

Correlation Between NoRedInk Performance and Achievement on the MAP Language Usage Assessment

Submitted to: NoRedInk

Bryan Matlen, Ph.D.
Min Huang, M.P.H.
Niufeng Zhu, M.S.

March 2017

Table of Contents

Executive Summary	2
Introduction	3
Method	3
Sample	3
Measures	5
Analytic Procedure	6
Results	6
Discussion	11
Limitations	11
Conclusions	11
References	12

Executive Summary

NoRedInk (www.noredink.com) is an adaptive learning platform that supports grammar and writing skills for students in grades 5-12. As many of the topics covered by NoRedInk are commonly found on standardized language assessments, a correlation is hypothesized between performance on NoRedInk and performance on language skills assessments.

WestEd (www.wested.org) – an educational research and service agency – partnered with NoRedInk to explore possible links between middle school students' language achievement and their performance on NoRedInk – specifically, the number of topics students master within the platform. To this end, WestEd examined correlations between students' experience on NoRedInk and their scores on the Northwest Evaluation Association's (NWEA; <https://www.nwea.org>) Measure of Academic Progress (MAP) Language Usage Assessment. The analysis consisted of 7,330 students from a large school district in the southeastern United States over the 2015-2016 school year.

Results indicated that, after controlling for students' demographic characteristics, baseline achievement, and the number of topics attempted on NoRedInk, mastery of NoRedInk topics (i.e., performance) was positively correlated with students' MAP language scores. Specifically, every 10 topics mastered in NoRedInk corresponded to 20% of students' expected annual language growth on average.

This study is exploratory and provides an initial step in understanding how NoRedInk experience relates to language achievement on a standardized assessment. Because this study did not include a comparison group of students not using NoRedInk, the results should not be used to draw causal inferences about students' performance on NoRedInk and language achievement on the MAP. Future studies could utilize a comparable control group (e.g., through random assignment) so that causal connections can be examined.

Introduction

NoRedInk (www.noredink.com) is an adaptive learning platform aimed at supporting students' grammar and writing skills. NoRedInk allows students to practice their language skills with authentic and engaging text. The platform provides immediate feedback and tracks student progress over time. NoRedInk has gained widespread popularity and adoption by teachers across the United States since its launch in 2012. Mastering grammar and writing topics in NoRedInk is hypothesized to result in improved language skills for students in grades 5-12.

A previous study conducted by Perez (2015) examined the correlation between 7th grade students' NoRedInk usage and performance and those same students' language achievement as measured by the Northwest Evaluation Association's Measure of Academic Progress (MAP) Language Usage Assessment. This study found a statistically significant positive correlation between students' NoRedInk performance and MAP language achievement, providing some support for the hypothesis that NoRedInk supports student language achievement. Though results of the Perez study show promise for a link between NoRedInk performance and students' language achievement, the results were limited to a single grade level (7th) in a single school ($N = 265$). Thus, it remains an open question whether the correlation between NoRedInk performance and language achievement will generalize to larger and more diverse samples.

The question we explore in the present study is: After controlling for students' baseline achievement and demographic characteristics, what is the relationship between mastery of NoRedInk topics and students' language growth over one academic year? To answer this question, we re-examined the hypothesis that NoRedInk supports students' language achievement, using procedures similar to Perez, but with a larger and more heterogeneous sample. Specifically, the study involves a sample of 7,330 students spread across three grade levels (6-8) from a large district in the southeastern United States. Mastery of NoRedInk topics in this study is defined by a proprietary algorithm developed by NoRedInk, which takes into consideration students' individual performance on specific subskills. We examine the correlation between the number of topics students mastered and their MAP language achievement scores, after controlling for the number of topics attempted, their baseline MAP score, and demographic variables.

Method

Sample

The sample for this study consisted of middle school students from grades 6 – 8. A description of the sample is provided in Table 1. Students came from 11 traditional, Title I schools, and one Education/Learning center in a large district in the southeastern United States (see Table 2).

