimum rate of readiness is calculated. MDE stated initially, in its 2014 annual report on the prekindergarten program:

> *The MDE will use data gathered from the Kindergarten Readiness Assessment to determine the cut point for kindergarten readiness. The MDE will also use the data to establish the percentage of students that must meet this criteria in order for each pre-K provider to remain eligible for pre-K program funds.*

The following year, in its 2015 annual report on the prekindergarten program, MDE staff stated that for collaboratives not meeting the minimum rate of readiness, MDE is researching the use of student growth as a determinant for continued funding eligibility. If MDE adds growth as a factor for only those collaboratives that fail to meet the minimum rate of readiness, then the effect will either be that the collaborative still fails to meet the minimum rate of readiness (after growth is applied) or that the collaborative grew enough to meet the minimum rate of readiness. Also, PEER notes that the statutory requirement is for MDE to establish a "minimum rate of readiness," not change in rate of readiness.

No Reporting on School Readiness by Collaborative and Provider

The Early Learning Collaborative Act requires that MDE evaluate the effectiveness of each participating collaborative and provider. Although MDE's 2015 annual report indicates that 59 percent of the collaboratives' students scores lie at or above the target score for exiting prekindergarten students, the report does not show the rates of school readiness for each collaborative and each provider. MDE plans to report this data for each collaborative, but did not provide a timeline for doing so.

According to MISS. CODE ANN. Section 37-21-51 (3) (c) (iv) (1972), MDE "will evaluate the effectiveness of each early childhood collaborative and each prekindergarten provider." CODE Section 37-21-51 (3) (h) (ii) (1972) defines the prekindergarten's program effectiveness as "the school readiness of participants."[5] As noted previously, CODE Section 37-21-51 (3) (g) (1972) requires that the department make an annual report to the Legislature and the Governor regarding the effectiveness of the program.

To comply with state law, MDE must include in its annual report the percentages of students that are ready for kindergarten (i. e., those who reached the target score) within each collaborative and by provider. However, rather than reporting the number or percentage of students reaching the target scores for each collaborative and provider (e. g., X percent of students reached/exceeded the

[5]According to MDE, the rate of readiness is the percentage of exiting prekindergarten students within a collaborative who meet or exceed the target score of 498 on the Kindergarten Readiness Assessment (post-test).

target score for a collaborative), MDE's report displays the overall *average* score for each collaborative (e.g., an average fall scaled score of 418 for a collaborative) and the overall percentage of students reaching the target score (i.e., 58.86%).

The average score for a collaborative does not provide the best information regarding the effectiveness of the program because it does not show how many students are ready for kindergarten in each collaborative. An average could be greatly affected by outlier scores, which are unusual compared to the rest of the scores in that they could be extremely low or extremely high. MDE also did not report the scores of individual providers within the collaboratives, as required by law.[6]

MDE stated in its 2015 annual report that the average scores of eight of eleven collaboratives exceeded the target score (498) for the Kindergarten Readiness Assessment. However, the department reported that only 58.86% of the collaboratives' students (as a whole) are at or above the target score for exiting prekindergarten students. For the three collaboratives that did not reach the target score, MDE stated that teaching staff will be required to attend a list of trainings specific to his or her area of need. These collaboratives will also have an assigned literacy coach to work with the teachers.

PEER inquired as to the percentages of students who reached the target score for each collaborative and provider. On July 24, 2015, MDE responded with the following:

> *Renaissance Learning and the MDE Office of Technology and Strategic Services are currently working to clean up data files from the Kindergarten Readiness Assessment. Final calculations of percentages will be determined once data files are finalized. The results for each collaborative will become publicly reported in a manner that does not violate FERPA and state regulation for privacy.*

Since MDE provided the kindergarten readiness scores to PEER upon request, PEER was able to calculate the percentages of students that reached the target score of 498 on the Kindergarten Readiness Assessment (post-test) by collaborative and by provider. For collaboratives, the percentage of students that met or exceeded the target score ranged from 39 percent to 70 percent. (See Appendix C, page 35, for the collaboratives' percentages.) For providers, the percentage of students who met or

[6]PEER notes that if MDE is not reporting scores of individual providers and collaboratives due to concerns regarding the Family Educational Rights and Privacy Act (FERPA), the department must still comply substantially with state law and must determine an alternative method of reporting.

exceeded the target score ranged from 19 percent to greater than 95 percent. (See Appendix D, page 36, for the providers' percentages.)

Awarding Funding to Collaboratives Utilizing a Curriculum Found to Have "No Discernable Effects" on Student Learning

MDE awarded funding to four collaboratives that utilized a prekindergarten curriculum found through rigorous research to have "no discernable effects" on student learning. For the next phase of funding beginning in FY 2017 (pending legislative approval), MDE indicated to PEER that it plans to require collaboratives to use evidence-based curricula.

The Early Learning Collaborative Act of 2013 (MISS. CODE ANN. Section 37-21-51 et seq.) requires that prekindergarten programs established by the act have curricula "based on the findings of current research" that have been "found to be effective in improving student learning." MDE's request for proposals issued to collaboratives on October 10, 2013, included in its participation criteria the following:

> *All participating pre-K programs must use a research-based curriculum that is designed to prepare students to be ready for kindergarten. The curriculum must place an emphasis on early literacy and be aligned with MDE's Early Learning Standards for Classrooms Serving Four-Year-Old Children.*

As shown in Exhibit 4, page 21, four collaboratives to which MDE awarded funding utilized a prekindergarten curriculum in 2014-15 found through rigorous research to have "no discernable effects" on student learning: Creative Curriculum for Preschool, Fourth Edition.

MDE's request for proposals to collaboratives applying for funding through the Early Learning Collaborative Act required the collaboratives to submit documentation of the research base for the curricula that would be utilized under the program. However, the RFP did not put forth specific requirements for such research. PEER believes that appropriate requirements for such research to meet an evidence-based standard would include that:

- multiple research studies should have been conducted with a high level of rigor;
- the research design should have included randomized controlled trials or rigorous quasi-experimental design; and,

Exhibit 4: Collaboratives' Curricula for School Year 2014-15 and Evidence of Effectiveness Based on the Results First Clearinghouse Database

Collaborative	Curriculum	Results First Clearinghouse Information
Clarke County Early Learning Partnership	Opening the World of Learning (OWL)	None
Coahoma County Pre-K Collaborative Initiative	Opening the World of Learning (OWL)	None
Corinth-Alcorn Prentiss Early Learning Collaborative	Galileo	None
DeSoto County Early Learning Collaborative	Creative Curriculum	No discernable effects
Gilmore Early Learning Initiative Collaborative	Opening the World of Learning (OWL)	None
Lamar County Early Learning Collaborative	Creative Curriculum	No discernable effects
McComb Community Collaborative for Early Learning Success	Big Day Curriculum	None
	Creative Curriculum	No discernable effects
Petal Early Learning Collaborative/Petal Excel by 5*	Frog Street	None
	HighScope**	Evidence of effectiveness
Picayune School District	Frog Street	None
Sunflower County Early Learning Collaborative	Creative Curriculum	No discernable effects
Tallahatchie Early Learning Alliance	Opening the World of Learning (OWL)	None

*The Petal Early Learning Collaborative/Petal Excel by 5 began using the Creative Curriculum in some of its classrooms (i. e., the Pearl River Valley Opportunity Head Start classrooms) on August 1, 2015.

**The HighScope curriculum will not be used in the 2015-16 school year; it was the curriculum used by a previous Head Start grantee.

SOURCE: MDE and the Results First Clearinghouse Database.

- the researcher should be independent (i. e., the researcher should not be the developer or vendor of the curriculum).

MDE should use all available resources to verify that the prekindergarten curricula funded through the Early Learning Collaborative Act have been found to be effective in improving student outcomes. One way to identify curricula that have been proven to be effective in improving student outcomes is to review research clearinghouse data. Research clearinghouses conduct literature reviews and rate programs and curricula based on rigorous evaluations. An example of a clearinghouse that addresses educational programs and curricula is SAMHSA's National Registry of Evidence-based Programs and Practices.

