

## **Title Page**

**Title:** Updates on the Effects of the Tennessee Prekindergarten Program on Children's Achievement and Behavior through Sixth Grade

**Authors:** Kelley Durkin, Mark Lipsey, and Dale Farran

# **Updates on the Effects of the Tennessee Prekindergarten Program on Children's Achievement and Behavior through Sixth Grade**

## **Background/Context:**

State-funded prekindergarten has expanded rapidly in the U.S., and there are expectations that such programs will have short-term benefits on school readiness and longer-term positive effects on later academic performance and behavioral outcomes (e.g., Phillips et al., 2017). There is strong evidence of pre-k effects on measures of kindergarten readiness (e.g., Gormley et al., 2005; Weiland & Yoshikawa, 2013); however, there is not strong evidence about longer-term effects, specifically for state-funded pre-k. This follow-up study through 6<sup>th</sup> grade of a large randomized control trial of the Tennessee Voluntary Pre-K Program (TN-VPK) continued to investigate the long-term effects of the program on children's achievement and behavior. To date, the findings of the study have shown significant, positive immediate effects of TN-VPK on academic outcomes, but null or negative effects on long-term academic and behavioral outcomes through 3<sup>rd</sup> grade (Lipsey et al., 2018). This is the first well-controlled experimental study of a state pre-k program with follow-up into middle school.

## **Purpose/Objective/Research Question:**

We investigated whether participation in TN-VPK had significant effects on achievement, grade retention, absenteeism, disciplinary infractions, and special education placement through 6<sup>th</sup> grade.

## **Setting:**

TN-VPK is a statewide voluntary pre-k program that prioritizes children eligible for the federal free or reduced price lunch programs and serves over 18,000 children with programs in all but a few of the school districts in the state.

## **Population/Participants/Subjects:**

The current study follows 2990 children randomly assigned to participate or not in TN-VPK in 2009 or 2010. All children were from low-income households; about 49% were male, 49% were white, 27% were Black, 23% were Hispanic, and 24% were non-native English speakers. Their mean age when they entered pre-k was 53.3 months.

## **Intervention/Program/Practice:**

TN-VPK is a typical pre-k program and requires a licensed teacher and an aide in each classroom, a maximum class size of 20, a state-approved curriculum, and a minimum instructional time of 5.5 hours per day during the school year. When the program began, it met 9 of the 10 standards advocated until recently revised by the National Institute of Early Education Research (Barnett et al., 2009).

## **Research Design:**

The current study involved 79 over-subscribed schools in which cohorts of pre-k applicants were randomized to offers of admission or waitlist status during school years 2009-10 and 2010-11. This resulted in 111 separate school-level randomizations and a total sample of

2990 children with 1852 in the intent-to-treat (ITT) group offered admission and 1138 in the control group of waitlisted children not ultimately offered admission. Of those assigned to VPK, 13.2% did not actually attend (no shows), and 34.2% assigned to the control group managed to attend VPK somewhere (crossovers). The treatment-on-treated (TOT) sample consists of 1997 VPK participants and 993 nonparticipants. Tests for baseline equivalence on the demographic variables showed no significant differences between the ITT treatment and control groups (Lipsey et al., 2018).

### **Data Collection and Analysis:**

Data drawn from the state database for the sample for each year through the 6<sup>th</sup> grade year included outcome variables for achievement, attendance, retention in grade, disciplinary actions, and special education designations. Scores on the state achievement tests were not available for the sample in 4<sup>th</sup> and 5<sup>th</sup> grades. Tennessee had a major breakdown in its state testing program that resulted in a loss of 4<sup>th</sup> grade state achievement test scores for Cohort 2 (began pre-k in 2010) and 5<sup>th</sup> grade scores for Cohort 1 (began pre-k in 2009).

Hierarchical linear models with children nested within randomized lists nested within districts were used to test the effects of TN-VPK, and demographic covariates were included in all models. ITT condition was used as a predictor in the primary analyses, and the results of each ITT analysis were then used to estimate TOT effects using two-stage least-squares instrumental variable regression models. Multiple imputation was used for missing data; analyses using observed data produced similar patterns of findings.

### **Findings/Results:**

TN-VPK continued to have a negative effect on children's achievement scores in Math and Science in 6<sup>th</sup> grade and a significant negative effect emerged for English Language Arts (Table 1). There were no significant effects on attendance across years (Table 2). There were also no significant effects on cumulative retention in grade through 6<sup>th</sup> grade after the initial effect in kindergarten (Table 3).

School disciplinary actions (suspensions and expulsions) were coded as minor offenses of violating school rules or more serious, major offenses (e.g., fighting, bullying, bringing a weapon to school). The frequency of offenses was low so the outcome variable used aggregated across the K-6<sup>th</sup> grade years (yes/no for any recorded actions during that period). VPK participants had marginally more violations of school rules ( $p = .10$ ) by 6<sup>th</sup> grade (Table 4). There were no significant differences on major offenses or all offenses, though VPK participants had more offenses through 6<sup>th</sup> grade than nonparticipants.

