

Multitudes v3 (2024/25) Prediction Models

Julian M. Siebert

2025-07-17

Cautionary Note

This document is a WORK IN PROGRESS and will be updated as new data (e.g., item recalibrations, additional demographic information) become available.

To Do

- ~~recompute scores for universal screening tasks with updated 2025 item parameters once available~~
- ~~filter out implausible RAN score~~
- ~~add 23/24 data~~
- ~~align grades between years~~
- ~~align tasknames between years~~
- ~~re-run 23/24 models~~
- ~~run 24/25 models~~
- ~~run combined model~~
- run two-year model
- recompute scores for remaining tasks with updated 2025 item parameters once available
- try out different definitions of risk (one/both lgs.)
- try different aggregations of tasks (max, mean, etc.) for skill-specific/language-independent predictors

Table of contents

1	PART I - Re-analysis of 2023/24 Data	4
1.1	Sample	4
1.2	Descriptives	6
1.2.1	Distributions of English Multitudes Tasks	6
1.2.2	Distributions of Spanish Multitudes Tasks	7
1.2.3	Distributions of Outcome Measures	8
1.2.4	Correlations	9
1.3	Task Selection	15
1.3.1	English-to-English (English to English)	15
1.3.2	Spanish-to-Spanish (Spanish to Spanish)	16
1.4	Models Evaluations - ROC Curves	17
1.4.1	English-to-English - Kindergarten	17
1.4.2	English-to-English - Grade 1	18
1.4.3	English-to-English - Grade 2	19
1.4.4	Spanish-to-Spanish - Kindergarten	20
1.4.5	Spanish-to-Spanish - Grade 1	21
1.4.6	Spanish-to-Spanish - Grade 2	22
1.5	Models Evaluations - Sensitivity and Specificity	23
2	PART II - Analysis of 2024/25 Data	24
2.1	Sample	24
2.2	Descriptives	26
2.2.1	Distributions of English Multitudes Tasks	26
2.2.2	Distributions of Spanish Multitudes Tasks	27
2.2.3	Distributions of Outcome Measures	28
2.2.4	Correlations	29
2.3	Task Selection	35
2.3.1	English-to-English (English to English)	35
2.3.2	Spanish-to-Spanish (Spanish to Spanish)	36
2.4	Models Evaluations - ROC Curves	37
2.4.1	English-to-English - Kindergarten	37
2.4.2	English-to-English - Grade 1	38
2.4.3	English-to-English - Grade 2	39
2.4.4	Spanish-to-Spanish - Kindergarten	40
2.4.5	Spanish-to-Spanish - Grade 1	41
2.4.6	Spanish-to-Spanish - Grade 2	42
2.5	Models Evaluations - Sensitivity and Specificity	43
3	PART III - Combined Analysis of 2023/24 and 2024/25 Data	44
3.1	Sample	44

3.2	Descriptives	46
3.2.1	Distributions of English Multitudes Tasks	46
3.2.2	Distributions of Spanish Multitudes Tasks	47
3.2.3	Distributions of Outcome Measures	48
3.2.4	Correlations	49
3.3	Task Selection	55
3.3.1	English-to-English (English to English)	55
3.3.2	Spanish-to-Spanish (Spanish to Spanish)	56
3.4	Models Evaluations - ROC Curves	57
3.4.1	English-to-English - Kindergarten	57
3.4.2	English-to-English - Grade 1	58
3.4.3	English-to-English - Grade 2	59
3.4.4	Spanish-to-Spanish - Kindergarten	60
3.4.5	Spanish-to-Spanish - Grade 1	61
3.4.6	Spanish-to-Spanish - Grade 2	62
3.5	Models Evaluations - Sensitivity and Specificity	63
4	Part IV - Evaluating 23/24 Model With 24/25 Data	64

1 PART I - Re-analysis of 2023/24 Data

1.1 Sample

Throughout, I only included tasks taken by at least 50 students per language per grade, so as to not distort the task selection analyses.

- $N = 2701$
- 28 schools
- 15 districts

Table 1: Demographic Characteristics of the AY 23/24 Sample (N = 2,701) by Grade and Screening Language

Characteristic	English			Spanish		
	K N = 723	G1 N = 546	G2 N = 704	K N = 354	G1 N = 233	G2 N = 254
Gender						
Female	373 (52%)	257 (47%)	356 (51%)	192 (54%)	120 (52%)	141 (56%)
Male	350 (48%)	289 (53%)	346 (49%)	162 (46%)	113 (48%)	113 (44%)
Unknown	0	0	2			
Race						
Asian	86 (13%)	67 (12%)	65 (9.3%)	12 (4.0%)	3 (1.3%)	5 (2.0%)
Black/African American	92 (14%)	54 (9.9%)	98 (14%)	2 (0.7%)	0 (0%)	2 (0.8%)
Not reported	157 (23%)	110 (20%)	142 (20%)	208 (69%)	154 (68%)	149 (60%)
Other	94 (14%)	74 (14%)	85 (12%)	6 (2.0%)	7 (3.1%)	6 (2.4%)
White	245 (36%)	238 (44%)	307 (44%)	75 (25%)	62 (27%)	88 (35%)
Unknown	49	3	7	51	7	4
Ethnicity						
Hispanic/Latin(o/a)	413 (58%)	281 (52%)	383 (55%)	330 (94%)	220 (94%)	237 (93%)
Not Hispanic/Latin(o/a)	299 (42%)	263 (48%)	316 (45%)	22 (6.3%)	13 (5.6%)	17 (6.7%)
Unknown	11	2	5	2	0	0
ELPD						
EL	182 (27%)	94 (17%)	122 (19%)	267 (76%)	178 (77%)	173 (69%)
EO	416 (62%)	393 (73%)	474 (72%)	24 (6.8%)	19 (8.3%)	29 (12%)
EP	51 (7.6%)	51 (9.5%)	61 (9.3%)	42 (12%)	33 (14%)	45 (18%)
TBD	20 (3.0%)	1 (0.2%)	1 (0.2%)	18 (5.1%)	0 (0%)	2 (0.8%)
Unknown	54	7	46	3	3	5
Ever IEP/504	52 (8.6%)	43 (9.0%)	57 (11%)	31 (9.5%)	21 (9.7%)	20 (10%)
Unknown	117	69	204	26	17	60
Home Language						
English	496 (69%)	411 (76%)	472 (72%)	42 (12%)	24 (10%)	34 (14%)
Other	64 (9.0%)	51 (9.4%)	63 (9.6%)	5 (1.4%)	2 (0.9%)	4 (1.6%)
Spanish	154 (22%)	81 (15%)	122 (19%)	299 (86%)	205 (89%)	209 (85%)
Unknown	9	3	47	8	2	7

1.2 Descriptives

1.2.1 Distributions of English Multitudes Tasks

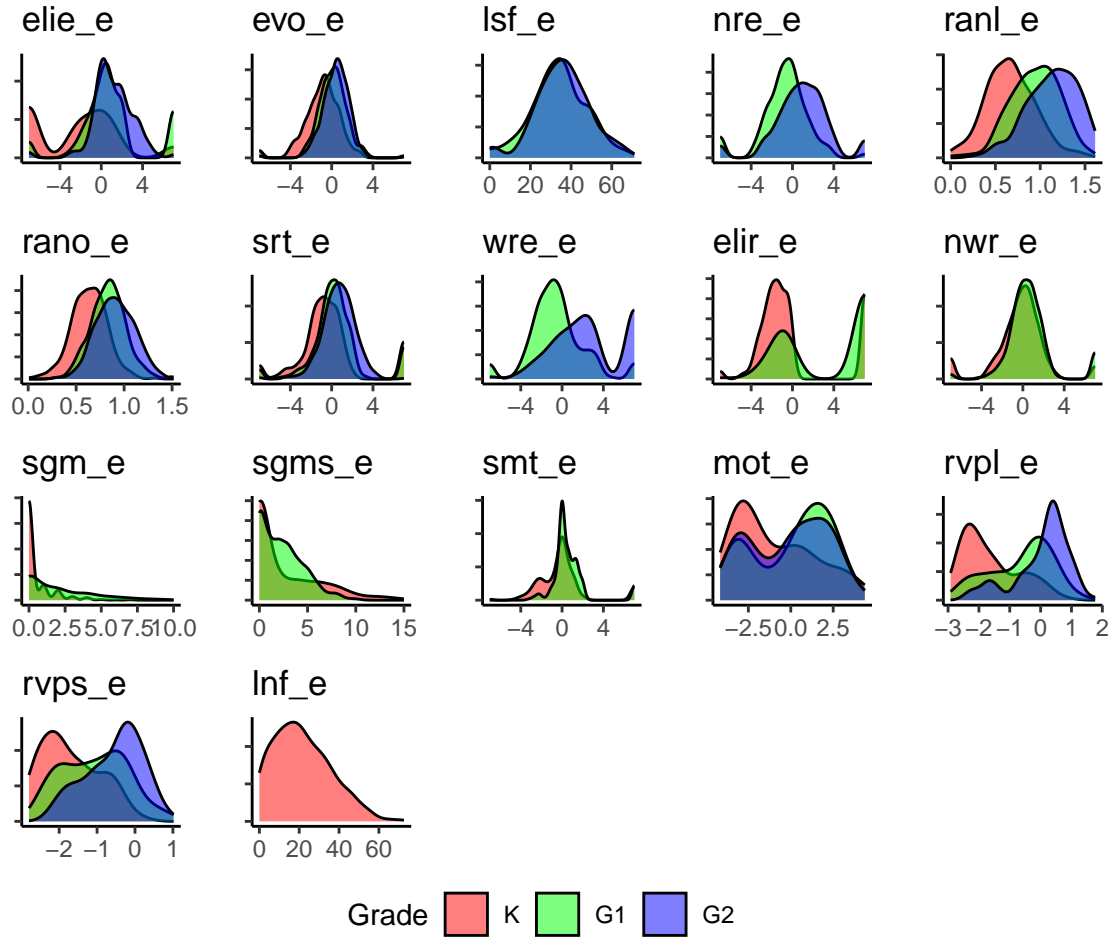


Figure 1: Distributions of English Multitudes Tasks by Grade (Fall 2023)

1.2.2 Distributions of Spanish Multitudes Tasks

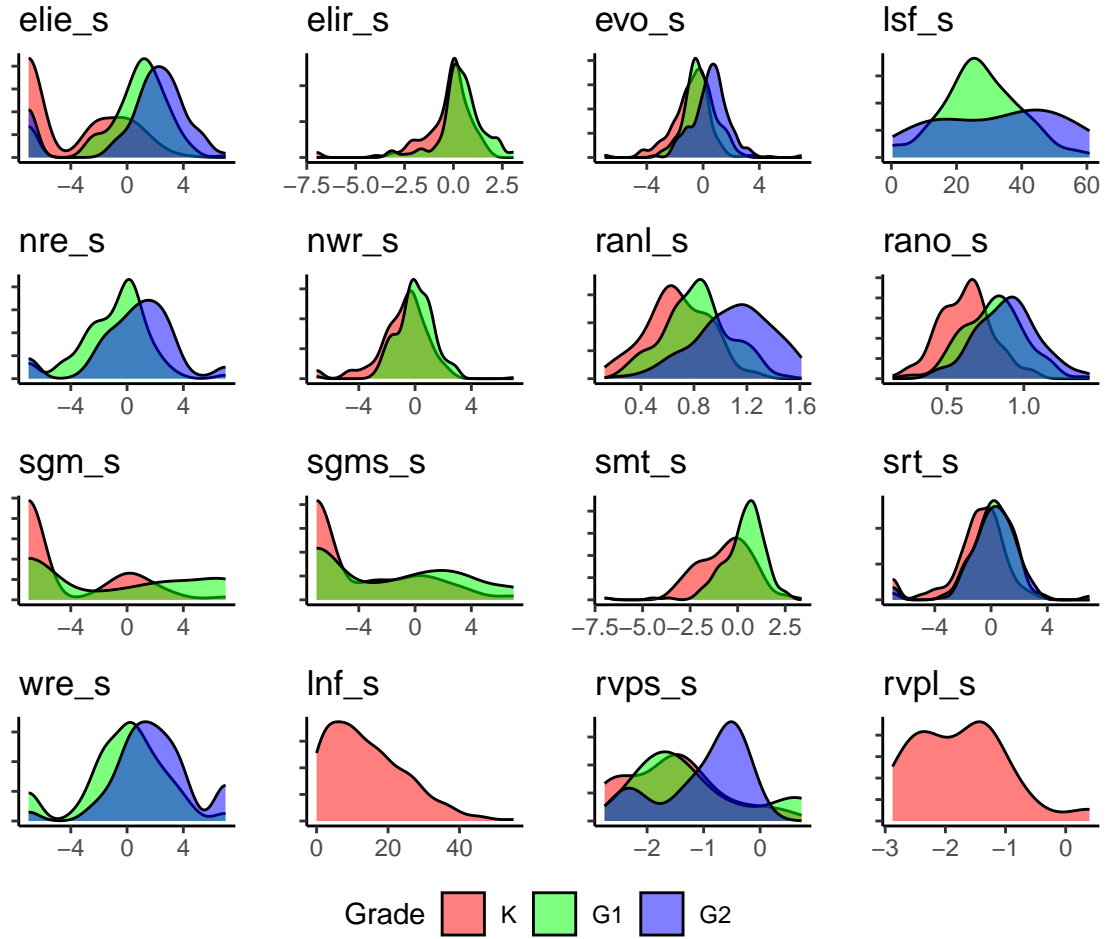


Figure 2: Distributions of Spanish Multitudes Tasks by Grade (Fall 2023)

Table 2: Comparison of Proportions of Struggling Readers (\leq 20th Percentile) in the Study Sample Based on Overall Multitudes Sample and on Published Woodcock-Johnson/Muñoz Norms (Spring 2024).

Reference	English			Spanish		
	K N = 754	G1 N = 590	G2 N = 780	K N = 375	G1 N = 224	G2 N = 211
Norm-based						
Not struggling	470 (62%)	415 (70%)	541 (69%)	169 (45%)	145 (65%)	106 (50%)
Struggling	284 (38%)	175 (30%)	239 (31%)	206 (55%)	79 (35%)	105 (50%)
Sample-based						
Not struggling	619 (82%)	539 (91%)	635 (81%)	300 (80%)	191 (85%)	162 (77%)
Struggling	135 (18%)	51 (8.6%)	145 (19%)	75 (20%)	33 (15%)	49 (23%)

1.2.3 Distributions of Outcome Measures

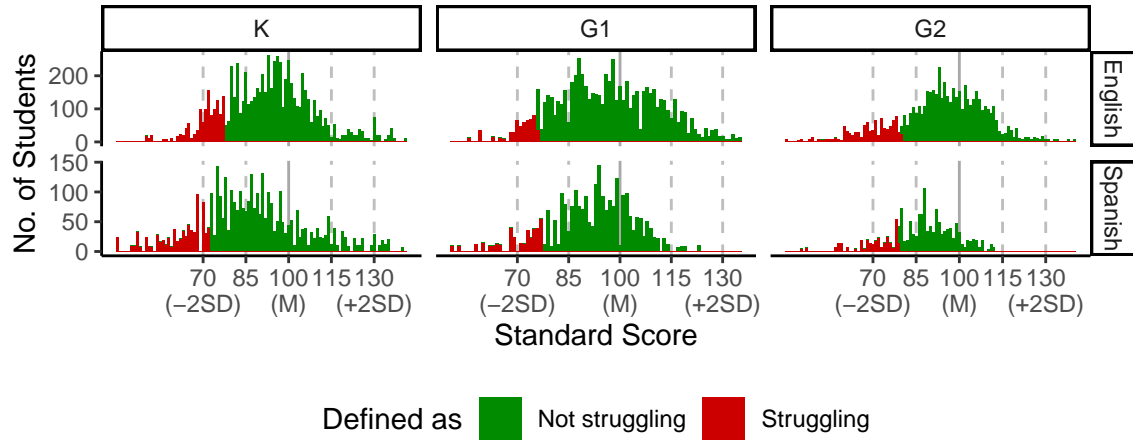
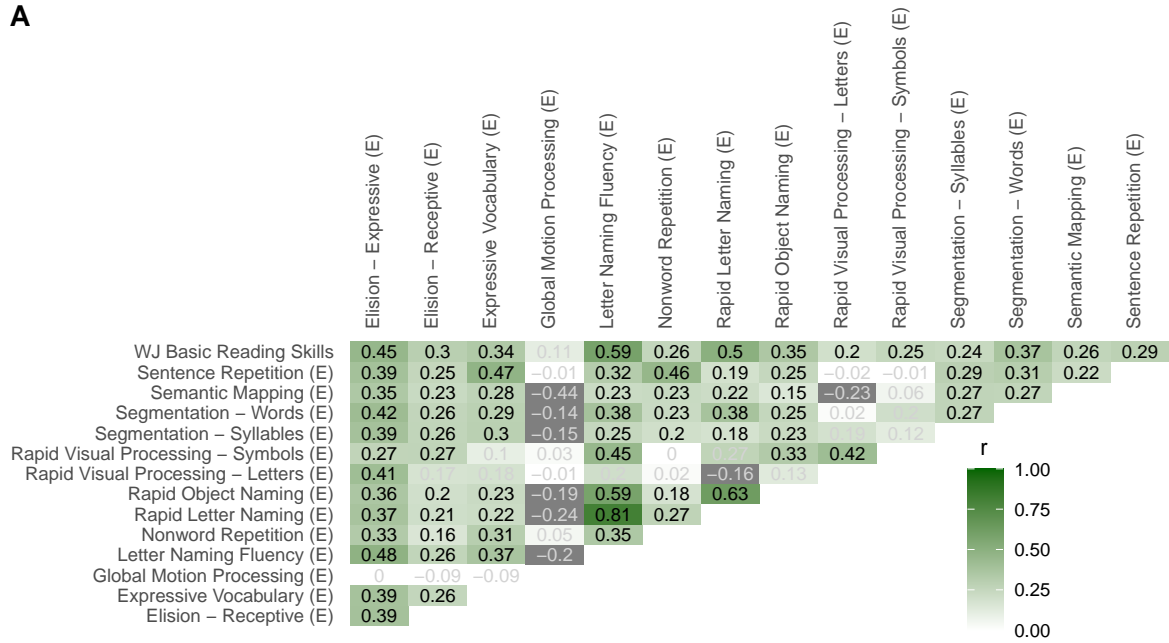


Figure 3: Distribution of Woodcock-Johnson/-Muñoz Basic Reading Skills Cluster (for Kindergarten) Broad Reading Cluster (for Grades 1 and 2) by Grade and Language with Struggling Readers (\leq 20th Percentile) Highlighted (Spring 2024).

1.2.4 Correlations

1.2.4.1 Kindergarten - English

A



B

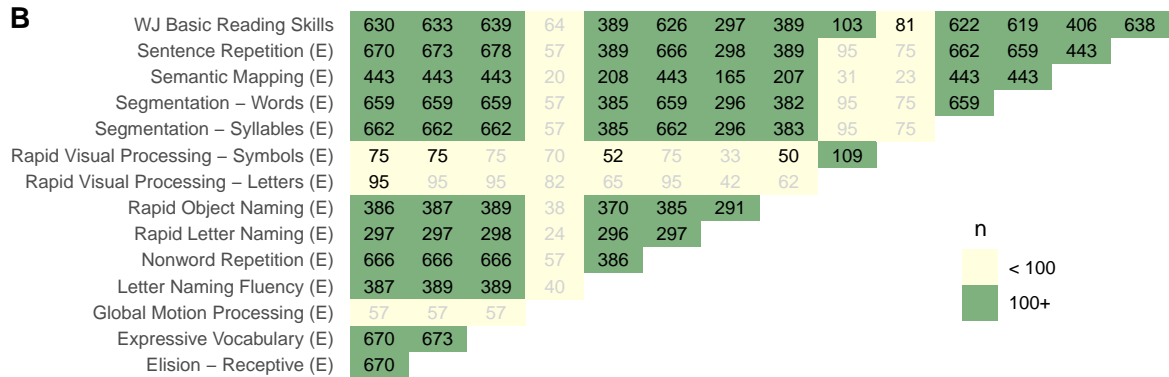
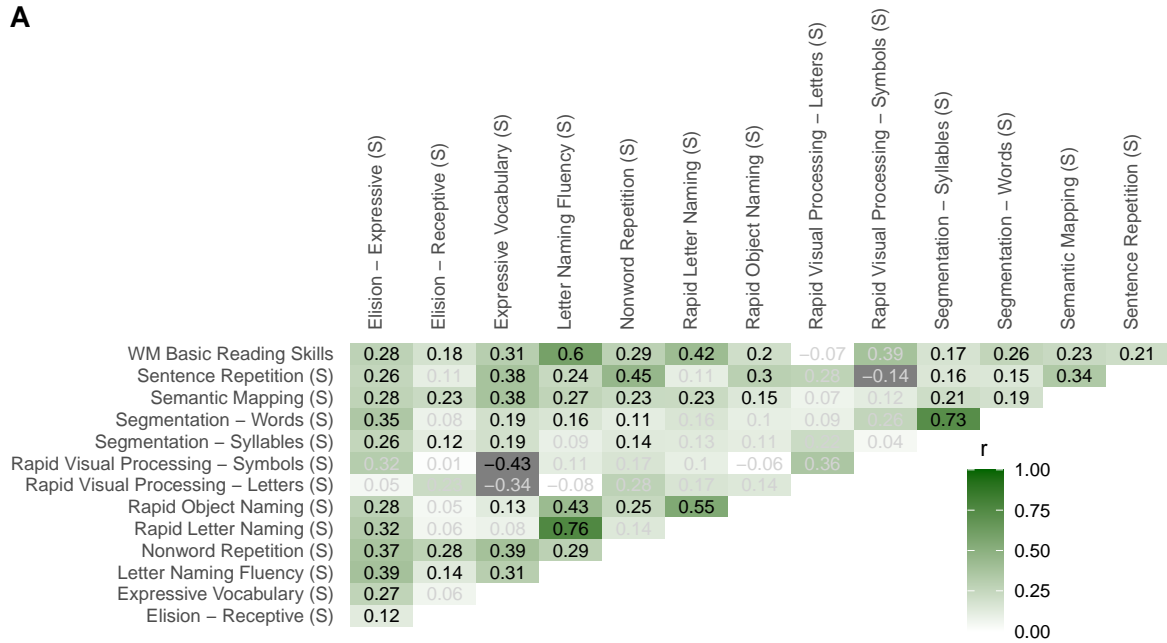


Figure 4: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Kindergarten; AY 2023/24).

1.2.4.2 Kindergarten - Spanish

A



B

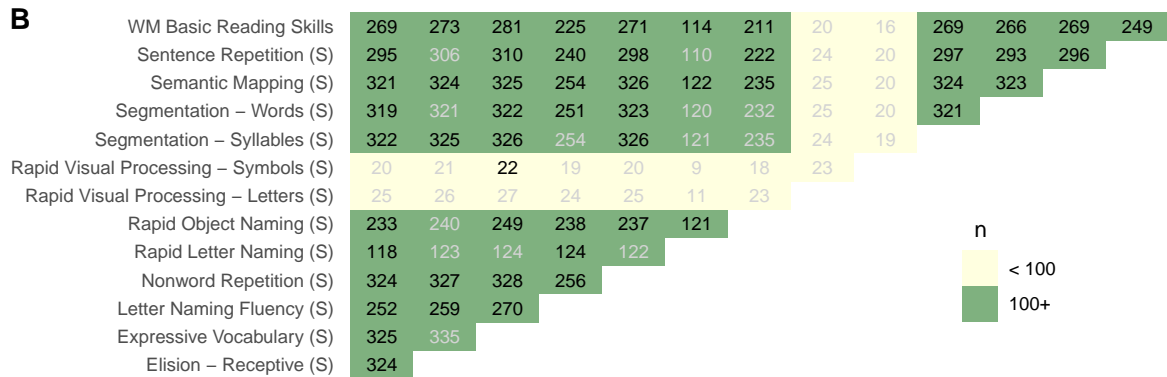
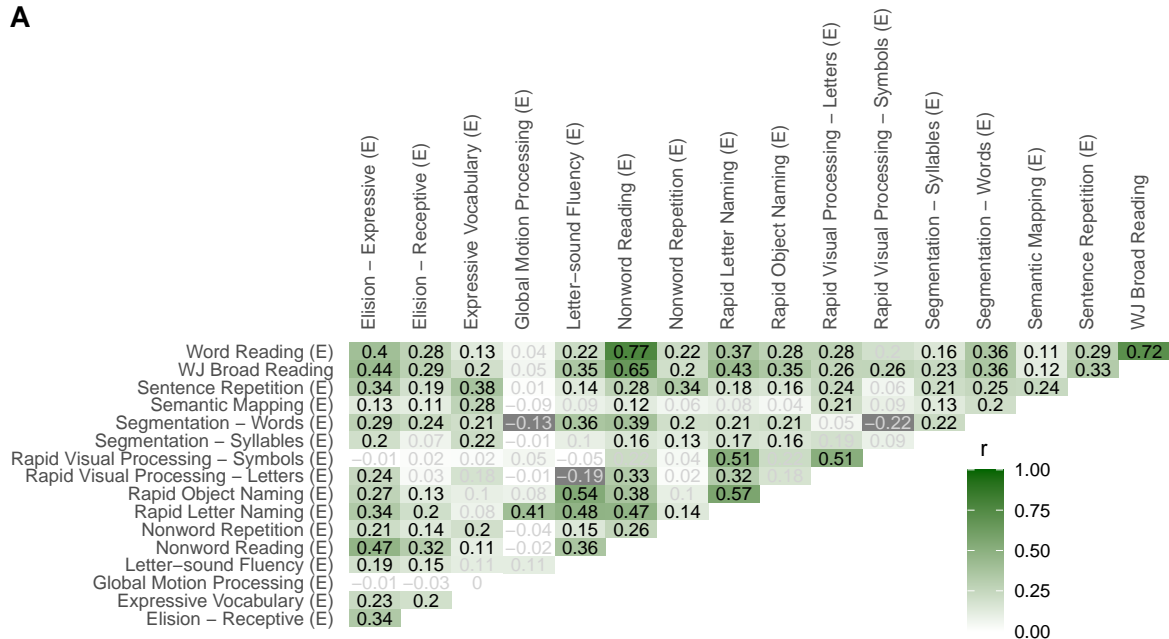


Figure 5: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Kindergarten; AY 2023/24).