Also, the Pew-McArthur Results First Initiative has compiled information from multiple clearinghouses into the Results First Clearinghouse Database. This database includes information from eight national research clearinghouses covering a range of policy areas. While some prekindergarten curricula might not be included in the Results First Clearinghouse Database, this database could be used, at a minimum, to prevent using curricula that research has shown to have no effect or have a negative effect on student outcomes.

To determine whether the curricula utilized by the collaboratives in the first full year of implementation of the Early Learning Collaborative Act had strong research bases that showed evidence of program effectiveness, PEER checked the Results First Clearinghouse Database. As shown in Exhibit 4, page 21, of the curricula utilized by the collaboratives in the 2014-15 school year:

- eight had no information in the database;
- four had research showing "no discernable effects" on student learning (all four used Creative Curriculum for Preschool, Fourth Edition); and,
- one had evidence of effectiveness.[7]

PEER notes that in order to achieve program effectiveness, programs must be implemented with fidelity to program design. PEER has no evidence regarding whether the curricula utilized by the eleven collaboratives were implemented according to design.

For the next phase of funding beginning in FY 2017 (pending legislative approval), MDE indicated to PEER that it plans to require collaboratives to use evidence-based curricula.

[7]As shown in Exhibit 4, two of the collaboratives utilized two different curricula during the 2014-15 school year, so the total number of curricula for the eleven collaboratives is thirteen.

No Assessment of Students' Progress on Some Content Areas in MDE's Early Learning Standards

Although MDE requires participating collaboratives to align their curricula with the Early Learning Standards for Classrooms Serving Four-Year-Old Children adopted by the State Board of Education, MDE's Kindergarten Readiness Assessment does not measure students' progress on five of the eight content areas in the early learning standards.

As noted previously, *Early Learning Standards for Classrooms Serving Four-Year-Old Children*, adopted by the State Board of Education prior to 2013, define what four-year-old children should understand and be able to do in the following content areas:

- English Language Arts (ELA);
- Mathematics;
- Approaches to Learning;
- Social and Emotional Development;
- Science;
- Physical Development;
- Creative Expression; and,
- Social Studies.[8]

According to MDE, the ELA and Math standards were updated to align with the Common Core State Standards for ELA and Math. The standards in the other six content areas were updated based on national standards for early childhood education. The request for proposals for participation in the Early Learning Collaborative Act required that collaboratives align their curricula with these standards and all participating collaboratives are expected to teach to the same standards.

MDE provided information to PEER regarding the Kindergarten Readiness Assessment's measurement of MDE's early learning standards that showed that the assessment does not measure students' progress on five of the content areas: Social and Emotional Development, Science, Physical Development, Creative Expression, and Social Studies. Within the content area of English Language Arts, the assessment does not assess writing or speaking and listening, presumably due to students being tested on the computer rather than with paper and pencil.

Because MDE required the collaboratives to align their curricula with the *Early Learning Standards for Classrooms*

[8]To review the *Early Learning Standards for Classrooms Serving Four-Year-Old Children*, see http://www.mde.k12.ms.us/ESE/EC.

Serving Four-Year-Old Children, the logical assumption would be that the assessment would measure children's progress on content areas within these standards as part of a measure of program effectiveness. According to MDE, collaboratives use checklists and other curriculum-based assessments during the school year to measure students' progress on the content areas not measured by the Kindergarten Readiness Assessment.

While not part of this review, PEER notes that an analysis of whether the MDE-adopted Kindergarten Readiness Assessment is the best tool to measure readiness for kindergarten based on MDE's standards for four-year-olds is an issue for further study.

PEER's Independent Evaluation of Program Effectiveness

PEER found that prekindergarteners in the collaboratives achieved the end-of-the-year target score or higher on the Kindergarten Readiness Assessment less often than those enrolled in other public prekindergarten programs. Students participating in two collaboratives that used the OWL curriculum achieved the end-of-the-year target score or higher on the Kindergarten Readiness Assessment significantly more often than students taught under the Creative Curriculum (which was determined by multiple, rigorous controlled studies to have no discernable effect on student learning).

As noted previously, MISS. CODE ANN. Section 37-21-51 (3) (g) (1972) requires that MDE make an annual report to the Legislature and the Governor regarding the effectiveness of the program. The department should base this evaluation of program effectiveness on statistical analysis of students' scores on the Kindergarten Readiness Assessment. The department should also use statistical analysis to set the target scores for the assessment. As noted on page 14, MDE did not identify research that supported the basis of the target scores or document why 70% is an appropriate measurement of mastery of early literacy skills.

Also, as noted previously, MISS. CODE ANN. Section 37-21-51 (3) (g) (1972) directs the PEER Committee to conduct "an independent evaluation of program operation and effectiveness" of the prekindergarten programs funded through the Early Learning Collaborative Act of 2013. PEER notes that this effectiveness evaluation is limited to the performance of the eleven collaboratives that received funding through the Early Learning Collaborative Act and does not address whether prekindergarten is the best investment of Mississippi tax dollars in comparison to other statewide priorities. Although research conducted in

New Mexico[9] and Oklahoma[10] has concluded that prekindergarten programs may have broad benefits for their participants, decisions regarding future investment of public funds in prekindergarten programs in Mississippi should be based on statistical analysis of the effectiveness of the programs established by the Early Learning Collaborative Act and on other research targeted to our state's specific needs.

In order to make a judgment regarding the effectiveness of programs offered by the eleven collaboratives, PEER analyzed data collected by MDE--i. e., results of the pre- and post-assessments given to prekindergarteners enrolled in the collaboratives as well as those enrolled in other public prekindergarten programs. (As stated on page 14, MDE required all prekindergarteners enrolled in public prekindergarten programs during the 2014-15 school year to take the Kindergarten Readiness Assessment.) The assessment data provided by MDE limits PEER's efforts in drawing conclusions regarding effectiveness of the collaboratives in preparing prekindergarteners to enter kindergarten due to the following:

- The department used its Kindergarten Readiness Assessment, an early literacy test, as its sole determinant as to whether prekindergarteners were ready to enter kindergarten.
- State law, as well as the request for proposals used by the department to select collaboratives to receive grant funds, required curriculum used by the collaboratives to teach to the department's *Early Learning Standards for Classrooms Serving Four-Year-Old Children*. The standards encompass concepts that are not assessed by the department's Kindergarten Readiness Assessment, such as creative expression, physical development, and social and emotional development. The standards are designed to affect many different areas of a child's education development that are not measured by the department's early literacy test. (See page 23 for further discussion.)
- Data from the Kindergarten Readiness Assessment do not include sufficient information with which to isolate confounding variables. For example, curriculum performance must be separated from student performance so that a collaborative with mediocre curriculum and high-performing students (who would most likely excel regardless of the curriculum under which they were taught) is not confused with a

[9]See Hustedt, J. T., Barnett, W. S., and Jung, K. (2007). The effects of the New Mexico PreK initiative on young children's school readiness. New Brunswick, NJ: National Institute for Early Education Research.

[10]See Gormley, W. T., Jr., Gayer, T., Phillips, D., and Dawson, B. (2005). The effects of universal pre-k on cognitive development. *Developmental Psychology*, 41, 872-884.

collaborative using a cutting-edge, highly functional curriculum capable of making previously low-performing students succeed. The effects of ordinary maturation must be separated from those of education. In its analysis, PEER employed mathematical means to control for such variables to the degree possible, but control through proper data collection planned in advance would be preferred.

- Data from the department's assessment does not address the fidelity of a collaborative's implementation of a particular curriculum, thereby limiting information with which to determine whether measurement of the effects of a program occurred due to the faithful use of the curriculum as intended by its creators. (See page 22.)

Given the limitations of the assessment data described above, PEER complied with its statutory mandate to assess effectiveness of the early learning collaboratives through two limited analyses of the assessment data provided by the department:

- comparison of the performance of prekindergarteners enrolled in the collaboratives to those enrolled in other public prekindergarten programs; and,
- comparison of the performance of collaboratives to each other via a baseline established through existing research.

