

State-Funded Pre-K and Children's Language and Literacy Development: The Case of COVID-19

Elizabeth Burke Hadley¹ , Siyu Liu¹, Eunsook Kim¹ , and Meaghan McKenna² 

Research on how COVID-19 school closures impacted the learning of young children is still sparse, and the broader implications of these findings have yet to be fully explored. In this study, we examine the impacts of COVID-19 school closures on language and literacy development in a state-funded pre-K program in a Florida school district. Using COVID closures as a natural experiment, we additionally explore the implications of these findings for the literature on the efficacy and dosage of state-funded pre-K programs. We used propensity score matching to compare children from the COVID cohort ($n = 1,211$) to children from a pre-COVID cohort ($n = 1,167$). Results revealed no significant difference on the Florida Kindergarten Readiness Screener (FLKRS) assessment between the two cohorts at kindergarten entry. The COVID cohort had significantly higher scores on each of the individual subdomains of the FLKRS than the pre-COVID cohort, with especially large positive effect sizes for constrained literacy skills. Results indicate that COVID-19 closures did not have significant negative impacts on pre-K children's language and literacy skills at kindergarten entry. Findings further suggest that state-funded pre-K programs with comparatively less funding and lower quality ratings may not provide adequate support for children's language and literacy outcomes.

Keywords: child development; early childhood; literacy; quasi-experimental analysis; regression analyses

In the spring of 2020, COVID-19 arrived in the United States, closing school buildings and disrupting the education of millions of students. Teachers rapidly pivoted to remote instruction with little time to prepare and meager resources. While researchers projected major impacts on educational outcomes for all children, the effects of closures on young children attending public pre-K programs were of particular concern. Public pre-K programs support a large percentage of 4-year-olds nationwide, with many states prioritizing enrollment for children from low-income families or those with disabilities. Public pre-K is intended to improve "school readiness," with a special focus on narrowing opportunity gaps. However, implementing pre-K remotely during school closures was challenging, with teachers reporting difficulties with student attendance and engagement (Bassok et al., 2021). School closures therefore may have had an especially negative impact on our youngest learners, but little research currently exists that examines the nature and extent of the academic impacts of COVID on pre-K learners.

Questions about what happens when state-funded pre-K classrooms close also have broader implications. In recent years,

debates have swirled over the efficacy of state-funded pre-K programs, including concerns that the impacts of large-scale programs are not as substantial as early studies conducted in smaller, higher quality pre-K settings (Meloy et al., 2019). There is also growing evidence that the positive impacts of state-funded pre-K "fade out" over time, with initial academic benefits for pre-K attenders disappearing or even reversing in elementary school (e.g., Durkin et al., 2022). In the present article, we use COVID-19 school closures as a natural experiment, comparing academic outcomes for a typical pre-K year to one in which children missed a quarter of the school year due to COVID-19, to further advance the literature on the dosage and efficacy of state-funded pre-K programs.

We address the following research questions:

Research Question 1: Were pre-K children's language and literacy outcomes at kindergarten (K) entry impacted by COVID-related school closures in spring 2020? Did

¹University of South Florida, Tampa, FL

²University of Connecticut, Storrs, CT

children who spent a full year in state-funded pre-K have better language and literacy scores than children who missed a quarter of the pre-K year?

Research Question 2: Did COVID-related impacts vary across different language and literacy skills?

Research Question 3: Were COVID-related impacts moderated by child demographic characteristics (i.e., race/ethnicity, socioeconomic status, and disability status)?

Background

Impacts of COVID on Young Children's Academic Outcomes

Research on the academic impacts of COVID-19 on young children in Fall 2020 is sparse.¹ While available reports have been crucially important for guiding decision-making, many of these analyses are not peer reviewed and have significant methodological weaknesses (Ho, 2021). Additionally, nearly all studies focus on older children, drawing on national databases of standardized tests that are unavailable for pre-K learners. Considered as a whole, these studies reveal that the academic impacts of COVID were somewhat smaller than initially projected, with an early meta-analysis finding that students' test scores were 0.15 *SD* lower across grades and subject areas than in previous years (Storey & Zhang, 2021). Student achievement in math declined more than in reading (Kuhfeld et al., 2022), with relatively small negative impacts on reading skills (e.g., 1.5% more students placing below grade level in reading; Curriculum Associates, 2020). In fact, some studies found that students' reading achievement was not negatively impacted by school closures in fall 2020 in the youngest grades (e.g., Renaissance Learning, 2020). High-poverty schools and those with larger percentages of minoritized students experienced the largest declines in reading (Curriculum Associates, 2020; Patarapichayatham et al., 2021).

To date, there are only a small handful of studies that specifically examine COVID impacts for pre-K children. One analysis found that 9.3% more Virginia kindergarteners were identified as at risk for reading difficulties in fall 2020 compared to fall 2019 (McGinty et al., 2021). Another report from Texas found that learning trajectories were flat from March 2020 to fall 2020 for rising kindergarteners (Patarapichayatham et al., 2021).

Studies conducted with fall 2020 data come with several important caveats and cautions. The population of students tested was often not comparable to prior years, with low participation rates overall and more missing data for students who were also more at risk of negative impacts due to COVID (e.g., students of color; Kuhfeld et al., 2022). Many students also took assessments remotely, and the validity of these data has been called into question, particularly for K–2 learners who likely received assistance from caregivers (Kilbride et al., 2022; Kuhfeld et al., 2020). One nationwide study found a clear “at-home advantage” for remote testing in fall 2020: Students testing in school scored two or more grade levels below their current grade, whereas students testing remotely actually outperformed pre-pandemic averages (Huff, 2020).

The present study addresses many of the limitations of prior work. Our test participation rate was high: 87% of the children

who attended Florida Voluntary Pre-K (FL-VPK) in our partner school district in 2019–2020 also took the kindergarten screening assessment in fall 2020 (compared to a 94% participation rate for a pre-COVID cohort). Missing data in our sample do not appear to disproportionately represent at-risk children, and the missing patterns are similar to pre-COVID data. Unlike previous descriptive studies, our study design and analytic approach allow for causal inferences: We compare a pre-COVID cohort to the COVID cohort using propensity scores to match children between cohorts and adjust for selection bias, thereby mimicking an experimental design. Finally, no remote testing was permitted in our sample, and therefore our data are more comparable to pre-COVID data and not subject to many of the biases found in remote testing data.