1.2.4.3 Grade 1 - English

A



B

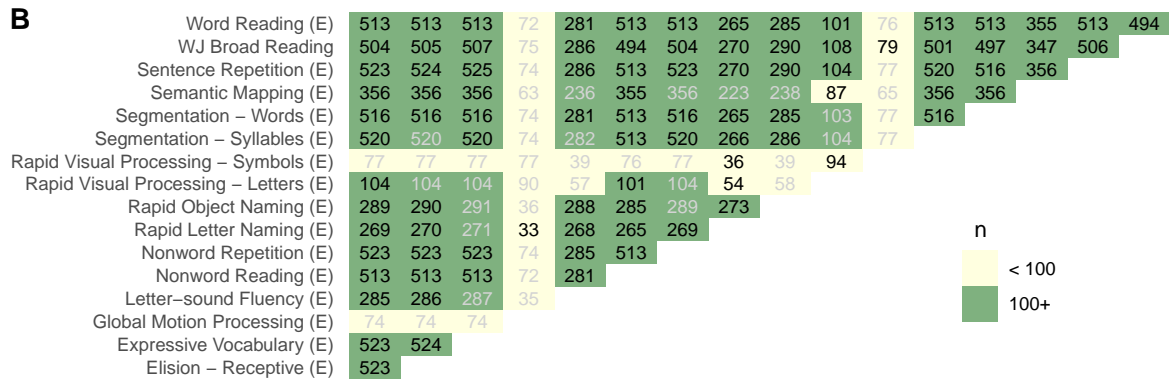
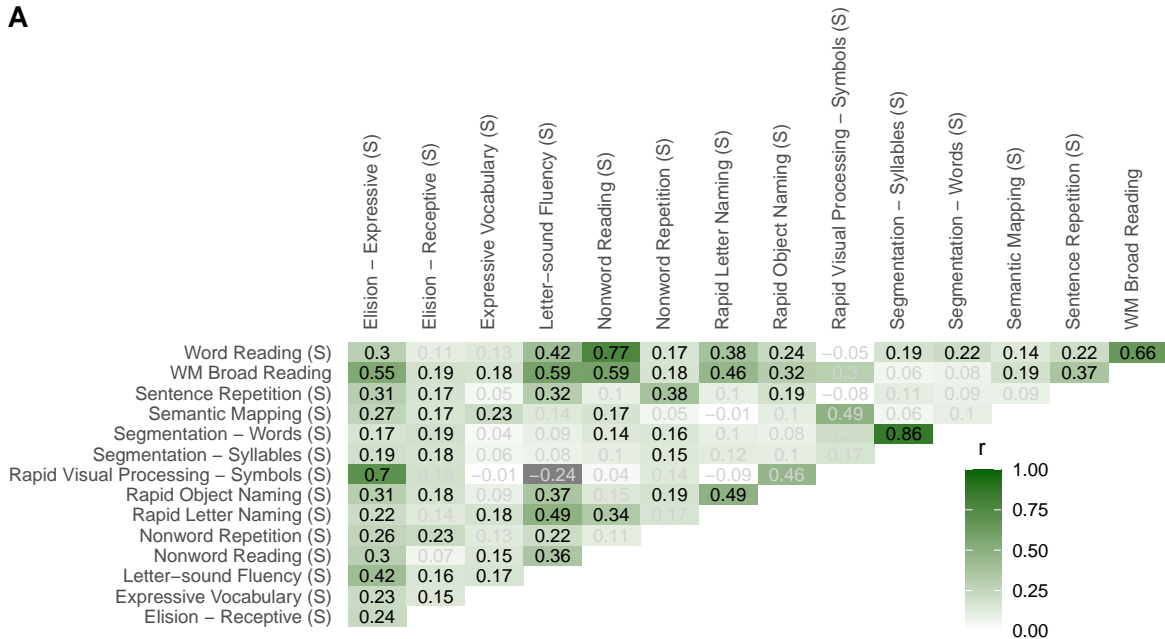


Figure 6: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Grade 1; AY 2023/24).

1.2.4.4 Grade 1 - Spanish

A



B

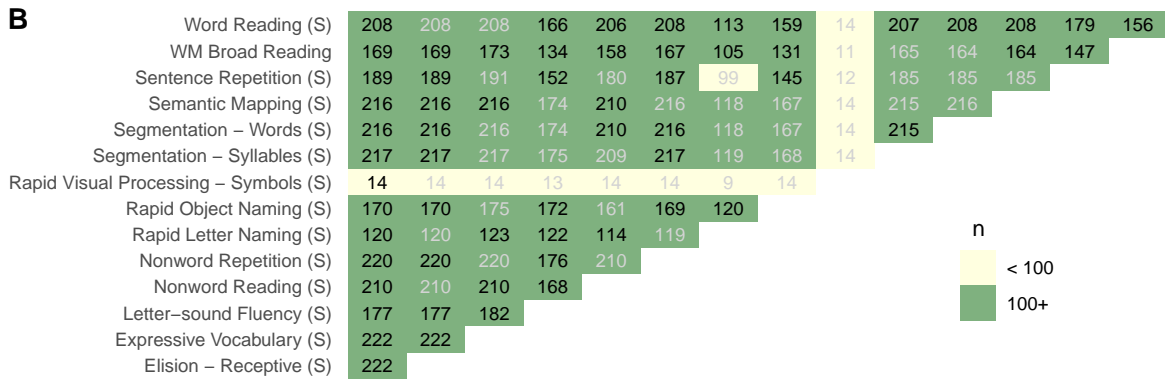
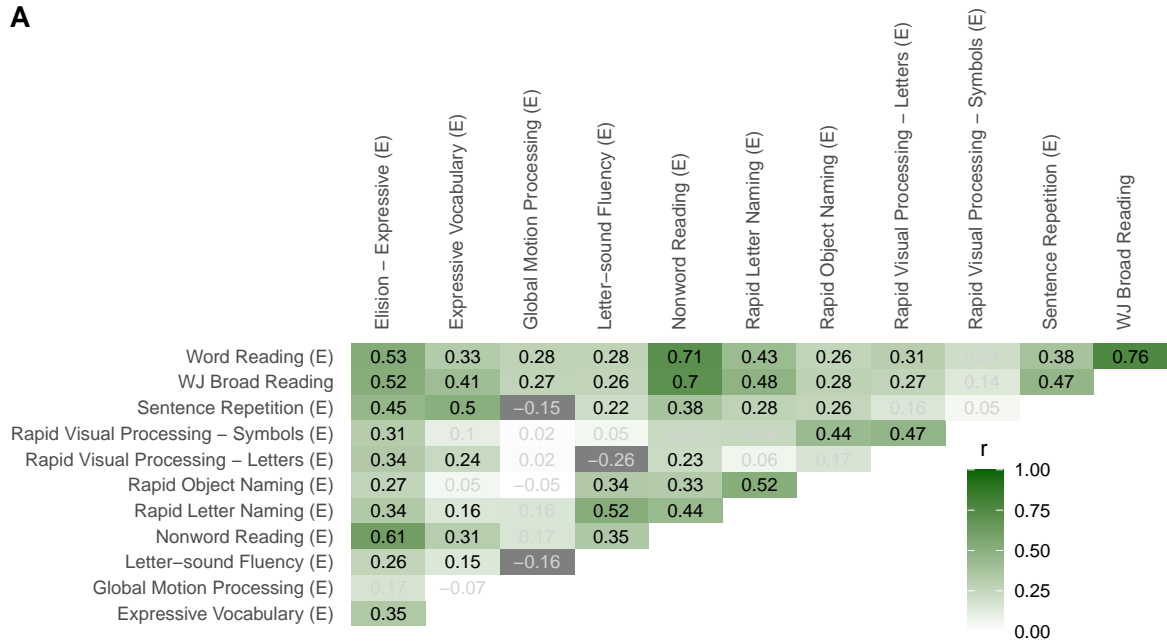


Figure 7: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Grade 1; AY 2023/24).

1.2.4.5 Grade 2 - English

A



B

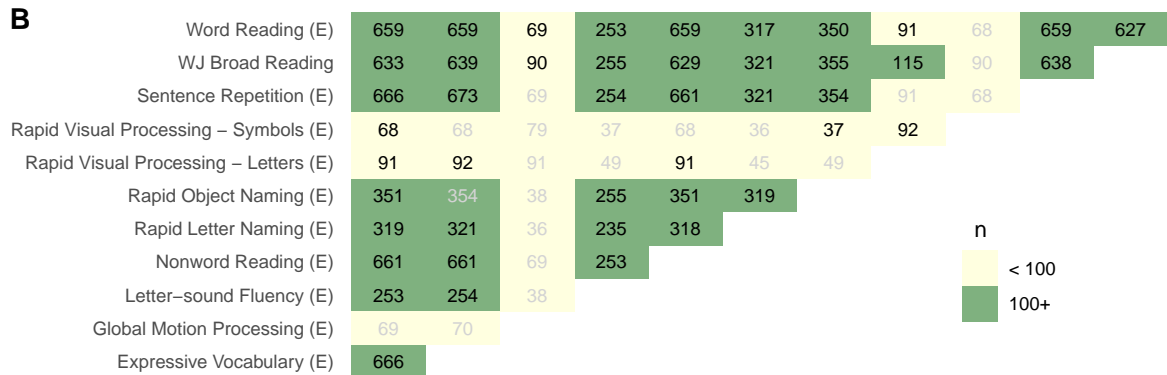
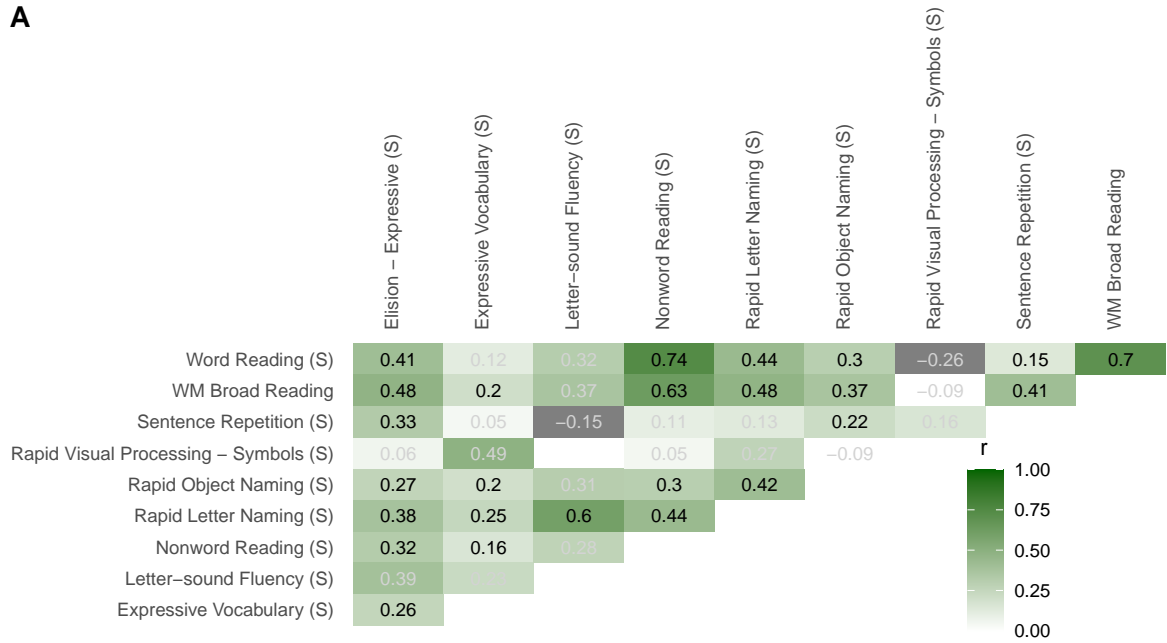


Figure 8: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Grade 2; AY 2023/24).

1.2.4.6 Grade 2 - Spanish

A



B

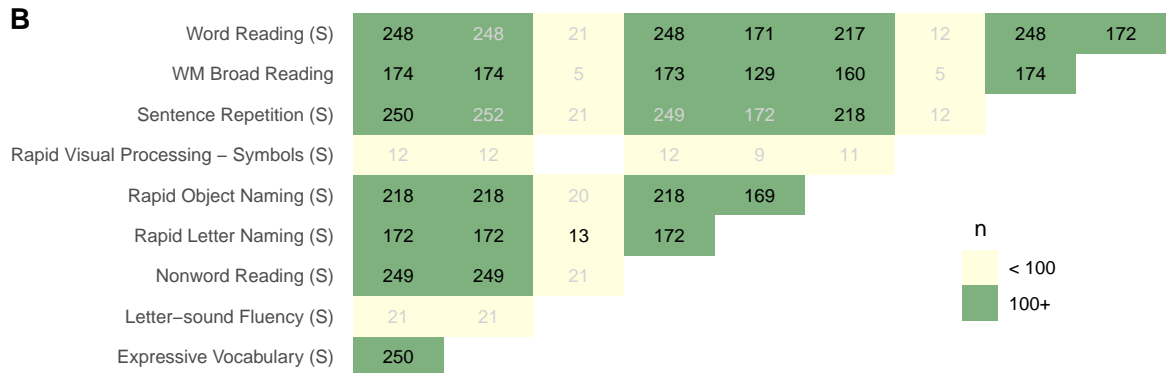


Figure 9: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Grade 2; AY 2023/24).

1.3 Task Selection

1.3.1 English-to-English (English to English)

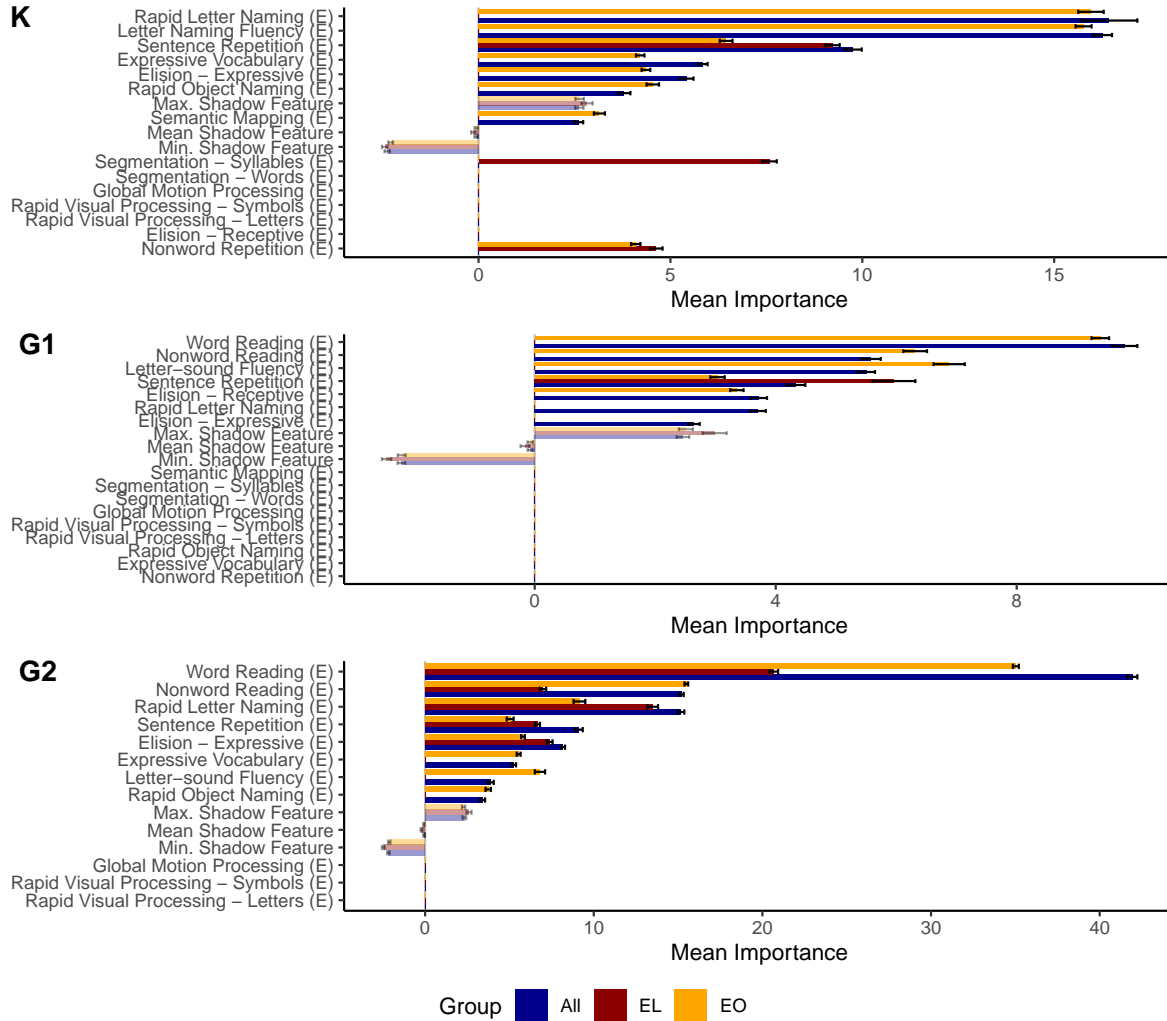


Figure 10: Mean Task Importance Obtained From Boruta Algorithm for English Predictions, with Results Disaggregated by English Proficiency Designation (AY 2023/24).

1.3.2 Spanish-to-Spanish (Spanish to Spanish)

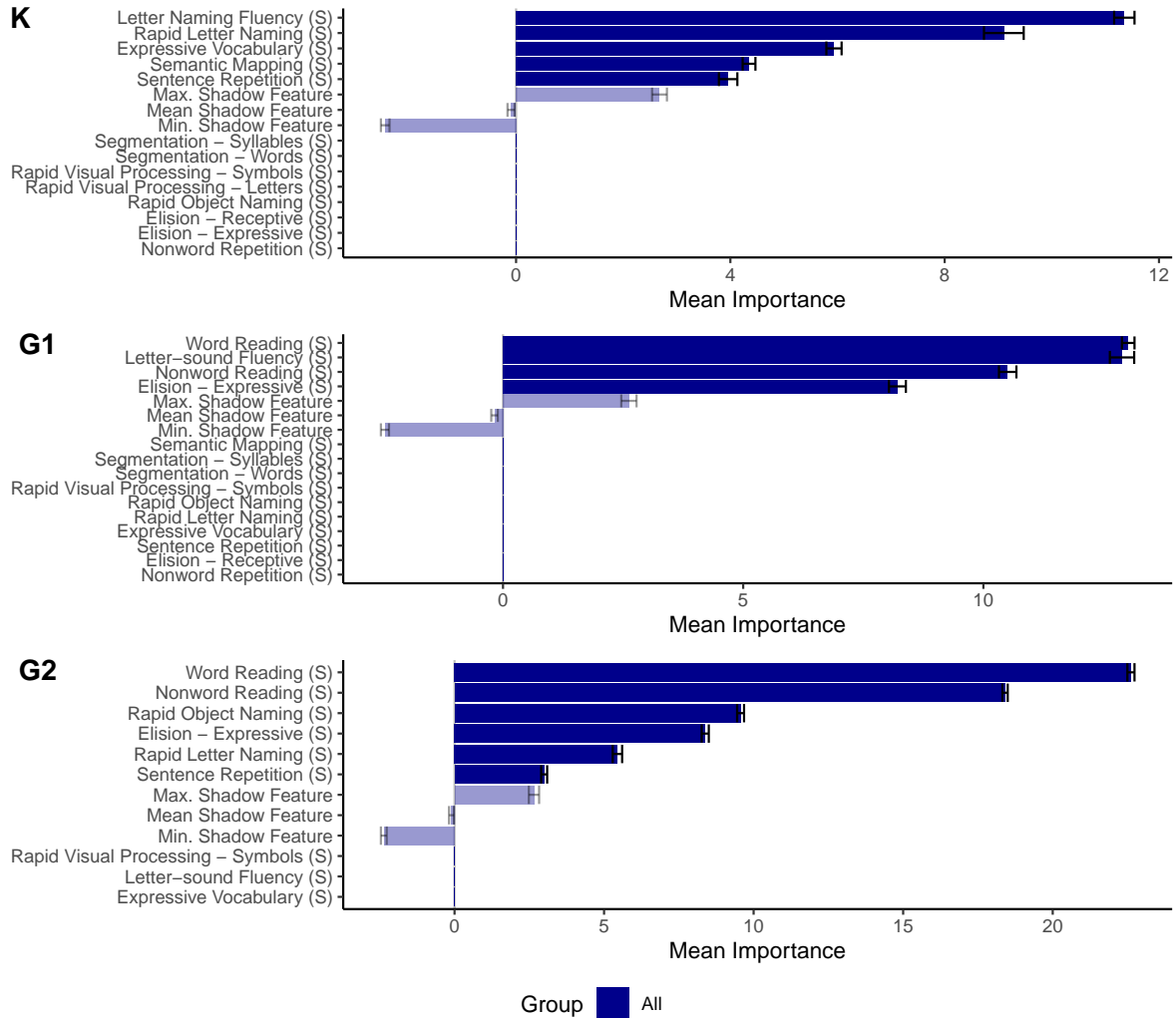


Figure 11: Mean Task Importance Obtained From Boruta Algorithm for Spanish Predictions (AY 2023/24).

1.4 Models Evaluations - ROC Curves

1.4.1 English-to-English - Kindergarten

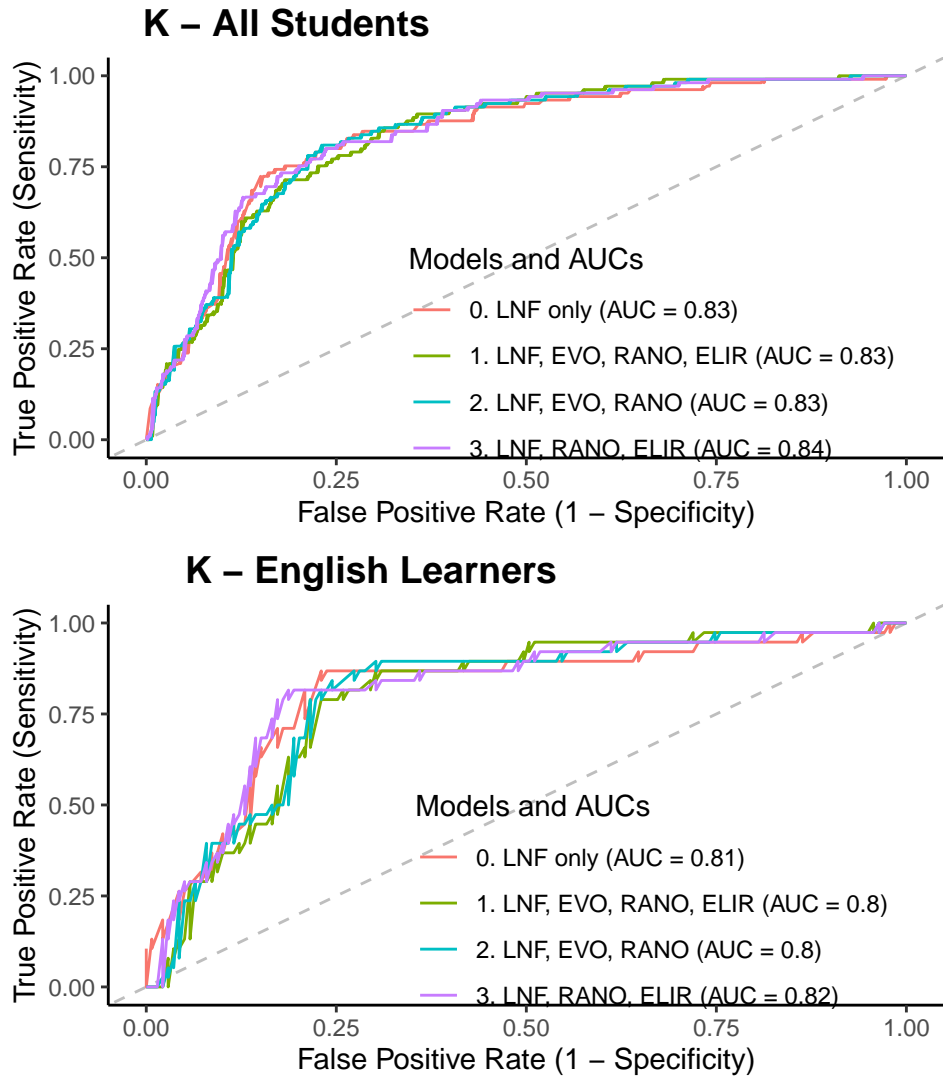


Figure 12: Receiver-operator Characteristic Curves For English-to-English Kindergarten Models (AY 2023/24).