Comparison of Prekindergarteners Participating in Collaboratives to Those Participating in Other Public Prekindergarten Programs

PEER found that prekindergarteners who had been participating in the Early Learning Collaborative Act through collaboratives achieved at least a 498 score (the end-of-the-year target score for exiting prekindergarteners) on the Kindergarten Readiness Assessment test less often than those enrolled in other public prekindergarten programs, after adjusting for initial score differences. Students in collaboratives had a 6% lower adjusted pass rate than the non-collaborative students, a distinction that is statistically significant ($p < 0.001$).

In order to arrive at this conclusion, PEER removed all students from the data who had not taken both tests, thus creating true cohorts in both the test (collaborative) and control (non-collaborative) groups. PEER chose to analyze pass rate rather than scaled score because of theoretical doubts about the scaled score's interval data properties and because the pass rate is important in the current context of student evaluation. PEER corrected for starting student ability by subtracting initial pass rate from final pass rate; this was done in order that factors affecting performance at the individual student level (e. g., possibly innate ability and socioeconomic status) would not bias

the final score of a student. PEER ensured that a national index of poverty was not statistically distinguishable between test and control groups to rule out socioeconomic factors as a cause of different rates of change among groups, thus controlling for that particular factor both at the level of the student and at the level of the group. Finally, PEER employed a permutation test of means in order to determine the significance of the observed differences; this test was chosen because of its nonparametric nature, exactitude, and suitability for non-sampled populations.

See the Technical Appendix, page 37, for further information regarding analytical procedures used by PEER to reach this conclusion.

Comparison of Performance of Collaboratives to a Baseline Established Through Existing Research

PEER found that students participating in the Gilmore Early Learning Initiative Collaborative and the Clarke County Early Learning Partnership (that both used the OWL curriculum) achieved at least a 498 score (the end-of-the-year target score for exiting prekindergarteners) significantly more often than a set of students taught under a curriculum determined by multiple, rigorous controlled studies to have no discernable effect (i. e., the Creative Curriculum). The Gilmore Collaborative students had an adjusted pass rate 21% higher than the baseline group (a significant effect at $p < 0.001$). The Clarke County collaborative students had an adjusted pass rate 20% higher than the baseline group (a significant effect at $p < 0.01$).

In order to reach the above conclusion, PEER followed fundamentally the same procedure as above, with the following addition: the baseline in this case was created from all collaborative students educated under Creative Curriculum. This means PEER compared collaborative students educated under Creative Curriculum to collaborative students not educated under that curriculum. This entailed splitting some collaboratives, an intentional effort to remove the confounding effect of a known non-performing curriculum from the performance of a collaborative in other respects. Thus, the test groups were cohorts formed from individual collaboratives, minus those students taught under Creative Curriculum. Finally, Holm multiple hypothesis correction was employed on all p-values (including that for the comparison of collaboratives to non-collaboratives) in order to ensure that the number of hypotheses tested in this study did not lead to an unacceptable rate of false positives. Corrected p-values are represented in the text above.

PEER found that the Gilmore Early Learning Initiative Collaborative and the Clarke County Early Learning Partnership achieved a 498 significantly more often than a set of students taught under a curriculum determined by

multiple, rigorous controlled studies to have no discernable effect (i. e., the baseline group). The Gilmore Collaborative students had an adjusted pass rate 21% higher than the baseline group (a significant effect at $p < 0.001$). The Clarke County collaborative students had an adjusted pass rate 20% higher than the baseline group (a significant effect at $p < 0.01$). No other collaboratives were distinguishable from the baseline group.

As stated on page 22, PEER notes that in order for the success of a program to be replicated, it should be implemented as designed. PEER has no evidence regarding whether the curricula utilized by the eleven collaboratives were implemented with fidelity to program design.

See the Technical Appendix, page 37, for further information regarding analytical procedures used by PEER to reach this conclusion.

Conclusion

Meta-analytic research on public preschool programs concludes that quality programs can have a positive impact on important outcomes that benefit society. Examples of these long-term outcomes include improved test scores and high school graduation rates as well as decreases in K-12 grade repetition, special education, and crime.[11]

In recognition of the potentially significant economic and social impact of outcomes associated with state-funded prekindergarten, the Mississippi Legislature passed the Early Learning Collaborative Act of 2013. The act established a mechanism for implementing state-funded four-year-old prekindergarten in Mississippi on a phased-in basis, with each phase conditioned on a review of the program's effectiveness in preparing students for kindergarten. The act contains numerous provisions intended to ensure the quality of the preschool programs established using state funds--e. g., requirements for coordination of early childhood education programs and services; adoption of curricula proven to be effective in improving student learning through rigorous research; ongoing monitoring of funded prekindergarten programs to ensure proper implementation of the chosen curriculum; ongoing evaluation and research of the effectiveness of each program provider; and technical assistance, informed by the department's evaluation findings, to improve program quality.

The Legislature appropriated a total of $9 million to fund Mississippi's four-year-old prekindergarten program during its first phase of operation (fiscal years 2014 through 2016). The 2014-2015 school year marked the first full year of operation of the program. During that year, the State Department of Education distributed $3,182,320 in state funds to fifty-three prekindergarten program providers in eleven collaboratives to serve approximately 1,580 students.

MDE's annual report on the effectiveness of the state-funded prekindergarten program following its first full year of implementation states that 58.86% of students in the collaboratives scored at or above the end-of-school-

[11]The Legislature of the State of Washington created the Washington State Institute for Public Policy (WSIPP) in 1983 to carry out practical, non-partisan research that answers relevant policy questions. WSIPP has developed a model to evaluate the costs and benefits of a wide range of public policies and programs, including Pre-K to 12 education, that affect many different outcomes. The institute follows a meta-analytic framework to assess systematically all relevant evaluations on a given topic (e. g., state and district early childhood education programs) meeting the institute's high-quality standards and computes an average effect on each measured outcome from all of the credible studies so identified. WSIPP's benefit-cost analyses, including estimated average effect sizes by outcome measured, are available on its website: http://www.wsipp.wa.gov/

year target score of 498 on the Kindergarten Readiness Assessment. MDE also reports that the average scaled scores at the end of the school year for eight of the eleven collaboratives exceeded the target score of 498.

While the department and the collaboratives should be recognized for their accomplishments in implementing a state-funded prekindergarten program aligned to the state's rigorous early childhood education standards, PEER's independent analysis of student assessment data for the 2014-2015 school year shows that there is significant room for improvement in the program's effectiveness as measured by student achievement. The following examples clearly illustrate this point.

- *The average performance of students in non-collaborative publicly funded prekindergarten programs was better than the average performance of students in collaborative prekindergarten programs by a statistically significant amount.* PEER determined that students served by publicly funded prekindergartens that are not part of a collaborative had a 6% higher average adjusted end-of-school-year pass rate on the Kindergarten Readiness Assessment (defined as scoring 498 or above) than students served by the collaboratives.
- *There is a wide range in the percentage of students scoring at or above the end-of-school-year target score.* The percentage of students scoring at or above the end-of-school-year target score of 498 on the Kindergarten Readiness Assessment, by provider, ranges from 19% to above 95%. With twenty- two providers reporting percentages below the state average of 58.86% and only one of the fifty-three providers achieving above 95% performance at or above the target score, there is significant opportunity for improving student readiness for kindergarten among schools in the collaboratives.
- *There is a wide range in average assessment scores, by collaborative.* PEER determined that students participating in certain collaboratives (e. g., Gilmore Early Learning Initiative) using certain curriculums (OWL) performed significantly better on the assessment than students taught in classrooms using The Creative Curriculum for Preschool, 4th Edition.
- *There is room for MDE to increase its target scores on the Kindergarten Readiness Assessment.* As discussed on page 14, MDE stated that the end-of-school-year target score of 498 for exiting preschool students indicates "70% mastery of early literacy skills." According to MDE, based on a four-year longitudinal study, 84% of students entering kindergarten with a scaled score of 530 or above on the Kindergarten Readiness Assessment (MDE's target score for

beginning-of-the-school year kindergarten students) met or exceeded the criterion for proficient reading at the end of the third grade. In light of the fact that MDE is advocating 100% student proficiency in reading by the end of the 3rd grade, current target scores fall short of setting all students on a path to achieve that goal.

