Title:

Recognizing the Growth Made by Students with Limited English Proficiency: A Latent Growth Curve Approach

Authors and Affiliations:

Shuai Zhang, Ph.D., Research Assistant Professor, Appalachian State University

Kay Wijekumar, Ph.D., Professor, Texas A&M University,

Sarah Fitzhugh, Ph.D., Director of Assessment, Lewisville Independent School District

Background

In Texas, students from Grades 3-12 are required to take the End-of-Year State of Texas Assessments of Academic Readiness (STAAR) reading test. The results from these assessments are used to measure the success of the school administration, effectiveness of instruction, and identify students at risk for learning disabilities (Haertel, 2018; Nichols et al., 2005). However, as these STAAR reading tests only report a raw score, little is known if these tests recognize the progress made by students during the school year. This is particularly important for students with limited English proficiency (LEP). Due to the difficulty in understanding English vocabulary and discourse, LEP students tend to perform lower on high-stakes tests as compared to their Non-LEP peers. However, LEP students may show comparable or faster reading growth rates than their peers. In this research, we used the Istation Indicators of Progress Early Reading (*Istation*) to assess the growth of both LEP and Non-LEP elementary students, and investigate the relation between growth factors of Istation measures and results from the STAAR-reading test.

Research questions

RQ 1. What are the growth trajectories of students' reading comprehension and vocabulary from Grade 1-3? Do LEP students grow slower on vocabulary and reading comprehension than their peers?

RQ 2. Can the fast growth rates of LEP (if any) help reduce their achievement gap with Non-LEP students on the Grade 3 STAAR reading test?

Participants

A total of 3,372 Grade 1 students from Lewisville, Texas were investigated for the current research. Around 26% students were identified as having LEP based on school records.

Research design

At each grade level, Istation data was collected at the beginning (BOY), middle (MOY) and end (EOY) of the academic year. A total of 9 time points of Istation data were available for analyses. The STAAR reading data was collected at the final time point (i.e., EOY of Grade 3). Missing data was handled by the full information maximum likelihood (FIML) method.

Istation reading comprehension consists of two subtests: 1) Matching Sentences and Pictures and 2) Sentence Completion. The task complexity was adjusted based on students' grade level, and items with poor item discrimination power, or too high/low item difficulty, were removed (Mathe et al., 2016). Mathes et al. reported good concurrent validity with other test scores and strong internal consistency of the scores.

The Istation vocabulary consists of 1) Picture Vocabulary 2) Synonym Matching subtests. the vocabulary test items and scores were also adjusted based on grade level. The minimal reliability of the vocabulary test was about .90, and the manual reports strong concurrent validity (Mathes et al., 2016).

The 2018 Grade 3 Reading STAAR contains 4 passages: 2 narrative passages, 1 expository passage, and 1 poem. Each passage was then followed by 8-9 multiple-choice questions for a total of 34 questions. The reported reliability of the Grade 3 students' reading test score was .89 (Texas Education Agency, n.d.).

Data analysis

To answer RQ 1, we first adopted the latent growth curve analysis (LGCA) to understand the growth trends of Istation reading comprehension and vocabulary, and then regressed the STAAR test results on the growth factors (intercept and slopes).

To answer RQ 2, we set LEP status as a direct predictor of the STAAR reading test result, and built indirect paths from LEP to the STAAR results via the slopes and intercepts of reading comprehension and vocabulary.

Results

As Figures 1a and 1b show, students did not show apparent improvement before and after summer breaks. Thus, we dropped the beginning of year data at Grade 2 and 3 to better represent the growth of students. The remaining 7 time points were used to construct the LGCA models.

After comparing a series of models (e.g., linear, quadradic), we found that breaking the slopes into three pieces drastically improved model fit. That is, the growths from Grade 1 BOY to Grade 1 EOY (Phase 1), Grade 1 EOY to Grade 2 EOY (Phase 2), and Grade 2 EOY to Grade 3 EOY (Phase 3) were set as three different phases, but the three phases shared a common intercept (initial status). This piecewise model was applied for both reading comprehension and vocabulary and received excellent fit (see Figure 2).