Table 1. Demographic information on student sample.

Students' characteristics, performance and experience on NoRedInk		Students who attempted at least one NoRedInk topic in 2015-16
		n=7,330
Gender (Percent)	Female	49.3
	Male	50.7
Grade (Percent)	Grade 6	34.4
	Grade 7	35.2
	Grade 8	30.4
Ethnicity (Percent)	White	67.1
	African American	17.8
	Hispanic	8.0
	Others	5.7
	Asian	1.4
Average MAP score (SD)	Overall fall 2015	216.1 (13.0)
	Goal 1 (PODRR) fall 2015	216.8 (14.8)
	Goal 2 (Grammar) fall 2015	215.4 (13.4)
	Goal 3 (Mechanics) fall 2015	216.0 (13.4)
Average number of NoRedInk topics (SD)	Attempted	65 (37.7)
	Mastered	34 (28.3)

* Others included Native Americans, Pacific Islanders, and multiracial

Table 2. Demographic information on participating schools (each school is represented as a row*).

White	Free or Reduced Price Lunch	Underrepresented Minority ¹	School Locale	# Students at School
85%	60%	12%	Rural	675
60%	68%	33%	Rural	671
62%	64%	33%	City	515
70%	57%	22%	Suburb	1,195
60%	72%	38%	Rural	622
53%	75%	43%	Rural	702
49%	71%	42%	City	1,090
61%	58%	30%	Rural	1,109
76%	43%	16%	Suburb	1,161
78%	47%	14%	Suburb	1,194
50%	83%	46%	City	804
<i>Averages</i>				
64%	63%	30%	NA	885

* One education/learning center also participated, and is not included in the table.

¹ Defined as Black, Hispanic, and American Indian/Alaskan students.

Measures

Data for the study consisted of the Northwest Evaluation Association's Measure of Academic Progress (MAP) Assessment in Language. We used NWEA's Rasch Unit scale (RIT) score as the primary outcome. The RIT is a score derived from Item-Response Theory that is on an equal interval scale (i.e., it can be compared across grade levels). The assessment was administered three times in the school year---fall, winter, and spring. The growth from fall to spring was used as the outcome measure, and fall score was used as baseline performance. In one of our models, we derived a measure of percentage of expected annual growth based on NWEA's projected growth from fall to spring for each student. We divided the actual observed growth by the NWEA's projected growth to determine whether each student met, surpassed, or fell short of their projected growth. The projection is based on the average growth of students in the same grade beginning with a similar RIT score in NWEA's large national normative sample ("2015 NWEA Measures of Academic Progress Normative Data", 2015). The MAP assessment also includes subscores that cover three areas: 1) *Writing: Plan, Organize, Develop, Revise, Research (PODRR)*, 2) *Language: Understand, Edit for Grammar, and Usage, (Grammar)* and 3) *Language: Understand, Edit Mechanics (Mechanics)*. Additional analyses looked at the relationship of NoRedInk performance with these subscores.

Within NoRedInk, students can attempt questions that are aligned with various language topics. Topics range from commonly confused words and capitalization rules to parallel structure and active vs. passive voice. NoRedInk tracks students' performance over time on each topic, and when students have demonstrated consistent accuracy on a given topic, they are considered to have mastered it.

The topics chosen for analysis in this study were those that aligned with the MAP language assessment. (See [NoRedInk NWEA Alignment](#) for the full list of topics included in this study.) Using NoRedInk's mastery criteria, performance on NoRedInk was defined as the number of aligned topics each student mastered

within the study period. Attempts were defined as the total number of aligned topics each student attempted on NoRedInk during the study period. NoRedInk user data was extracted from NoRedInk databases and matched with NWEA data. Students with multiple accounts in NoRedInk during the study period were aggregated to produce one attempt measure and performance score for each student.

Student demographic data consisted of gender (male or female), grade (6th, 7th, or 8th) and ethnicity (White, Asian, African American, Hispanic, or other).