The purpose of the Early Learning Collaborative Act of 2013 is to help ensure that all children have access to quality early childhood education and development services. While meta-analysis of the research on state funded prekindergarten programs indicates a positive return on investment, the chance that program benefits will exceed costs is not absolute.[12]

Through careful research and analysis of Mississippi's publicly funded prekindergarten programs, including analysis of the quality of the curricula selected and the fidelity of providers in implementing each program according to its research-based design, MDE must learn to replicate those factors most associated with positive student outcomes and ensure that they are provided to all students enrolled in the state's early learning collaboratives.

[12]According to meta-analysis conducted by the Washington State Institute for Public Policy, the chance that the monetized benefits of a state-funded prekindergarten program will exceed its costs is 83%.

Appendix A: Criteria for Prekindergarten Program Funding

According to MISS. CODE ANN. Section 37-21-51 (3) (d) (1972), prekindergarten program funds shall be awarded to early childhood collaboratives whose proposed programs meet the following program criteria:

- Voluntary enrollment of children
- Collaboration among pre-K providers and other early childhood programs through establishment of an early learning collaborative
- Qualifications of master teachers, teachers and assistants
- At least fifteen hours of annual professional development for program instructional staff, including professional development in early literacy
- The use of state-adopted comprehensive early learning standards
- The use of a research-based curriculum that is designed to prepare students to be ready for kindergarten, with emphasis in early literacy, and is aligned with the comprehensive early learning standards
- The use of age-appropriate assessments aligned to the comprehensive early learning standards
- Teacher/child ratios of 1:10 with a maximum of twenty and minimum of five children per classroom
- The provision of at least one meal meeting state and federal nutrition guidelines
- Plans to screen and/or refer children for vision, hearing, and other health issues
- Parent involvement opportunities
- Plans to serve children with disabilities
- Number of instructional hours to be provided, which shall be no less than 540 instructional hours per school year for half-day programs and 1,080 for full-day programs
- A budget detailing the use of funds for allowed expenses

SOURCE: MISS. CODE ANN. Section 37-21-51 (3) (d) (1972).

Appendix B: Literacy Classification Levels Based on Scores on the Kindergarten Readiness Assessment

Emergent Reader (scores between 300 – 674)

- Early Emergent Reader (300 – 487): Student is beginning to understand that printed text has meaning. The student is learning that reading involves printed words and sentences and that print flows from left to right and from the top to the bottom of the page. The student is also beginning to identify colors, shapes, numbers, and letters.

- Late Emergent Reader (488 – 674): Student can identify most of the letters of the alphabet and can match most of the letters to their sounds. The student is also beginning to "read" picture books and familiar words around the home. Through repeated reading of favorite books with an adult, students at this stage are building their vocabularies, listening skills, and understandings of print.

Transitional Reader (scores between 675 – 774)

Student has mastered alphabet skills and letter-sound relationships. The student can identify many beginning and ending consonant sounds and long and short vowel sounds and is probably able to blend sounds and word parts to read simple words. The student is also likely using a variety of strategies to figure out words, such as pictures, story patterns, and phonics.

Probable Reader (scores between 775 – 900)

Student is becoming proficient at recognizing many words, both in and out of context. The student spends less time identifying and sounding out words and more time understanding what was read. Probable readers can blend sounds and word parts to read words and sentences more quickly, smoothly, and independently than students in the other stages of development.

SOURCE: MDE.

Appendix C: Percentages of Students That Met or Exceeded Target Score on the Kindergarten Readiness Assessment (Post-test) by Collaborative

Collaborative (Random Number Assignment*)	Percentage of Students That Met or Exceeded the Target Score of 498 on Post-test
2	70%
1	69%
4	67%
7	61%
11	59%
10	59%
3	58%
9	50%
5	42%
6	42%
8	39%

*PEER assigned a random number to each collaborative in order to avoid any potential Family Educational Rights and Privacy Act (FERPA) violation involving personally identifiable information.

SOURCE: PEER analysis of student test data.

Appendix D: Percentages of Students That Met or Exceeded Target Score on the Kindergarten Readiness Assessment (Post-test) by Provider

Provider (Random Number Assignment)*	Percentage of Students That Met or Exceeded the Target Score of 498 on Post-test
9	Greater than 95%
36	93%
12	90%
17	89%
39	88%
32	87%
19	86%
4	84%
21	84%
29	84%
14	82%
35	80%
53	80%
25	79%
50	76%
44	75%
26	72%
27	71%
5	70%
20	70%
38	70%
10	69%
43	68%
6	67%
15	67%
2	64%
33	61%

Provider (Random Number Assignment)*	Percentage of Students That Met or Exceeded the Target Score of 498 on Post-test
8	61%
13	60%
37	60%
31	59%
51	56%
45	56%
3	54%
7	53%
47	51%
28	51%
46	50%
49	50%
40	47%
41	47%
16	47%
42	44%
23	43%
34	43%
52	42%
22	41%
11	38%
30	36%
48	33%
24	32%
18	28%
1	19%

*PEER assigned a random number to each provider in order to avoid any potential Family Educational Rights and Privacy Act (FERPA) violation involving personally identifiable information.

SOURCE: PEER analysis of student test data.

Technical Appendix

Introduction to the Technical Appendix

The purpose of this technical appendix is to provide information on specific methodological choices made during the course of the analytic portion of this study (pages 24 through 28). This technical appendix contains discussions of the raw data, the data preprocessing, the analytic procedures employed, and the results.

PEER notes two caveats about this appendix and about this study's analytic procedures. Regarding this appendix, much of the detail that would normally be present in a technical appendix to a report of this nature has been omitted in order to comply with requirements of the Family Educational Rights and Privacy Act regarding personally identifiable information. Also, as noted in the text of the report, PEER believes that a more detailed study, including methods and data collection planned specifically for the purpose (see pages 25-26) would be necessary to evaluate prekindergarten programs in Mississippi in the most thorough manner.

The goal of this study was to deliver the most useful evaluation possible within the constraints of time, rigor, and the available data. While PEER could not assess the effects of prekindergarten programs compared to the absence of prekindergarten due to data limitations (see pages 25-26), it was possible to assess the effects of the programs chosen for the prekindergarten collaboratives relative to one another (and thus to a baseline); it was also possible to assess the performance of the collaboratives relative to that of non-collaboratives.

Regardless of whether prekindergarten is a valuable public investment compared to other uses of public dollars, if the Legislature is going to invest in it, it should invest in a way that will have the greatest effect. Thus, it is important to know the results of the implementation of the Early Learning Collaborative Act's prekindergarten program in terms of relative effectiveness. Also, if the collaborative organizational structure itself is intended as a novel approach to the problem of early childhood education, it is important to know whether it has any effects relative to other such structures.

Thus pages 24 through 28 of the report set out to answer two questions: Which of the collaboratives performed better than expected for a baseline prekindergarten curriculum to a greater than chance degree? And how did collaboratives in general perform relative to non-collaborative publicly funded prekindergartens?

PEER notes that this analysis was essentially descriptive and only minimally inferential. That is, it was not designed to make generalizations about either collaboratives or curricula; rather, it was designed to assess the collaboratives' actual performance in the particular context of initial implementation of the prekindergarten program established by the Early Learning Collaborative Act.

Since PEER has no data on rigorous fidelity studies conducted as a part of Mississippi's initial implementation of the Early Learning Collaborative Act, this study could not distinguish a useful curriculum poorly implemented from a useless curriculum faithfully implemented. However, even with these caveats, these results can be useful in informing and shifting the burden of proof. For instance, if PEER found that a collaborative performs no better than the baseline, the assertion that its curriculum had a sound research base demands evidence that its implementation was somehow flawed. And if, on the other hand, PEER had assurance of faithful curriculum implementation at a school that performed no better than chance, then a different curriculum should be considered at that school.