Impacts of State-Funded Pre-K on Children's Language and Literacy Development

Investigating COVID-related impacts on pre-K learners also has implications for research on the efficacy of public pre-K programs. Public pre-K is typically, but not exclusively, funded by individual states, with wide variations nationwide in both the funding and the quality of programs (Friedman-Krauss et al., 2020). The present study uses data from Florida, which offers half-day pre-K to all 4-year-olds through their FL-VPK program, which operates in both public schools and community-based settings. The FL-VPK program is notable for its wide reach, with free pre-K available to all 4-year-old Florida children, and an actual participation rate of 72% (Friedman-Krauss et al., 2020). However, the FL-VPK program also meets fewer quality standards (e.g., professional development, curriculum supports) than nearly any other state-funded pre-K program (Friedman-Krauss et al., 2020). Research on FL-VPK's long-term efficacy found that attending VPK had no impact on the likelihood of children completing third grade on time (Miller & Bassok, 2019). Other rigorous evidence on the shorter term academic impacts of FL-VPK is scant (e.g., Meloy et al., 2019). Therefore, it remains an open question whether an investment in universal access to pre-K such as Florida's leads to better academic outcomes for children at kindergarten entry if quality safeguards are not also put into place.

Studies that have examined the impact of other state pre-K programs on language and literacy have nearly all found positive impacts in the short term, with children attending pre-K outperforming nonattenders upon program completion (Barnett et al., 2018; Meloy et al., 2019). Larger positive impacts are more consistently found for code-based, constrained skills (e.g., letter identification; $d = 1.10$, Barnett et al., 2018) compared to language-based, unconstrained skills (e.g., vocabulary; $d = 0.24$, Barnett et al., 2018). However, the magnitude of effects varies considerably based on the quality of programs (van Huizen & Plantenga, 2018), with factors such as the comprehensiveness of instructional support associated with more positive outcomes (Johnson et al., 2016). The beneficial effects of public pre-K on academic skills also vary by subgroup, with larger impacts for Hispanic/Latino and Black children and children with disabilities (Huang et al., 2012).

The strongest design for evaluating program efficacy is a randomized control trial (RCT) in which children are randomly assigned to attend either the pre-K program being evaluated or a comparison group. However, RCTs are often not feasible, especially for universal pre-K programs like Florida's. The main body of pre-K evaluation research uses regression discontinuity designs based on age cutoffs, but this design has been critiqued as subject to multiple biases (Lipsey et al., 2015). In the present article, we use COVID school closures as a natural experiment to estimate the impact of FL-VPK on language and literacy. We use the pre-COVID cohort, who received a full year of pre-K, as our pre-K group and the COVID cohort, who missed a quarter of the pre-K year, as our comparison group. COVID school closures approximate a true experiment because the assignment of children to either group (i.e., a regular year or a year interrupted by COVID) was random (Dunning, 2012). School closures resulted from the exogenous factor of the COVID virus, and participants could not influence their "assignment" to attend pre-K during either a typical year or a COVID-interrupted year. Therefore, differences in outcomes between the two groups can be attributed to school closures. A careful comparison of pre- and post-COVID data, then, presents a unique opportunity to investigate the impacts of in-person FL-VPK. Based on the average effect size for state-funded pre-K programs for literacy ($d = 1.10$) and language ($d = 0.20$; Barnett et al., 2018), we hypothesized that the COVID cohort would show a reduction of at least $0.275 SD$ in literacy skills (one quarter of average effect sizes) and a small reduction in language effect sizes ($d = 0.05$) compared to the pre-COVID cohort.

Data and Method

The current article draws on data collected by a large school district ("Sunnyside") in west-central Florida with schools in both urban and suburban areas. Our project examines children who attended VPKs located in public elementary schools. Using VPK data from public elementary schools presents several advantages. Approximately 60% of all state-funded pre-K programs are situated within public elementary schools (Johnson et al., 2022), and as such, the public VPKs we examine here may be comparable to those programs with similar profiles (see Friedman-Krauss et al., 2020). Additionally, while states typically cannot connect pre-K data to elementary school data, VPKs located in public elementary schools assign children an ID number at pre-K entry that is used into elementary school. These longitudinal data allow us to more accurately estimate the impacts of COVID school closures. Finally, FL-VPKs situated in public schools are often regarded as higher quality than those situated in community settings, making our sample a "best-case scenario" for FL-VPK classrooms.

Sample

Our data set includes two cohorts of children from 89 schools. The pre-COVID cohort attended FL-VPK in the Sunnyside school district in 2016–2017 ($n = 1,211$), and the COVID cohort attended FL-VPK in Sunnyside in 2019–2020 ($n = 1,167$). All children who attended FL-VPK in Sunnyside schools

during these years and who had VPK assessment data were included in our sample (see Table 1).

Sunnyside schools were closed on March 16, 2020, due to COVID-19 and remained closed for the remainder of the school year. Children in the COVID cohort missed 47 of 180 instructional days, approximately 26% of the pre-K school year. Remote instruction began at the end of March, was conducted for the remainder of the school year, and included optional live online sessions with teachers, recorded lessons, and worksheets/learning resources sent home.

Pre-K and Kindergarten Literacy Skills

Voluntary Pre-K assessment. The state of Florida requires that voluntary pre-K (VPK) providers administer the VPK assessment, a progress monitoring measure for early academic skills, at the beginning (AP1), middle (AP2), and end (AP3) of the pre-K year. The VPK assessment is aligned with Florida early learning standards and measures skills in four domains: oral language, phonological awareness, print knowledge, and mathematics. Teachers administer the VPK assessment one-on-one with children. Because the present project focuses on language and literacy development, we use data only for each of the first three domains. Children are assigned a separate score for each domain, with 1 point given for each item answered correctly: Print Knowledge (0–12), Phonological Awareness (0–14), and Oral Language/Vocabulary (0–22). The VPK assessment has moderate to high internal consistency and test-retest reliability (for our study sample, McDonald's $\omega = .74$ for AP1, $.78$ for AP2; Cronbach's $\alpha = .71$ for AP1, $.77$ for AP2). Although these alpha values are within an acceptable range (i.e., over $.70$; Tavakol & Dennick, 2011), readers should be aware of the moderate reliability of the VPK assessment.