1.4.2 English-to-English - Grade 1

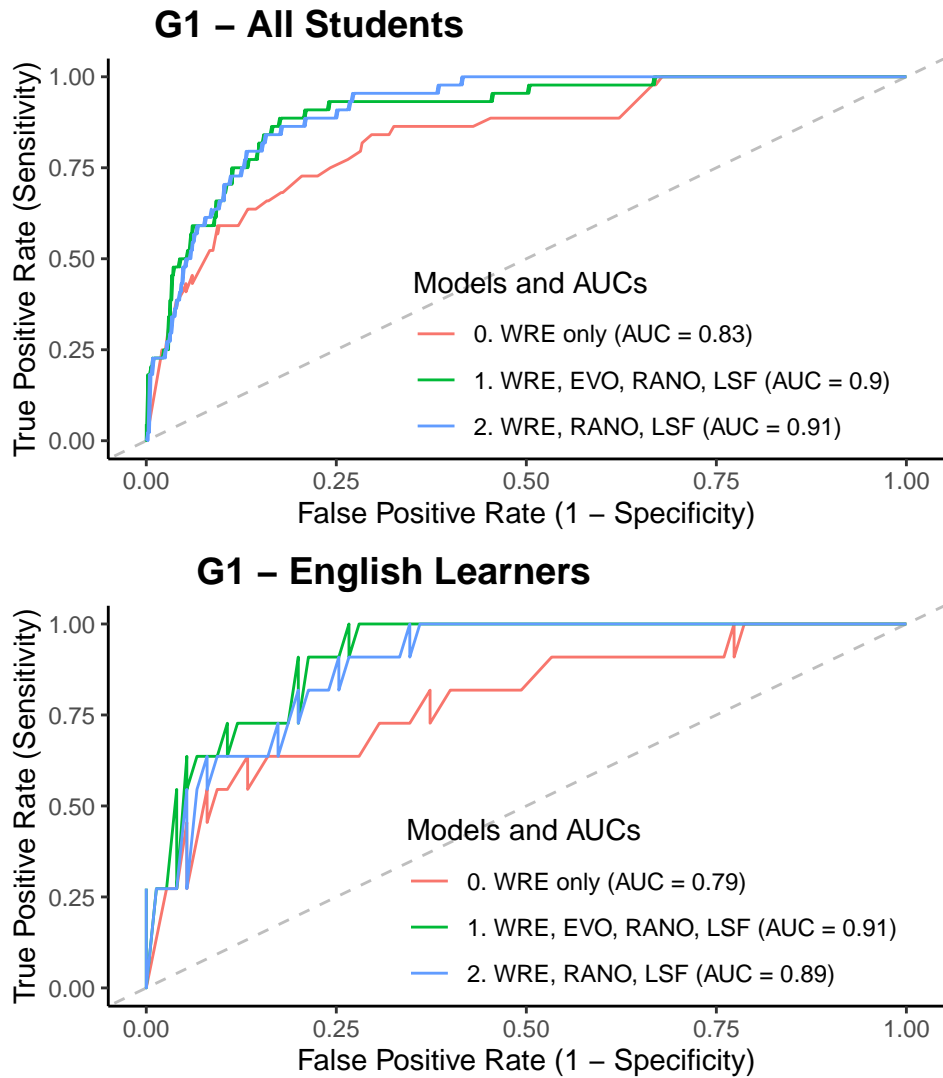


Figure 13: Receiver-operator Characteristic Curves For English-to-English Grade 1 Models (AY 2023/24).

1.4.3 English-to-English - Grade 2

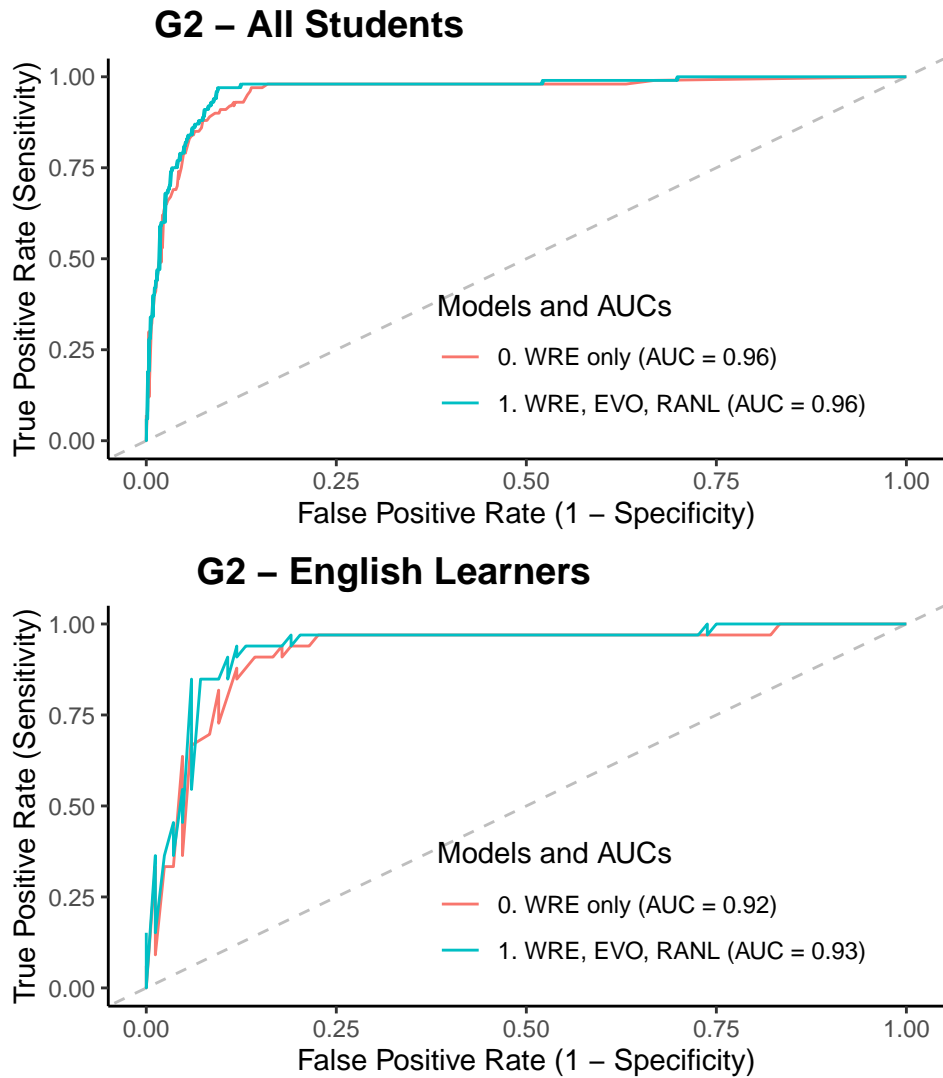


Figure 14: Receiver-operator Characteristic Curves For English-to-English Grade 2 Models (AY 2023/24).

1.4.4 Spanish-to-Spanish - Kindergarten

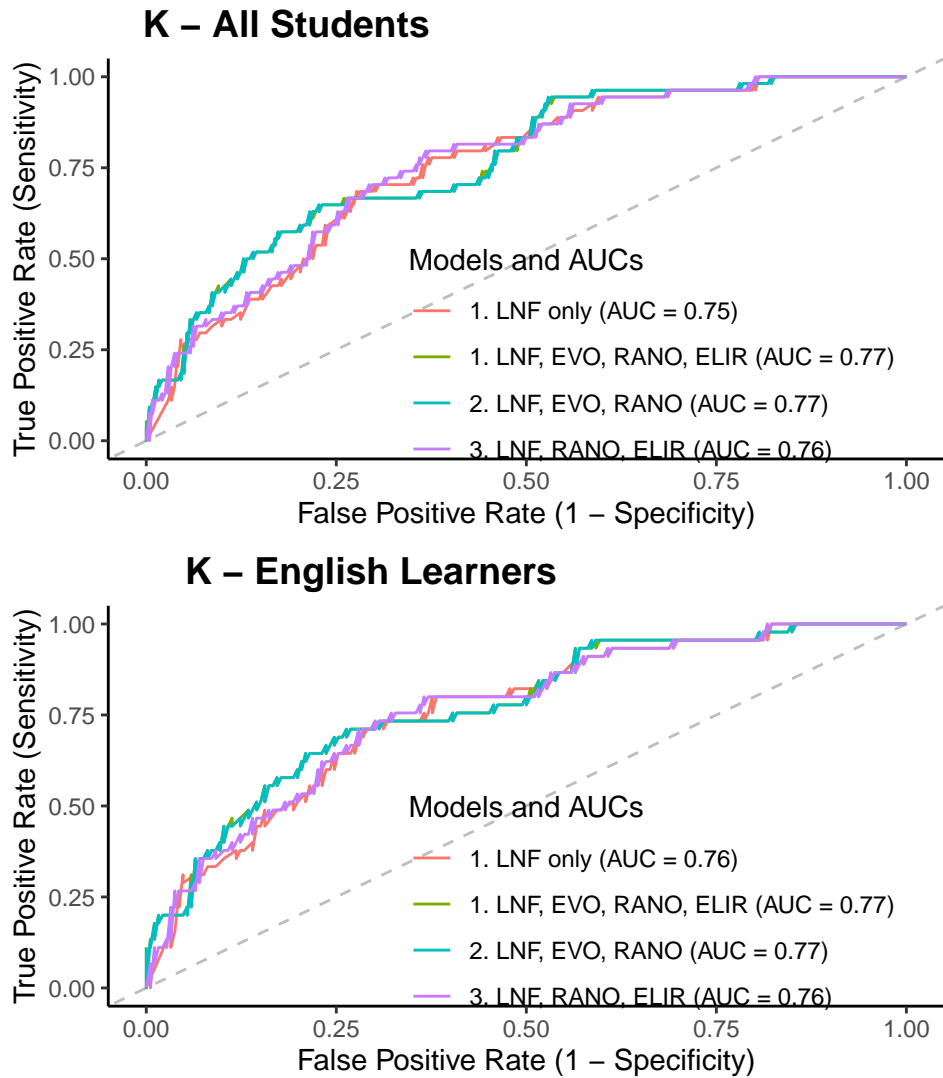


Figure 15: Receiver-operator Characteristic Curves For Spanish-to-Spanish Kindergarten Models (AY 2023/24).

1.4.5 Spanish-to-Spanish - Grade 1

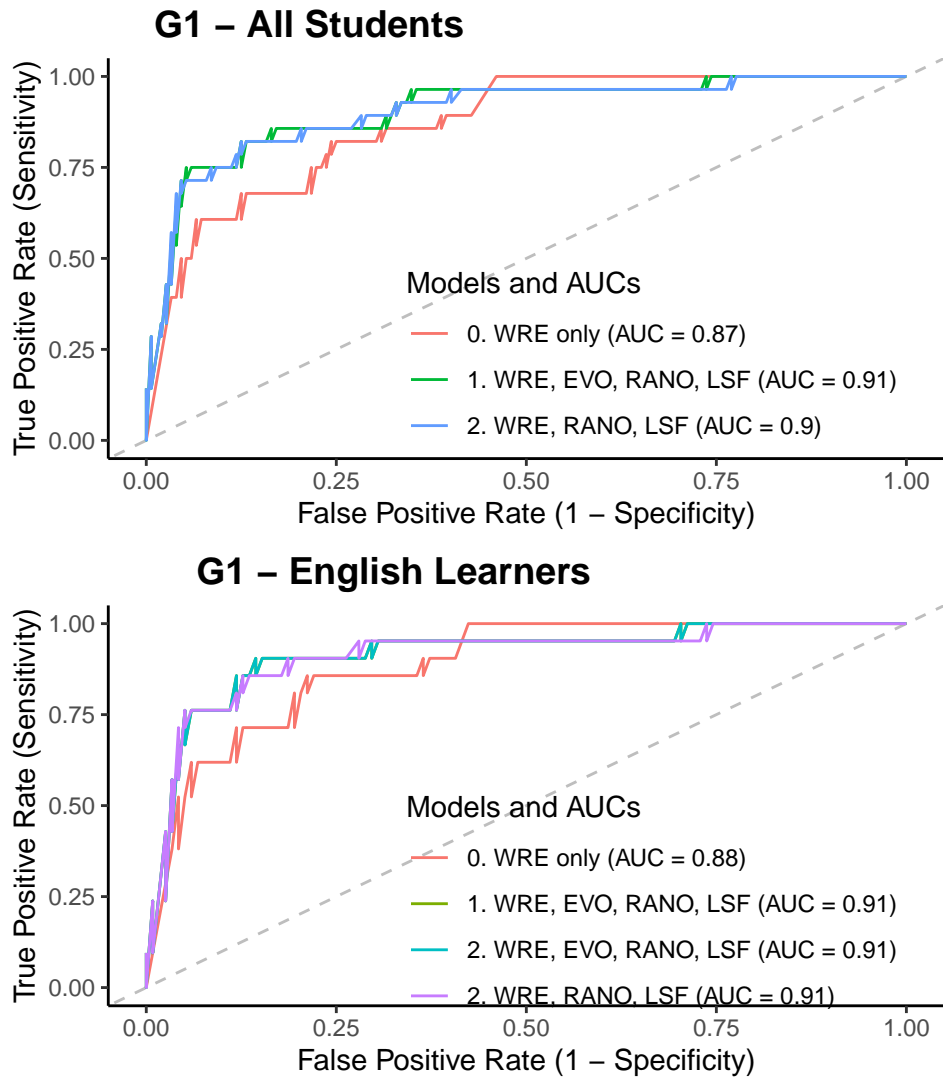


Figure 16: Receiver-operator Characteristic Curves For Spanish-to-Spanish Grade 1 Models (AY 2023/24).

1.4.6 Spanish-to-Spanish - Grade 2

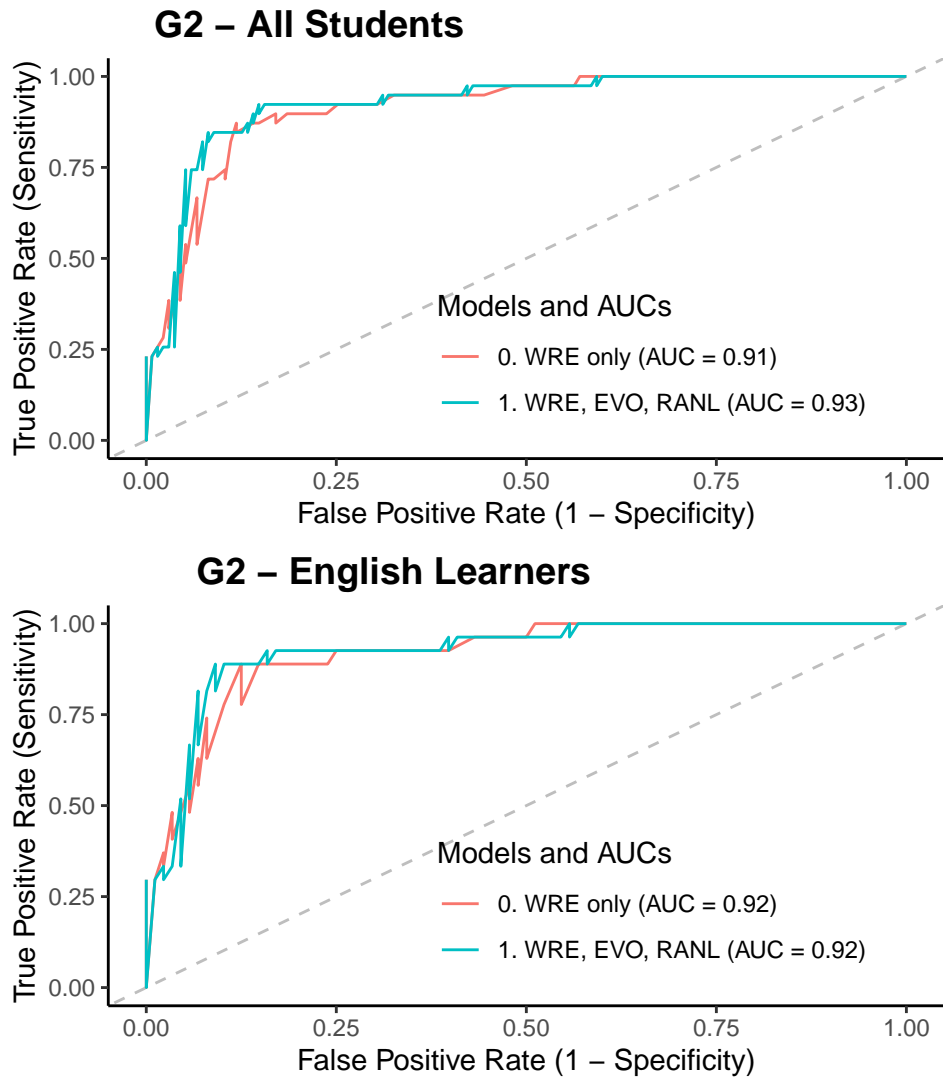


Figure 17: Receiver-operator Characteristic Curves For Spanish-to-Spanish Grade 2 Models (AY 2023/24).

Table 3: Summary of Best English and Spanish Prediction Models (AY 2023/24).

Grade	Model	English												Spanish			
		Sensitivity			Specificity			Accuracy			Balanced Acc.			Sens.	Spec.	Acc.	BA
		All	EL	EO	All	EL	EO	All	EL	EO	All	EL	EO	All	All	All	All
K	LNF only	0.7333	0.8684	0.6727	0.8311	0.7266	0.8659	0.8165	0.7822	0.7571	0.7975	0.8402	0.7693	0.6667	0.7231	0.7128	0.6949
K	LNF, RANO, ELIR	0.7905	0.8421	0.8182	0.7642	0.6475	0.7877	0.7681	0.7773	0.6893	0.7448	0.7918	0.8029	0.7778	0.6322	0.6588	0.705
K	LNF, RANO, ELIR, EVO	0.8476	0.8947	0.8727	0.6873	0.5324	0.7179	0.7112	0.7675	0.6102	0.7136	0.7385	0.7953	0.6296	0.7727	0.7466	0.7012
K	LNF, RANO, EVO	0.8	0.8947	0.7455	0.7692	0.6331	0.8101	0.7738	0.7846	0.6893	0.7639	0.8015	0.7778	0.6296	0.7727	0.7466	0.7012
G1	WRE only	0.8182	0.8182	0.8182	0.7161	0.6267	0.7254	0.7247	0.7671	0.6512	0.7224	0.7335	0.7718	0.7857	0.7566	0.7611	0.7711
G1	WRE, RANO, LSF	0.8409	0.8182	0.8485	0.8225	0.7467	0.8295	0.8241	0.8317	0.7558	0.7824	0.8311	0.839	0.7857	0.875	0.8611	0.8304
G1	WRE, RANO, LSF, EVO	0.8636	1	0.8182	0.8246	0.72	0.8382	0.8279	0.8441	0.7558	0.86	0.8364	0.8282	0.7143	0.9474	0.9111	0.8308
G2	WRE only	0.96	0.9697	0.9677	0.8635	0.7619	0.8638	0.878	0.9117	0.8205	0.8658	0.878	0.9157	0.8462	0.8815	0.8736	0.8638
G2	WRE, RANL	0.97	0.9697	0.9839	0.8812	0.7976	0.8817	0.8946	0.9256	0.8462	0.8837	0.8958	0.9328	0.8974	0.8444	0.8563	0.8709
G2	WRE, RANL, EVO	0.96	0.9697	1	0.906	0.8095	0.91	0.9142	0.933	0.8547	0.8896	0.9224	0.955	0.8974	0.8519	0.8621	0.8746

Note. green: >= .8; black: >= .7; yellow: >= .6; red otherwise

1.5 Models Evaluations - Sensitivity and Specificity

2 PART II - Analysis of 2024/25 Data

2.1 Sample

Throughout, I only included tasks taken by at least 50 students per language per grade, so as to not distort the task selection analyses.

- $N = 1390$
- 26 schools
- 12 districts

Table 4: Demographic Characteristics of the AY 24/25 Sample (N = 1,390) by Grade and Screening Language

Characteristic	English			Spanish		
	K N = 438	G1 N = 569	G2 N = 336	K N = 272	G1 N = 305	G2 N = 154
Gender						
Female	180 (47%)	301 (56%)	145 (45%)	98 (44%)	159 (58%)	73 (48%)
Male	202 (53%)	237 (44%)	175 (55%)	125 (56%)	115 (42%)	80 (52%)
Unknown	56	31	16	49	31	1
Race						
Asian	24 (7.7%)	47 (9.3%)	33 (10%)	8 (4.6%)	9 (3.6%)	3 (2.0%)
Black/African American	41 (13%)	65 (13%)	40 (12%)	3 (1.7%)	3 (1.2%)	1 (0.7%)
Not reported	51 (16%)	101 (20%)	41 (13%)	45 (26%)	74 (29%)	44 (30%)
Other	36 (12%)	58 (12%)	35 (11%)	22 (13%)	23 (9.1%)	8 (5.4%)
White	159 (51%)	233 (46%)	179 (55%)	95 (55%)	144 (57%)	93 (62%)
Unknown	127	65	8	99	52	5
Ethnicity						
Hispanic/Latin(o/a)	298 (81%)	433 (78%)	234 (70%)	228 (98%)	295 (98%)	147 (96%)
Not Hispanic/Latin(o/a)	71 (19%)	124 (22%)	98 (30%)	4 (1.7%)	5 (1.7%)	6 (3.9%)
Unknown	69	12	4	40	5	1
ELPD						
EL	215 (59%)	251 (49%)	124 (39%)	211 (89%)	232 (83%)	117 (77%)
EO	123 (34%)	211 (41%)	155 (49%)	0 (0%)	3 (1.1%)	6 (4.0%)
EP	19 (5.2%)	40 (7.8%)	37 (12%)	17 (7.2%)	35 (13%)	27 (18%)
TBD	8 (2.2%)	11 (2.1%)	0 (0%)	8 (3.4%)	8 (2.9%)	1 (0.7%)
Unknown	73	56	20	36	27	3
Ever IEP/504	32 (9.2%)	33 (7.1%)	30 (10.0%)	23 (10%)	12 (4.7%)	5 (3.5%)
Unknown	92	102	35	49	49	13
Home Language						
English	156 (36%)	237 (42%)	170 (51%)	0 (0%)	1 (0.3%)	2 (1.3%)
Other	11 (2.6%)	16 (2.9%)	15 (4.5%)			
Spanish	264 (61%)	306 (55%)	146 (44%)	272 (100%)	304 (100%)	152 (99%)
Unknown	7	10	5			

2.2 Descriptives

2.2.1 Distributions of English Multitudes Tasks

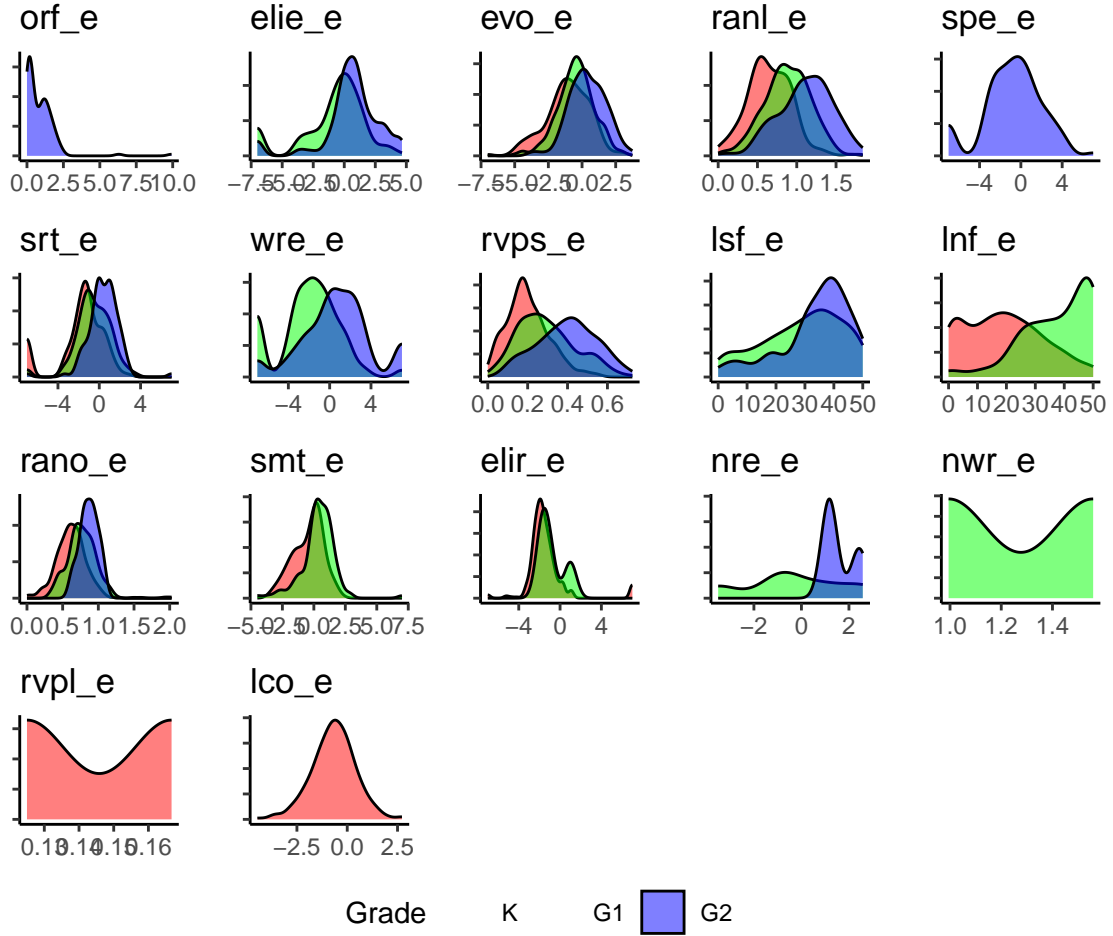


Figure 18: Distributions of English Multitudes Tasks by Grade and Language (Fall 2024).