Raw Data

The data used for this study was the set of all 2014-15 results of the Kindergarten Readiness Assessment. The data included the results of tests given both in the fall of 2014 and in the spring of 2015; many, but not all, students were tested on both occasions. The data included results from both students in the collaboratives currently under evaluation and students in non-collaborative prekindergarten programs.

Out of concern for the potential exposure of personally identifiable information, this technical appendix will not report the size of the data set or of any important subpopulations (e. g., collaboratives or schools) or the usual summary statistics; some summary statistics are reported in the main text with personally identifying information removed.

Data Preprocessing

The first step in testing the collaboratives was to ensure that the data represented a genuine cohort. Since not every student who took the first test was present for the second, and not every student who took the second test was present for the first, it was necessary to pare down the data to only those students who took both tests. To fail to do so would be to leave open the possibility that a change in scores over time was due to student attrition or acquisition rather than to any sort of learning.

For purposes of testing the collaboratives against the noncollaboratives, no further preprocessing was necessary.[13]

The next step in testing the collaboratives against baseline prekindergarten performance was to establish that baseline. One of the curricula in use in the prekindergarten pilot program--Creative Curriculum for Preschool, Fourth Edition--has a solid evidence base of studies[14] establishing that it has no effect on student performance above that of prekindergarten in general. In other words, one could safely rule out the curriculum as the cause of any improvement in students taught under this edition of Creative Curriculum.

Thus, all students taught under Creative Curriculum in the prekindergarten pilot program were aggregated into a control group.[15] This procedure meant that some collaboratives were split (e. g., the McComb Community Collaborative for Early Learning Success had some students taught under Creative Curriculum and others taught under Big Day Curriculum) for purposes of the study.

Analysis

Because of the absence of random samples and known distributions of traits of interest, this study used a permutation method comparison of means as its primary analytical mechanism. This test delivers meaningful p-values for nonrandom samples--indeed, for populations that are not samples--and is entirely nonparametric (i. e., it does not depend on background assumptions about the distribution of any

[13]PEER utilized an extra analytic step at this stage, not important to the final conclusions of the study, but worth reporting. Although the adjusted-pass-rate method discussed in the analytic section should control for socioeconomic effects on student ability as manifested in the variable of interest, it was nonetheless deemed of interest to make a determination of whether the collaboratives and noncollaboratives served socioeconomically distinct populations. While PEER lacked direct information about students' socioeconomic status, it was possible to employ a data imputation strategy whereby each student was assigned the average poverty rate of his or her school district. As such, each collaborative had an assigned poverty rate equal to a weighted average of the rates of the school districts to which each of its students belonged. The collaborative and noncollaborative groups were compared with the same permutation test described in the main text; there was no significant difference between them. This method should, of course, not be used for making inferences about individual students, but at the level of aggregates of thousands gives reasonably plausible results. Information about district poverty rates was derived from: U. S. Census Bureau (2015*). School District Poverty Estimates*. Retrieved from https://www.census.gov/did/www/schooldistricts/data/poverty.html

[14]U. S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, March). *Early Childhood Education Intervention Report: The Creative Curriculum® for Preschool, Fourth Edition*. Retrieved from http://whatworks.ed.gov. The studies aggregated by the What Works Clearinghouse examine Creative Curriculum's effect on the skill set that is (ex hypothesi) tested by the Kindergarten Readiness Assessment, among other skills. Thus, the evidence from these studies generalizes to the current case.

[15]Information about which schools in the collaboratives used Creative Curriculum was gained from the Mississippi Department of Education. Students outside of the pre-k pilot program were not used for the control group because of concerns about the accuracy of available information on curriculum use.

particular characteristic of interest). The p-value delivered by this test is also highly intuitive: it is simply the probability of achieving a result equal to or greater than the observed difference in means, assuming that there is no real difference between the populations and thus all variation in the characteristic of interest is by chance.[16]

First proposed by Fisher,[17] this method has many advantages, but is impractical for some purposes because of its extreme computational demands. However, Dwass[18] proposed that a Monte Carlo simulation could be used to greatly reduce these demands; the Dwass method effectively takes a sample from the very large space of permutations that would otherwise be required, rendering the test practical at the cost of introducing a small, but precisely measurable, amount of uncertainty into the p-value. Depending on the amount of uncertainty acceptable in the study and the sensitivity desired, a smaller or larger set of simulations might be used. For purposes of this study, 10,000 simulations was considered to be the number that would provide an initially acceptable balance of certainty, practicality, and sensitivity.[19] This number of simulations is well within standard practice in the literature on such methods.[20] The confidence intervals resulting from this practice are discussed on pages 26-28 in the section on results.

The variable compared under the permutation method test was the gain in pass rate from fall to spring--that is, the percentage of students scoring at or above the 498 scaled score benchmark after prekindergarten minus the percentage of students scoring at or above that benchmark before prekindergarten. This procedure corrects for confounding variables expressing themselves in the distribution of student scores at the beginning of the process.

It would have been possible to have used gain score--spring score on the reading test, minus fall score on the reading test--as the variable for comparison across collaboratives; in theory, the scaled score is on an interval scale. However, there were several reasons not to take this option. One is that the gain

---

[16]The p-values delivered by a Monte Carlo sample from a permutation test like this one must be interpreted slightly differently from traditional p-values obtained by random-sampling methods. This "achieving results" is not a matter of sampling--pulling a sample with the appropriate characteristics from a universe in which the null hypothesis is true--but a matter of the likelihood of the specific members of the study population performing as they did if the null hypothesis were true.

[17]Fisher, R. A. (1935). *The Design of Experiments*. Oliver and Boyd, Edinburgh.

[18]Dwass, M. (1957). Modified Randomization Tests for Nonparametric Hypotheses. *Annals of Mathematical Statistics* 28, 181-187.

[19]For two comparisons, 10,000 simulations failed to capture the p-value, instead returning a p-value of zero. The true p-value in a permutation test will never be zero; a detected p-value of zero almost invariably indicates that the test employed too few simulations to capture adequately the rarity of the event under examination's occurrence under the null hypothesis. For one of the comparisons (collaboratives versus noncollaboratives), 100,000 simulations was adequate, while for the other (Gilmore versus baseline), 1,000,000 simulations captured a nonzero p-value.

[20]Burton, A., Altman, D. G., Royston, P. and Holder, R. L. (2006). The Design of Simulation Studies in Medical Statistics. *Statistics in Medicine* 25, 4279-4292.

score is necessarily more noisy than any individual score and in context of the current study, PEER lacked the information necessary to estimate the amount of noise. Another is that regression to the mean will affect the relative degree of change among low or high initial scorers and those who initially score closer to the average, rendering inferences about relative degrees of change problematic. But for this purpose the most serious consideration was that there is some doubt that the scaled score on a test of this sort actually qualifies as interval scale data.[21]

The method used reduces the data to an ordinal scale relative to a pragmatically (and, ideally, educationally) important benchmark, trading informational content for rigor.[22] In order to conduct the permutation test, students were assigned a binary value depending on whether they had met or exceeded the 498 benchmark: 1 for yes, 0 for no. The mean value of the difference between each student's fall and spring binary scores (treated as base-ten numeric) is necessarily equal to the mean difference in pass rates and this mean value can be analyzed using a standard permutation test.

Because this project involved multiple hypothesis testing, it was necessary to employ a correction factor to prevent an unacceptably high possibility of false positives. Because in this context a false positive was considered to have more serious pragmatic consequences than a false negative, PEER chose to control the familywise error rate, which minimizes the probability of any type I error. In order to minimize false negatives, PEER employed Holm's method rather than Bonferroni's, since the former admits fewer false negatives than the latter while admitting no more false positives.[23]

Results and Discussion

As discussed in the text of the report, the collaboratives as a whole performed worse than the noncollaboratives, with a 6% lower adjusted pass rate. This difference was statistically significant, with an unadjusted p-value of 0.00012 and a 95% confidence interval on that p-value of +/- 0.0000679. The midpoint of the Holm-adjusted p-value is 0.00096.