Florida Kindergarten Readiness Screener. All kindergarteners in Florida public schools are assessed for school readiness within the first 30 days of school, a process termed the Florida Kindergarten Readiness Screener (FLKRS). The Star Early Literacy assessment is currently used for this screening process (FLKRS/Star) and is an online, adaptive measure that children complete independently. Children receive a scaled score ranging from 300 to 900 based on the difficulty of items and their responses. A score of 500 or above indicates that children are academically "ready" for kindergarten. The FLKRS/Star consists of 10 subdomains, nine of which measure language and literacy skills (see Table 3) and one that measures early numeracy. Each of these subdomains contributes to the overall scaled score. We examine the nine language and literacy subdomains separately in our analyses for Research Question 2. The FLKRS/Star has high reliability (Cronbach's $\alpha = .96$, McDonald's $\omega = .98$ with the study sample).

All children in both the pre-COVID cohort and the COVID cohort took the FLKRS/Star assessment at school: No remote administration was permitted. Children in the COVID cohort were given the option to attend kindergarten either virtually or in person in fall 2020. Children who attended kindergarten in person ($n = 667$) took the FLKRS/Star as part of their typical school day. Children who attended kindergarten remotely ($n =$

500) were asked to come in to take the FLKRS/Star at school. A small percentage of remote learning kindergarteners did not take the FLKRS/Star (11% of remote learners), with a slightly larger percentage of in-person kindergarteners missing the test (14% of in-person learners): The difference was not statistically significant, $\chi^2(1) = 1.65, p = .22$. In addition, mean FLKRS/Star scores were not different between in-person and online learners, $t(978) = 0.65, p = .51$.

Child Demographic Characteristics

We include several child-level variables in our analysis, including gender, race/ethnicity, the primary home language of the child's parent(s), free/reduced price lunch (FRPL) status, disability status, 504 Plan eligibility (i.e., qualifying for academic accommodations), and children designated as dual language learners (DLLs) by their school. Demographic variables in our data set were collected during the VPK year and updated in kindergarten if necessary. For the variables of disability status and FRPL, some children's categories changed over time, so we report variables at two time points (VPK and end of kindergarten). All VPK child variables were used as covariates in the propensity score analysis.

Analytic Plan

Before we compared the two cohorts with respect to their language and literacy skills, we conducted missing data analysis by cohort to investigate whether COVID-19 made any impact on the number and demographic makeup of students taking the FLKRS/Star. We also conducted a series of preliminary analysis (e.g., examining differences in demographics and test scores between cohorts using chi-square tests and t tests). Under the assumption of missing at random (missingness as a function of the observed data but not the missing data; Rubin, 1976), multiple imputation by chained equations was performed 30 times ($m = 30$) using the *mice* package (Van Buuren, 2021) in R to impute the missing data in the covariates and the FLKRS/Star scores.

To address the research questions, which investigate the impact of disruptions to state-funded pre-K on language and literacy development at kindergarten entry, we used propensity score matching (PSM; Rosenbaum & Rubin, 1983). PSM allows us to compare the COVID cohort with their counterfactual cohort who had not experienced school closures (the pre-COVID cohort). We matched the COVID cohort and pre-COVID cohort children on demographics and all available pre-K literacy assessment (VPK) scores and then compared matched children with respect to their language and literacy scores at kindergarten entry (FLKRS/Star). Ideally, PSM reduces selection bias and yields an unbiased estimate of treatment effect (Thoemmes & Kim, 2011).

We first estimated the propensity scores (the probability that a child belonged to the treatment condition) using a logistic regression with the cohort membership as a binary outcome and VPK covariates and their interactions as predictors (Green & Stuart, 2014). Furthermore, we included dummy-coded school variables as fixed effects to account for the clustered structure of the data (Arpino & Cannas, 2015). We then implemented the propensity scores through full matching with a caliper of 0.03

(Hansen, 2004), which has previously been shown to be effective at reducing bias due to confounding (Stuart & Green, 2008). We further specified the average treatment effect on the treated as the target estimand because we are interested in the impact of school disruption on the random COVID cohort student. We diagnosed the covariate balance through a visual evaluation of the area of common support using a kernel density plot, and absolute standardized mean differences with a cutoff value of 0.25 SD (Stuart, 2010; Thoemmes & Kim, 2011). Lastly, we estimated the cohort difference and its standard error with matched data sets using a linear regression model. We built a doubly-robust model (Shadish & Steiner, 2010) in which we included the demographic and VPK assessment variables to make an additional covariance adjustment. Note that we conducted PSM with each of 30 imputed data sets and pooled the results using the *MatchThem* package (Pishgar et al, 2021) in R to arrive at a single set of coefficient and standard error estimates from the imputed data sets.

To address Research Question 2, we compared the two cohorts for each of nine subscales of the FLKRS/Star. The treatment effect was estimated in the same way as in Research Question 1. To investigate whether preschool disruptions made differential impacts on subgroups of children such as those who received FRPL (Research Question 3), we tested the interaction effects between these demographic variables and cohort (Green & Stuart, 2014).

Results

Preliminary Analysis

No substantial differences were found in the patterns of missingness between pre-COVID and COVID cohorts (for the results of missing data analysis, see Table S1, available on the journal website). Child demographics were generally also comparable between cohorts (Table 1), although the number of children with disabilities who were given an individualized education program (IEP) showed substantial variation between cohorts, with 89% of students in the COVID cohort categorized as "no disability" in kindergarten compared to 66% in the pre-COVID cohort. Student test scores were then compared between cohorts without any adjustment (no covariates controlled for). The independent sample t tests suggested that there were statistically significant mean differences between the pre-COVID cohort and COVID cohort across all VPK scores and FLKRS/Star scores (Table 2) favoring the COVID cohort. In addition, several two-way analyses of variance were performed to evaluate the interaction effect between cohort membership and the demographic variables on the FLKRS/Star scores without controlling for any covariates. None of the interaction effects was found to be statistically significant. In other words, the difference in cohorts did not depend on students' demographics.