2.2.2 Distributions of Spanish Multitudes Tasks

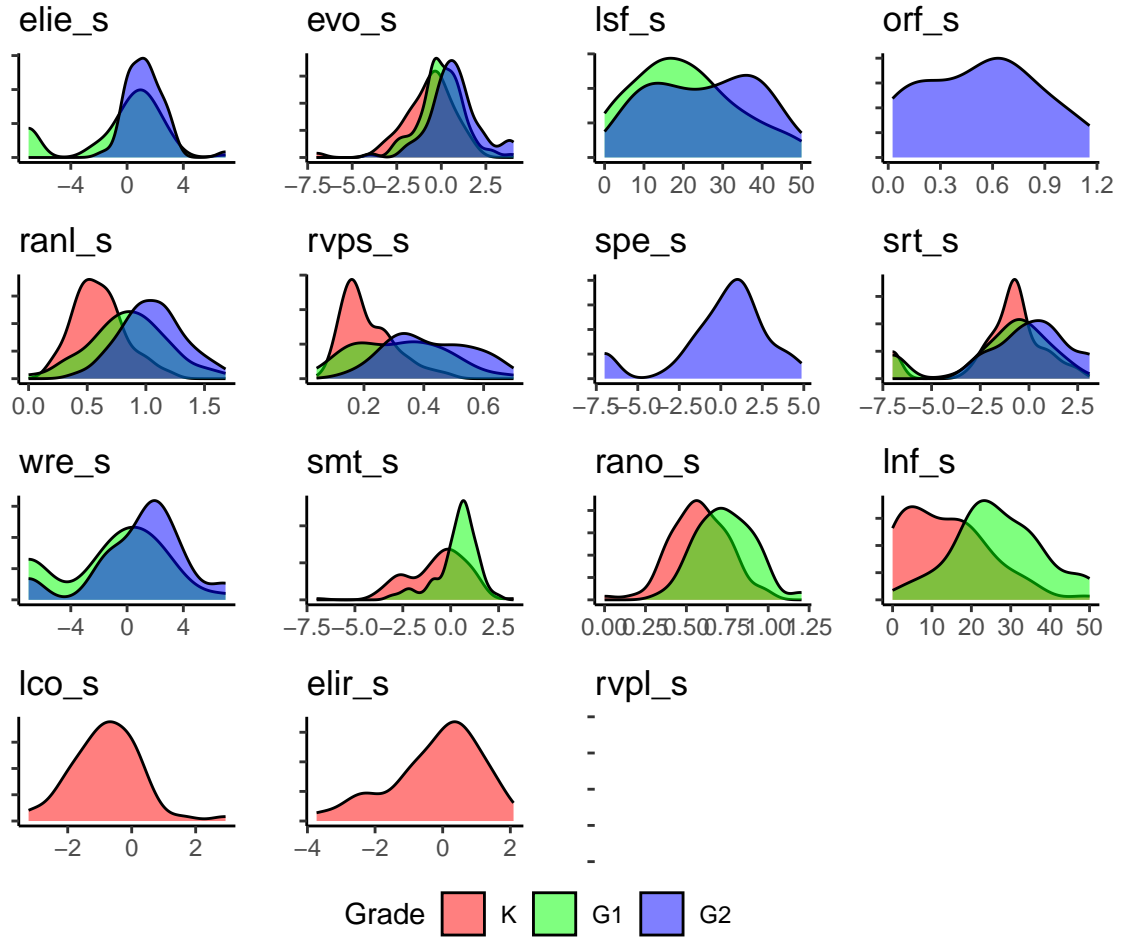


Figure 19: Distributions of Spanish Multitudes Tasks by Grade and Language (Fall 2024).

Table 5: Comparison of Proportions of Struggling Readers (\leq 20th Percentile) in the Study Sample Based on Overall Multitudes Sample and on Published Woodcock-Johnson/Muñoz Norms (Spring 2025).

Reference	English			Spanish		
	K N = 443	G1 N = 561	G2 N = 359	K N = 129	G1 N = 134	G2 N = 86
Norm-based						
Not struggling	231 (52%)	251 (45%)	211 (59%)	58 (45%)	66 (49%)	48 (56%)
Struggling	212 (48%)	310 (55%)	148 (41%)	71 (55%)	68 (51%)	38 (44%)
Sample-based						
Not struggling	336 (76%)	361 (64%)	249 (69%)	99 (77%)	88 (66%)	61 (71%)
Struggling	107 (24%)	200 (36%)	110 (31%)	30 (23%)	46 (34%)	25 (29%)

2.2.3 Distributions of Outcome Measures

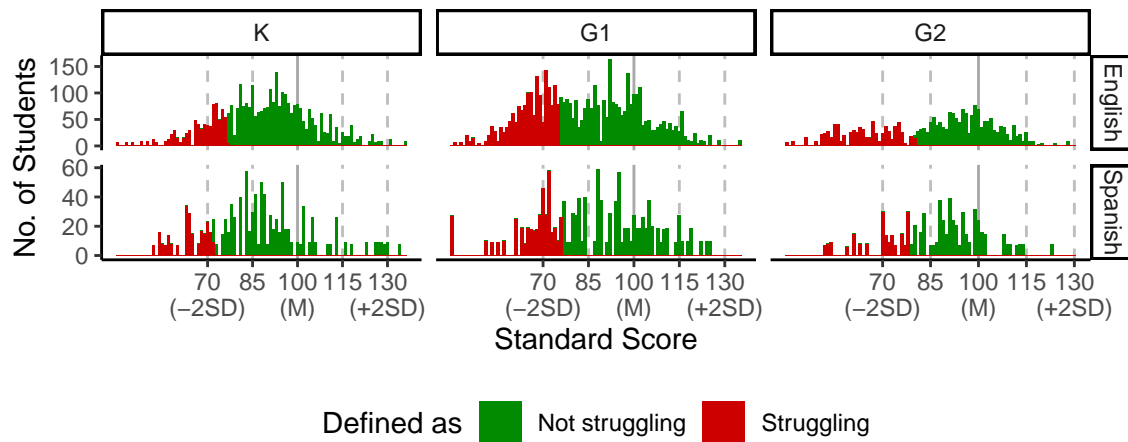
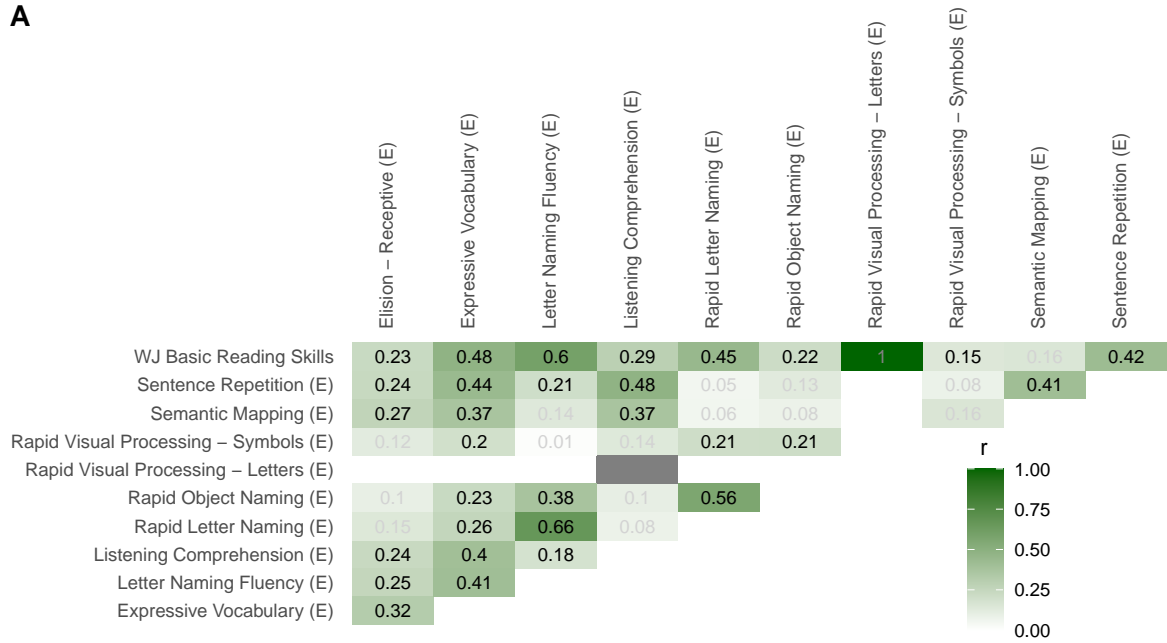


Figure 20: Distribution of Woodcock-Johnson/-Muñoz Basic Reading Skills Cluster (for Kindergarten) Broad Reading Cluster (for Grades 1 and 2) by Grade and Language with Struggling Readers (\leq 20th Percentile) Highlighted (Spring 2025).

2.2.4 Correlations

2.2.4.1 Kindergarten - English

A



B

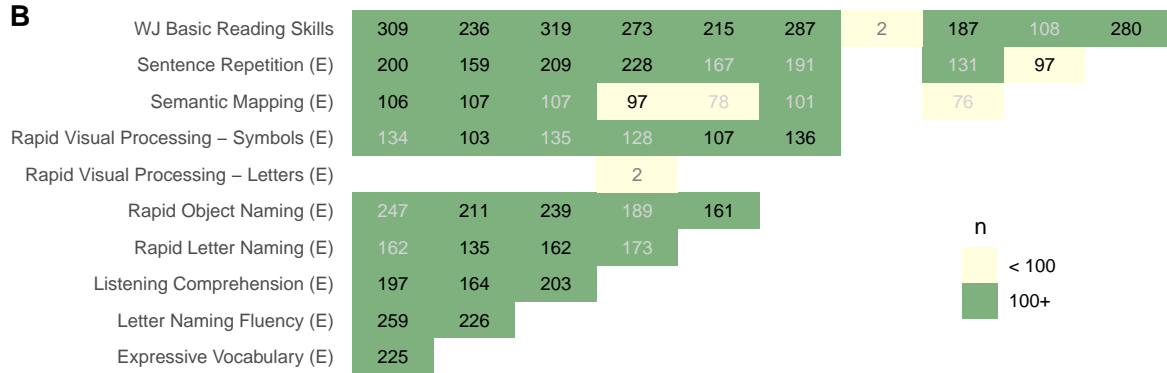
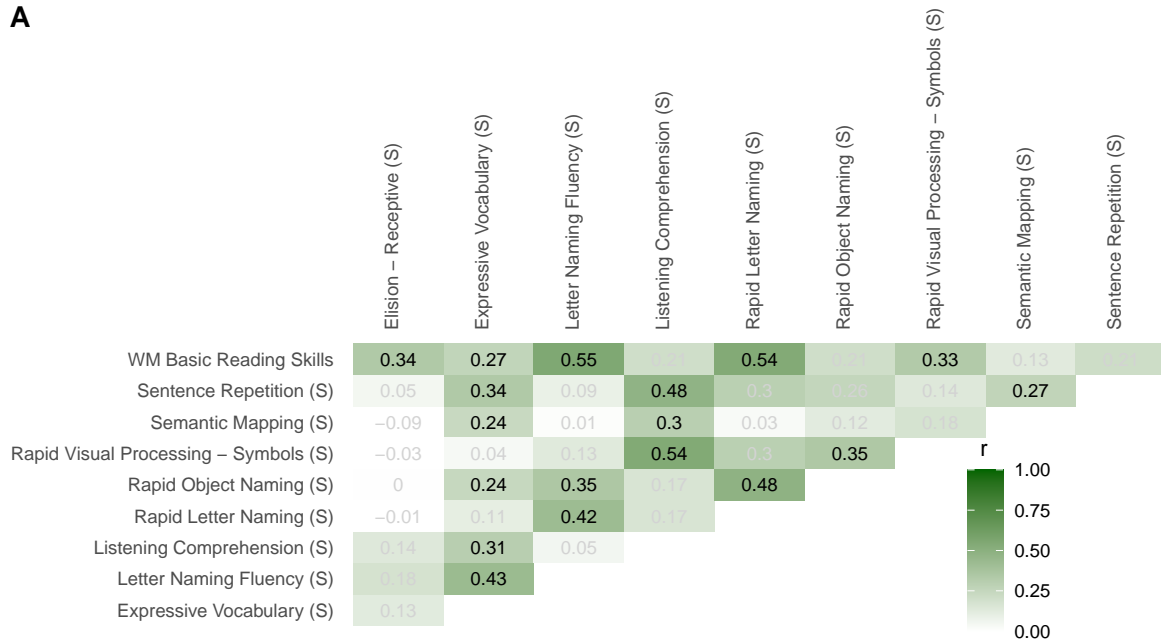


Figure 21: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Kindergarten; AY 2024/25).

2.2.4.2 Kindergarten - Spanish

A



B

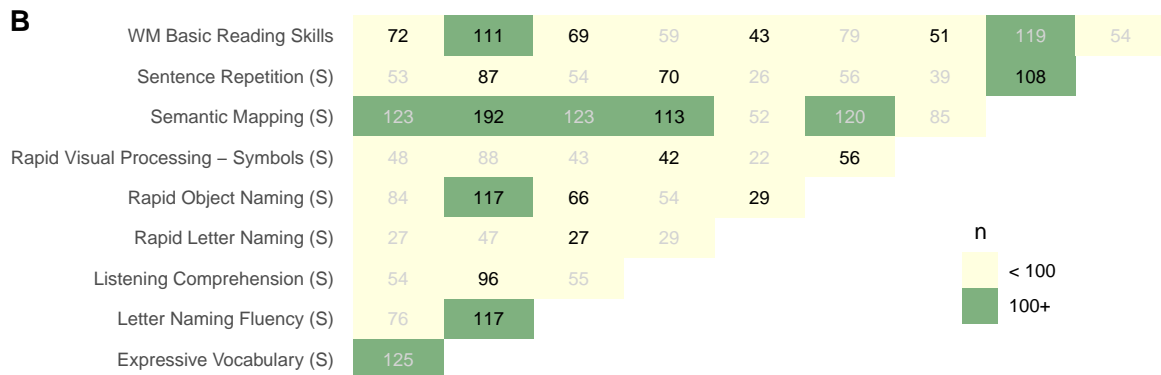
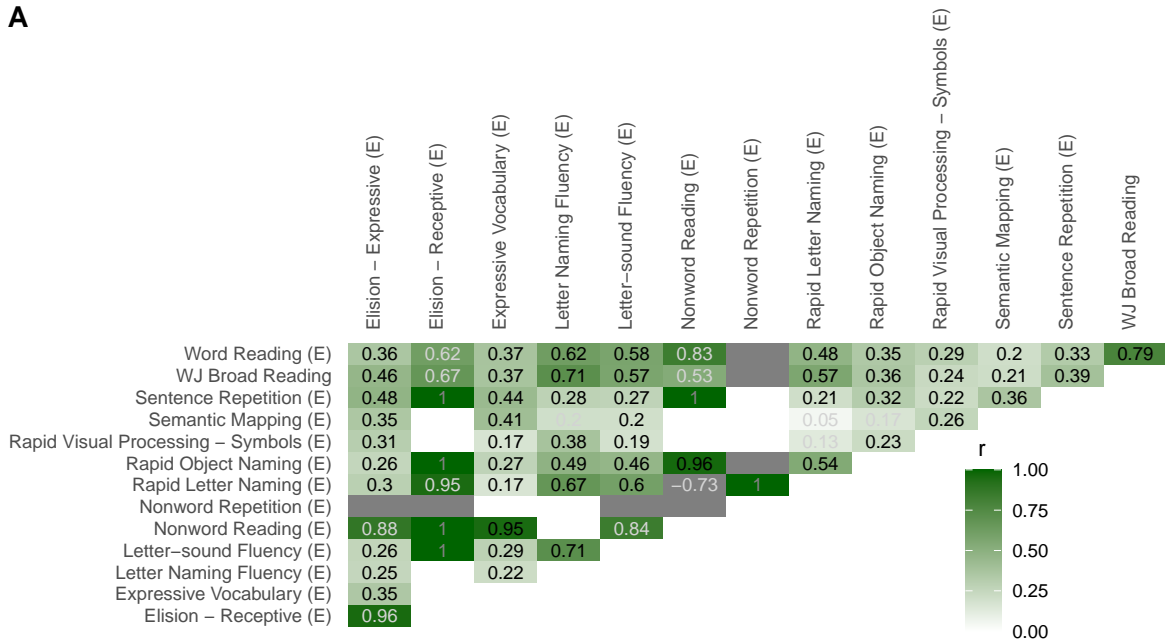


Figure 22: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Kindergarten; AY 2024/25).

2.2.4.3 Grade 1 - English

A



B

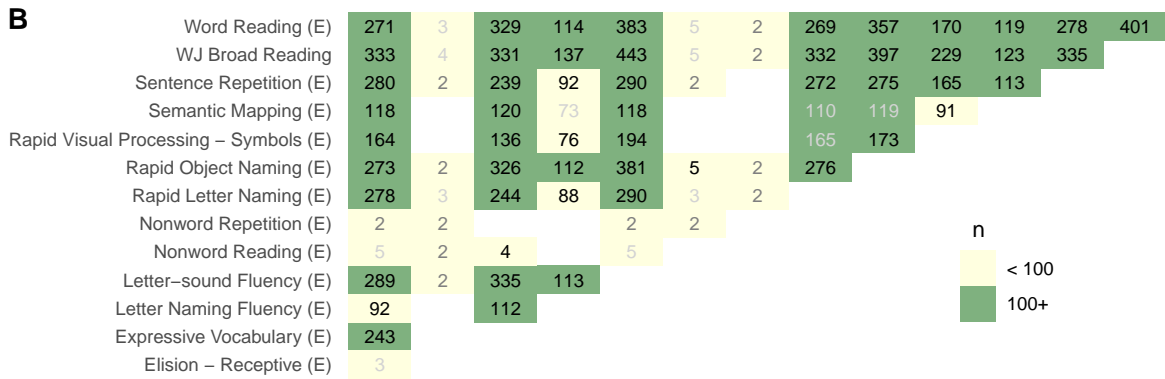
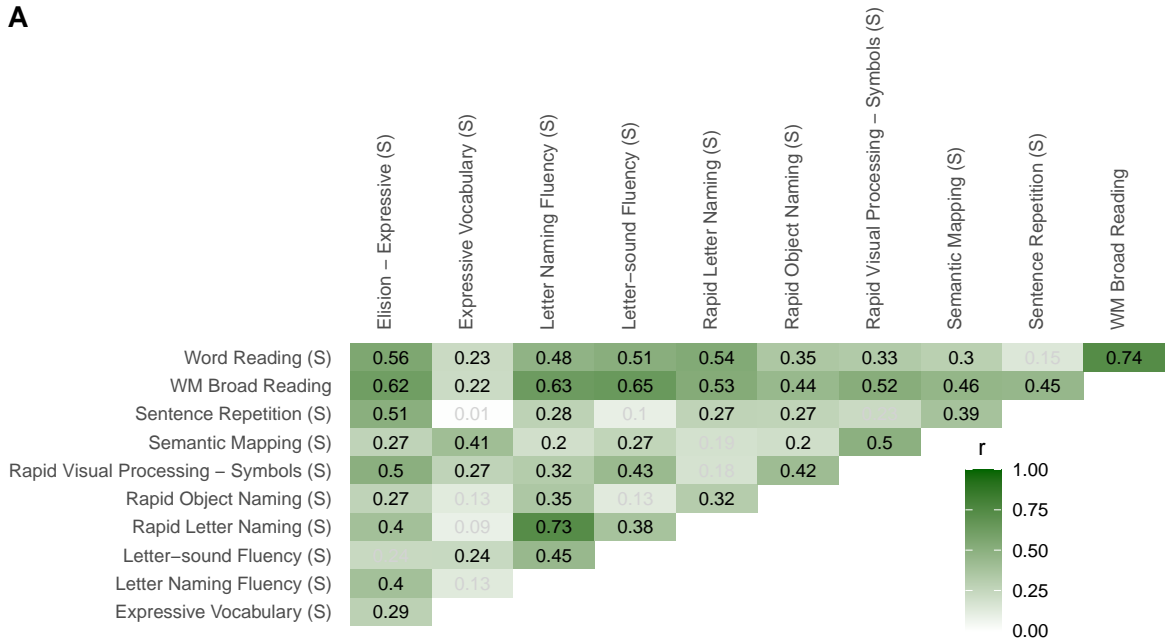


Figure 23: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Grade 1; AY 2024/25).

2.2.4.4 Grade 1 - Spanish

A



B

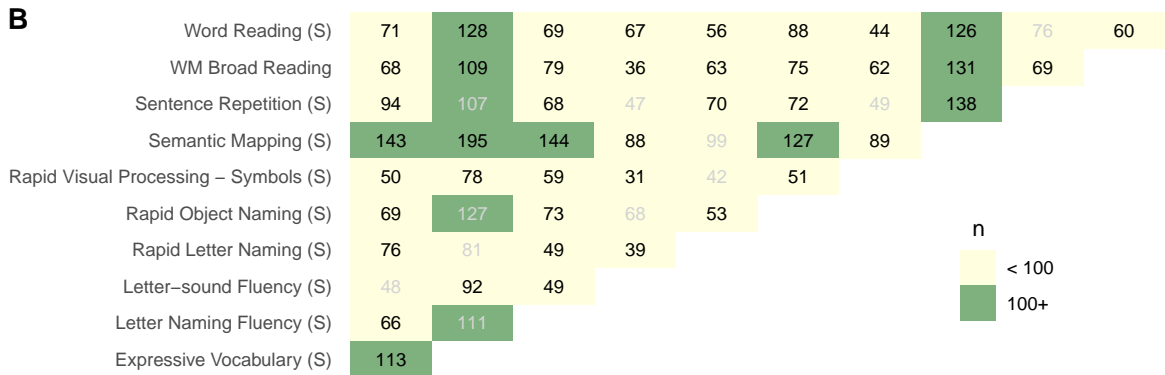


Figure 24: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Grade 1; AY 2024/25).