Analytic Procedure

The primary analytic procedure involved fitting multivariate linear regression models. For each school year and sample, we conducted analysis by regressing the growth in between spring and fall RIT scores onto students' usage and performance in NoRedInk, controlling for fall RIT score, gender, grade, and ethnicity.

In the model above, represents the outcome (e.g., growth in between the spring and fall MAP scores) for the i -th student, and are numeric values that represent a) the total number of NoRedInk topics attempted by each student and b) the total number of NoRedInk topics considered "mastered" by each student, respectively. is the students' fall MAP language RIT score (e.g., overall score or subscore), and , , , , and are dummy variables, with male, white, and 6th grade students used as the reference categories. is a residual term. The relationship between NoRedInk performance and the growth in MAP score is represented by .

Results

We constructed several analytic models to understand the relationship between NoRedInk and MAP performance, using different outcomes and predictors (see Table 3). In the main model (Model 7), we analyzed the correlation between each student's NoRedInk performance (topics mastered) and whether they met their expected annual growth on the MAP. With percentage of expected annual MAP growth as the outcome measure, this model yielded a statistically significant coefficient for NoRedInk performance at 0.02, suggesting that every 10 topics mastered on NoRedInk corresponds to achieving 20% of expected annual growth on average (i.e., $0.02 * 10 * 100 = 20$). This is the case regardless of baseline (fall) MAP score, grade level, gender, and race, as we controlled for those factors in the model. In other words, students who mastered more topics on NoRedInk showed more growth on the MAP on average (Fig. 1).