Special education placement was coded for whether or not there was an Individualized Education Program (IEP) other than for intellectually gifted or physical disability. In every year K to 6<sup>th</sup> grade, more VPK participants had IEPs than nonparticipants (Table 5), a pattern that began when VPK participants were identified for special education services in pre-k.

### **Conclusions:**

The current findings continue the pattern of null or negative long-term effects for the TN-VPK program. This first and, to date, only randomized study of a state pre-k program

provides a cautionary tale about what should be expected from such programs, especially regarding long-term effects. To understand these long-term effects, we are investigating subgroups of children. For instance, children living in neighborhoods with the highest concentration of poverty showed positive effects of TN-VPK on 3<sup>rd</sup> grade achievement scores (Pearman, in press). Additionally, we are investigating whether experience with high quality schools and teachers after pre-k provides sustaining environments for pre-k effects (Pearman et al., 2019). VPK participants also have earlier identification of special education needs, and future work is exploring the longer-term effects of this early identification.

## References

- Barnett, S., Epstein, D. J., Friedman, A. H., Sansanelli, R. A., & Hustedt, J. T. (2009). *The state of preschool 2009: State preschool yearbook (Tennessee pp. 132-133)*. The National Institute for Early Education Research (NIEER). Rutgers University <http://nieer.org/wp-content/uploads/2016/10/TN.pdf>
- Gormley, W. T., Jr., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-k on cognitive development. *Developmental Psychology, 41*(6), 872-884. doi: 10.1037/0012-1649.41.6.872
- Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee prekindergarten program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly, 45*, 155-176. <https://doi.org/10.1016/j.ecresq.2018.03.005>
- Pearman, F. A. (in press). The moderating effect of neighborhood poverty on preschool effectiveness: Evidence from the Tennessee Voluntary Prekindergarten Experiment. *American Educational Research Journal*.
- Pearman, F. A., Springer, M., Lipsey, M., Lachowicz, M., Farran, D., & Swain, W. (2019). Teachers, schools, and pre-k effect persistence: An examination of the sustaining environment hypothesis. (EdWorkingPaper: 19-85). Retrieved from Annenberg Institute at Brown University: <http://www.edworkingpapers.com/ai19-85>
- Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., Duncan, G. J., Dynarski, M., Magnuson, K. A., & Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects. A consensus statement*. Washington, DC: The Brookings Institution.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84*(6), 2112-2130. <https://doi.org/10.1111/cdev.12099>

Table 1  
*Intent-to-treat (ITT) and treatment-on-treated (TOT) impact estimates for 6<sup>th</sup> grade state achievement tests*

	ITT			Coefficient <sup>c</sup> for T-C difference	Effect Size <sup>d</sup>	p-value <sup>e</sup>	TOT			Effect size <sup>d</sup>
	Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Pooled SD <sup>b</sup>				Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Coefficient <sup>c</sup> for T-C difference	
English	319.9	323.2	30.61	<b>-3.32*</b>	<b>-.109</b>	<b>.007</b>	318.4	324.7	<b>-6.31*</b>	<b>-.206</b>
Mathematics	317.1	322.7	37.13	<b>-5.55*</b>	<b>-.149</b>	<b>.000</b>	314.6	325.2	<b>-10.54*</b>	<b>-.284</b>
Science	749.5	753.4	40.26	<b>-3.87*</b>	<b>-.096</b>	<b>.022</b>	747.7	755.1	<b>-7.36*</b>	<b>-.182</b>
	<i>N</i> = 1852	<i>N</i> = 1138		<i>N</i> = 2990			<i>N</i> = 1997	<i>N</i> = 993	<i>N</i> = 2990	

\* $p < .05$ , † $p < .10$  for coefficients; significant coefficients and related estimates are also bolded.

Notes: State achievement scaled scores. Only students who had all 3 test scores are included in these analyses. Students below, at, or above expected grade levels are all included in these analyses.

<sup>a</sup> Covariate-adjusted means generated by the multilevel analysis models with covariates set at the grand means for the sample.

<sup>b</sup> Pooled treatment and control group standard deviations. There are minor variations between the pooled SDs for ITT and TOT; the mean is presented here but effect sizes are computed on the exact values.

<sup>c</sup> Coefficients for treatment-control differences from OLS multilevel models with children nested in R-Lists and R-Lists nested in districts. Covariates are age, male, Black, Hispanic, and non-native English.

<sup>d</sup> Effect size: coefficient for the treatment-control difference divided by the pooled standard deviation.

<sup>e</sup> The 2SLS analysis model yields p-values for statistical significance that are the same for the ITT and TOT coefficients.