2.2.4.5 Grade 2 - English

A



B

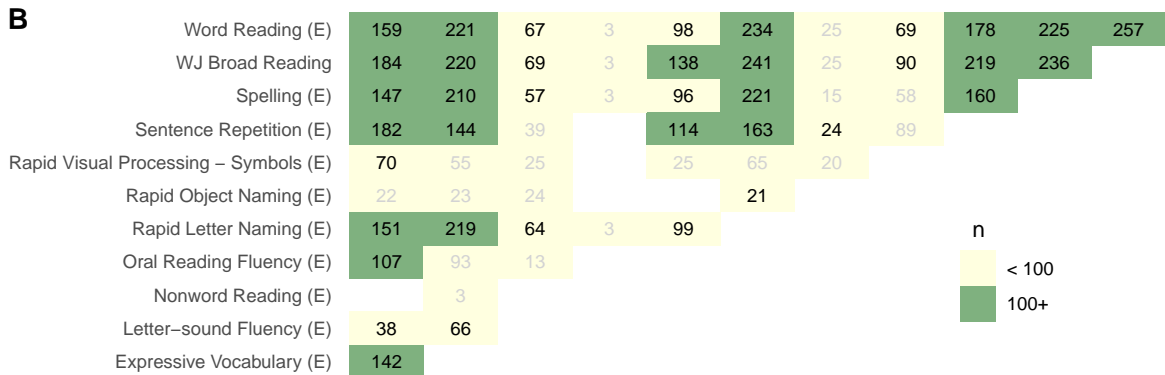
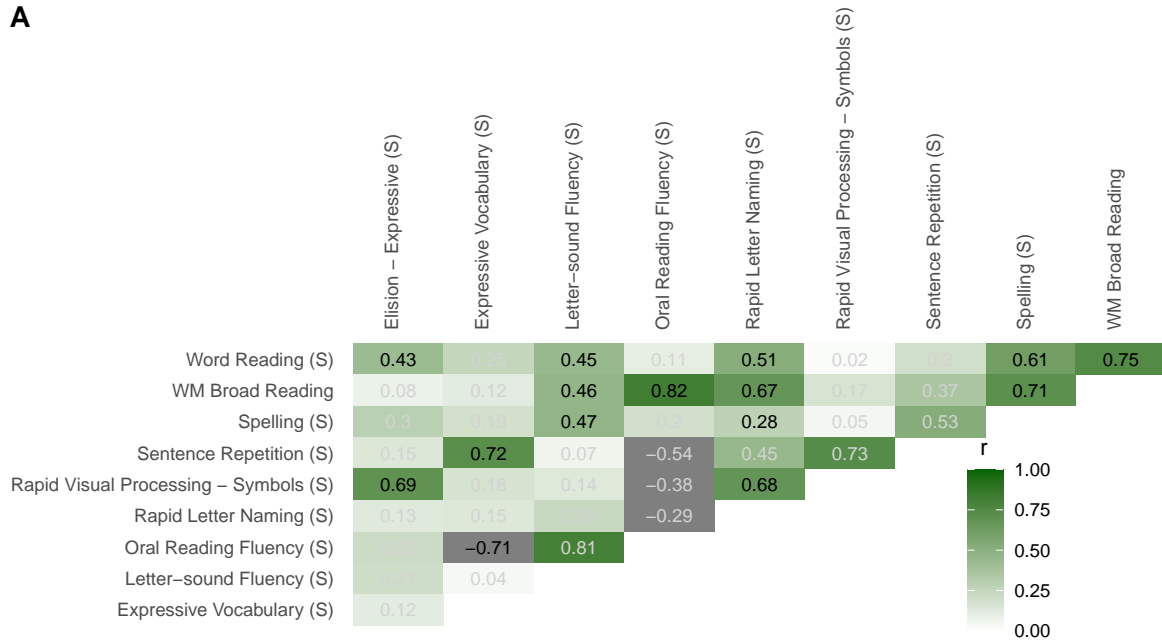


Figure 25: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Grade 2; AY 2024/25).

2.2.4.6 Grade 2 - Spanish

A



B

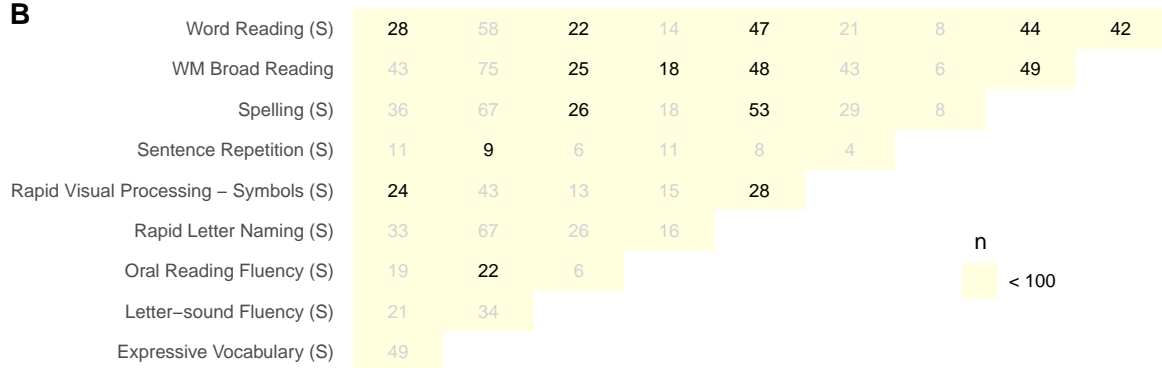


Figure 26: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Grade 2; AY 2024/25).

2.3 Task Selection

2.3.1 English-to-English (English to English)

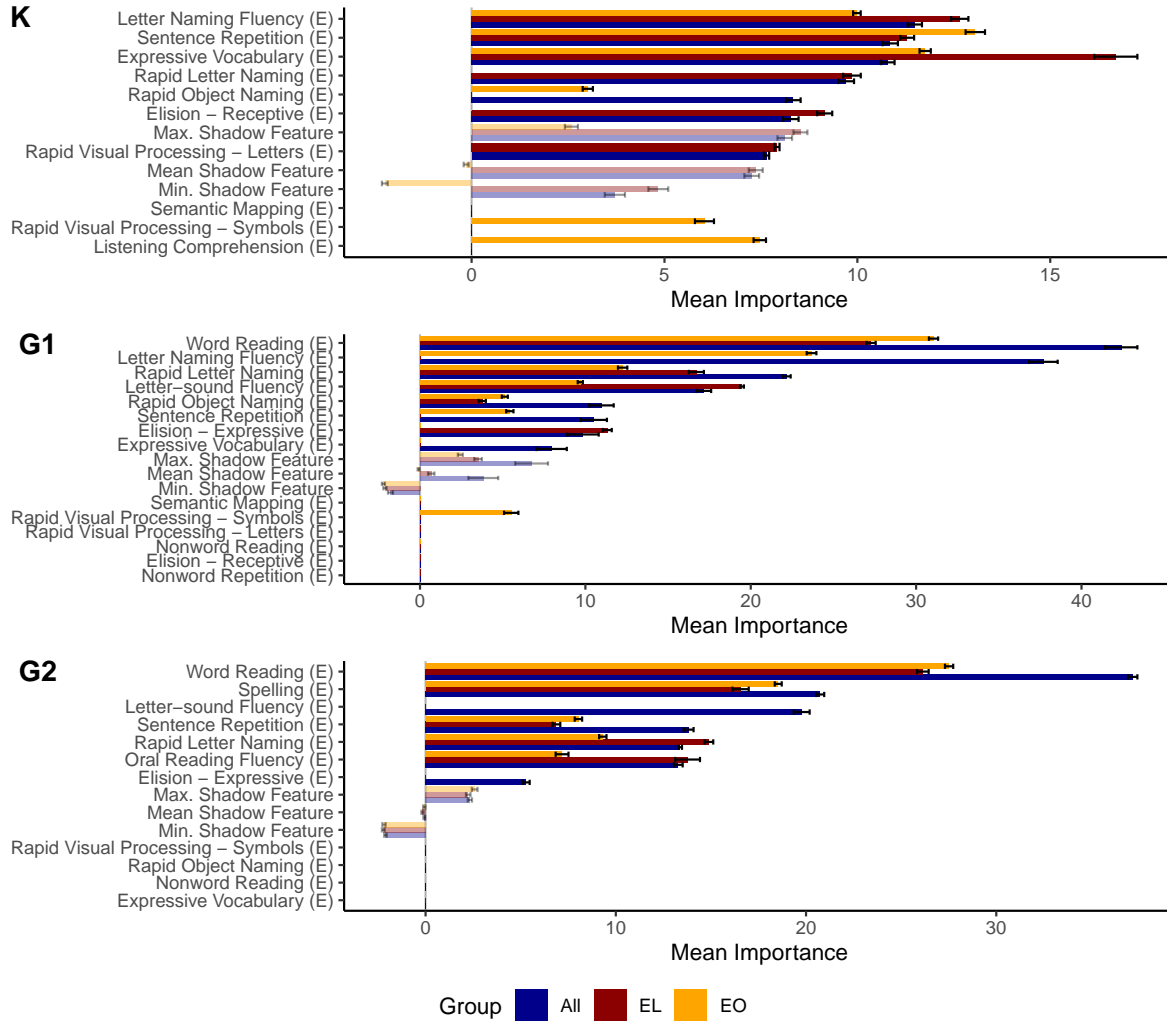


Figure 27: Mean Task Importance Obtained From Boruta Algorithm for English-to-English Predictions, with Results Disaggregated by English Proficiency Designation (AY 2024/25).

2.3.2 Spanish-to-Spanish (Spanish to Spanish)

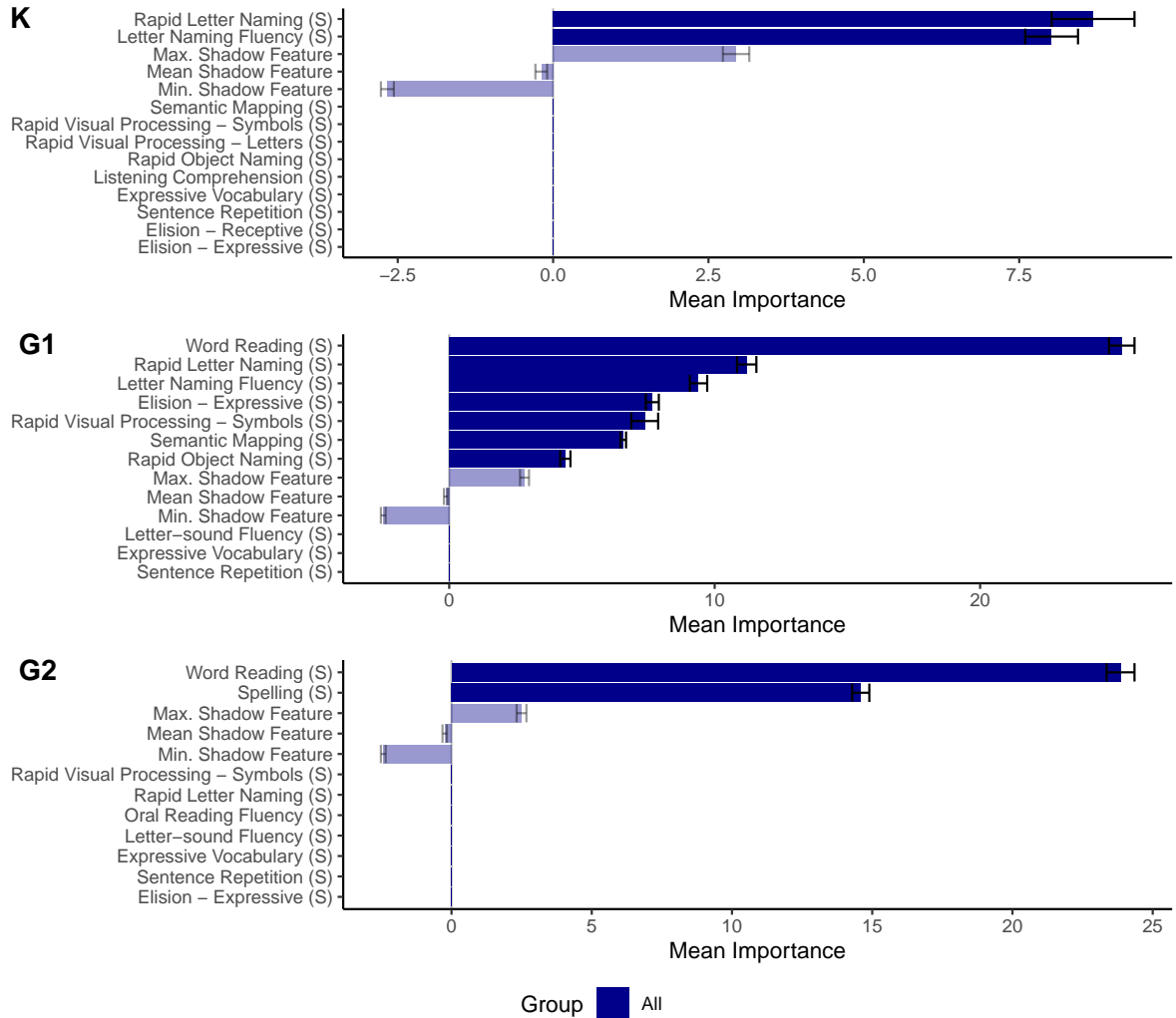


Figure 28: Mean Task Importance Obtained From Boruta Algorithm for Spanish-to-Spanish Predictions (AY 2024/25).

2.4 Models Evaluations - ROC Curves

2.4.1 English-to-English - Kindergarten

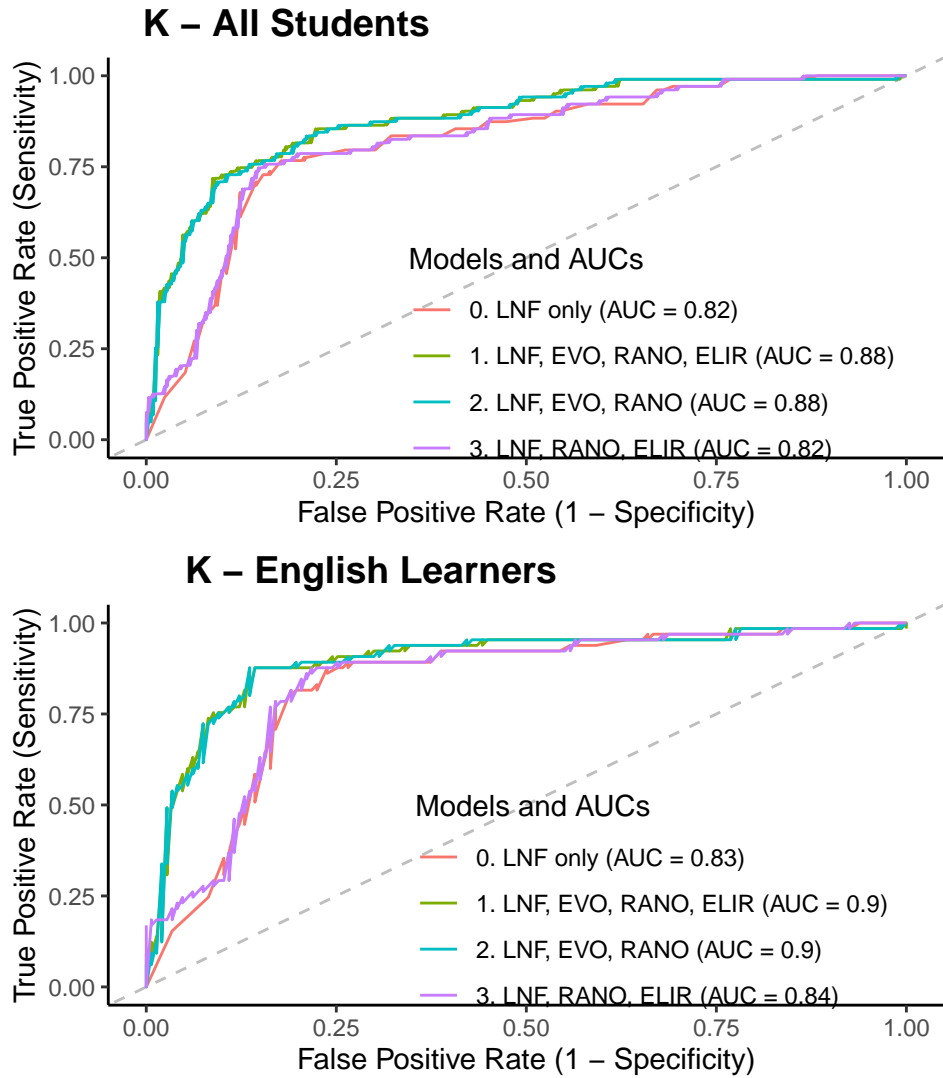


Figure 29: Receiver-operator Characteristic Curves For English-to-English Kindergarten Models (AY 2024/25).

2.4.2 English-to-English - Grade 1

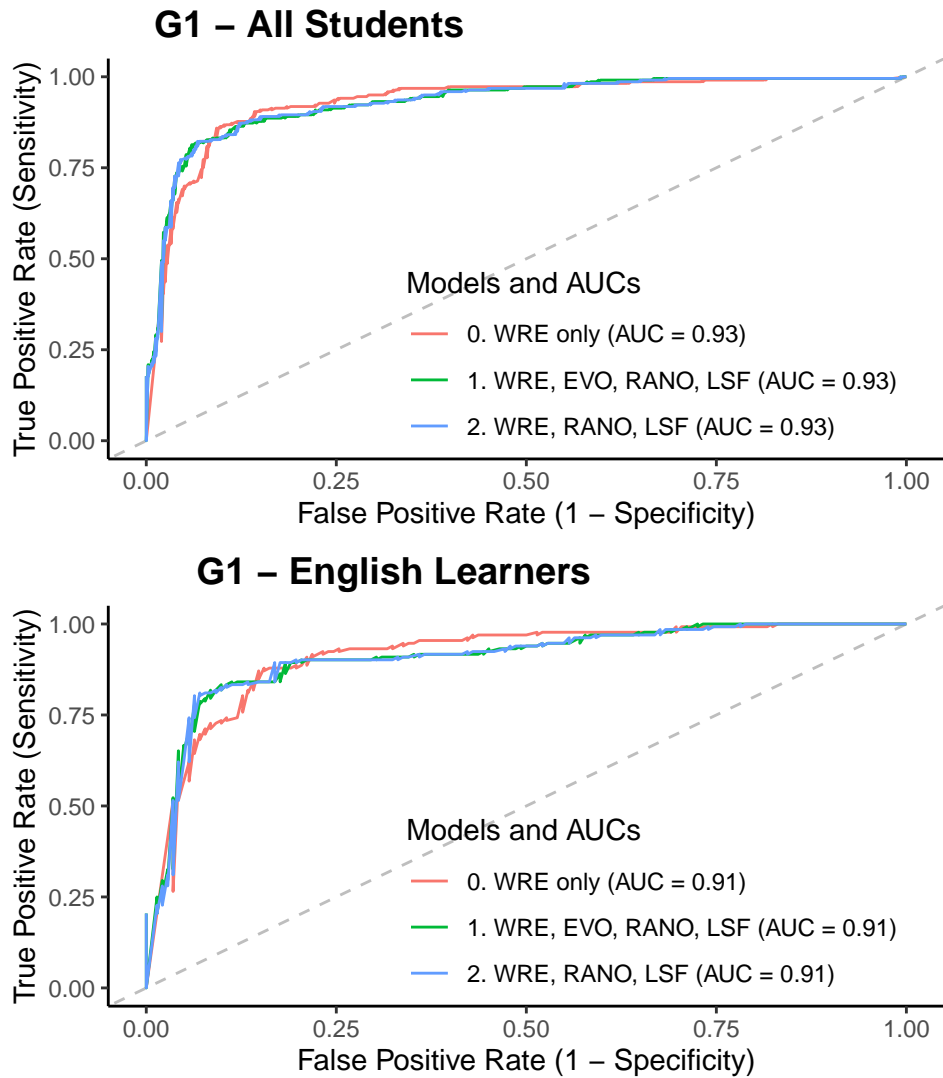


Figure 30: Receiver-operator Characteristic Curves For English-to-English Grade 1 Models (AY 2024/25).

2.4.3 English-to-English - Grade 2

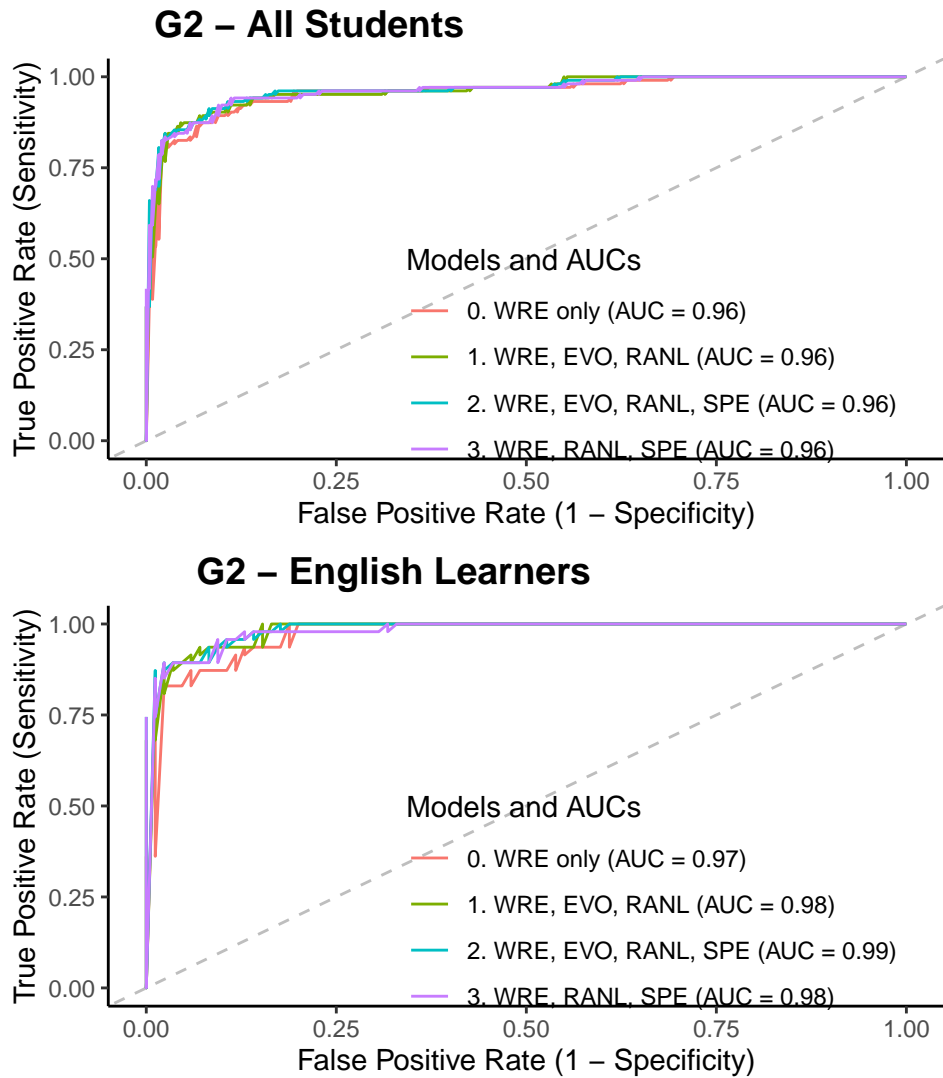


Figure 31: Receiver-operator Characteristic Curves For English-to-English Grade 2 Models (AY 2024/25).

2.4.4 Spanish-to-Spanish - Kindergarten

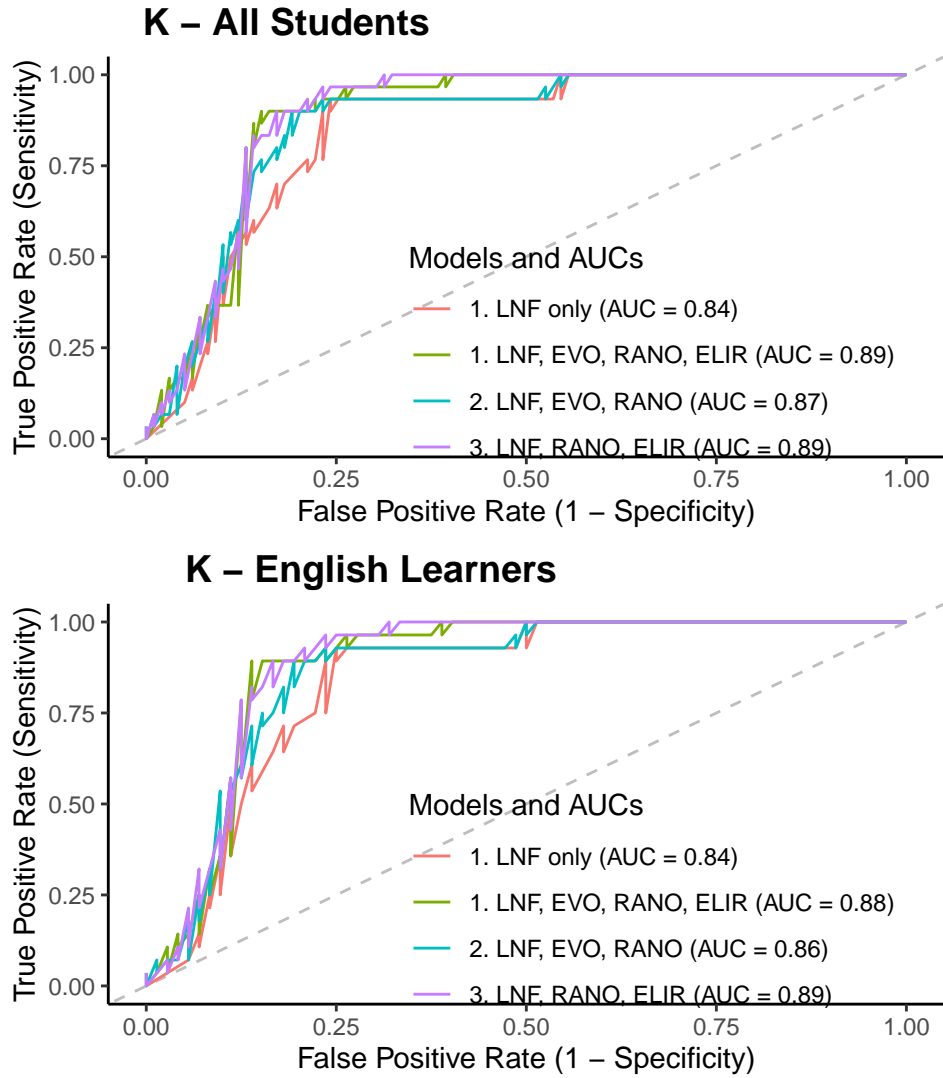


Figure 32: Receiver-operator Characteristic Curves For Spanish-to-Spanish Kindergarten Models (AY 2024/25).