Propensity Score Matching

After propensity score estimation, all the covariates were balanced between cohorts. For detailed results of balance and common support, see Table S2 and Figure S1, available on the journal website. A linear regression model on the matched

Table 1
Sample Size of Demographic Subgroups by Cohort

Variables	Pre-COVID cohort (n = 1,211)	COVID cohort (n = 1,167)	Diff. in sample size
Learning mode (K)			
Remote	—	500 (43%)	—
Face-to-face	—	667 (57%)	—
Gender			
Male	677 (56%)	610 (52%)	-4%
Female	534 (44%)	557 (48%)	+4%
Race/ethnicity			
White	684 (56.5%)	595 (51%)	-5.5%*
Black	240 (19.8%)	199 (17%)	-2.8%
Latino/Hispanic	178 (14.7%)	212 (18.2%)	+3.5%
Asian	68 (5.6%)	70 (6%)	+0.4%
Multiracial	41 (3.4%)	91 (7.7%)	+4.3%**
Parent language			
English	961 (79.4%)	931 (80%)	+0.3%
Spanish	130 (10.7%)	119 (10%)	-0.5%
Other	120 (10%)	117 (10%)	0%
FRPL (VPK)			
Not eligible	316 (26%)	322 (28%)	+2%
Eligible	895 (74%)	845 (72%)	-2%
FRPL (K)			
Not eligible	304 (25%)	347 (30%)	+5%*
Eligible	907 (75%)	820 (70%)	-5%
Disability status (VPK)			
No disabilities	769 (63.5%)	1,053 (90%)	+26.5%**
Disabilities	442 (36.5%)	114 (10%)	-26.5%
Speech disability	84 (6.9%)	71 (6.1%)	-0.8%
Language disability	160 (13.2%)	16 (1.4%)	-11.8%**
Autism spectrum disorder	36 (3%)	1 (.1%)	-2.9%**
Other disabilities	162 (13.4%)	26 (2.2%)	-11.2%**
Disability status (K)			
No disabilities	797 (65.8%)	1,036 (89%)	+23.2%**
Disabilities	414 (34.2%)	131 (11%)	-23.2%
Speech disability	106 (8.8%)	90 (7.4%)	-1.4%
Language disability	182 (15%)	27 (2.3%)	-12.7%**
Autism spectrum disorder	51 (4.2%)	3 (.3%)	-3.9%**
Other disabilities	75 (6.2%)	11 (.9%)	-5.3%**
504 Plan status			
Not eligible	1,192 (98.5%)	1,163 (99.7%)	+1.2%*
Eligible	19 (1.5%)	4 (.3%)	-1.2%
DLL			
Not applicable	1,108 (91.6%)	1,056 (90.4%)	-1.2%
Eligible	103 (8.4%)	111 (9.6%)	+1.2%

Note. Diff = proportional difference in sample size between cohorts; K = data collected in kindergarten; VPK = data collected during voluntary pre-K year; parent language = primary home language of parent; FRPL = child qualifies for free or reduced price lunch based on family income; disability status = children with disabilities are defined as those with an individualized education program; 504 plan = child is eligible under Section 504 of Rehabilitation Act to receive accommodations to ensure their academic success; DLL = dual language learner. The naïve match rate (Ho, 2021) of COVID cohort to pre-COVID cohort is 96% (=1,167/1,211); the match rate between VPK and K (Ho, 2021) of COVID cohort is 100% (= 1,167/1,167) because all children in VPK (before COVID) were included in kindergarten (after COVID). * $p < .05$. ** $p < .01$.

sample (pre-COVID: $n = 1,131$; COVID: $n = 1,115$) with the cohort membership as a predictor showed the nonsignificant effect of COVID-related disruptions on children’s language and literacy outcomes at kindergarten entry: The difference in the

FLKRS/Star scores between cohorts was not statistically significant ($\beta = 8.04$, $SE = 6.92$, 95% CI = [-5.56, 21.64], $p = .25$). Children who spent a full year in FL-VPK did not perform any better on language and literacy measures at kindergarten entry

Table 2
Descriptive Information for Continuous Variables by Cohort

Variables	Pre-COVID cohort (n = 1,211)							COVID cohort (n = 1,167)						
	N	M	SD	Missing	Skew	Kurt	N	M	SD	Missing	Skew	Kurt	Cohen's d	
VPK assessment (pre-K)														
Print Knowledge AP1	1,090	6.6	3.7	121 (10%)	0.1	-1.3	1,107	6.4	3.5	60 (5%)	0.2	-1.2	0.03	
Phonological Awareness AP1	1,090	6.5	3.4	121 (10%)	0.3	-0.7	1,107	6.8	3.4	60 (5%)	0.3	-0.7	0.11	
Oral Language/Vocabulary AP1	1,090	15.6	4.7	121 (10%)	-1.0	0.1	1,107	16.1	4.3	60 (5%)	-1	0.8	0.16*	
Print Knowledge AP2	1,126	9.1	3.5	85 (7%)	-0.9	-0.6	1,148	9.2	3.2	19 (1.6%)	-0.9	-0.4	0.07	
Phonological Awareness AP2	1,126	8.9	3.7	85 (7%)	-0.4	-0.9	1,148	10.0	3.3	19 (1.6%)	-0.8	-0.2	0.32**	
Oral Language/Vocabulary AP2	1,126	18.9	4.2	85 (7%)	-1.9	3.9	1,148	19.4	3.2	19 (1.6%)	-1.8	4.9	0.19**	
Print Knowledge AP3	1,149	9.9	3.1	62 (5%)	-1.4	0.8	—	—	—	—	—	—	—	
Phonological Awareness AP3	1,149	10.3	3.6	62 (5%)	-0.9	-0.3	—	—	—	—	—	—	—	
Oral Language/Vocabulary AP3	1,149	19.4	3.9	62 (5%)	-1.9	3.9	—	—	—	—	—	—	—	
FLKRS/Star assessment (K)														
Scale score	1,138	526.9	101	73 (6%)	0.4	-0.3	1,017	548.8	107	150 (13%)	0.5	-0.1	0.21**	
Alphabetic Principle	1,138	61.0	18	—	-0.4	-0.4	1,017	76.2	15	—	-1.1	1.3	0.92**	
Concept of Word	1,138	62.2	18	—	-0.5	-0.3	1,017	71.2	16	—	-0.8	0.5	0.53**	
Visual Discrimination	1,138	70.9	17	—	-0.8	0.3	1,017	77.7	15	—	-1.2	1.7	0.44**	
Phonemic Awareness	1,138	39.5	17	—	0.3	-0.3	1,017	45.0	18	—	0.3	-0.3	0.32**	
Phonics	1,138	37.5	17	—	0.4	-0.2	1,017	44.6	19	—	0.4	-0.2	0.40**	
Structural Analysis	1,138	30.1	16	—	0.7	0.3	1,017	36.7	19	—	0.7	0.3	0.38**	
Vocabulary	1,138	40.4	17	—	0.3	-0.3	1,017	45.6	18	—	0.3	-0.2	0.30**	
Sentence Comprehension	1,138	32.3	17	—	0.6	0.1	1,017	38.9	19	—	0.6	0	0.36**	
Paragraph Comprehension	1,138	31.0	15	—	0.7	0.5	1,017	36.2	17	—	0.8	0.5	0.32**	