As shown in Figure 2, students with LEP had lower intercept than Non-LEP students, but the two groups had comparable reading comprehension growth rates across the three phases of growth. For vocabulary, LEP students had lower initial level and slower Phase 1 growth rate than Non-LEP students. At Phase 2, their vocabulary growth rate was faster than Non-LEP students.

From Figure 3 and 4, we observe that those who had a lower reading comprehension level at the initial phase and a slower reading comprehension growth rate at Phase 1 tended to perform lower on the Grade 3 STAAR reading test. Similar findings were identified from the vocabulary growth model.

When examining the mediation effect of LEP and STAAR, we found that the intercept of Istation reading comprehension was the only significant predictor of the discrepancy between LEP and Non-LEP students (see Figure 3). Thus, the fast growth rates of LEP students on vocabulary did not help reduce this discrepancy.

Conclusion

The slower growth of reading comprehension and vocabulary skills in Grade 1 may be risk factors for lower performance on Grade 3 STAAR reading test, which suggests teachers need to provide early and intensive supports to low-performing students.

For policymakers, it may not be plausible to use only high-stakes test results to measure students, teachers and schools, as the growth of LEP students could not be recognized through STAAR test results.

However, since our study only used Istation to track students' growth, the results and conclusions may be restricted to the scores collected from these measurements. All the tests from Istation are multiple-choice based, and may lead to mono-operation and mono-method bias during assessment (Shadish et al., 2002).

In summary, the current research proposes to use growth-based methods instead of static test scores to recognize the growth of low achievement and LEP students.
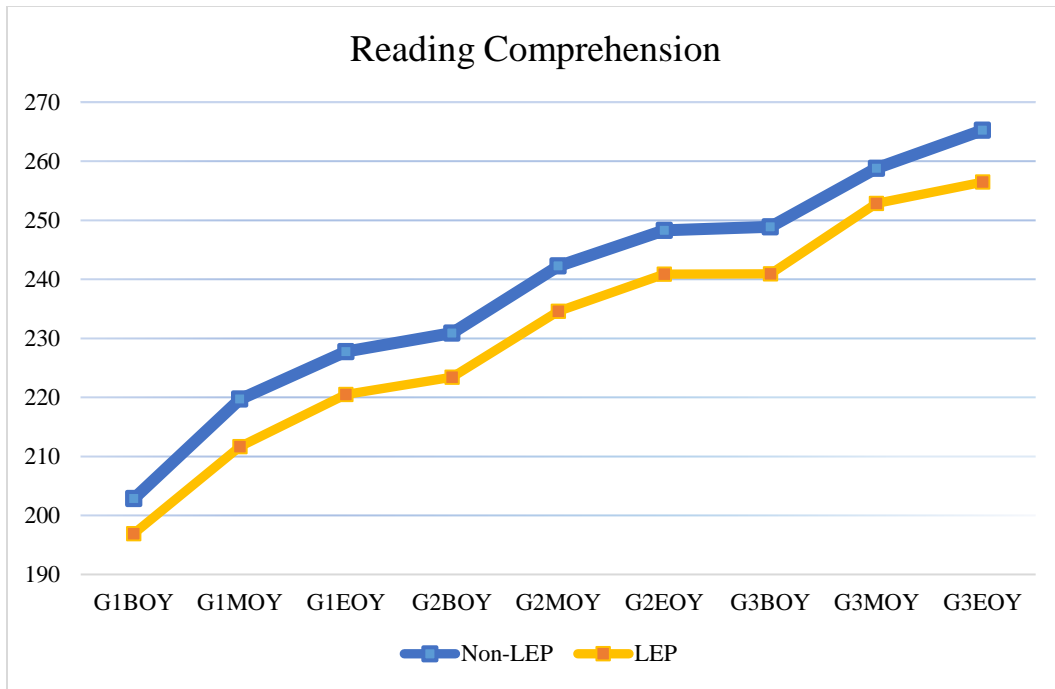
[Word count: 1,000]

