Comparison of Screening Methods for Computer Adaptive Tests

Emily Forcht & Ethan R. Van Norman PhD

Lehigh University, Bethlehem, PA

INTRODUCTION

- Schools may implement preventive frameworks such as multi-tiered systems of support (MTSS) or Response to Intervention (RTI) in education to identify students who are at risk for academic failure and then provide those students with supplementary support.
- The first step in preventive frameworks is universal screening where educators administer brief assessments to all students in a building (Berkeley et al., 2009).
- Unfortunately, the screening process is not perfect, and errors do occur.
- To decrease the number of screening errors, research has examined the idea of using multiple measures to inform screening outcomes (Fuchs et al., 2012). Typically, using multiple screening measures improves the accuracy of the screening process (Van Meveren et al., 2020).
- One method: Gated Screening Framework → multiple phases of screening
  - First phase, all students are administered a screener and performance is evaluated.
  - Second phase, students who are determined as "at risk" are administered a second screener.
- The purpose of the current study is to compare the diagnostic accuracy of using a single computer adaptive test (CAT), Star Reading or Star Math, or a combination of the two, in a gated screening framework, to predict end-of-year proficiency on reading and math tests.
- **Research Questions:** (1) What is the difference in diagnostic accuracy between a single screening method or a gated screening method to predict end-of-year proficiency in reading? (2) What is the difference in diagnostic accuracy between a single screening method or a gated screening framework to predict end-of-year proficiency in math?

METHOD

Participants

- Extant dataset managed by Renaissance Learning
- 13,009 students in Grades 3-8
  - 2058 Grade 3, 2081 Grade 4, 2074 Grade 5, 1991 Grade 6, 2417 Grade 7, and 2388 Grade 8

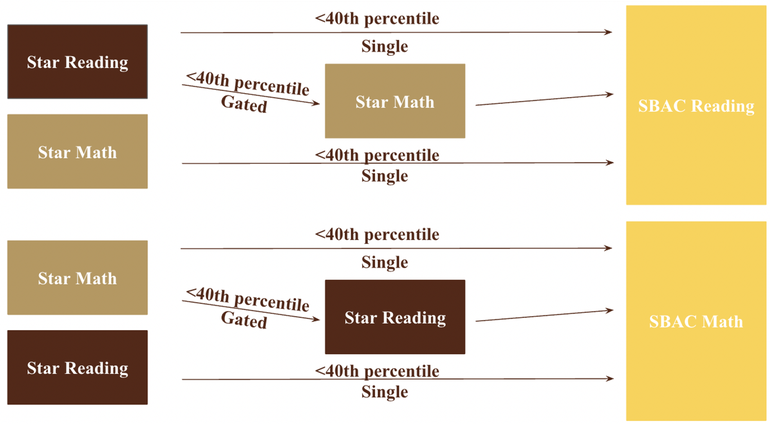
Measures

- Star Reading (SR): CAT designed to assess reading achievement, developed by Renaissance Learning
- Star Math (SM): CAT designed to assess mathematical skills, developed by Renaissance Learning
- Smarter Balanced Assessment Consortium (SBAC): end-of-year statewide test that measures proficiency in reading and math, during the school year

ANALYSIS

- **Single Screening Method:** SR predicting proficiency on SBAC Reading; SR predicting proficiency on SBAC Math; SM predicting proficiency on SBAC Math; SM predicting proficiency on SBAC Reading
- **Gated Screening Method**: SR and SM to predict SBAC Reading; SM and Star Reading to predict SBAC Math.
  - If a case scores below the 40th percentile on the first CAT, performance on the second CAT will be considered.
  - The case will then be considered at risk if they score below the 40th percentile on the additional screening measure.
- The distribution of cases in four categories, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) were displayed in a contingency table.
- Sensitivity (Sn): (TP)/ (TP+FN), Specificity (Sp): (TN) /(TN+FP), and Base Rate (BR): (TP+FN)/(TP+FN+FP+TN) were estimated based on the contingency tables.