Note. Mean score difference between cohorts was tested by running *t* tests. Cohen's *d* was also calculated to represent the effect size. The Pre-COVID cohort missed AP3 of the VPK assessment due to school closures. FLKRS/Star subdomain scores represent the percentage of items a student would be expected to answer correctly in that area. Skew = skewness; kurt = kurtosis; VPK = voluntary pre-K assessments; AP = assessment period; AP1 = fall; AP2 = winter; AP3 = spring; FLKRS/Star = Florida Kindergarten Readiness Screener/Star Early Literacy assessment, administered in fall of kindergarten year.

* $p < .05$. ** $p < .01$.

Table 3
Linear Regression Models Based on the Nine Subdomains of Florida Kindergarten Readiness Screener/Star Early Literacy Assessment Using Cohort Membership as Predictor

Predictor	β	SE	Cohen's <i>d</i>
Alphabetic Principle	12.84***	1.14	0.23
Concept of Word	6.63***	1.18	0.11
Visual Discrimination	4.59***	1.11	0.08
Phonemic Awareness	3.21***	1.15	0.06
Structural Analysis	4.41***	1.11	0.08
Vocabulary	3.23*	1.18	0.05
Sentence Comprehension	4.22***	1.17	0.07
Paragraph Comprehension	3.07***	1.04	0.06
Phonics	4.74***	1.17	0.08

* $p < .05$. ** $p < .001$,

than those who missed a quarter of the pre-K year. In addition, we investigated the effect of the COVID disruptions on FLKRS/Star's nine subdomains. The results showed statistically significant differences in all subdomains, with the COVID cohort having higher mean scores for each (Table 3). The largest mean difference (15.2 points; $d = 0.23$) was on the subdomain test for Alphabetic Principle; the smallest mean differences (5.2 points) were on Paragraph Comprehension ($d = 0.06$) and Vocabulary tests ($d = 0.05$). Finally, the effect of COVID-related disruptions did not vary by students' demographic characteristics (e.g., race/ethnicity, socioeconomic status, and disability status): None of the interaction effects between demographic variables and cohort was statistically significant. We also investigated the interaction effects using the nine subdomains. Similar to the FLKRS/Star scale scores, the effect of COVID-related disruptions did not vary across different subgroups for the nine subdomains. For detailed results of the estimates for each subgroup, see Table S3, available on the journal website.

Discussion

The present article provides one of the first examinations of how spring 2020 school closures impacted language and literacy development for pre-K children. Our key finding is that COVID school closures had no negative impacts on pre-K children's language and literacy achievement at kindergarten entry as measured by standardized tests. In fact, the COVID cohort of children scored slightly *better* on kindergarten screening measures than their pre-COVID peers, although the difference was not statistically significant once PSM was applied. We also did not find significantly negative impacts on the standardized test scores of children from marginalized groups. Considered from one perspective, this is excellent news: Contrary to early projections, pre-K children in our sample did not suffer major "learning losses" from school closures. Considered from another perspective, however, this finding raises pressing questions about the efficacy of some state-funded pre-K programs. Put plainly, why did missing a quarter of the in-person pre-K year essentially not matter for children's language and literacy test scores? We explore the implications of our findings in the following.

Impacts of COVID on Language and Literacy

There are few studies on the impacts of COVID on pre-K children, and our results both accord with and converge from available reports. Similar to our results, a descriptive study in Texas found that pre-K children did not experience learning losses in reading at kindergarten entry (Patarapichayatham et al., 2021), but this study also excluded all remote testing data from their analysis, potentially biasing results. An analysis of Virginia kindergarteners found that COVID school closures *did* negatively impact pre-K children's literacy learning, with about 9% more kindergarteners in fall 2020 scoring below literacy test benchmarks compared to fall 2019 (McGinty et al., 2021). While this study, like ours, compared a pre- and post-COVID cohort in fall of kindergarten, there was no pre-K data that could be used to adjust for baseline differences between cohorts, and fewer children tested overall in the COVID cohort.

Several studies with elementary school students (first to third grades) report that COVID had small negative impacts on reading in fall 2020 (e.g., Curriculum Associates, 2020). Others corroborate our results, finding no negative impacts on reading in fall 2020 for the youngest grades (e.g., Kilbride et al., 2022; Renaissance Learning, 2020), but these studies also include remote testing data that may have overestimated children's scores. Drawing firm conclusions about the impacts of COVID on literacy and language is difficult given questions about data quality, but on the whole, impacts for young children do appear to have been smaller than projected.

There are several potential reasons why school closures had little impact on language and literacy achievement in our sample. First, teachers' and parents' efforts during remote learning may have compensated for school closures. Recent studies, however, suggest that remote instruction was of generally low intensity and quality: Pre-K teachers reported limited amounts of "live" online teaching, difficulty keeping children engaged, and lower quality interactions with children (McKenna et al., 2021; Weiland et al., 2021). Most pre-K children participated less than once per week in remote learning activities during school closures (Barnett & Jung, 2020).

