



Validity and Accuracy of Istation Screener for English Language Learners

Jennifer Morgan, Florina Erbeli, Marcia Montague

Department of Educational Psychology

INTRODUCTION

- Universal screening is federally mandated to identify students who may be at risk for falling below grade level expectations in reading as early as kindergarten.
- There is a gap in reading proficiency between English Language Learners (ELLs) and non-ELLs indicating a need for an accurate screener.
- Computer Adaptive Tests (CATs), such as Istation's Indicators of Progress-Early Reading (ISIP-ER), are widely used by school districts to screen students despite insufficient, independent peer-reviewed research.
- Research is limited on the long-term predictive validity and diagnostic accuracy of ISIP-ER for ELLs.

RESEARCH QUESTIONS

RQ1: What is the long-term predictive validity of ISIP-ER overall reading score in spring of kindergarten to predict scores on the STAAR reading assessment three years later, in the spring of third grade, for English Language Learners?

RQ2: What is the classification accuracy of an ISIP-ER kindergarten cut-score for predicting who will meet reading expectations on the STAAR assessment three years later, in the spring of third grade, for English Language Learners?

RQ3: What cut-score on the ISIP-ER overall reading maximizes classification accuracy for English Language Learners from the school district participating in the current study?

METHOD

Participants:

- 99 English Language Learners (ELLs) from 15 elementary schools in a suburban, central Texas school district

Measures:

ISIP-ER

- Computer Adaptive Test designed to progress monitor students at multiple times throughout the school year. It can be administered to a whole group in 40 minutes.
- Assesses listening comprehension, phonemic awareness, and letter knowledge (kindergarten) and generates an overall reading score.

STAAR

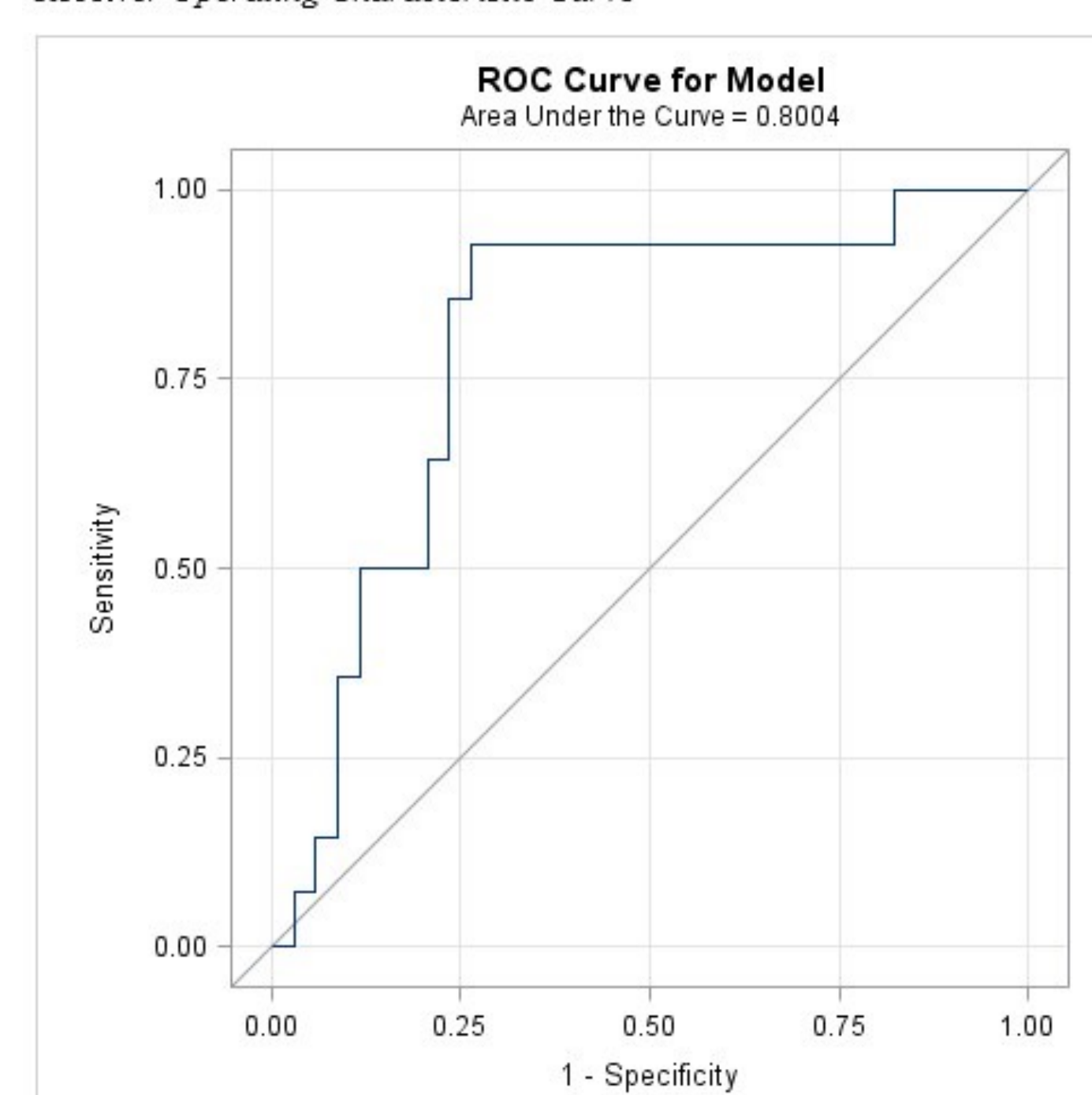
- State-wide reading comprehension assessment based on Texas state standards. It consists of four text passages followed by multiple-choice questions.