The Clarke collaborative had a 20% higher adjusted pass rate than the baseline, a statistically significant difference at p = 0.0004 +/- 0.00039191 at 95% confidence. The midpoint of the Holm-adjusted p-value is 0.0028.

---

[21]See, e.g., Yen, W. (1986). The Choice of Scale for Educational Measurement: An IRT Perspective. *Journal of Educational Measurement*, 23(4), 299-325.

[22]The method used is in line with the theoretical stance, though not identical to the methods, of Ballou, D. (2008). Test Scaling and Value-Added Measurement. Retrieved from https://my.vanderbilt.edu/performanceincentives/files/2012/10/200823_Ballou_TestScaling11.pdf

[23]Aickin, M. and Gensler, H. (1996). Adjusting for Multiple Testing When Reporting Research Results: The Bonferroni vs. Holm Methods. *American Journal of Public Health* 86(5), 726-728.

The Gilmore collaborative had a 21% higher adjusted pass rate than the baseline, a statistically significant difference at p = 0.000006 +/- 0.0000048 at 95% confidence. The midpoint of the Holm-adjusted p-value is 0.000054.

No other collaboratives were statistically distinguishable from the baseline in the current study.

As noted previously, PEER's study cannot answer the question of whether funding prekindergarten to any particular degree is an advisable use of public resources. It only examines the relative performance of existing collaboratives corrected for the pass rate of the student population.

This study demonstrates that as currently instantiated, most prekindergarten curricula in use in collaboratives implemented through the Early Learning Collaborative Act were performing no better than prekindergartens with no particular curriculum. Since there are evidence-based prekindergarten curricula available, this fact strongly suggests that Mississippi should change its prekindergarten curricula. At the very least, this study serves to shift the burden of proof; if one wished to claim that a given curriculum that failed to achieve results in this study was still valuable, one would need to show that it was not implemented properly or that other confounding factors were present.

Furthermore, this study demonstrates that collaboratives as instantiated in the Early Learning Collaborative Act were performing worse than other prekindergarten programs to a non-chance degree. There may be many explanations for this fact, but it should be addressed in future policy decisions.

Above all, this study suggests that more, and more relevant, data are necessary to evaluate rigorously the prekindergarten programs in Mississippi.

MISSISSIPPI DEPARTMENT OF EDUCATION

Carey M. Wright, Ed.D.
State Superintendent of Education

October 16, 2015

James Barber, Executive Director
Joint Legislative PEER Committee
Woolfolk Building
501 North West Street, Suite 301A
Jackson, Mississippi 39201

Dear Mr. Barber,

The Mississippi Department of Education (MDE) has reviewed the PEER Committee report "The Early Learning Collaborative Act of 2013: Evaluation of the Operations and Effectiveness of the Program." This PEER Committee report has not followed the Joint Committee on Standards for Educational Evaluation, which are the national standards for evaluating educational programs.

The MDE has prepared the following responses to each PEER critique.

Section I

PEER Critique: No Minimum Rate of Readiness Set for Continued Funding—"MISS. CODE ANN. Section 37-21-51 (3)(c)(iv)(1972) requires that MDE adopt a minimum rate of readiness on the Kindergarten Readiness Assessment that each prekindergarten provider must meet in order to remain eligible for program funds. As of July 24, 2015, MDE had not adopted a minimum rate of readiness."

MDE Response:

MDE has not yet set a minimum rate of readiness for determining whether providers are eligible for continued funding due to circumstances specific to the Kindergarten Readiness Assessment. These circumstances are explained below.

POST OFFICE BOX 771 • JACKSON, MISSISSIPPI 39205 • (601) 359-3512 • FAX (601) 359-3242

What concerns MDE more are the motives that PEER ascribes to MDE for this delay, which are not supported by any evidence. In its explanation, PEER states that MDE's failure to establish the minimum rate of readiness prior to program implementation opens the possibility that MDE may set an artificially low rate in order to justify future funding requests to the Legislature. Furthermore, PEER claims that MDE's statement that it will consider the role of growth in calculating the readiness rate is yet another way that MDE will introduce "leniency" in the metric. Both of these critiques are faulty. We can only assume that PEER is operating under at least one of four misperceptions:

1) PEER conflates the readiness benchmark with the minimum rate of readiness. The readiness benchmark represents what children should know and be able to do to be ready for success in Kindergarten. Currently, this benchmark is based on a scale score of 498 on the Mississippi K-3 Assessment Support System, which is primarily a literacy test. A score of 498 at the end of a child's four-year-old pre-Kindergarten year is a rigorous benchmark, as explained in Section II. Had MDE set this benchmark low, this would indeed be cause for concern because then MDE could claim that an artificially high number of children were meeting the mark. Not only did MDE **not** do this, this score was set prior to the first full year of implementation of the collaboratives. More importantly, the readiness benchmark is not the same as the minimum rate of readiness, which refers to the percentage of children in a particular collaborative or provider's program who meet the readiness benchmark at the end of the pre-Kindergarten year.

2) PEER misunderstands the purpose of the minimum rate of readiness. The purpose of the minimum rate of readiness is to hold collaboratives and providers accountable for quality program implementation—not to build a case for expanding the program. The plain language of the text—"minimum rate"—implies that MDE should set this rate at the lowest acceptable percentage of children meeting the readiness benchmark. The crafters of the *Early Learning Collaborative Act* modeled the minimum rate of readiness language off of similar language in Florida's law establishing its Voluntary Pre-Kindergarten program. Because funding is not available to establish an early learning collaborative in every community that seeks one, this language is intended to ensure that precious state dollars are given to the programs that are most effective. If the minimum rate of readiness is set too high, MDE risks eliminating effective programs; if the minimum rate of readiness is too low, MDE risks missing the opportunity to fund more effective programs. To try to use the success of funded collaboratives on this measure as an indicator of the effectiveness of the legislation as PEER's critique states would indeed create strange incentives. Such a usage is neither expressed nor implied anywhere in the legislation.

3) PEER lacks a basic understanding of the role of baseline data in setting accountability metrics. PEER's suggestion that MDE could have set a defensible minimum rate of readiness prior to establishing baseline data for a new assessment belies a lack of knowledge about the responsible creation of accountability systems. The

process of setting a minimum rate of readiness is a complicated one that required MDE to first establish a measurement of school readiness. Prior to 2013, Mississippi had neither a school readiness assessment nor even a statewide definition of school readiness for any of the five accepted domains of early childhood learning.

In 2013, the Mississippi Legislature passed both the *Early Learning Collaborative Act* (ELCA) and *the Literacy-Based Promotion Act* (LBPA). The LBPA required MDE to implement a statewide K-3 literacy assessment for the purpose of determining whether K-3 students were on track for meeting the state's reading benchmark by the end of third grade. This K-3 literacy assessment had to incorporate a Kindergarten readiness assessment for literacy to produce baseline data for each student.

The ELCA required MDE to adopt a minimum rate of readiness for Kindergarten if MDE adopts a "statewide Kindergarten screening that assesses the readiness of each student for Kindergarten." It does not require that this assessment be the same as the assessment implemented under the LBPA. However, due to both time and resource constraints, MDE issued a single RFP for a K-3 literacy assessment that would also be robust enough to serve as a Kindergarten readiness assessment. The selected assessment is the STAR Early Literacy assessment, now called the Mississippi K-3 Assessment Support System (MKAS^2).

In the absence of an adopted school readiness definition and within the context of having only the MKAS^2 as a statewide Kindergarten entry assessment, MDE chose the only measurable benchmark available as a school readiness indicator—a score on the MKAS^2 that was supported by research to indicate a child's chance of school success. A committee of practitioners chose a scale score of 498 as the readiness benchmark in summer 2014, prior to the first full year of implementation for collaboratives. (More on this decision in Section II.)