Remote learning may have bolstered certain skills more than others. We found that while the COVID cohort outperformed

pre-COVID children in all domains, they did so most substantially on code-based skills (76% correct on the Alphabetic Principle domain in 2020; 61% in 2019). Many of the apps and resources sent home during remote learning focused on these code-based, constrained skills (Hadley et al., 2022; McKenna et al., 2021), meaning that children may have practiced skills like letter knowledge more intensively during remote learning.

Children may have also gained an extra learning “boost” from interacting with both teachers and parents as they worked together to facilitate remote learning. Available data, however, paint a mixed picture of parents’ availability during a very stressful time. A nationally representative survey found that while parents frequently read with and sang songs to pre-K children in spring 2020, rates slightly declined from 2019 (Barnett & Jung, 2020). Most pre-K parents reported that they felt overwhelmed by the responsibility of facilitating remote learning (Barnett & Jung, 2021). Parents relied on screen time during school closures, with low socioeconomic status kindergarteners spending an average of 6.6 hours a day using technology both for entertainment and instruction (Dore et al., 2021).

In considering our findings, it is important to note that our data do suggest serious negative consequences of school closures that are not captured by standardized test scores. We found striking differences in the number of children with IEP plans in the COVID cohort ($n = 114$, 10%) versus the pre-COVID cohort ($n = 442$, 36.5%) at the end of the VPK year, and these differences persisted through the end of kindergarten. This suggests that many children with disabilities were not identified during school closures and therefore did not receive needed services during crucial early intervention years. Other studies corroborate this finding, with lower rates of identification of pre-K children with disabilities and difficulties providing services for children with existing IEPs during COVID (Barnett & Jung, 2021). Additionally, emerging research reveals negative impacts from COVID-19 on other developmental domains, with studies showing disruptions to sleep patterns, more food insecurity, and increased anxiety, depression, and behavioral issues in young children (Naff et al., 2022; Niles et al., 2020). Because nonacademic skills like social-emotional competencies underpin children’s long-term academic success, children whose development was impacted by COVID-19 may face future academic difficulties as they progress through school. Therefore, the extent of negative academic impacts of COVID-19 may have yet to be fully captured. Newer research suggests that test scores further declined during the 2020–2021 school year (Kuhfeld et al., 2022). These declines may be partially attributable to continued remote learning but may also indicate later emerging impacts from related domains such as social-emotional development.

Efficacy of FL-VPK for Language and Literacy Development

Our study serves as a natural experiment, allowing us to estimate the impact of missing a quarter of a year of state-funded pre-K. Based on prior literature on effect sizes for state-funded pre-K (Barnett et al., 2018), we hypothesized that children in the COVID cohort would have substantially lower literacy skills (a reduction of approximately $d = 0.275$) and slightly lower language skills than

the pre-COVID cohort. This hypothesis was not supported: There was no significant difference in overall kindergarten entry scores between cohorts, and the COVID cohort had *higher* scores than the pre-COVID cohort in every FLKRS/Star subdomain, with effect sizes ranging from $d = 0.23$ to $d = -0.08$ on literacy skills and $d = 0.05$ on language skills (see Table 3).


These findings are surprising given that nearly all pre-K efficacy studies have found positive impacts of pre-K on literacy and language in the short term, although effects for individual subdomains are sometimes small or null (e.g., Meloy et al., 2019). It is important to note the counterfactual condition in these prior studies: The contrast is not between pre-K and no pre-K but, rather, children who attended public pre-K versus those assigned to a comparison group, who attended another preschool or childcare center, or stayed home (Meloy et al., 2019). Our study, on the other hand, examines the contrast between (a) children who attended a full year of state-funded pre-K versus (b) children who attended three-quarters of a year and had remote support for the remaining quarter. In both cases, a full year of state-funded pre-K is contrasted with a comparison condition in which many children receive at least some exposure to learning. However, in the present study, the dosage of state pre-K was reduced by only one-fourth, which may not have been enough to move the needle on standardized tests.


However, it is striking that the COVID cohort actually outperformed the pre-COVID cohort on individual subdomains. Staying home for a quarter of the year with caregivers was counterintuitively *more* beneficial for children in individual subdomains, especially in terms of constrained literacy skills, than attending a full year of FL-VPK. This suggests that FL-VPK had little impact on children’s language and literacy skills when compared to a counterfactual, a finding that is not wholly unanticipated by prior literature on the long-term impacts of FL-VPK (e.g., Miller & Bassok, 2019). FL-VPK has been critiqued as a program that prioritizes access over quality (e.g., Miller & Bassok, 2019) because it serves nearly three-quarters of children in Florida but meets only two out of 10 of the National Institute for Early Education Research’s benchmarks for early childhood program quality (Friedman-Krauss et al., 2020). For example, FL-VPK does not have systematic support for selecting and implementing curricula, teachers are not required to have a bachelor’s degree, and no continuous quality improvement system is in place (Friedman-Krauss et al., 2020). FL-VPK also ranks amongst the lowest states in terms of per-pupil funding for pre-K. In contrast, pre-K studies with large effect sizes involve full-day programs that have made substantial investments in program quality. For example, the Boston Public Schools (BPS) pre-K program is the outcome of a decade of concerted leadership efforts, partnerships, and funding from private and public sources. BPS requires teachers to have a bachelor’s degree, earn a master’s degree, use a research-based curriculum, and be coached by a master educator (Weiland & Yoshikawa, 2013). In comparison with the counterfactual, in which two-thirds of children attended another form of center-based care, BPS children had substantially better language and literacy skills (e.g., $d = 0.44$ for vocabulary, $d = 0.62$ for print awareness).


In our study, a lower quality, minimally funded state pre-K program did not outperform a counterfactual condition (i.e., missing a quarter of the in-person school year). Prior research

suggests that a certain threshold of quality is necessary to impact academic outcomes in pre-K and that programs must provide sufficient levels of supportive teacher-child interactions and skilled instruction to be effective (Burchinal et al., 2010). To date, however, there are few studies that report on the efficacy of state pre-K programs that have comparatively less funding and are rated as lower quality. Such research, in addition to research on exemplary programs, is important because it allows us to test both the upper and lower bounds of pre-K quality: What are the nature and extent of supports that are needed to ensure quality instruction at scale? Our findings suggest that instruction provided by state-funded programs without adequate funding and key quality safeguards may not be enough to move the needle on children's language and literacy outcomes. For future research, there would be much to be learned from implementing cycles of iterative improvement and assessment in these programs, such as testing the impact of introducing factors like research-based curricula, pre-K coaches, and higher pay for teachers (Bassok & Engel, 2019). Prior research has shown that well-funded, high-quality pre-K programs can and do make a difference for families and children across a range of developmental domains, including language and literacy. As a matter of equity, children nationwide deserve access to early childhood programs that set them on a path to thrive in school.