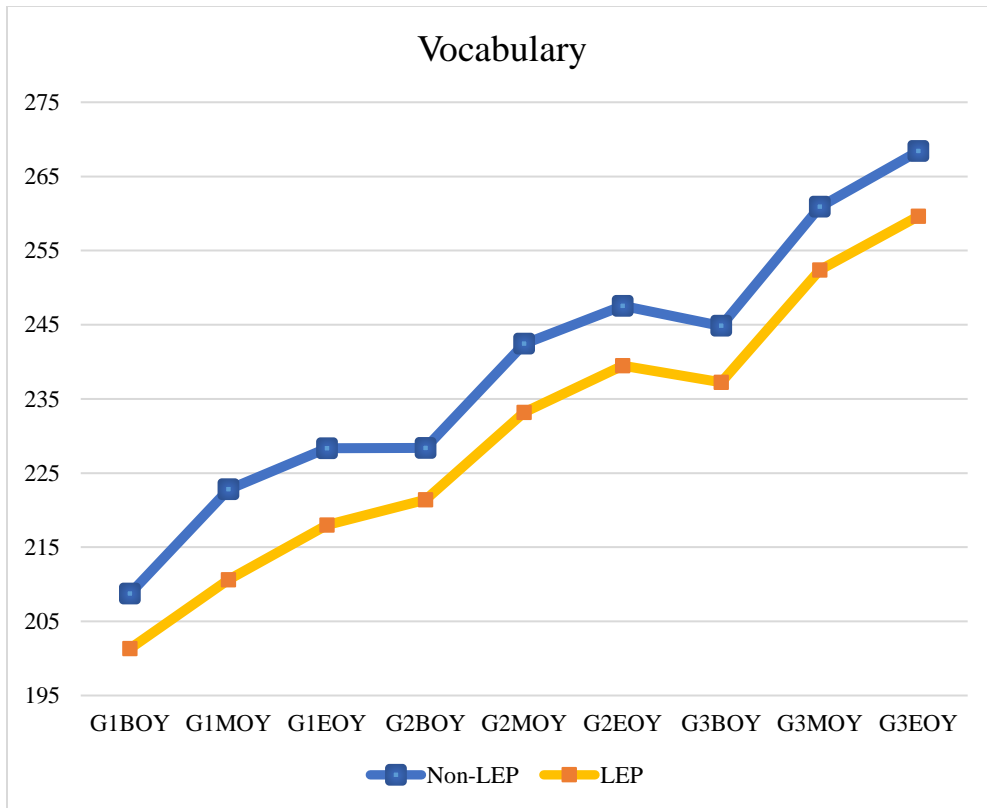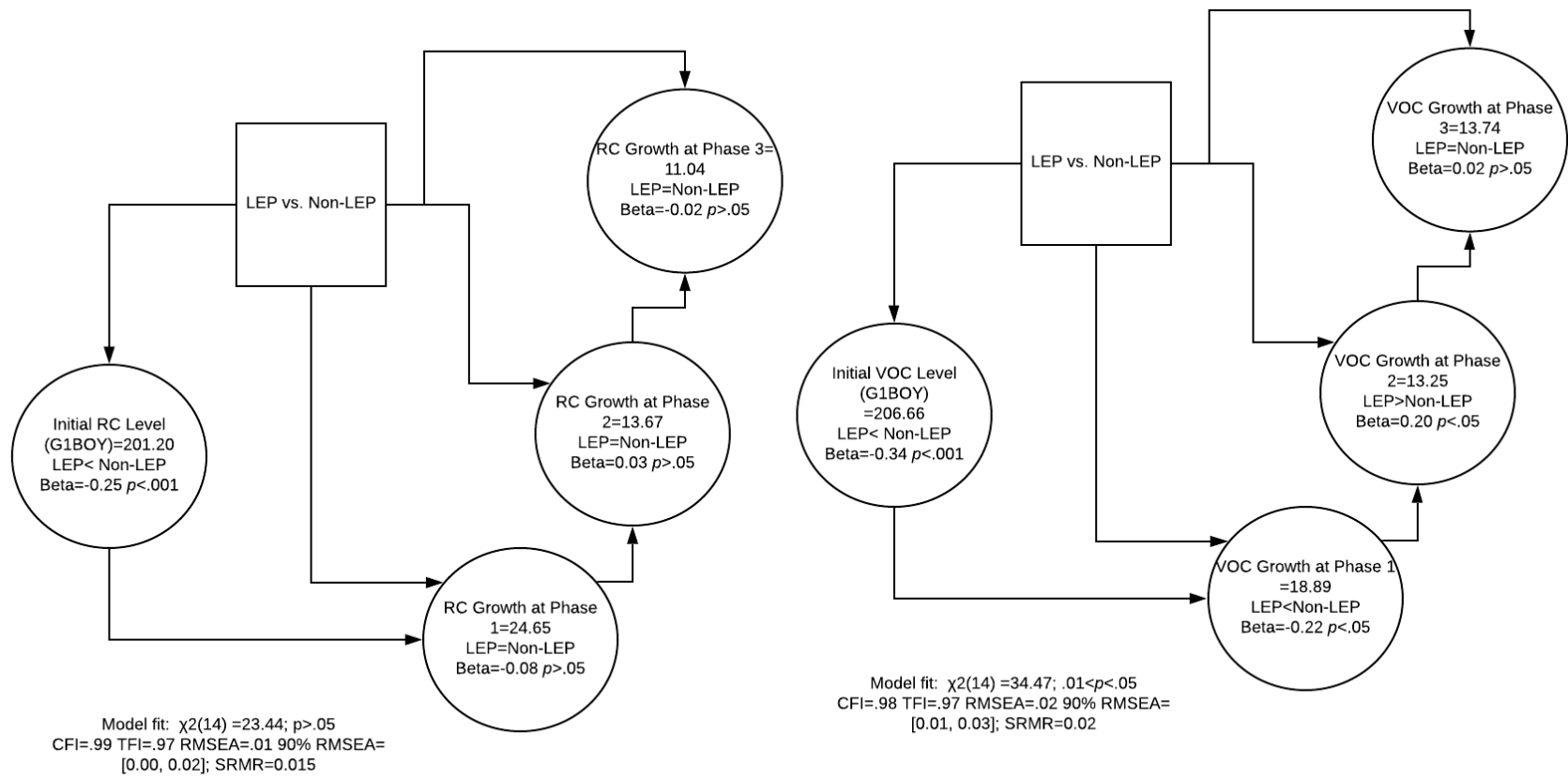