Table 2

*Intent-to-treat (ITT) and treatment-on-treated (TOT) impact estimates for attendance across years*

	ITT			Coefficient <sup>c</sup> for T-C difference	Effect Size <sup>d</sup>	<i>p</i> -value <sup>e</sup>	TOT			Effect size <sup>d</sup>
	Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Pooled SD <sup>b</sup>				Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Coefficient <sup>c</sup> for T-C difference	
K	.947	.949	.042	-.002	-.051	.192	.946	.950	-.004	-.096
1 <sup>st</sup> grade	.954	.955	.038	.000	-.010	.792	.954	.955	-.001	-.019
2 <sup>nd</sup> grade	.958	.959	.035	-.002	-.045	.259	.957	.960	-.003	-.085
3 <sup>rd</sup> grade	.961	.963	.042	-.001	-.028	.509	.960	.964	-.002	-.053
4 <sup>th</sup> grade	.975	.975	.037	-.001	-.016	.734	.975	.975	-.001	-.030
5 <sup>th</sup> grade	.972	.971	.031	.001	.028	.540	.972	.971	.002	.053
6 <sup>th</sup> grade	.972	.974	.029	-.002	-.068	.120	.971	.975	-.004	-.129
	<i>N</i> = 1852	<i>N</i> = 1138		<i>N</i> = 2990			<i>N</i> = 1997	<i>N</i> = 993	<i>N</i> = 2990	

\* $p < .05$ , <sup>†</sup> $p < .10$  for coefficients; significant coefficients and related estimates are also bolded.

<sup>a</sup> Covariate-adjusted means generated by the multilevel analysis models with covariates set at the grand means for the sample.

<sup>b</sup> Pooled treatment and control group standard deviations. There are minor variations between the pooled SDs for the ITT and TOT; the mean is presented here but effect sizes are computed on the exact values.

<sup>c</sup> Coefficients for the treatment-control differences from OLS multilevel multiple imputation models with children nested in R-Lists and R-Lists nested in districts. Covariates are age, male, Black, Hispanic, and non-native English.

<sup>d</sup> Effect size: coefficient for the treatment-control difference divided by the pooled standard deviation.

<sup>e</sup> The 2SLS analysis model yields *p*-values for statistical significance that are the same for the ITT and TOT coefficients

Table 3

*Intent-to-treat (ITT) and treatment-on-treated (TOT) impact estimates for cumulative retention across years*

	ITT			TOT						
	Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Pooled SD <sup>b</sup>	Coefficient <sup>c</sup> for T-C difference	Effect Size <sup>d</sup>	<i>p</i> -value <sup>e</sup>	Treatment group mean	Control group mean	Coefficient <sup>c</sup> for T-C difference	Effect size <sup>d</sup>
K	.052	.067	.227	<b>-.016<sup>†</sup></b>	<b>-.068</b>	<b>.080</b>	.045	.074	<b>-.029<sup>†</sup></b>	<b>-.130</b>
K-1 <sup>st</sup>	.110	.101	.298	.009	.029	.457	.114	.097	.016	.055
K-2 <sup>nd</sup>	.125	.122	.317	.003	.009	.821	.126	.121	.005	.017
K-3 <sup>rd</sup>	.132	.129	.328	.003	.009	.816	.133	.128	.006	.017
K-4 <sup>th</sup>	.139	.132	.333	.007	.022	.587	.142	.129	.014	.041
K-5 <sup>th</sup>	.144	.133	.337	.011	.032	.423	.149	.128	.020	.060
	<i>N</i> = 1852	<i>N</i> = 1138		<i>N</i> = 2990			<i>N</i> = 1997	<i>N</i> = 993	<i>N</i> = 2990	

\*  $p < .05$ , <sup>†</sup>  $p < .10$  for coefficients; significant coefficients and related estimates are also bolded.

Notes: Cumulative values across the grades indicated; e.g., K-2<sup>nd</sup> refers to retention in any grade from K through 2nd.

<sup>a</sup> Covariate-adjusted means generated by the multilevel analysis models with covariates set at the grand means for the sample.

<sup>b</sup> Pooled treatment and control group standard deviations. There are minor variations between the pooled SDs for the ITT and TOT; the mean is presented here but effect sizes are computed on the exact values.

<sup>c</sup> Coefficients for the treatment-control differences from OLS multilevel models with children nested in R-Lists and R-Lists nested in districts. Covariates are age, male, Black, Hispanic, and non-native English.

<sup>d</sup> Effect size: coefficient for the treatment-control difference divided by the pooled standard deviation.

<sup>e</sup> The 2SLS analysis model yields *p*-values for statistical significance that are the same for the ITT and TOT coefficients.