ORCID IDS

Elizabeth Burke Hadley  <https://orcid.org/0000-0002-2880-6175>

Eunsook Kim  <https://orcid.org/0000-0003-1054-1735>

Meaghan McKenna  <https://orcid.org/0000-0001-9430-415X>

NOTE

The research reported in this article was made possible by grants from the Spencer Foundation (No. 02200049) and the American Educational Research Association Division E Doctoral Student and Early Career Scholar Seed Grant Program. The views expressed are those of the authors and do not necessarily reflect the views of the Spencer Foundation or AERA.

¹Note that this section focuses on the immediate impacts of school closures in fall 2020 because our study examines this time period.

REFERENCES

- Arpino, B., & Cannas, M. (2016). Propensity score matching with clustered data. An application to the estimation of the impact of caesarean section on the Apgar score. *Statistics in Medicine*, 35(12), 2074–2091. <https://doi.org/10.1002/sim.6880>
- Barnett, W. S., & Jung, K. (2020). *Young children's home learning and preschool participation experiences during the pandemic*. https://nieer.org/wp-content/uploads/2020/11/NIEER_Tech_Rpt_July2020_Young_Childrens_Home_Learning_and_Preschool_Participation_Experiences_During_the_Pandemic-AUG2020.pdf
- Barnett, W. S., & Jung, K. (2021). *Seven impacts of the pandemic on young children and their parents: Initial findings from NIEER's December 2020 preschool learning activities survey*. https://nieer.org/wp-content/uploads/2021/02/NIEER_Seven_Impacts_of_the_Pandemic_on_Young_Children_and_their_Parents.pdf
- Barnett, W. S., Jung, K., Friedman-Krauss, A., Frede, E. C., Nores, M., Hustedt, J. T., Howes, C., & Daniel-Echols, M. (2018). State prekindergarten effects on early learning at kindergarten entry: An analysis of eight state programs. *AERA Open*, 4(2). <https://doi.org/10.1177/2332858418766291>
- Bassok, D., & Engel, M. (2019). Early childhood education at scale: Lessons from research for policy and practice. *AERA Open*, 5(1). <https://doi.org/10.1177/2332858419828690>
- Bassok, D., Weisner, K., Markowitz, A. J., & Hall, T. (2021). *Teaching young children during COVID-19: Lessons from early educators in Virginia*. EdPolicyWorks at the University of Virginia. http://bit.ly/see_partnerships_teaching_covid
- Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-Kindergarten programs. *Early Childhood Research Quarterly*, 25(2), 166–176. <https://doi.org/10.1016/j.ecresq.2009.10.004>
- Curriculum Associates (2020). *Understanding student needs: Early results from fall assessments*. <https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/i-ready-diagnostic-results-understanding-student-needs-paper-2020.pdf>
- Dore, R. A., Purtell, K. M., & Justice, L. M. (2021). Media use among kindergarteners from low-income households during the COVID-19 shutdown. *Journal of Developmental & Behavioral Pediatrics*, 42(8), 672–676.
- Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. Cambridge University Press.
- Durkin, K., Lipsey, M. W., Farran, D. C., & Wiesen, S. E. (2022). Effects of a statewide pre-kindergarten program on children's achievement and behavior through sixth grade. *Developmental Psychology*, 58(3), 470–484. <https://doi.org/10.1037/dev0001301>
- Friedman-Krauss, A., Barnett, W. S., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., & Gardiner, B. A. (2020). *The state of preschool 2020*. https://nieer.org/wp-content/uploads/2021/04/YB2020_Full_Report.pdf
- Green, K. M., & Stuart, E. A. (2014). Examining moderation analyses in propensity score methods: Application to depression and substance use. *Journal of Consulting and Clinical Psychology*, 82(5), 773–783. <https://doi.org/10.1037/a0036515>
- Hadley, E. B., McKenna, M., & Soto-Boykin, X. (2022, April). *Early childhood teachers' use of digital apps for language and literacy instruction during COVID-19* [Poster presentation]. Annual Meeting of the American Educational Research Association, San Diego, CA.
- Hansen, B. B. (2004). Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association*, 99(467), 609–618. <https://doi.org/10.1198/016214504000000647>
- Ho, A. (2021). *Three test-score metrics that all states should report in the COVID-19-affected spring of 2021*. <https://scholar.harvard.edu/files/andrewho/files/threemetrics.pdf>
- Huang, F. L., Invernizzi, M. A., & Drake, E. A. (2012). The differential effects of preschool: Evidence from Virginia. *Early Childhood Research Quarterly*, 27(1), 33–45. <https://doi.org/10.1016/j.ecresq.2011.03.006>
- Huff, K. (2020). *National data quantifies impact of COVID learning loss; raises questions about at-home testing*. <https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/ca-impact-of-covid-learning-loss-fall-2020.pdf>
- Johnson, A. D., Markowitz, A. J., Hill, C. J., & Phillips, D. A. (2016). Variation in impacts of Tulsa pre-K on cognitive development in kindergarten: The role of instructional support. *Developmental Psychology*, 52(12), 2145–2158. <https://doi.org/10.1037/dev0000226>
- Kilbride, T., Hopkins, B., Strunk, K. O., & Yu, D. (2022). *Michigan's fall 2021 benchmark assessments*. https://epicedpolicy.org/wp-content/uploads/2022/04/Benchmark_Report_April2022.pdf