2.4.5 Spanish-to-Spanish - Grade 1

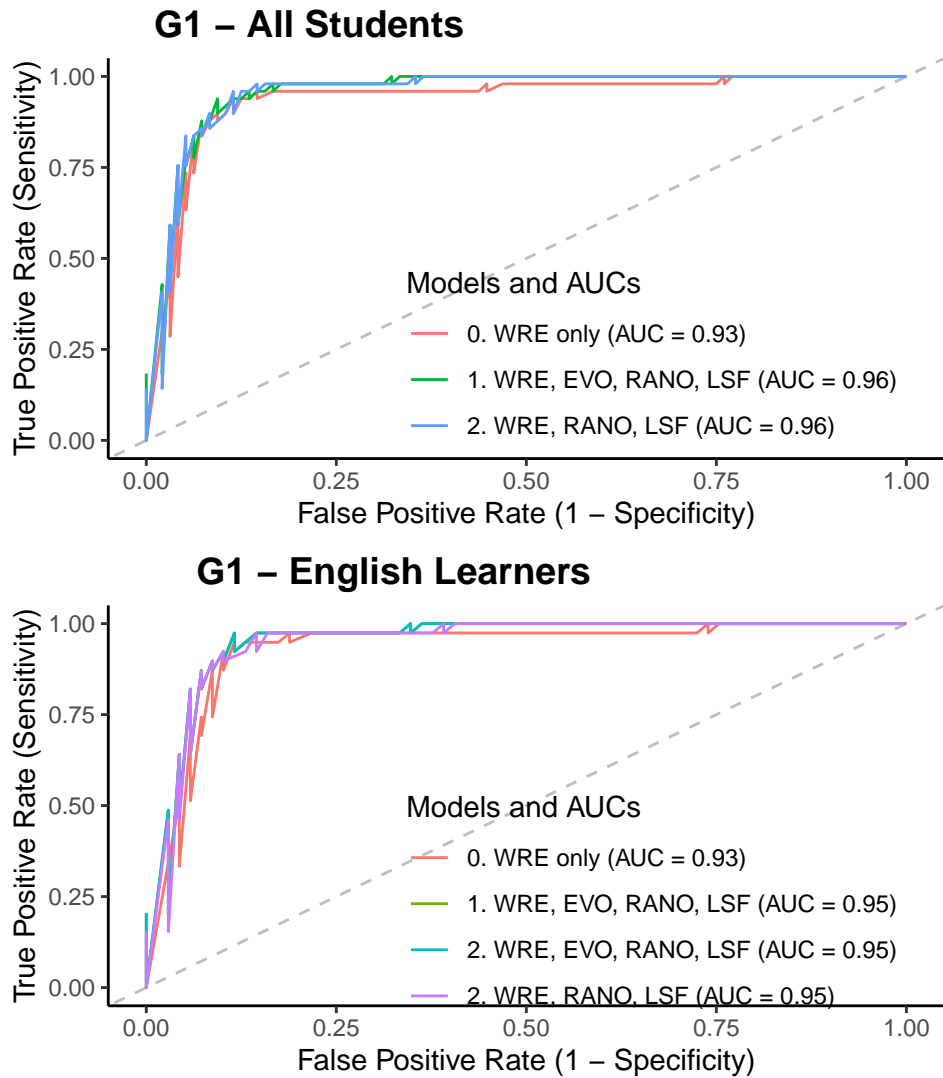


Figure 33: Receiver-operator Characteristic Curves For Spanish-to-Spanish Grade 1 Models (AY 2024/25).

2.4.6 Spanish-to-Spanish - Grade 2

No data collected.

Table 6: Summary of Best English and Spanish Prediction Models (AY 2024/25).

Grade	Model	English												Spanish			
		Sensitivity			Specificity			Accuracy			Balanced Acc.			Sens.	Spec.	Acc.	BA
		All	EL	EO	All	EL	EO	All	EL	EO	All	EL	EO	All	All	All	All
K	LNF only	0.7573	0.8615	0.48	0.8223	0.7619	0.9082	0.8069	0.7898	0.7925	0.8117	0.8211	0.6941	0.9	0.7576	0.7907	0.8288
K	LNF, RANO, ELIR	0.7476	0.8615	0.44	0.8434	0.7823	0.9286	0.8207	0.7955	0.8066	0.8219	0.8293	0.6843	0.9333	0.7677	0.8062	0.8505
K	LNF, RANO, ELIR, EVO	0.8447	0.9077	0.68	0.7771	0.7075	0.898	0.7931	0.8109	0.7689	0.8076	0.8537	0.789	0.8667	0.8485	0.8527	0.8576
K	LNF, RANO, EVO	0.8252	0.9077	0.6	0.7892	0.7143	0.9184	0.7977	0.8072	0.7736	0.811	0.8537	0.7592	0.8667	0.8081	0.8217	0.8374
G1	WRE only	0.8545	0.8712	0.807	0.9075	0.8521	0.9217	0.8887	0.881	0.8613	0.8617	0.8924	0.8644	0.9184	0.8854	0.8966	0.9019
G1	WRE, RANO, LSF	0.8182	0.8333	0.7368	0.9325	0.8873	0.9518	0.8919	0.8753	0.8613	0.8603	0.8969	0.8443	0.9388	0.8854	0.9034	0.9121
G1	WRE, RANO, LSF, EVO	0.8091	0.8333	0.7368	0.94	0.8944	0.9639	0.8935	0.8745	0.865	0.8638	0.9058	0.8503	0.898	0.9062	0.9034	0.9021
G2	WRE only	0.8641	0.8723	0.8776	0.9298	0.9412	0.9151	0.9101	0.8969	0.9167	0.9068	0.9032	0.8963	NA	NA	NA	NA
G2	WRE, RANL	0.8641	0.9149	0.8367	0.9587	0.9529	0.9434	0.9304	0.9114	0.9394	0.9339	0.9097	0.8901	NA	NA	NA	NA
G2	WRE, RANL, EVO	0.8641	0.8936	0.8367	0.9545	0.9647	0.9245	0.9275	0.9093	0.9394	0.9292	0.8968	0.8806	NA	NA	NA	NA

Note. green: >= .8; black: >= .7; yellow: >= .6; red otherwise

2.5 Models Evaluations - Sensitivity and Specificity

3 PART III - Combined Analysis of 2023/24 and 2024/25 Data

3.1 Sample

Throughout, I only included tasks taken by at least 50 students per language per grade, so as to not distort the task selection analyses.

- 4091 observations from 3533 unique students (as some took part in the study in both years)
- 42 schools
- 20 districts

Table 7: Demographic Characteristics of the Combined AY 23/24 and 24/25 Sample (N = 3,533) by Grade and Screening Language

Characteristic ¹	English			Spanish		
	K N = 1,161 ¹	G1 N = 1,114 ¹	G2 N = 1,040 ¹	K N = 625 ¹	G1 N = 538 ¹	G2 N = 408 ¹
Gender						
Female	553 (50%)	558 (52%)	501 (49%)	290 (50%)	279 (55%)	214 (53%)
Male	552 (50%)	525 (48%)	521 (51%)	286 (50%)	228 (45%)	193 (47%)
Unknown	56	31	18	49	31	1
Race						
Asian	110 (11%)	114 (11%)	98 (9.6%)	20 (4.2%)	12 (2.5%)	8 (2.0%)
Black/African American	133 (14%)	119 (11%)	138 (13%)	5 (1.1%)	3 (0.6%)	3 (0.8%)
Not reported	208 (21%)	211 (20%)	183 (18%)	252 (53%)	228 (48%)	193 (48%)
Other	130 (13%)	132 (13%)	120 (12%)	28 (5.9%)	30 (6.3%)	14 (3.5%)
White	404 (41%)	470 (45%)	486 (47%)	170 (36%)	206 (43%)	181 (45%)
Unknown	176	68	15	150	59	9
Ethnicity						
Hispanic/Latin(o/a)	711 (66%)	713 (65%)	617 (60%)	557 (96%)	515 (97%)	384 (94%)
Not Hispanic/Latin(o/a)	370 (34%)	387 (35%)	414 (40%)	26 (4.5%)	18 (3.4%)	23 (5.7%)
Unknown	80	14	9	42	5	1
ELPD						
EL	397 (38%)	345 (33%)	246 (25%)	477 (81%)	410 (81%)	290 (73%)
EO	539 (52%)	603 (57%)	629 (65%)	24 (4.1%)	22 (4.3%)	35 (8.8%)
EP	70 (6.8%)	91 (8.7%)	98 (10%)	59 (10%)	68 (13%)	72 (18%)
TBD	28 (2.7%)	12 (1.1%)	1 (0.1%)	26 (4.4%)	8 (1.6%)	3 (0.8%)
Unknown	127	63	66	39	30	8
Ever IEP/504	84 (8.8%)	76 (8.1%)	87 (11%)	53 (9.6%)	33 (7.0%)	25 (7.5%)
Unknown	209	171	239	75	66	73
Home Language						
English	652 (57%)	647 (59%)	642 (65%)	42 (6.8%)	25 (4.7%)	36 (9.0%)
Other	75 (6.6%)	67 (6.1%)	78 (7.9%)	5 (0.8%)	2 (0.4%)	4 (1.0%)
Spanish	418 (37%)	387 (35%)	268 (27%)	570 (92%)	509 (95%)	361 (90%)
Unknown	16	13	52	8	2	7

¹Test

3.2 Descriptives

3.2.1 Distributions of English Multitudes Tasks

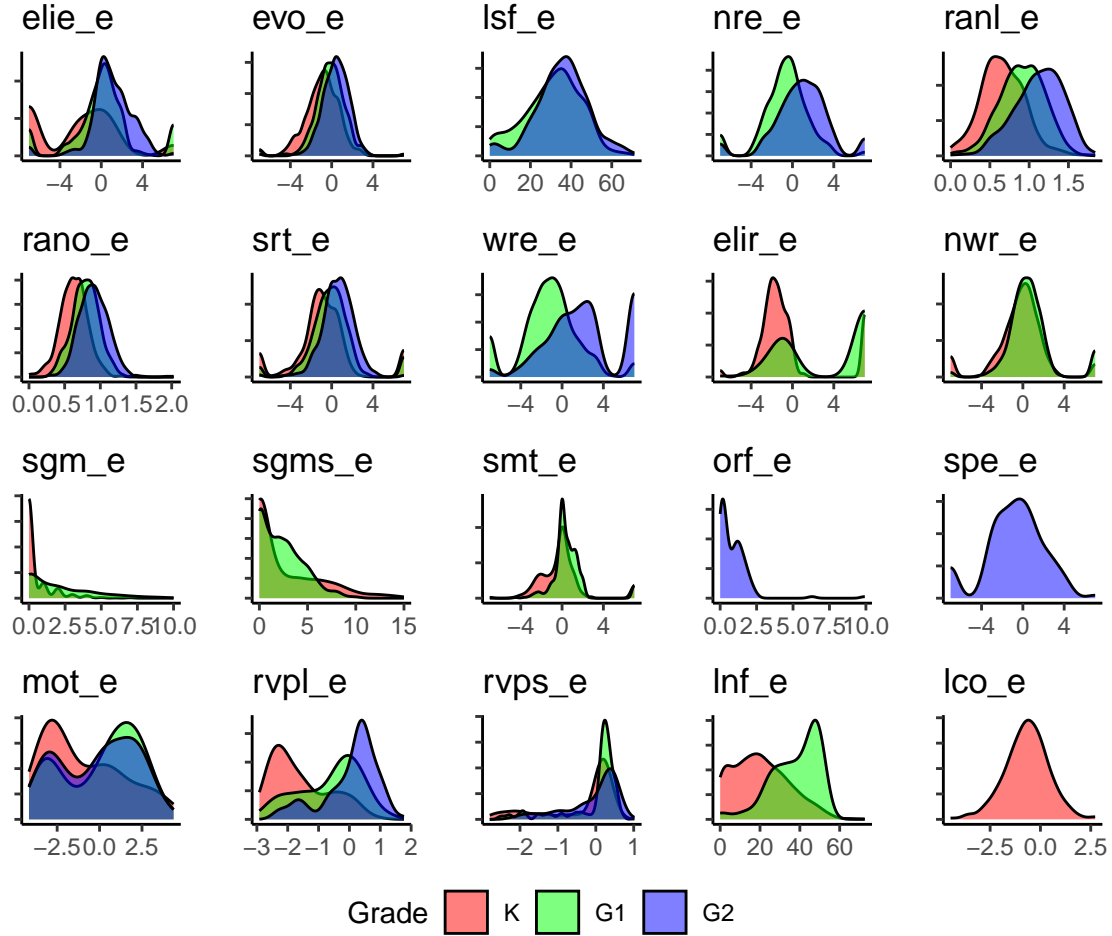


Figure 34: Distributions of English Multitudes Tasks by Grade (Fall 2023 and 2024 Combined).

3.2.2 Distributions of Spanish Multitudes Tasks

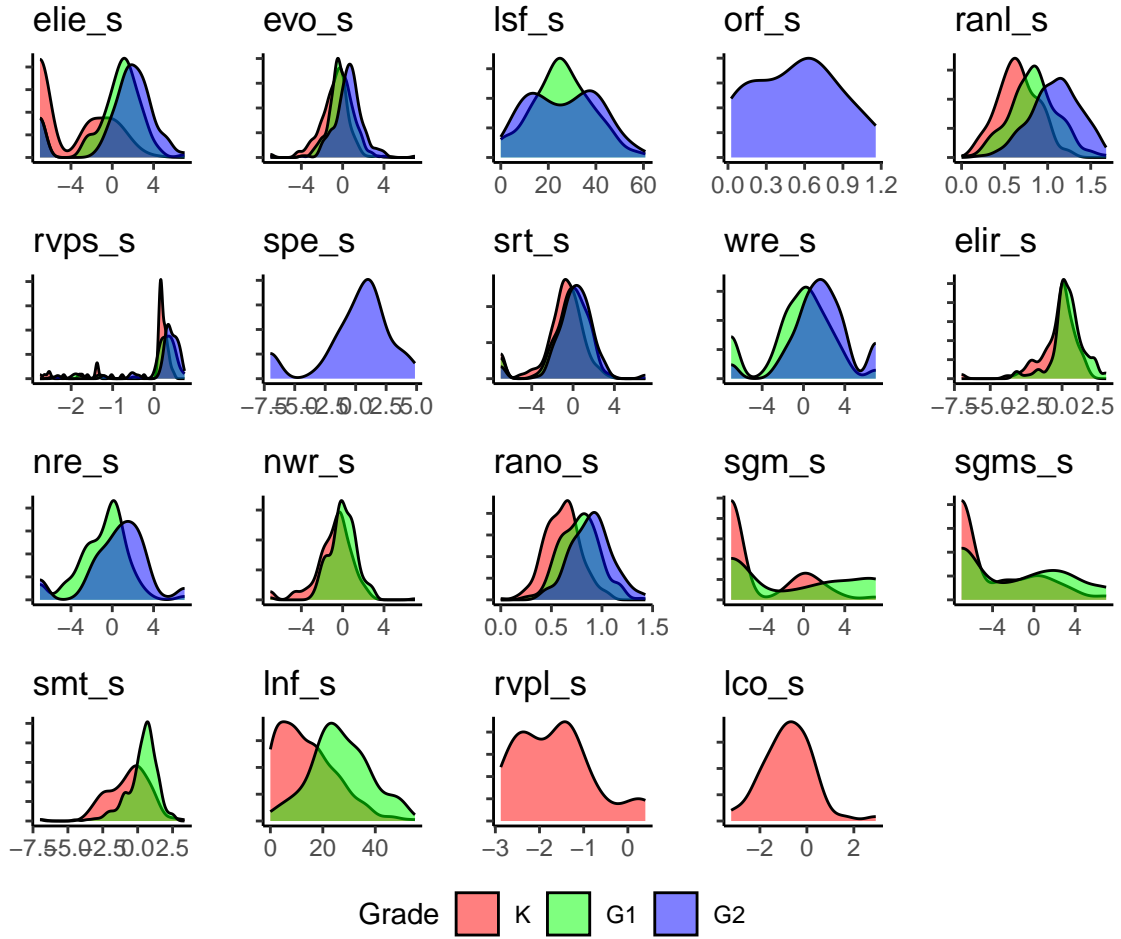


Figure 35: Distributions of Spanish Multitudes Tasks by Grade (Fall 2023 and 2025 Combined).

Table 8: Comparison of Proportions of Struggling Readers (\leq 20th Percentile) in the Study Sample Based on Overall Multitudes Sample and on Published Woodcock-Johnson/Muñoz Norms (Spring 2024 and 2025 Combined).

Reference	English			Spanish		
	K N = 1,196	G1 N = 1,150	G2 N = 1,139	K N = 504	G1 N = 358	G2 N = 297
Norm-based						
Not struggling	701 (59%)	666 (58%)	752 (66%)	227 (45%)	211 (59%)	154 (52%)
Struggling	495 (41%)	484 (42%)	387 (34%)	277 (55%)	147 (41%)	143 (48%)
Sample-based						
Not struggling	955 (80%)	900 (78%)	884 (78%)	399 (79%)	279 (78%)	223 (75%)
Struggling	241 (20%)	250 (22%)	255 (22%)	105 (21%)	79 (22%)	74 (25%)

3.2.3 Distributions of Outcome Measures

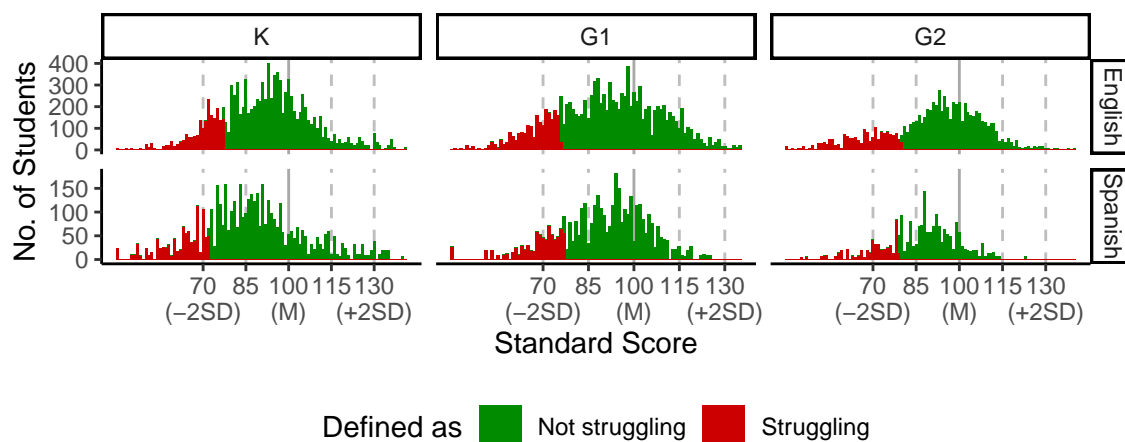
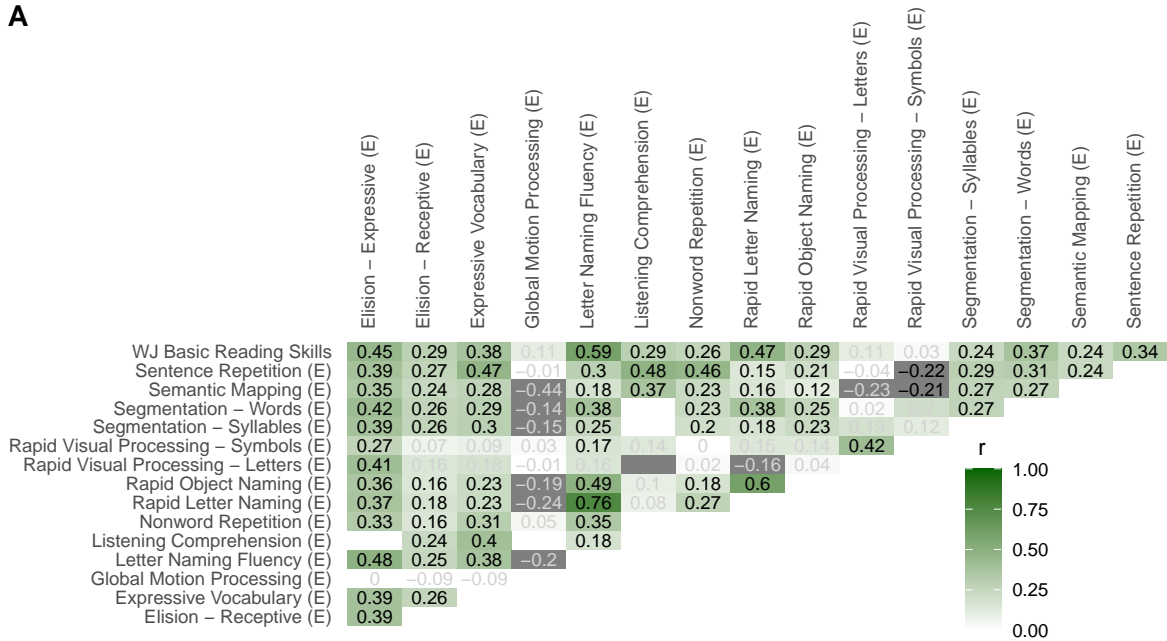


Figure 36: Distribution of Woodcock-Johnson/-Muñoz Basic Reading Skills Cluster (for Kindergarten) Broad Reading Cluster (for Grades 1 and 2) by Grade and Language with Struggling Readers (\leq 20th Percentile) Highlighted (Spring 2024 and 2025 Combined)

3.2.4 Correlations

3.2.4.1 Kindergarten - English

A



B

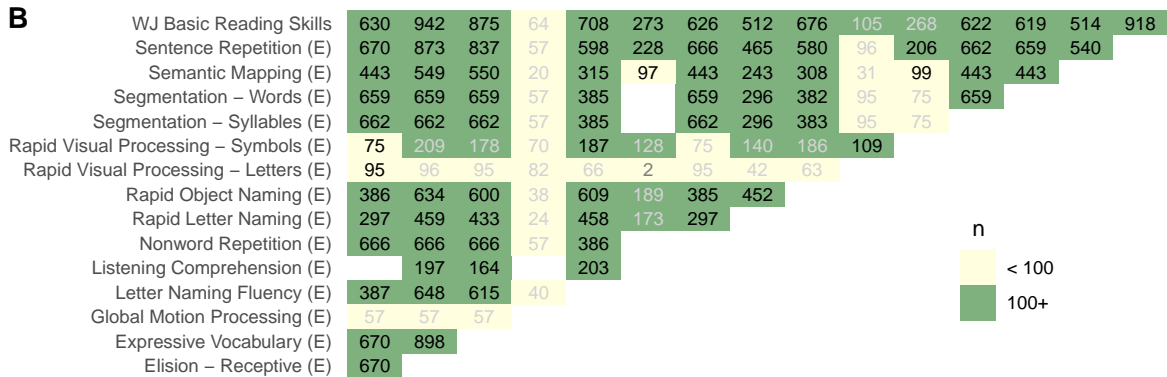
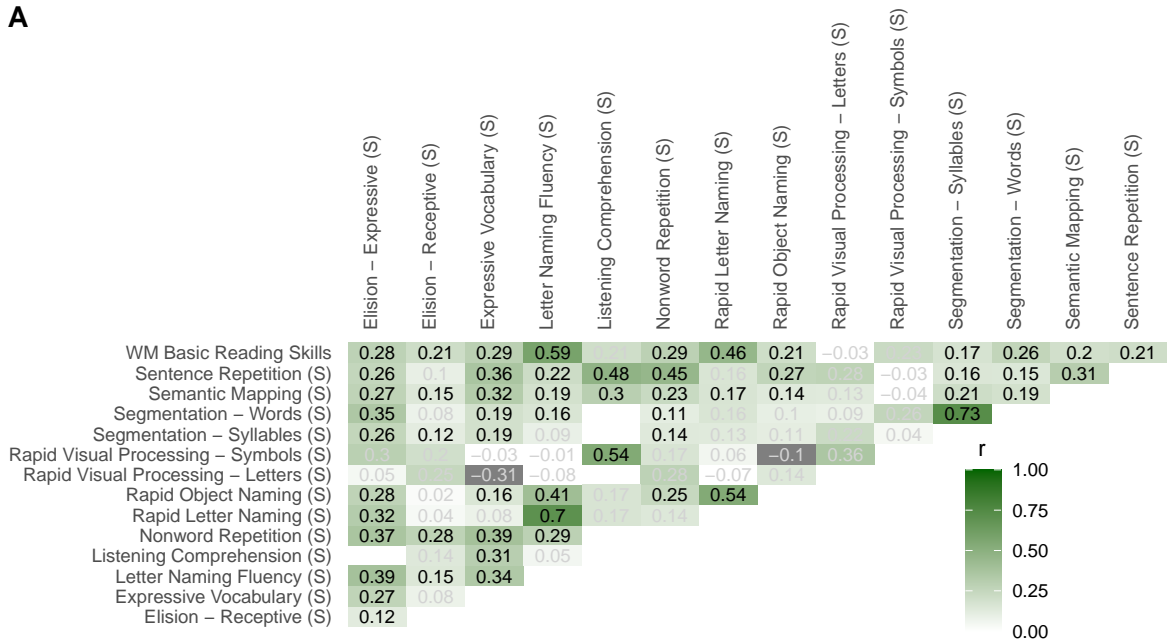


Figure 37: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Kindergarten; AYs 2023/24 and 2024/25 Combined).