After MDE selected a school readiness benchmark, it could then determine what percentage of ready children each collaborative and each provider must meet in order to receive continued funding. Setting accountability requirements is necessarily a matter of values, research, and judgment as accountability metrics are only meaningful in the context in which they are designed and used. Accepted practice in education is to establish a baseline of scores on any assessment used for accountability purposes prior to setting accountability requirements, such as the high-stakes minimum rate of readiness referenced in the ELCA.

MKAS^2 was first piloted in the spring of the 2013-14 school year for K-3 students. Because so few collaboratives were operational at this point in time (the collaboratives were selected in December 2013), very few collaborative classrooms (only 6 of the 111 classrooms funded) participated in this pilot; participation was the result of circumstance rather than any sampling design that could yield meaningful baseline test score data. Consequently, the field test among collaborative classrooms was not sufficient to set a baseline of expected performance for programs. Instead, the field test

enabled MDE and the collaboratives to understand the process needed to test every child on a computer-based exam when not every child was in a facility with the technology to take the exam.

The first year in which a sufficient number of collaborative classrooms administered a fall and spring readiness assessment was 2014-15. As a result, MDE is relying on best practice in education to use assessment results from 2014-15 to set the minimum rate of readiness on the MKAS[2]. A standards validation committee underwent a process to support the benchmark score of 498 on the Star Early Literacy Assessment as the desired score for students exiting Pre-K. A score of 498 at the end of Pre-K means that the student has mastered 70% of the reading indicator skills needed and supports that the student is on track to meet the end-of-grade 3 reading proficiency expectations. This Kindergarten Readiness Assessment data is used in the beginning of the year to inform teachers to implement interventions and strategies to help children to strengthen literacy and a few numeracy skills. The data at the end of the year is used to determine their growth over the year in literacy and numeracy skills.

This work is expected to conclude within the next year. By that time, MDE also hopes to have adopted a statewide definition of school readiness as well as multiple, aligned measures that can assess each enrolled child's readiness. MDE expects to incorporate the finalized minimum rate of readiness for funding decisions beginning in the 2016-17 school year.

4) PEER does not understand the importance of growth in measuring educational effectiveness. A central assertion of the report authors is that a growth-based metric confounds measuring a program's effectiveness. A well-designed and fair system should include multiple indicators, including both status and growth/improvement. These designs assist in determining the degree to which the program is being effective in both specific levels of performance, and improvement both to and beyond the targeted level. Furthermore, to the extent that equity (or improving performance of at-risk students) is a priority outcome, indicators must be included that signal such progress and reward ongoing efforts. To this point, the Council of Chief State School Officers (CCSSO) document entitled *CCSSO Roadmap for Next-Generation State Accountability Principles (CCSSO, 2011)*, developed by a representative group of state education leaders and experts in the field of assessment and accountability, explicitly promotes multi-indicator designs.

Section II

PEER Critique: MDE awarded funding to four collaboratives that utilized a prekindergarten curriculum found through rigorous research to have "no discernable effects" on student learning.

MDE Response:

Each collaborative selected a researched-based curriculum and assessment based on the children's needs in the community, that had a literacy focus, and was aligned to all areas of the Early Learning Standards. All collaboratives met the Early Learning Collaborative Act of 2013 requirements in their curricula and assessment selections. With 2014-15 being the first full year of ELCA implementation, administrators and teachers have been working toward student mastery of the Early Learning Standards through the selected curriculum.

The National Association of the Education of Young Children describes indicators of effective curriculum likely to promote positive outcomes for all young children. These indicators are: children are active and engaged; goals are clear and shared by all; curriculum is evidenced based; valued content is learned through investigation, play, and focused, intentional teaching; curriculum builds on prior learning and experiences; curriculum is comprehensive; professional standards validate the curriculum's subject-matter content; and the curriculum is likely to benefit children (NAEYC & NAECS/SDE (2003).

PEER referenced the absence of Creative Curriculum in the Results First Clearinghouse Database. OWL, which is identified in the PEER critique as "the curriculum used by classes where the students performed better," was also absent in this Database. Collaboratives met the Request for Proposal's criteria for selection of both of the curricula.

Gilmore Early Learning Collaborative and Clarke County Early Learning Collaborative use the OWL Curriculum as well as Hatch, which is a technology-based program focused on literacy and numeracy. Hatch is used in the classroom on a daily basis and evaluates the students through computer games and creates reports for the teacher on the accomplishments of the students and their learning. Because of this natural relationship of students using the Hatch program, they could be more accustomed to using a computer. Therefore, the students could be more prepared to succeed on a computer-based assessment than those students who do not use Hatch on a daily basis.

Section III

PEER Critique: PEER found that pre-Kindergarteners in the collaboratives achieved the end-of-the-year target score or higher on the Kindergarten Readiness Assessment less than those enrolled in other public pre-Kindergarten programs. Students participating in two collaboratives that used the Owl curriculum achieved the end-of-the-year target score or higher on the Kindergarten Readiness Assessment significantly more often than students taught under the Creative Curriculum (which was determined by multiple, rigorous controlled studies to have no discernable effect on student learning).

MDE Response:

The range of target scores takes into account the variation of the collaboratives with different sites as each collaborative's composition of classrooms includes school district, and/or Head Start, and/or a childcare center. Childcare and Head Start centers have migrated to following testing protocols of school districts. For most collaborative classrooms, curriculum and assessment selections prior to fall 2014 were not governed by MDE. As collaboratives began serving children in fall 2014, this was the first interaction of coordination between partners following grant requirements (curriculum sharing, professional learning communities, conducting screeners and assessments, resource sharing, etc.). All collaboratives had previous agreements working towards an early childhood focused community plan to increase awareness of early childhood education. Many administrators and educators in the collaboratives spent the first portion of the 2014-15 school year blending curricula and formative assessments in instruction as well as ensuring the following of their licensing agency (e.g., childcare, Head Start, MDE) guidelines AND the ELCA.

Title I Pre-K classrooms have been established for many years in districts and for the most part have been in the practice of blending curricula and assessment in instruction for quite some time, ensuring the following of the MDE guidelines only. Collaborative partners (including childcare and Head Start centers) are administering MKAS^2 testing and screening tools for the first time as a collective group. The 2014-15 school year was the first year the MKAS^2 was formally administered. Through this administration, there will be data points that will help form literacy instruction in the classroom for pre-K and Kindergarten classrooms. Keep in mind that these scores are based on the performance of a four-year-old child in a 20-minute window at the end of the school year. The assessment scores were designed in the Early Learning Act language as a way to assess readiness for Kindergarten. After the summative assessment in the fall, teachers use the information to influence the intervention provided through activities in the classroom to strengthen skills.

The expected end-of-year scale score for four-year-old students is 498. This scale score at the end of Pre-K means that the student has mastered 70 percent of the targeted early literacy skills needed and supports that the student is on track to meet the end of grade 3 reading proficiency expectations. A scale score value of 498 falls within the Late Emergent Reader Classification Level. The 2014-15 Kindergarten Readiness Assessment results can be used to determine if each student has met the expected performance target of 498 scale scored points. This is a proxy for "basic" mastery of early literacy and should be seen as the minimal score needed to enter Kindergarten.

The scores that were set as the target score for pre-K and Kindergarten were identified as a score on the projection of growth where the child would be a successful reader by the third grade. The score indicated for the third grade target is 926 and is in the progression line with the percentages set for pre-K and K target scores.

Section IV

Technical Issues

MDE Response:

The method outlined in page 2 of the PEER report immediately suggests the report's title has no relationship with the actual method used to draw those conclusions reported on pages 29-31 of the report. The PEER "evaluators" did not conduct a mixed-method design (Tashakkori & Teddlie, 2003) necessary to collect independent, replicable data on the operations of the collaborative. The Joint Committee on Standards for Educational Evaluation (JCSEE) (Yarbrough, et al., 2011) clearly states within the Accuracy Standards, "the procedures of the evaluations should be monitored and described in enough detail, so that they can be identified and assessed." No information was provided on how the evaluators planned, conducted, and analyzed data associated with the implementation of different early learning curriculum. The evaluators conducted a desktop review of literature in lieu of designing a scientifically-based approach sufficient to support their comparative analysis.