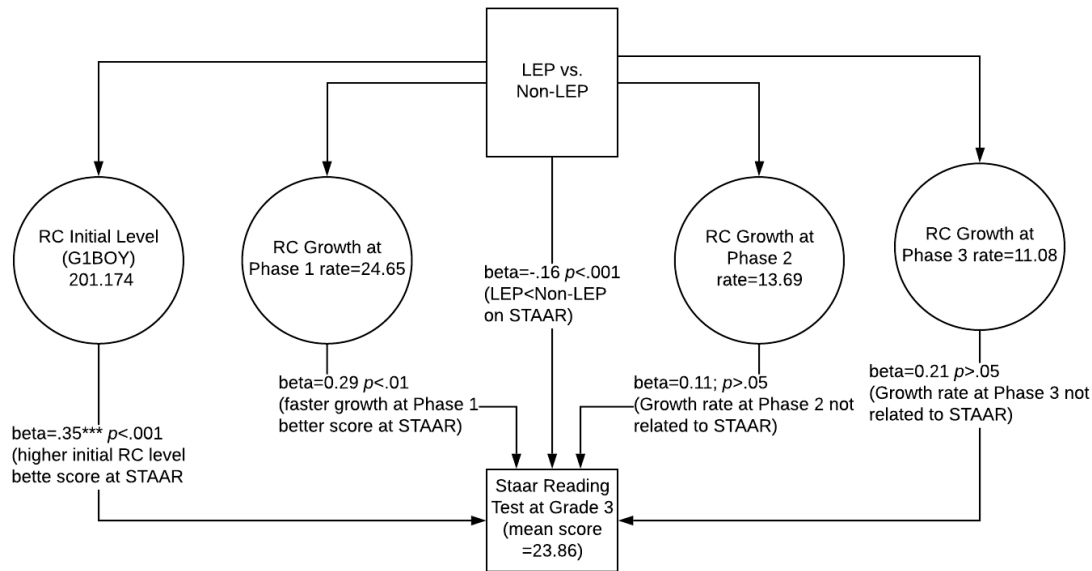*Figure 1a.* Reading comprehension growths of LEP and Non-LEP