Table 4

*Intent-to-treat (ITT) and treatment-on-treated (TOT) impact estimates for cumulative disciplinary offenses through sixth grade*

	ITT			Coefficient for T-C difference <sup>c</sup>	Effect Size <sup>d</sup>	<i>p</i> -value <sup>e</sup>	TOT			Effect size <sup>d</sup>
	Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Pooled SD <sup>b</sup>				Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Coefficient for T-C difference <sup>c</sup>	
School rules	.235	.207	.413	.028	.069	.103	.248	.194	.054	.131
Major offenses	.137	.125	.346	.012	.082	.450	.142	.120	.022	.061
All offenses	.279	.263	.445	.016	.035	.391	.286	.256	.030	.066
	<i>N</i> = 1852	<i>N</i> = 1138		<i>N</i> = 2990			<i>N</i> = 1997	<i>N</i> = 993	<i>N</i> = 2990	

\* $p < .05$ , <sup>†</sup> $p < .10$  for coefficients; significant coefficients and related estimates are also bolded.

Notes: School rules: violations of school rules or other administrative issues; major offenses: fighting, bullying, weapon in school, and the like; all offenses: total across school rule and major offenses categories. These are coded for whether there is any infraction recorded in school records (1 = yes, 0 = no) cumulatively from K through the 3rd grade year.

<sup>a</sup> Covariate-adjusted means generated by the multilevel analysis models with covariates set at the grand means for the sample.

<sup>b</sup> Pooled treatment and control group standard deviations. There were minor variations between the pooled SDs for the ITT and TOT; the mean is presented here but effect sizes were computed on the exact values.

<sup>c</sup> Coefficients for the treatment-control differences from OLS multilevel models with children nested in R-Lists and R-Lists nested in districts. Covariates are age, male, Black, Hispanic, and non-native English.

<sup>d</sup> Effect size: coefficient for the treatment-control difference divided by the pooled standard deviation.

<sup>e</sup> The 2SLS analysis model yields *p*-values for statistical significance that are the same for the ITT and TOT coefficients.

Table 5

*Intent-to-treat (ITT) and treatment-on-treated (TOT) impact estimates for IEPs not including gifted or physical disability*

	ITT			Coefficient <sup>c</sup> for T-C difference	Effect size <sup>d</sup>	p-value <sup>e</sup>	TOT			Effect size <sup>d</sup>
	Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Pooled SD <sup>b</sup>				Treatment group mean <sup>a</sup>	Control group mean <sup>a</sup>	Coefficient <sup>c</sup> for T-C difference	
K	.129	.095	.304	<b>.034*</b>	<b>.110</b>	<b>.004</b>	.144	.080	<b>.064*</b>	<b>.211</b>
1 <sup>st</sup> grade	.138	.107	.320	<b>.031*</b>	<b>.097</b>	<b>.012</b>	.152	.093	<b>.059*</b>	<b>.184</b>
2 <sup>nd</sup> grade	.138	.116	.329	<b>.022†</b>	<b>.067</b>	<b>.088</b>	.148	.106	<b>.042†</b>	<b>.128</b>
3 <sup>rd</sup> grade	.134	.108	.327	<b>.026*</b>	<b>.078</b>	<b>.049</b>	.146	.096	<b>.049*</b>	<b>.149</b>
4 <sup>th</sup> grade	.125	.102	.317	<b>.023†</b>	<b>.073</b>	<b>.069</b>	.135	.092	<b>.044†</b>	<b>.138</b>
5 <sup>th</sup> grade	.120	.095	.316	<b>.025*</b>	<b>.081</b>	<b>.045</b>	.132	.083	<b>.048*</b>	<b>.154</b>
6 <sup>th</sup> grade	.114	.081	.305	<b>.033*</b>	<b>.107</b>	<b>.009</b>	.129	.066	<b>.062*</b>	<b>.203</b>
	<i>N</i> = 1852	<i>N</i> = 1138		<i>N</i> = 2990			<i>N</i> = 1997	<i>N</i> = 993	<i>N</i> = 2990	

\* $p < .05$ , †  $p < .10$  for coefficients; significant coefficients and related estimates are also bolded.

Note: IEP = Individualized Educational Program as the formal special education designation.

<sup>a</sup> Covariate-adjusted means generated by the multilevel analysis models with covariates set at the grand means for the sample.

<sup>b</sup> Pooled treatment and control group standard deviations. There are minor variations between the pooled SDs for the ITT and TOT; the mean is presented here but effect sizes are computed on the exact values.

<sup>c</sup> Coefficients for the treatment-control differences from OLS multilevel models with children nested in R-Lists and R-Lists nested in districts. Covariates are age, male, Black, Hispanic, and non-native English.

<sup>d</sup> Effect size: coefficient for the treatment-control difference divided by the pooled standard deviation.

<sup>e</sup> The 2SLS analysis model yields p-values for the statistical significance that are the same for the ITT and TOT coefficients.