RESULTS

	N	Mean	SD	Range	Sk
ISIP-ER overall reading score Kindergarten (2014-2015)	99	182.54	15.03	141.58-213.26	-0.46
STAAR reading scale score Third grade (2017-2018)	99	1407.50	137.35	1165-1890	0.73

RQ1: Relationship between kindergarten ISIP-ER overall reading score and third grade STAAR reading was significant: Moderate correlation estimate of $r = .48$ (95% CI [.25, .66]).

Receiver Operating Characteristic Curve



	Tier 3 Cut-Score	Local Cut-Score
True Positive (ISIP-ER at-risk, STAAR fail)	38	32
True Negative (ISIP-ER not at risk, STAAR pass)	31	40
False Positive (ISIP-ER at risk, STAAR pass)	27	18
False Negative (ISIP-ER not at risk, STAAR fail)	3	9
Sensitivity	0.93	0.78
Specificity	0.53	0.69
Overall Accuracy	0.70	0.73

RQ2 and RQ3: Sensitivity of the ISIP-ER kindergarten overall reading score in predicting meeting expectations on the STAAR test three years later was 0.93, and specificity was 0.53. Local cut-score was calculated to correspond to ISIP-ER Tier 3 cut-score (<190, below 20th percentile) that would result in sensitivity level around 0.80 and specificity around 0.70.

DISCUSSION

- **RQ1:** Results suggest moderate predictive validity of ISIP-ER in K to third-grade STAAR for ELLs.
- **RQ2:** Classification accuracy estimates were above recommended benchmarks for sensitivity and below recommended benchmarks for specificity. The sensitivity value of 0.93 means that 7% of the students who would go on not to pass the third-grade STAAR were not identified as at-risk at the end of K. The specificity value of 0.53 means that 47% of the students were identified as at-risk when, in fact, they were not.
- **RQ3:** Results suggest that the local school district needs to prioritize achieving a balance of sensitivity and specificity when using Istation for ELLs.

Limitations:

- This study did not account for intervention received by ELLs identified at-risk in K.
- Sample of ELLs was limited to one district to students who remained in the district from K through third grade.

Implications:

- Results support the utility of Istation for identifying ELL students at-risk for falling below grade-level expectations in reading.
- Results indicated a need to supplement Istation with additional screeners for ELLs who were not found to be at-risk.

Contact:

Email: jthacker06@tamu.edu

Twitter: @JennT_SpEdTeach

Acknowledgements: This research was supported, in part, by a grant from the U.S. Department of Education, Office of Special Education Programs (H325D210073). Views expressed herein are those of the authors and have neither been reviewed nor approved by the granting agencies.