3.2.4.2 Kindergarten - Spanish

A



B

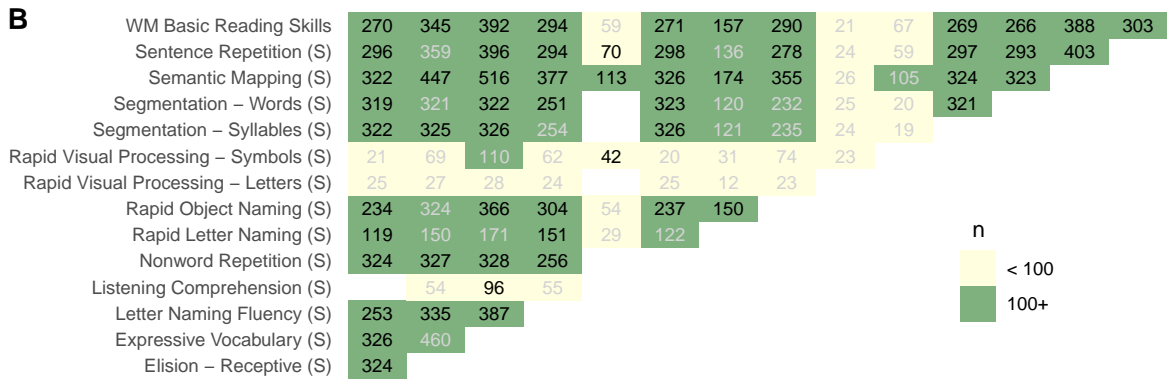
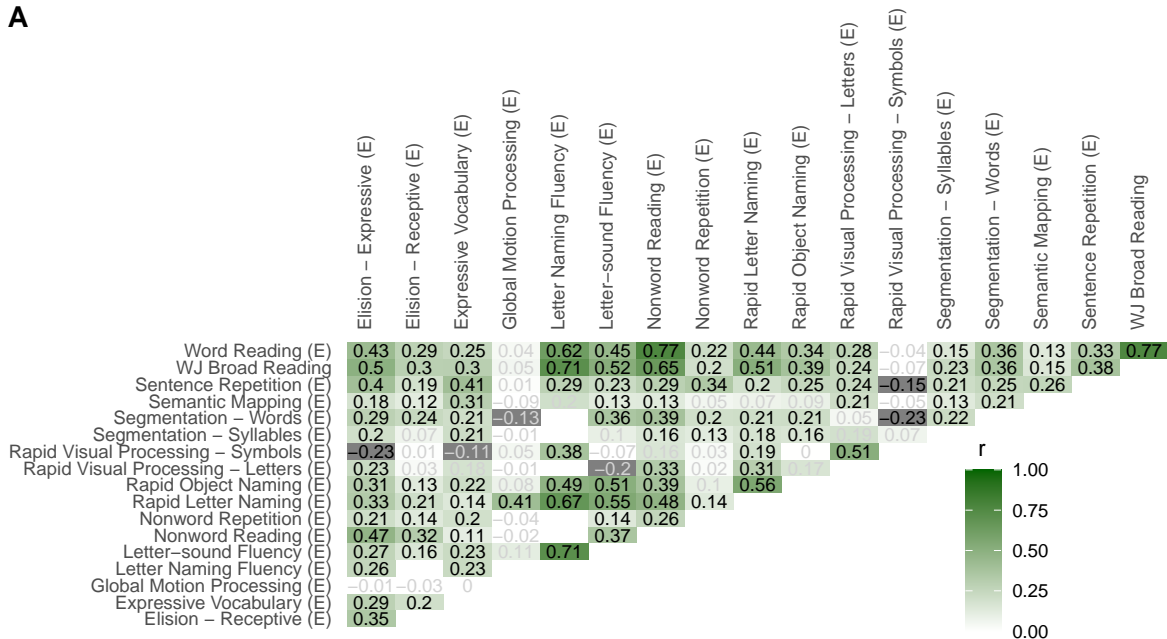


Figure 38: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Kindergarten; Ays 2023/24 and 2024/25 Combined).

3.2.4.3 Grade 1 - English

A



B

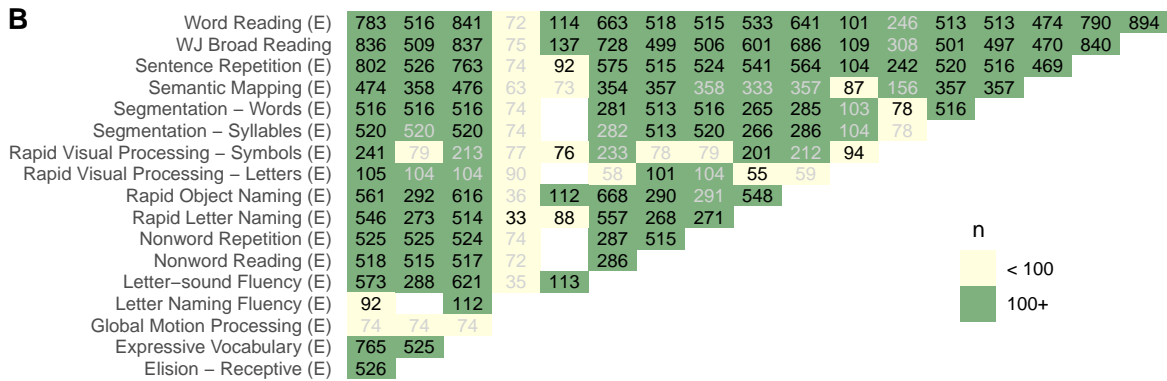
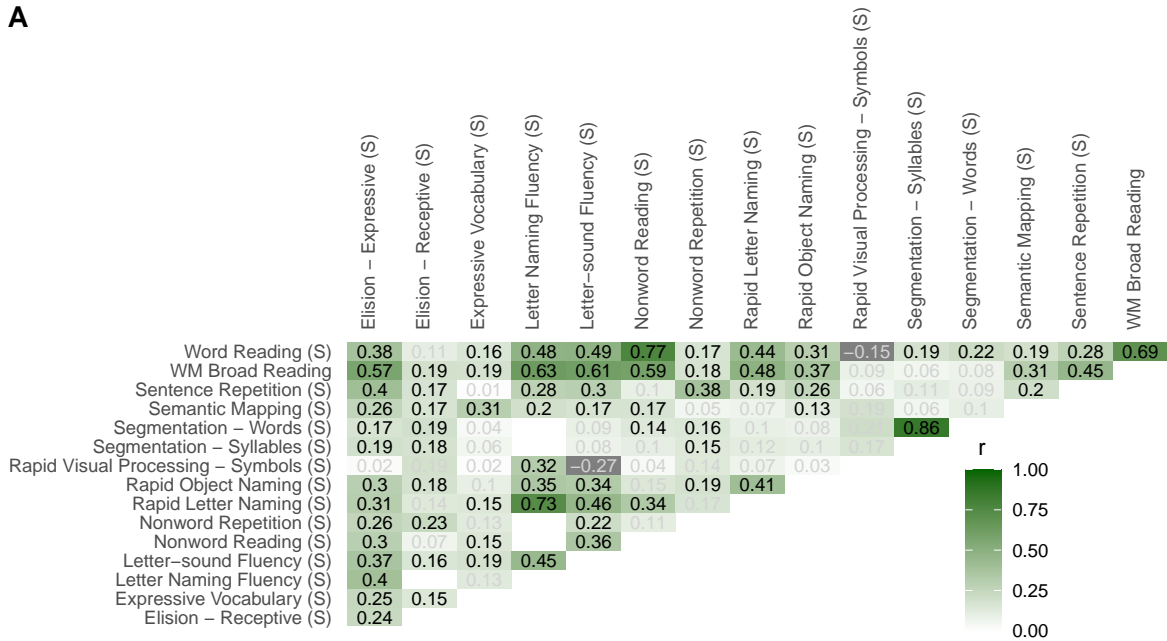


Figure 39: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Grade 1; AYs 2023/24 and 2024/25 Combined).

3.2.4.4 Grade 1 - Spanish

A



B

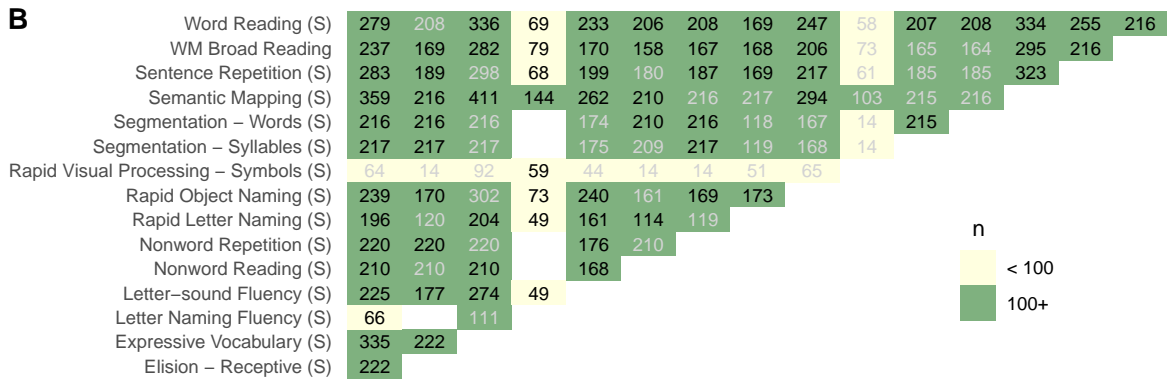
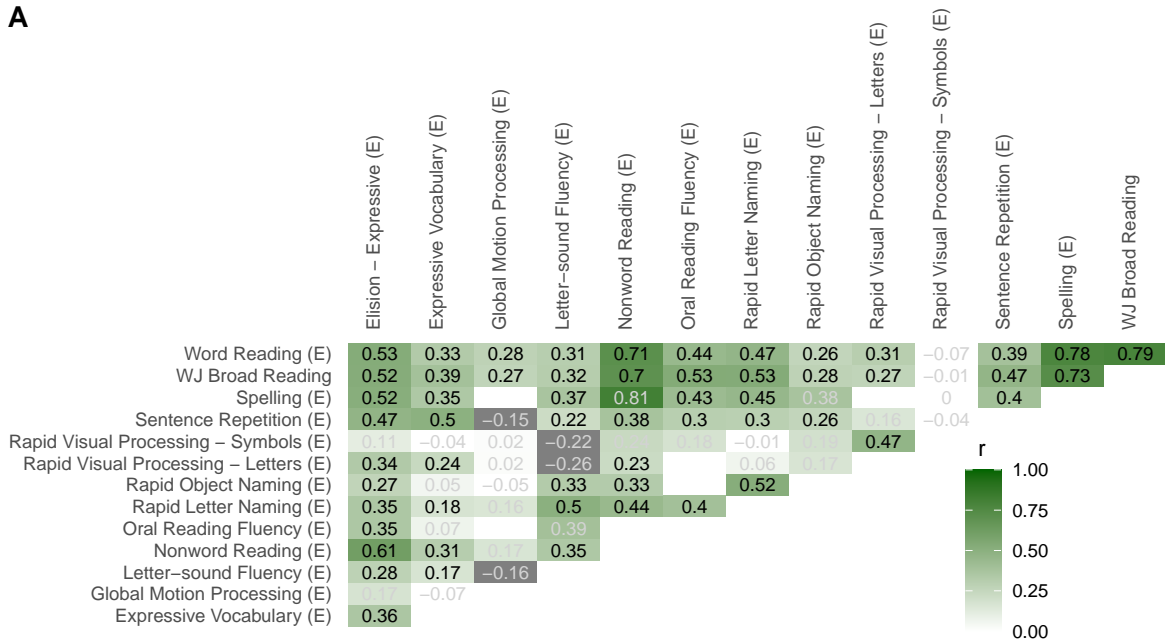


Figure 40: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Grade 1; AYs 2023/24 and 2024/25 Combined).

3.2.4.5 Grade 2 - English

A



B

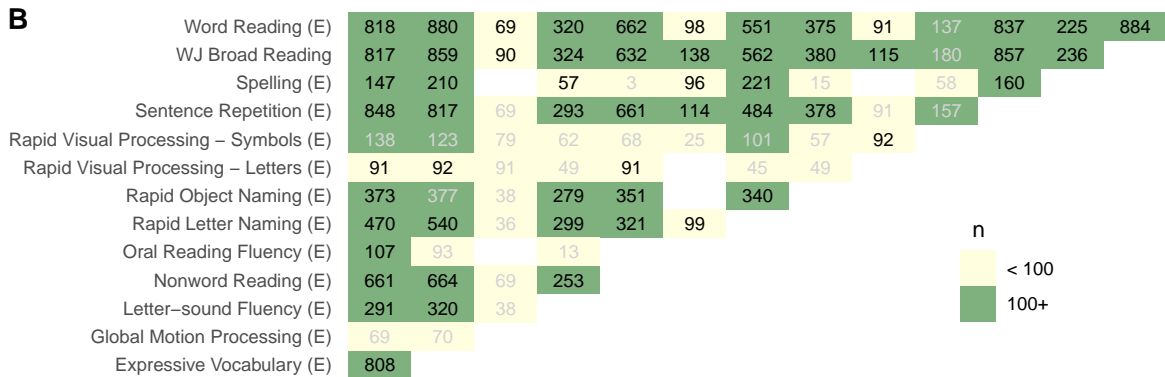
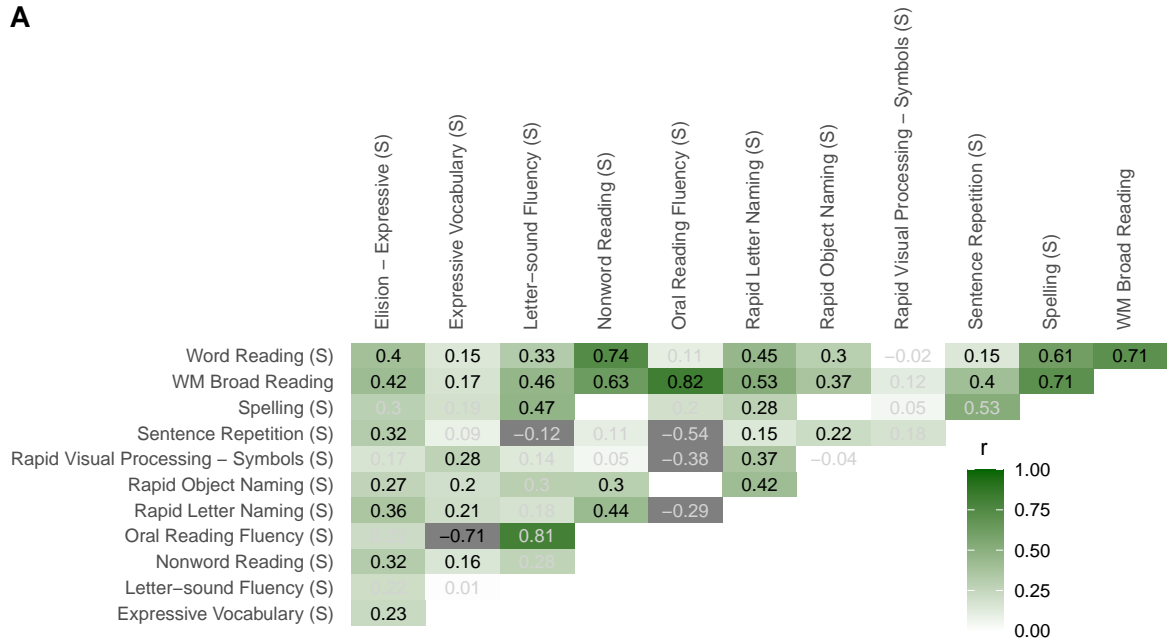


Figure 41: Correlations Between English Multitudes Tasks and English Reading Outcome Scores (Grade 2; AYs 2023/24 and 2024/25 Combined).

3.2.4.6 Grade 2 - Spanish

A



B

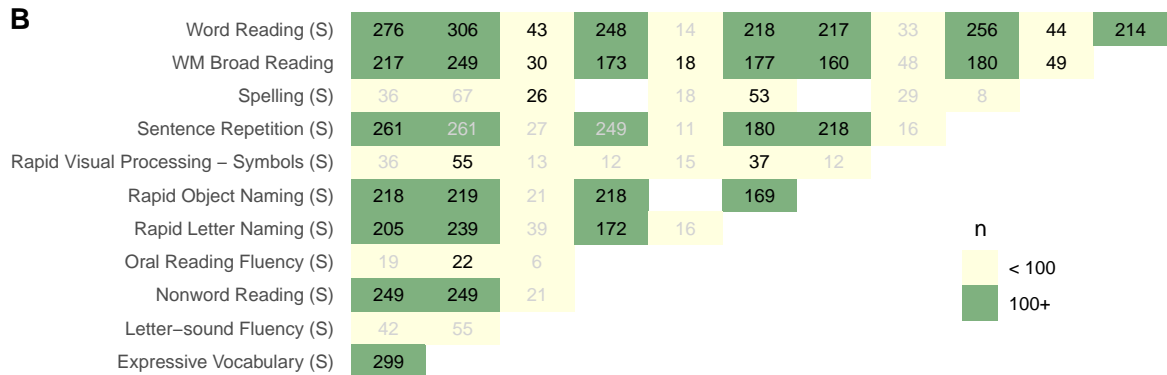


Figure 42: Correlations Between Spanish Multitudes Tasks and Spanish Reading Outcome Scores (Grade 2; AYs 2023/24 and 2024/25 Combined).

3.3 Task Selection

3.3.1 English-to-English (English to English)

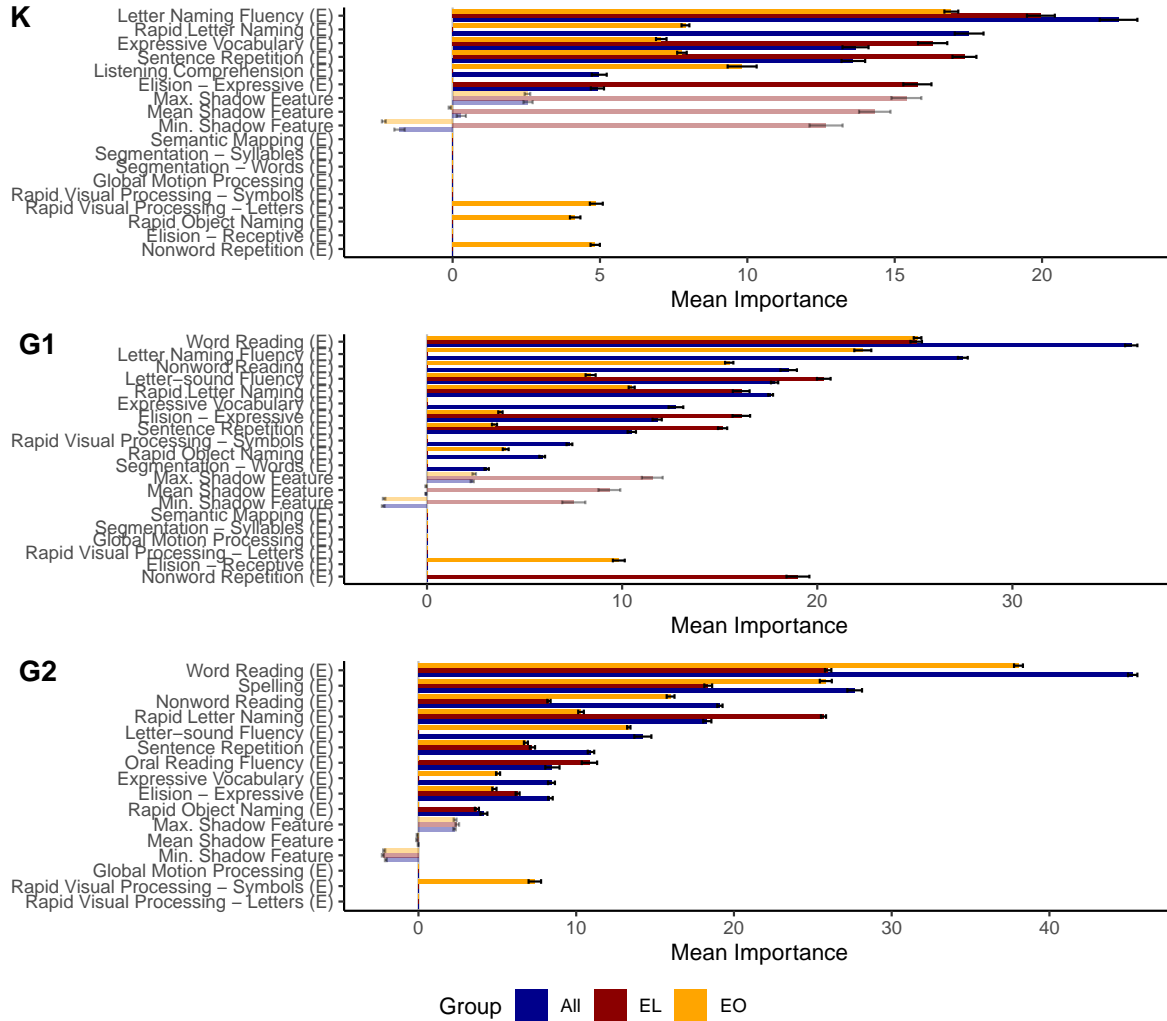


Figure 43: Mean Task Importance Obtained From Boruta Algorithm for English-to-English Predictions, with Results Disaggregated by English Proficiency Designation (AYs 2023/24 and 2024/25 Combined).

3.3.2 Spanish-to-Spanish (Spanish to Spanish)

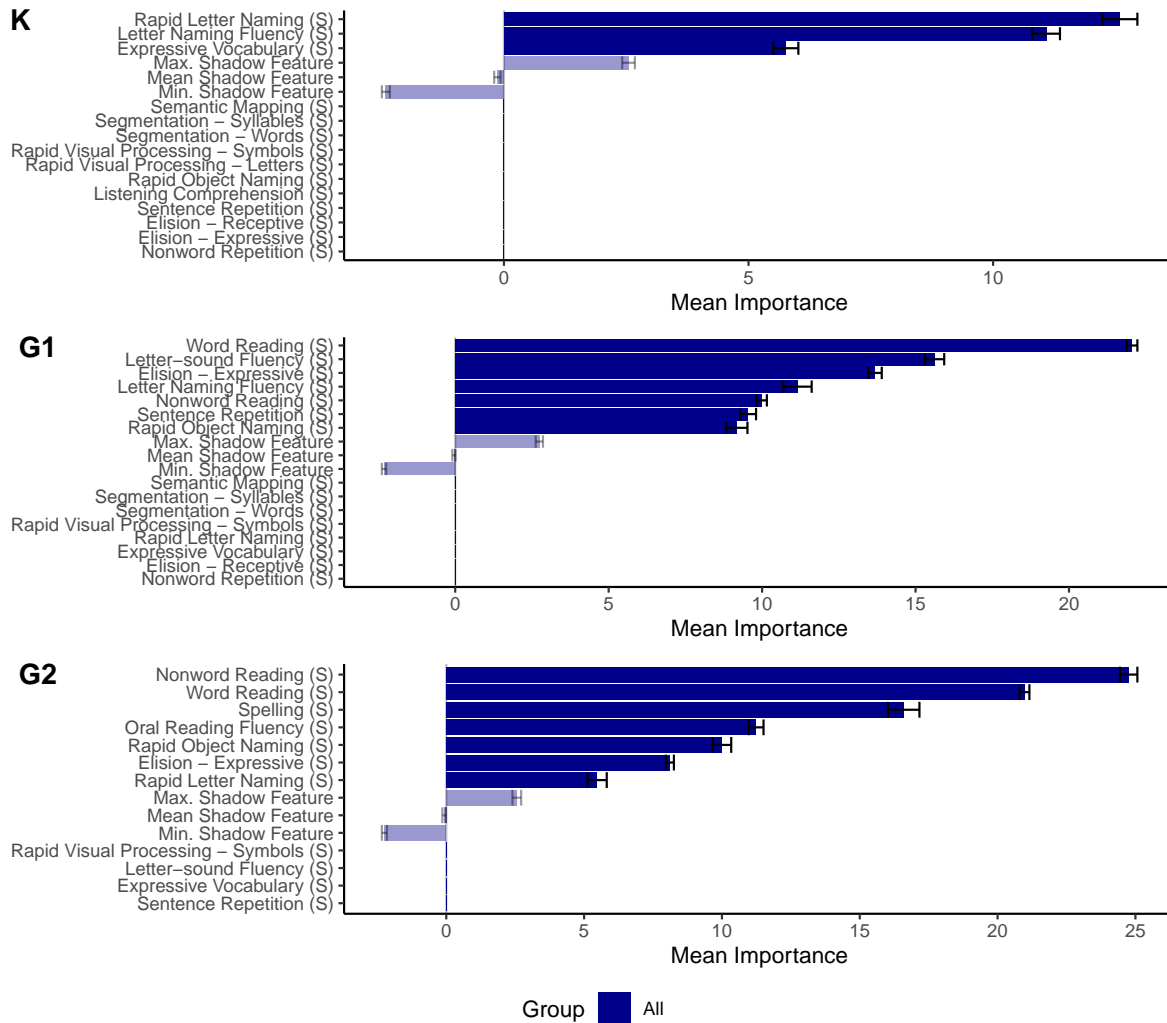


Figure 44: Mean Task Importance Obtained From Boruta Algorithm for Spanish-to-Spanish Predictions (AYs 2023/24 and 2024/25 Combined).