Figure 1b. Vocabulary growths of LEP and Non-LEP

Figures 2. Predictive power of LEP status on growth factors of Istation reading comprehension (left) and vocabulary (right); RC=Reading comprehension; VOC=Vocabulary; Non-LEP= Proficient English Speakers; LEP= Students with limited English proficiency; BOY=Beginning of the year

LEP vs. Non-LEP

RC Initial Level (G1BOY) 201.174

RC Growth at Phase 1 rate=24.65

beta=-.16 $p<.001$ (LEP<Non-LEP on STAAR)

RC Growth at Phase 2 rate=13.69

RC Growth at Phase 3 rate=11.08

beta=0.29 $p<.01$ (faster growth at Phase 1 better score at STAAR)

beta=0.11; $p>.05$ (Growth rate at Phase 2 not related to STAAR)

beta=0.21 $p>.05$ (Growth rate at Phase 3 not related to STAAR)

beta=.35*** $p<.001$ (higher initial RC level bette score at STAAR

Staar Reading Test at Grade 3 (mean score =23.86)

Model fit: $\chi 2(17) =26.14$; $p>.05$
CFI=.99 TFI=.98 RMSEA=.01 90%
RMSEA= [0.00, 0.02]; SRMR=0.02

	Beta	S.E.	*P* values
Sum of indirect effect for Reading comprehension	-0.27	0.015	<.001
LEP → RC Intercept→ STAAR	-0.08	0.03	<.01
LEP → RC Growth G1 BOY to G1 EOY→ STAAR	-0.02	0.055	0.72
LEP → RC Growth G1 EOY to G2 EOY→ STAAR	-0.013	0.054	0.804
LEP → RC Growth G2 BOY to G3 EOY→ STAAR	-0.157	0.077	0.041