The general finding that variance in performance outcomes was observed within the collaboratives is supported by the outcome data. The Department of Education, using unattentuated scale scores, observed a range of performance when the unit of analysis was fixed at the collaborative level. This variance suggests, in the aggregate, that eight (72.7%) of the eleven grant recipients met or exceeded the benchmark established by the MDE. Thus, the statement in the conclusion on page 30 can be corroborated in the outcome evidence. Unfortunately, as the PEER conclusions clearly violate JCSEE standards U1 Evaluator Credibility, U6 Meaningful Processes and Products, P4 Clarity and Fairness, A1 Justified Conclusions and Decisions, A6 Sound Designs and Analysis, A8 Communication and Reporting, and EA2 Internal Metaevaluation, any inferences about the Early Learning Collaborative effectiveness articulated in this report should be disregarded.

The below table enumerates the aforementioned statements:

JCSEE Standard	**Technical Violation**
U1 Evaluator Credibility	The evaluators fail to have the professional qualifications in the area of educational program evaluation, publications in scholarly journals associated with evaluating early childhood programs, utilization of subject matter experts in early childhood, and evaluating the psychometric characteristics of computer-adapted testing necessary to meet this standard.

JCSEE Standard	Technical Violation
U6 Meaningful Processes and Products	The evaluators failed to utilize socio-economic data of the actual students but rather relied on proxy data. No data was collected on the operational fidelity of the curricula implement by the subgrantees. No information was analyzed on the psychometrics of the assessment, the key dependent variable cited in the report. No comprehensive literature review demonstrating the scope and selection of the evaluation is supported within the larger body of research conducted throughout the world on early childhood programs. No theoretical framework or theory of action was used to support the evaluation's design. No data assumptions and limitations were comprehensively identified by the evaluators. No technical testing was conducted to support the analytical model used by the evaluators.
P4 Clarity and Fairness	The evaluators' scope is too narrow to support the conclusions; alternative explanations to the cited differences in performances are not stated.
A1 Justified Conclusions and Decisions	The evaluators' conclusion about the comparative group is unsupported due to failure by the evaluators to ensure equivalency of known exogenous factors impacting measures of latent constructs (achievement); conclusions associated with curriculum effectiveness are confounded by construct irrelevant characteristics both omitted and exacerbated by the use of proxy data that does not fit the actual dataset. The evaluators' use of a "four year longitudinal study" reported by the MDE is inaccurately cited.
A6 Sound Designs and Analysis	The evaluators' dismissal of the assessment's psychometric evidence is not based upon documented research; the evaluators' attenuation of interval level data distorts the results on effectiveness; the method used to develop a "pseudo" control group for comparative purposes did not establish equivalence of the groups on key independent factors; the evaluators failed to examine the dependent variable, which supports the use of more sophisticated and robust parametric analytics [reducing Type 1 errors]; the evaluators failed to use econometrics in making return-on-investment (ROI) statements.
A8 Communication and Reporting	The evaluators' purpose fails to align with the conclusions; the evaluators failed to enumerate the research question and subordinate questions prior to the beginning of the activity; the evaluators refused to accept MDE assistance in reviewing the methodology prior to the evaluation; the evaluators refused to disclose the statistical software planned for use during the evaluation; the evaluators did not identify the SMEs to be utilized; the evaluators did not request permission to conduct qualitative reviews; the evaluators' representative made no claim of an intent to conduct a fair, impartial evaluation.

JCSEE Standard	Technical Violation
EA2 Internal Metaevaluation	The evaluators failed to conduct sufficient review of the literature to support their efforts and findings.

References

NAEYC & NAECS/SDE (2003). *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8. Joint position statement.* Retrieved from www.naeyc.org/cape.

Tashakkori, A. & Teddlie, C. (2003). *Handbook of Mixed Methods in Social & Behavioral Research.* Thousand Oaks: Sage.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Conclusion

This PEER report is the second PEER Committee report produced in 2015 that has not followed the Joint Committee on Standards for Educational Evaluation, which are the national standards for evaluating educational programs. Therefore, any conclusions or recommendations contained in this PEER report lack merit.

Sincerely,

Carey M. Wright, Ed.D.
State Superintendent of Education

PEER's Response to the Department of Education's Response

The Mississippi Department of Education submitted a response to the PEER Committee's report entitled *The Early Learning Collaborative Act of 2013: Evaluation of the Operations and Effectiveness of the Program*. PEER strongly disagrees with many of the statements made in MDE's response and the Committee has chosen to release this response to MDE's response.

In the department's response, MDE claims that because PEER did not follow standards promulgated by the Joint Committee on Standards for Educational Evaluation (JCSEE), PEER's report conclusions lack merit. Following the JCSEE standards is neither necessary to nor sufficient for a rigorous evaluation. The Committee would note the following:

- PEER is a joint committee of the Legislature. State law directed PEER to conduct an independent review of the program operation and effectiveness of the Early Learning Collaborative Act of 2013. The Committee has the power to conduct evaluations of all state or local entities in Mississippi that receive public funds "in any manner and at any time deemed appropriate." (See MISS. CODE ANN. Section 5-3-51 et seq.)

- PEER was modeled on the federal government's General Accounting Office (now the Government Accountability Office [GAO]), which was designed to support Congress in helping to improve the performance and ensure the accountability of the federal government. As does GAO for Congress, PEER serves as Mississippi's legislative oversight committee and provides objective, timely information for decisionmakers. PEER's evaluations are designed to serve the needs of the Legislature, not the agency reviewed.

- Persons familiar with the profession of legislative auditing know that the work involves evaluating a variety of types of programs. Because many types of standards exist in many professional fields, legislative auditing staff must be proficient in applying the basics of disciplined inquiry. To evaluate the operations and effectiveness of the prekindergarten program, PEER's staff need not be experts in the standards for educational evaluation; rather, PEER's staff must have the skills necessary to conduct an objective review.

MDE's response specifically addressed concerns related to minimum rate of readiness. PEER notes the following regarding this issue:

- MDE does not agree that it should have set a minimum rate of readiness in the early phases of program planning. PEER contends that the minimum rate of readiness should have been set in program planning, as it would have been a more objective assessment of performance expectations. Setting the rate after test results are available increases the opportunity to use the test results to set the rate, regardless of the actual performance expectations.

- MDE does not agree that adding growth into the readiness rate could introduce leniency. However, it is a mathematical fact that such a method, as described by MDE, will either have no effect or will be more lenient because the growth standard will only be applied to those collaboratives that fail to meet the minimum rate standard. Thus, if a collaborative fails to meet the minimum rate

of readiness, the collaborative has a chance to qualify for funding after growth is applied. There is no scenario in which the minimum rate of readiness would be met initially and then not met after growth was applied.

Regarding MDE's concerns about "technical issues," PEER is confident that its findings would stand up in the court of expert judgment without additional defense.

In conclusion, PEER stands by this report and believes that it does not contain unsupported conclusions. PEER believes that this report is a more detailed and controlled analysis of the Early Learning Collaborative Act of 2013 than MDE's own assessment of the program.

PEER Committee Staff

James A. Barber, Executive Director

Legal and Reapportionment
Ted Booth, General Counsel
Ben Collins
Barton Norfleet

Administration
Barbara Hamilton
Larry Landrum
Rosana Slawson
Gale Taylor

Quality Assurance and Reporting
Ava Welborn, Director
Tracy Bobo

Performance Evaluation
Lonnie Edgar, Team Coordinator
David Pray, Team Coordinator
Jennifer Sebren, Team Coordinator
Kim Cummins
Matthew Dry
Matthew Holmes
Michael Surratt
Jenell Ward
Jade Watters
Sarah Williamson
Julie Winkeljohn
Ray Wright

Performance Accountability
Linda Triplett, Director
Kirby Arinder
Alicia Davis
Anna Johnson
Jessica Kelly
MeriClare Steelman