3.4 Models Evaluations - ROC Curves

3.4.1 English-to-English - Kindergarten

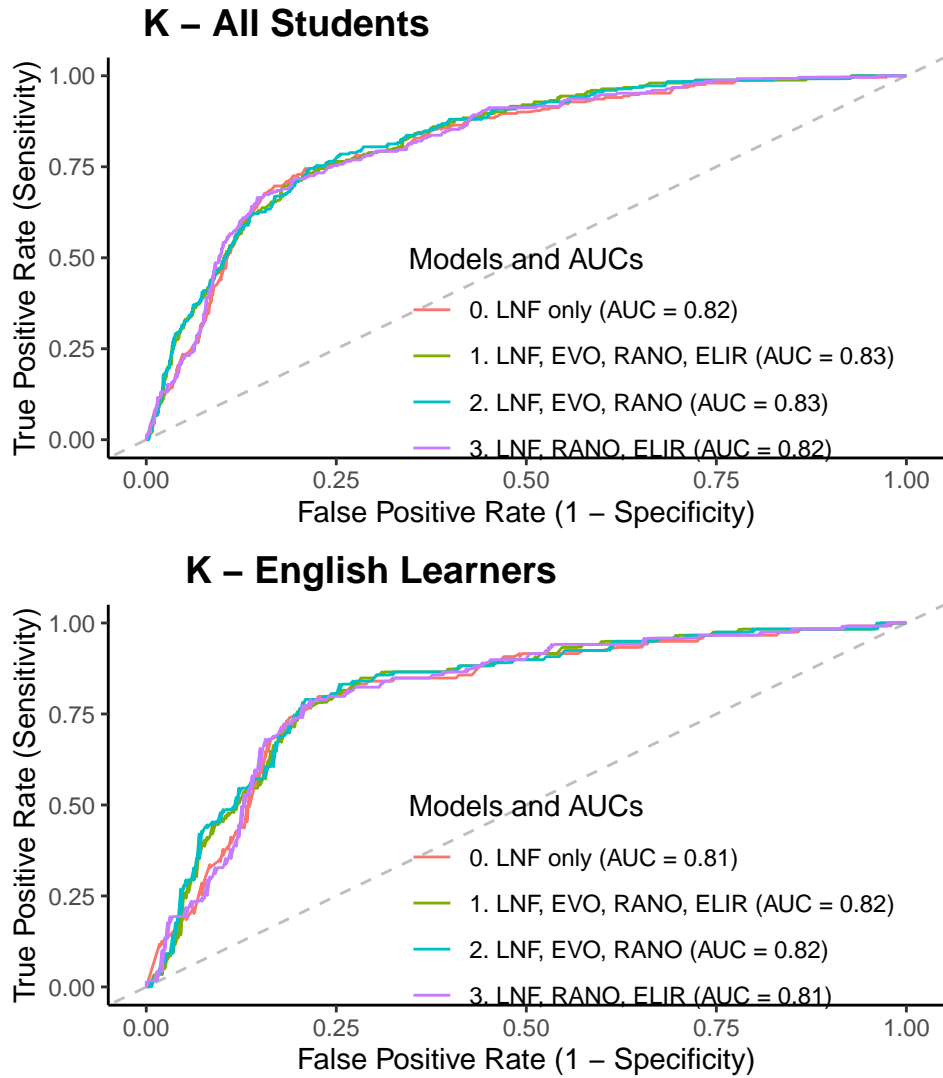


Figure 45: Receiver-operator Characteristic Curves For English-to-English Kindergarten Models (AYs 2023/24 and 2024/25 Combined).

3.4.2 English-to-English - Grade 1

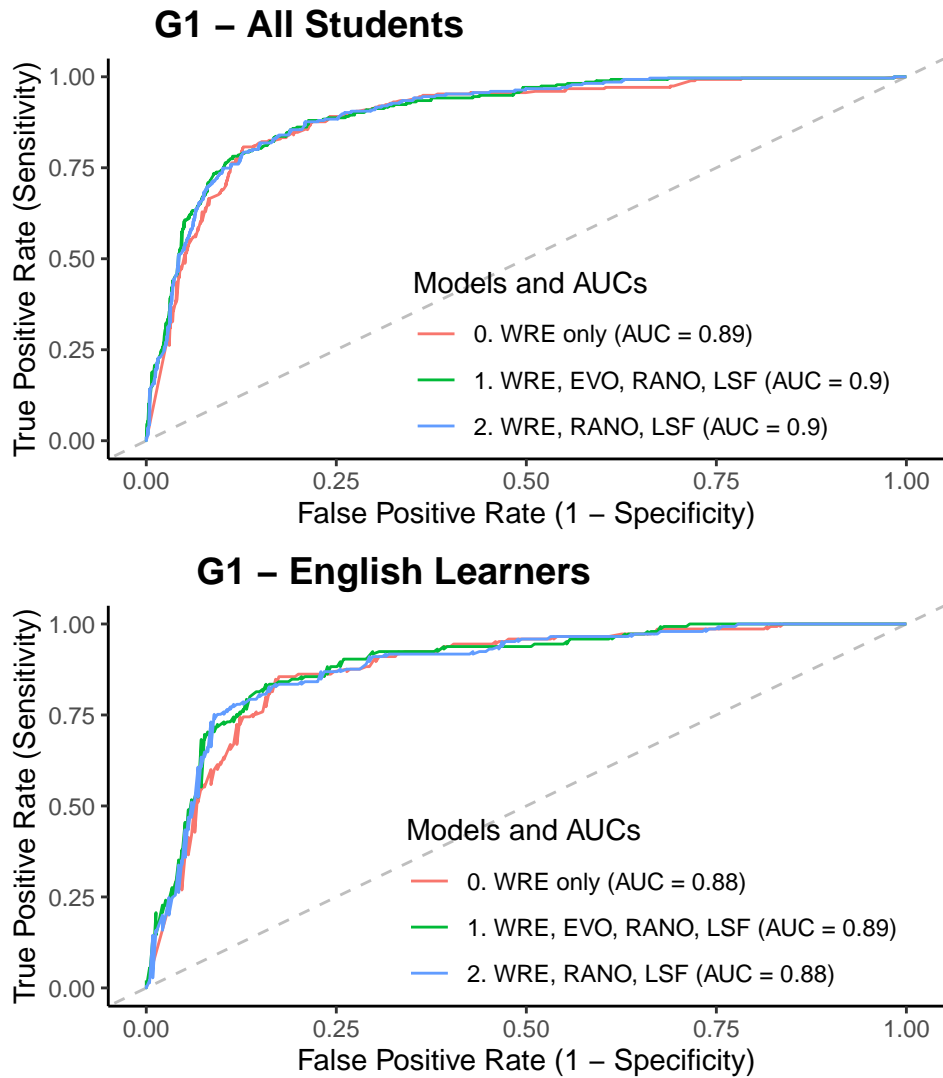


Figure 46: Receiver-operator Characteristic Curves For English-to-English Grade 1 Models (AYs 2023/24 and 2024/25 Combined).

3.4.3 English-to-English - Grade 2

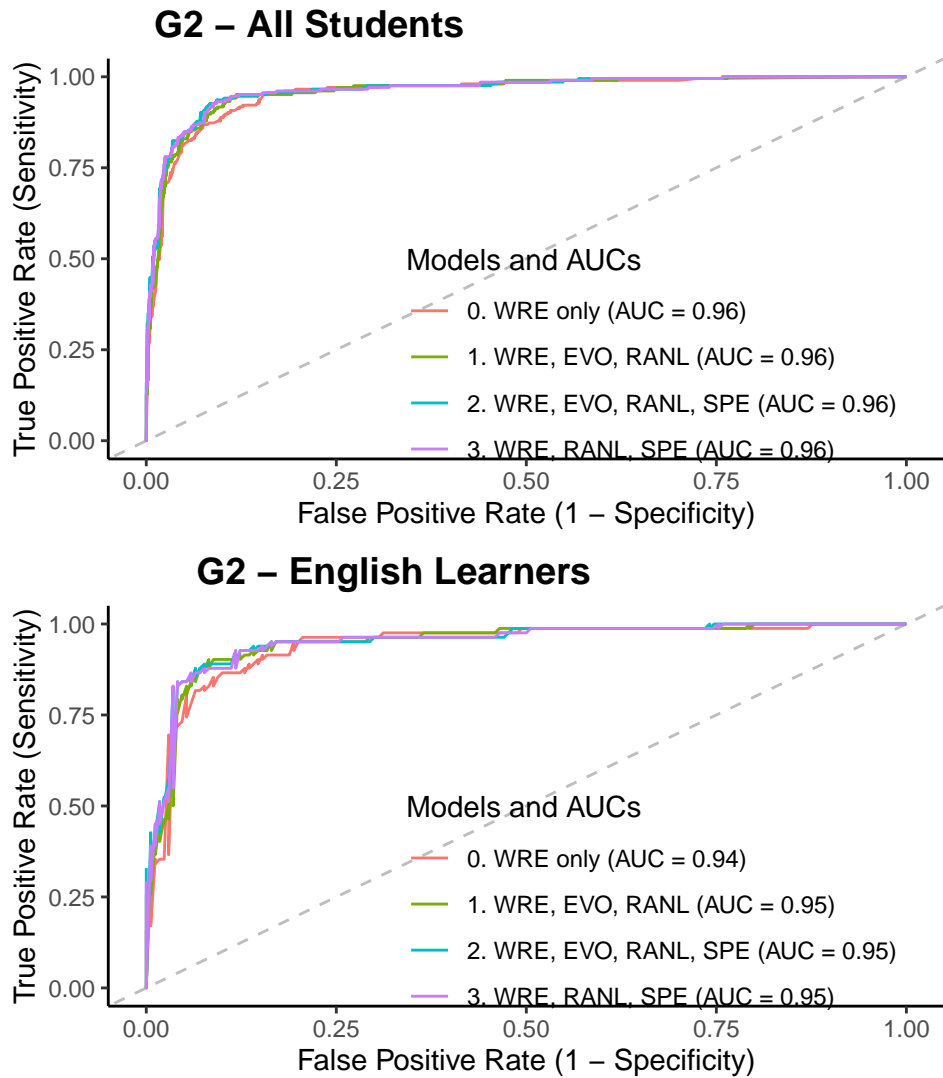


Figure 47: Receiver-operator Characteristic Curves For English-to-English Grade 2 Models (AYs 2023/24 and 2024/25 Combined).

3.4.4 Spanish-to-Spanish - Kindergarten

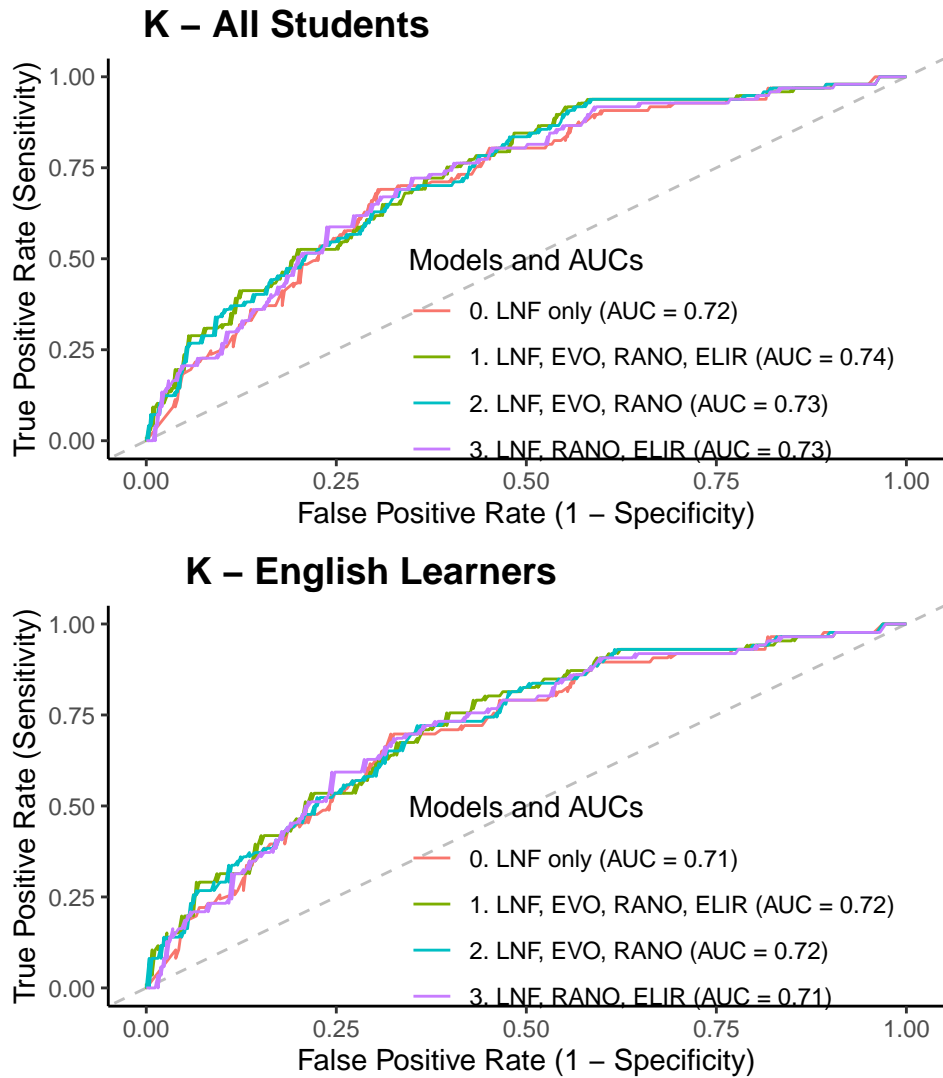


Figure 48: Receiver-operator Characteristic Curves For Spanish-to-Spanish Kindergarten Models (AYs 2023/24 and 2024/25 Combined).

3.4.5 Spanish-to-Spanish - Grade 1

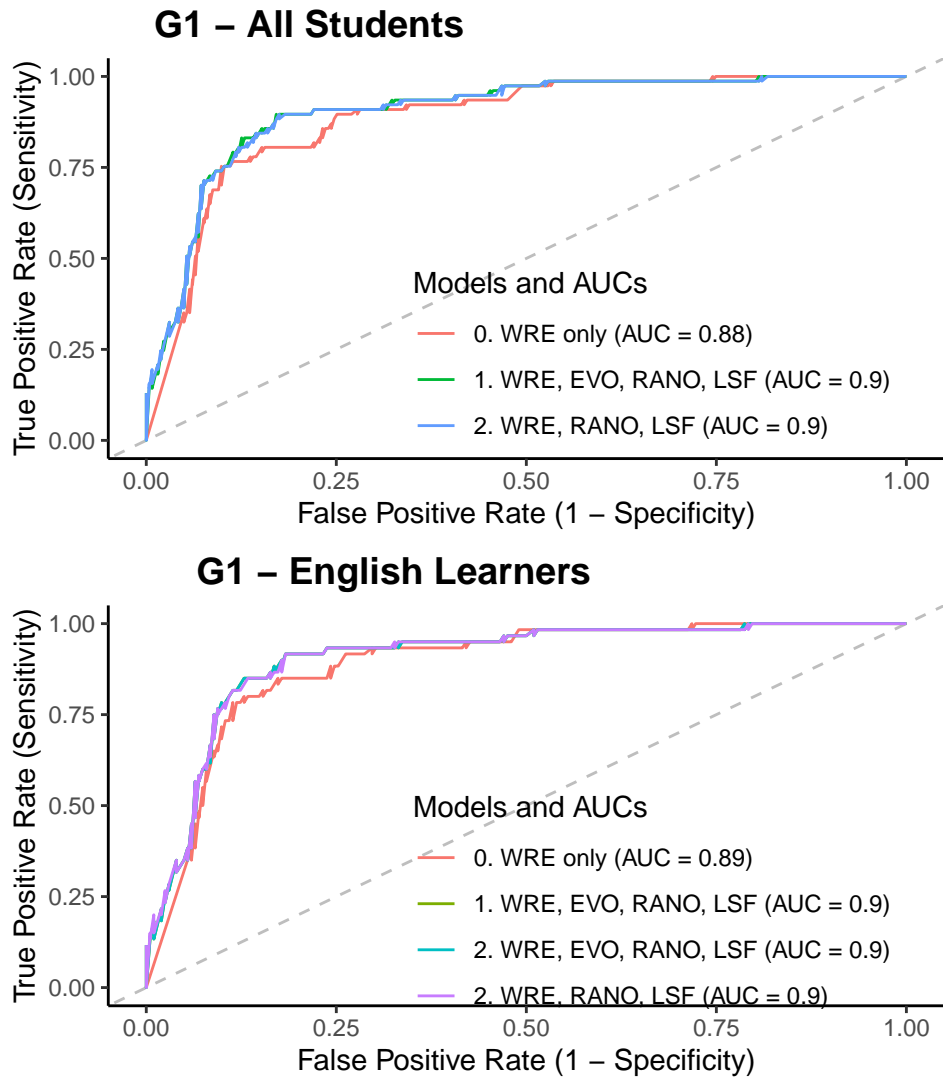


Figure 49: Receiver-operator Characteristic Curves For Spanish-to-Spanish Grade 1 Models (AYs 2023/24 and 2024/25 Combined).

3.4.6 Spanish-to-Spanish - Grade 2

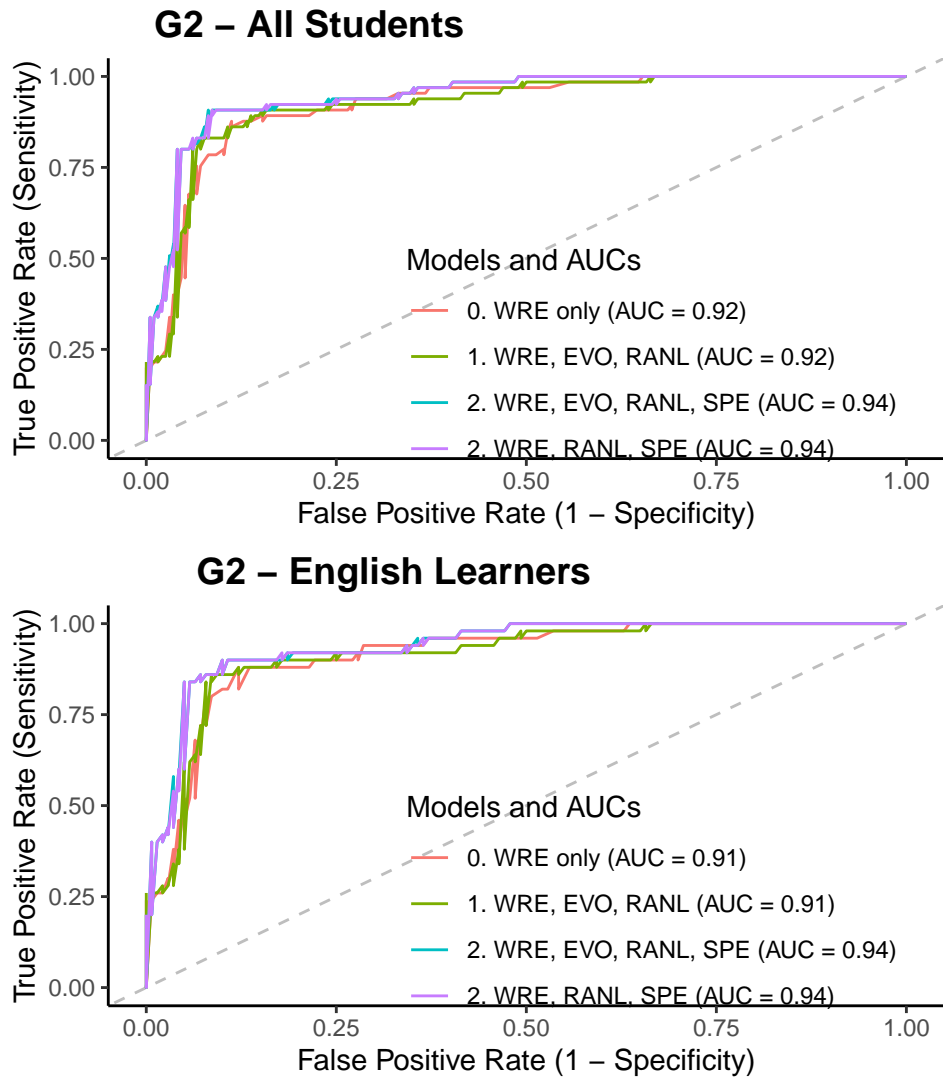


Figure 50: Receiver-operator Characteristic Curves For Spanish-to-Spanish Grade 2 Models (AYs 2023/24 and 2024/25 Combined).

Table 9: Summary of Best English and Spanish Prediction Models (AYs 2023/24 and 2024/25 Combined).

Grade	Model	English												Spanish			
		Sensitivity			Specificity			Accuracy			Balanced Acc.			Sens.	Spec.	Acc.	BA
		All	EL	EO	All	EL	EO	All	EL	EO	All	EL	EO	All	All	All	All
K	LNF only	0.741	0.8403	0.6471	0.7921	0.7073	0.8355	0.7813	0.7666	0.7463	0.7738	0.8014	0.7413	0.6804	0.695	0.6918	0.6877
K	LNF, RANO, ELIR	0.7131	0.8151	0.6176	0.806	0.7282	0.842	0.7864	0.7596	0.7537	0.7717	0.8014	0.7298	0.7113	0.651	0.6644	0.6812
K	LNF, RANO, ELIR, EVO	0.7251	0.8655	0.598	0.7921	0.676	0.8528	0.778	0.7586	0.7315	0.7708	0.8067	0.7254	0.9072	0.4487	0.5502	0.6779
K	LNF, RANO, EVO	0.741	0.8655	0.6176	0.7857	0.676	0.8442	0.7763	0.7634	0.7315	0.7708	0.8032	0.7309	0.6804	0.6686	0.6712	0.6745
G1	WRE only	0.8036	0.8483	0.7245	0.8734	0.8298	0.8783	0.8574	0.8385	0.8368	0.839	0.8544	0.8014	0.7532	0.8897	0.8588	0.8215
G1	WRE, RANO, LSF	0.8145	0.8345	0.7653	0.8506	0.7957	0.8652	0.8424	0.8326	0.8105	0.8151	0.8497	0.8152	0.8831	0.8251	0.8382	0.8541
G1	WRE, RANO, LSF, EVO	0.7782	0.8414	0.6735	0.8864	0.8213	0.9045	0.8616	0.8323	0.8289	0.8313	0.8687	0.789	0.8831	0.8289	0.8412	0.856
G2	WRE only	0.9512	0.9634	0.955	0.8444	0.7882	0.8454	0.866	0.8978	0.8452	0.8758	0.8654	0.9002	0.8615	0.8878	0.8812	0.8746
G2	WRE, RANL	0.9317	0.9268	0.9459	0.8901	0.8529	0.8855	0.8985	0.9109	0.877	0.8899	0.8966	0.9157	0.8769	0.8673	0.8697	0.8721
G2	WRE, RANL, EVO	0.9415	0.9268	0.964	0.8815	0.8353	0.8815	0.8936	0.9115	0.8651	0.8811	0.8966	0.9227	0.8154	0.9286	0.9004	0.872

Note. green: >= .8; black: >= .7; yellow: >= .6; red otherwise

3.5 Models Evaluations - Sensitivity and Specificity

4 Part IV - Evaluating 23/24 Model With 24/25 Data

Table 10: Re-evaluation of 2023/24 Prediction Model Using 2024/25 Data.

Grade	Model	English												Spanish			
		Sensitivity			Specificity			Accuracy			Balanced Acc.			Sens.	Spec.	Acc.	BA
		All	EL	EO	All	EL	EO	All	EL	EO	All	EL	EO	All	All	All	All
K	LNF, EVO, ELIR, RANO	0.7429	0.8182	0.7222	0.8196	0.6269	0.8632	0.8122	0.7812	0.6538	0.7225	0.851	0.7927	0.6	0.7811	0.7577	0.6905
G1	WRE, EVO, LSF, RANO	0.8462	NA	0.8462	0.7689	0.6667	0.797	0.7726	0.8075	0.6667	NA	0.8	0.8216	0.7059	0.95	0.9145	0.8279
G2	WRE, EVO, RANL	0.9767	1	1	0.8718	0.75	0.8837	0.8861	0.9243	0.8043	0.875	0.8984	0.9419	0.75	0.9327	0.8984	0.8413

Note. green: >= .8; black: >= .7; yellow: >= .6; red otherwise