- Kuhfeld, M., Lewis, K., Meyer, P., & Tarasawa, B. (2020). *Comparability analysis of remote and in-person MAP Growth testing in fall 2020*. <https://www.nwea.org/content/uploads/2020/11/Technical-brief-Comparability-analysis-of-remote-and-in-person-MAP-Growth-testing-in-fall-2020-NOV2020.pdf>
- Kuhfeld, M., Soland, J., Lewis, K., Ruzek, E., & Johnson, A. (2022). The COVID-19 school year: Learning and recovery across 2020–2021. *AERA Open*, 8. <https://doi.org/10.1177/23328584221099306>
- Lipsey, M. W., Weiland, C., Yoshikawa, H., Wilson, S. J., & Hofer, K. G. (2015). The prekindergarten age-cutoff regression-discontinuity design: Methodological issues and implications for application. *Educational Evaluation and Policy Analysis*, 37(3), 296–313.
- McGinty, A., Gray, A., Partee, A., Herring, W. A., & Soland, J. (2021). *Examining early literacy skills in the wake of COVID-19 spring 2020 school disruptions: Virginia Fall 2020 statewide screening findings and implications*. https://literacy.virginia.edu/sites/g/files/jsddwu1006/files/2022-03/PALS_Fall_2020_Data_Report_5_18_final.pdf
- McKenna, M., Soto-Boykin, X., Cheng, K., Haynes, L., Osorio, A., & Altshuler, J. (2021). Initial development of a national survey on remote learning in early childhood during COVID-19: Establishing content validity and reporting successes and barriers. *Early Childhood Education Journal*, 49, 815–827. <https://doi.org/10.1007/s10643-021-01216-y>
- Meloy, B., Gardner, M., & Darling-Hammond, L. (2019). *Untangling the evidence on preschool effectiveness: Insights for policymakers*. https://learningpolicyinstitute.org/sites/default/files/product-files/Untangling_Evidence_Preschool_Effectiveness_REPORT.pdf
- Miller, L. C., & Bassok, D. (2019). The effects of universal preschool on grade retention. *Education Finance and Policy*, 14(2), 149–177. https://doi.org/10.1162/edfp_a_00248
- Naff, D., Williams, S., Furman-Darby, J., & Yeung, M. (2022). The mental health impacts of COVID-19 on PK–12 students: A systematic review of emerging literature. *AERA Open*, 8. <https://doi.org/10.1177/23328584221084722>
- Niles, M. T., Bertmann, F., Belarmino, E. H., Wentworth, T., Biehler, E., & Neff, R. (2020). The early food insecurity impacts of COVID-19. *Nutrients*, 12(7), Article 2096. <https://doi.org/10.3390/nu12072096>
- Patarapichayatham, C., Locke, V. N., & Lewis, S. (2021). *COVID-19 learning loss in Texas*. https://www.istation.com/Content/downloads/studies/COVID-19_Learning_Loss_Texas.pdf
- Pishgar, F., Greifer, N., Leyrat, C., & Stuart, E. (2021). MatchThem: Matching and weighting after multiple imputation. *R Journal*, 13(2), 292–305. <https://doi.org/10.32614/RJ-2021-073>
- Renaissance Learning. (2020). *How kids are performing: Tracking the impact of COVID-19 on reading and mathematics achievement*. <https://renaissance.widen.net/s/wmjtlxkhbm>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581–592. <https://doi.org/10.1093/biomet/63.3.581>
- Shadish, W. R., & Steiner, P. M. (2010). A primer on propensity score analysis. *Newborn and Infant Nursing Reviews*, 10(1), 19–26. <https://doi.org/10.1053/j.nainr.2009.12.010>
- Storey, N., & Zhang, Q. (2021). *A meta-analysis of COVID learning loss*. EdArXiv. <https://doi.org/10.35542/osf.io/qekw2>
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1–21. <https://doi.org/10.1214/09-STS313>
- Stuart, E. A., & Green, K. M. (2008). Using full matching to estimate causal effects in nonexperimental studies: Examining the relationship between adolescent marijuana use and adult outcomes. *Developmental Psychology*, 44(2), 395–406. <https://doi.org/10.1037/0012-1649.44.2.395>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research*, 46, 90–118. <https://doi.org/10.1080/00273171.2011.540475>
- van Buuren, S. (2021). *Multivariate imputation by Chained Equations*. R Documentation. <https://www.rdocumentation.org/packages/mice/versions/3.13.0>
- van Huizen, T., & Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review*, 66, 206–222.
- Weiland, C., Greenberg, E., Bassok, D., Markowitz, A., Guerrero Rosada, P., Luetmer, G., Abenavoli, R., Gomez, C., Johnson, A., Jones-Harden, B., Maier, M., McCormick, M., Morris, P., Norens, M., Phillips, D., & Snow, C. (2021). *Historic crisis, historic opportunity: Using evidence to mitigate the effects of the COVID-19 crisis on young children and early care and education programs*. <https://edpolicy.umich.edu/sites/epi/files/uploads/EPI-UI-Covid%20Synthesis%20Brief%20June%202021.pdf>
- 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>

AUTHORS

ELIZABETH BURKE HADLEY, PhD, is an assistant professor at the University of South Florida, 4202 E. Fowler Ave. EDU105 Tampa FL 33620; hadleye@usf.edu. Her research focuses on oral language and literacy development in pre-K, with a special focus on helping teachers create equitable and enriching classroom language environments for children.

SIYU LIU, MS, is a doctoral student at University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620; sliu7@usf.edu. Her research focuses on applying propensity score methods in latent class/latent profile analysis.

EUNSOOK KIM, PhD, is a professor at the University of South Florida, 4202 E. Fowler Ave. EDU105 Tampa FL 33620; ekim3@usf.edu. Her research focuses on quantitative research methodology and psychometrics, including structural equation modeling, multilevel modeling, mixture modeling, and measurement invariance testing.

MEAGHAN MCKENNA, PhD, is an assistant research professor at the University of Connecticut in the Department of Educational Psychology in the Neag School of Education Charles B. Gentry Building 249 Glenbrook Road U-3064, Storrs, CT 06269; meaghan.mckenna@uconn.edu. Her research focuses on engaging in collaborative partnerships to solve educational problems of practice and enhancing how early literacy is defined, assessed, and supported within a multi-tiered system of supports.

Manuscript received August 16, 2022
Revisions received February 9, 2023, and April 18, 2023
Accepted April 30, 2023