SCREENING METHODS DECISION TREES



Diagnostic Accuracy Statistics for Reading Outcomes

Grade	Method	BR	TP	TN	FP	FN	Sn	Sp
3	Star Reading	.47	679	949	142	288	**.70**	**.87**
	Star Math	.47	526	1014	77	441	.54	.93
	SR + SM	.47	449	1056	35	518	.46	.97
4	Star Reading	.42	555	1096	119	311	.64	.90
	Star Math	.42	425	1102	113	441	.49	.91
	SR + SM	.42	338	1181	34	528	.39	.97
5	Star Reading	.39	582	1098	171	223	**.72**	**.87**
	Star Math	.39	386	1168	101	419	.48	.92
	SR + SM	.39	335	1224	45	470	.42	.96
6	Star Reading	.40	602	1006	184	199	**.75**	**.85**
	Star Math	.40	437	1069	121	364	.55	.90
	SR + SM	.40	375	1142	48	426	.47	.96
7	Star Reading	.39	705	1256	216	240	**.75**	**.85**
	Star Math	.39	533	1337	135	412	.56	.91
	SR + SM	.39	450	1424	48	495	.48	.97
8	Star Reading	.39	706	1219	242	221	**.76**	**.83**
	Star Math	.39	516	1287	174	411	.56	.88
	SR + SM	.39	438	1394	67	489	.47	.95

RESULTS

- The single screening method did not show extreme differences between screening across topic areas and grades.
- The only condition to yield sufficient Sn and Sp (>.70) was when a single reading screener predicted reading outcomes in all grades except grade 4.
- No conditions for math outcomes met sufficient Sn and Sp (>.70). However, using a single reading screener showed slightly better results for Sn than a single math screener.
- The gated screening method yielded lower Sn and higher Sp values than the single screening method across grades and topic areas.
- The gated screening method when predicting both reading and math outcomes yielded extremely high Sp values across all grades (range = .95-.99). On average, the Sp values were slightly higher for math outcomes (Sp = .98) than for reading outcomes (Sp = .96).

Diagnostic Accuracy Statistics for Math Outcomes

Grade	Method	BR	TP	TN	FP	FN	Sn	Sp
3	Star Math	.45	524	1055	79	400	.57	.93
	Star Reading	.45	617	930	204	307	.67	.82
	SM + SR	.45	438	1088	46	486	.47	.96
4	Star Math	.44	469	1105	69	438	.52	.94
	Star Reading	.44	512	1012	162	395	.56	.86
	SM + SR	.44	338	1140	34	569	.37	.97
5	Star Math	.51	459	985	28	602	.43	.97
	Star Reading	.51	635	895	118	426	.60	.88
	SM + SR	.51	367	1000	13	694	.35	.99
6	Star Math	.48	506	988	52	445	.53	.95
	Star Reading	.48	621	875	165	330	.65	.84
	SM + SR	.48	404	1021	19	547	.42	.98
7	Star Math	.49	626	1180	42	569	.52	.97
	Star Reading	.49	747	1048	174	448	.63	.86
	SM + SR	.40	480	1204	18	715	.40	.99
8	Star Math	.54	628	1037	62	661	.49	.94
	Star Reading	.54	793	944	155	496	.62	.86
	SM + SR	.54	487	1081	18	802	.38	.98

DISCUSSION

- Overall, the optimal screening method was using a single reading screener to predict reading proficiency as this is the only method that yielded sufficient Sn and Sp (>.70).

Single or Gated Screening Method?

- A single reading screener produced more optimal results than the gated screening method when predicting reading proficiency. However, when predicting math proficiency, the single and gated screening methods produced similar results.
- The gated screening method may be more beneficial over the single screening method if the goal of screening is to inform "rule-out" decisions and to identify students who are at low risk for academic difficulties.
- Educators must decide the value in administering additional screening measures to predict proficiency in reading and math and if the gated screening method is worth the additional resources.

Predicting Reading Proficiency

- Using a single reading screener is best. Using a single math screener may be beneficial if the goal of screening is to inform "rule-out" decisions.

Predicting Math Proficiency

- Using both a single math screener and a single reading screener produced similar results. Slightly more ideal results were seen when using a single reading screener.

Questions? Please email Emily Forcht erf220@lehigh.edu