Figure 3. Predictive power of growth factors of Istation reading comprehension on Grade 3 STAAR reading results; RC=Reading comprehension; LEP= Students with limited English proficiency; BOY=Beginning of the year; EOY=End of the year.
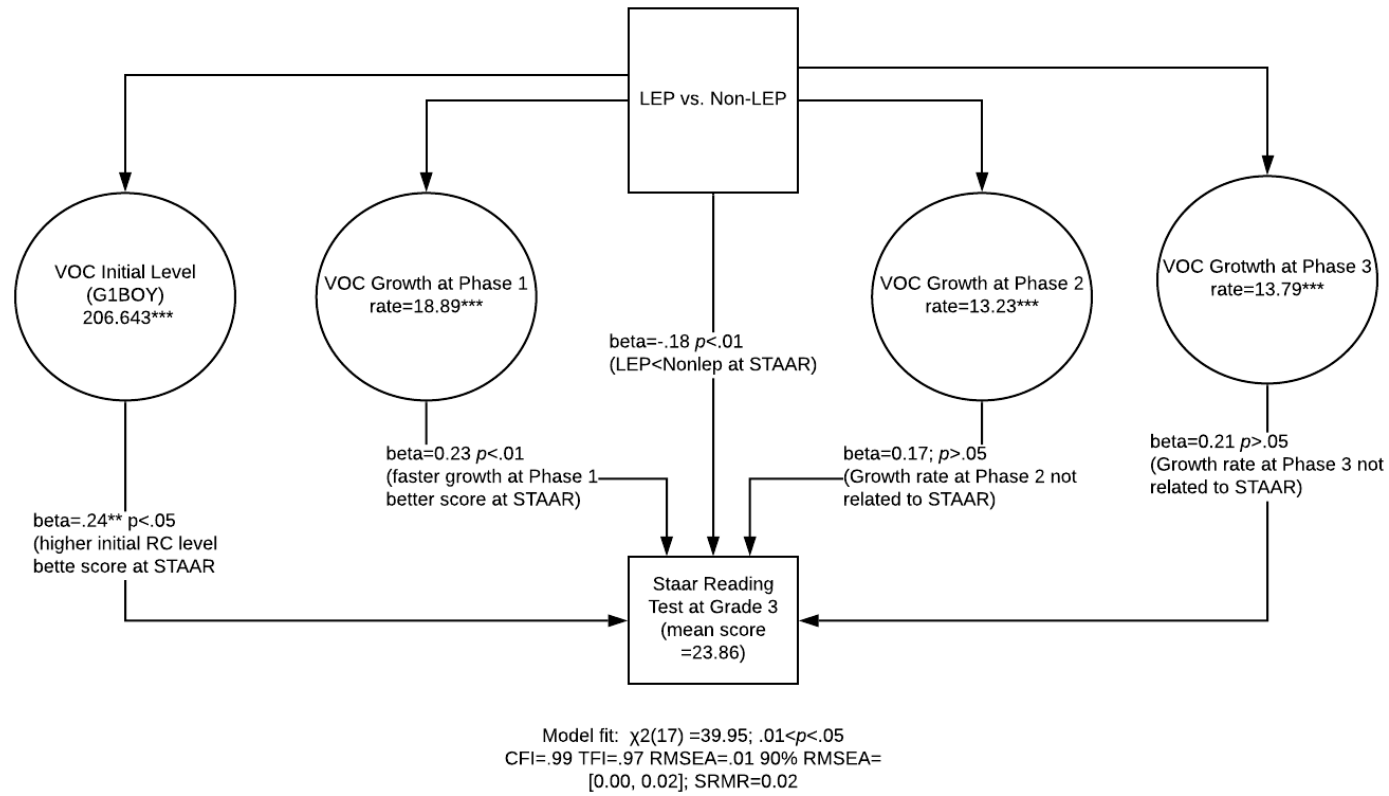
Sum of indirect effect for vocabulary			
	-0.09	0.203	0.642
LEP → VOC Intercept→ STAAR	-0.041	0.023	0.077
LEP → VOC Growth G1 BOY to G1 EOY→ STAAR	-0.07	0.042	0.098
LEP → VOC Growth G1 EOY to G2 EOY→ STAAR	0.022	0.215	0.918
LEP → VOC Growth G2 BOY to G3 EOY→ STAAR	-0.005	0.028	0.851

Figures 4. Predictive power of growth factors of Istation vocabulary on Grade 3 STAAR reading results; RC=Reading comprehension; LEP= Students with limited English proficiency; BOY=Beginning of the year; EOY=End of the year.

References

Haertel, E. H. (2018). Tests, Test Scores, and Constructs. *Educational Psychologist, 53*, 203-216.

Nichols, S. L., Glass, G. V., & Berliner, D. C. (2005). *High-stakes testing and student achievement: Problems for the No Child Left Behind Act* (No. EPSL-0509- 105-EPRU). Tempe: Education Policy Research Unit, Education Policy Studies Laboratory, College of Education, Division of Educational Leadership and Policy Studies, Arizona State University

Kieffer, M. J., Lesaux, N. K., Rivera, M., & Francis, D. J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*, 1168-1201.

Mathes, P., Torgesen, J., & Herron, J. (2016). *Computer adaptive testing system for continuous progress monitoring of reading growth for students pre-k through grade 3.*. Retrieved from: https://www.istation.com/Content/downloads/studies/er_technical_report.pdf

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Texas Education Agency (n.d.) *The Independent Evaluation of the Validity and Reliability of the State of Texas Assessments of Academic Readiness (STAAR) Assessments in grades 3-8, Parts 1 and 2*. Retrieved from https://tea.texas.gov/student.assessment/reports/