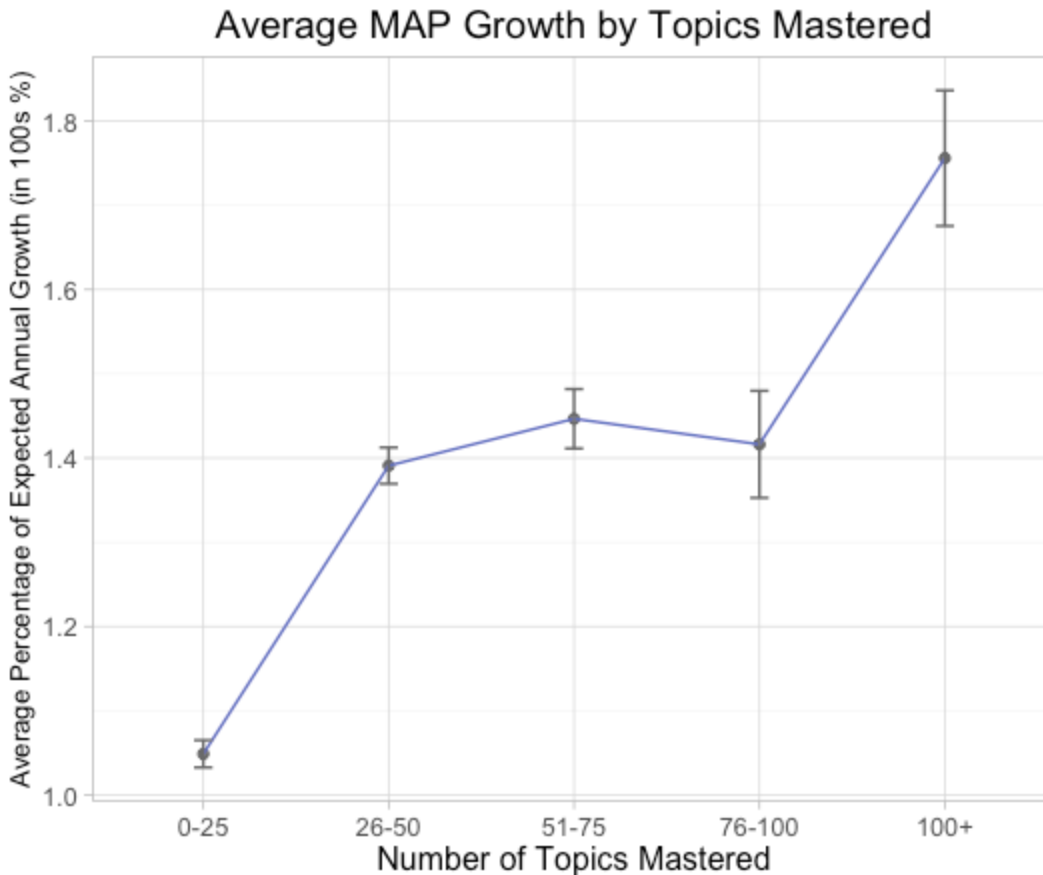


Figure 1. Average percentage of expected annual growth on the MAP for students segmented by number of topics mastered on NoRedInk. Error bars represent the standard error of the mean.

Figure 2 shows the distribution of the percentage of expected annual growth in the student sample by from fall to spring. We performed several additional analyses to understand the relationship between MAP growth and NoRedInk performance. We first looked to tease apart the effects of NoRedInk usage (topics attempted) compared with NoRedInk performance (topics mastered) in Models 1 and 2, respectively. Coefficients in both models were statistically significant and positive. After adding both NoRedInk metrics into a single model (Model 3), we found that the number of topics mastered in NoRedInk was more positively correlated with MAP growth than the number of topics attempted. The adjusted R-square values indicate that Model 3 is a better fit for the data compared to Models 1 and 2, suggesting that NoRedInk performance, controlling for topics attempted, is a better predictor of MAP growth than either metric alone.

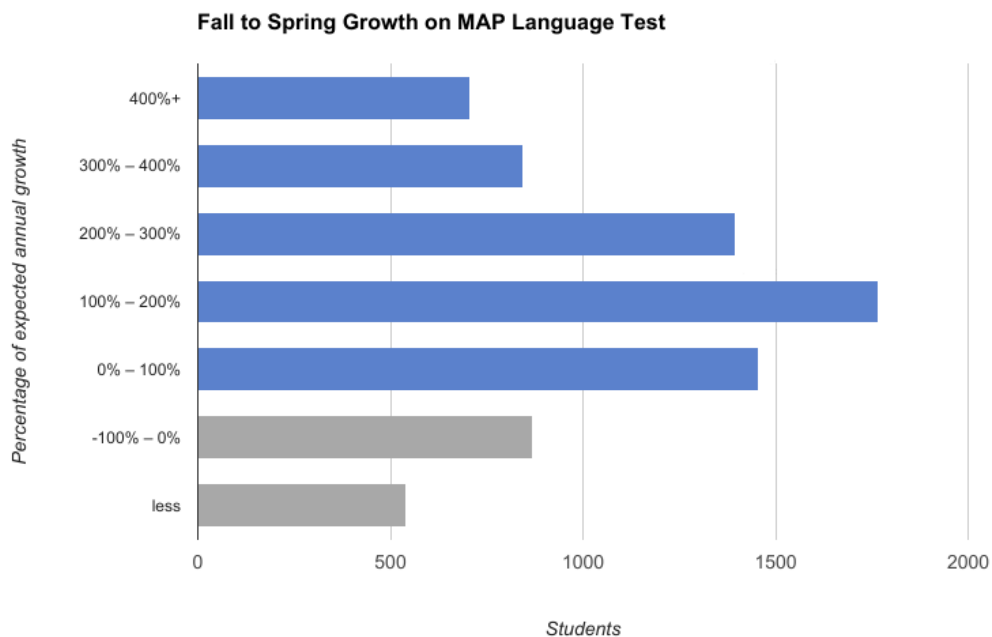


Figure 2. Percentage of expected annual growth of the student sample from fall to spring.

Models 4 through 6 examined correlations between NoRedInk performance and growth on each of the MAP subscales. Performance was positively correlated with all three MAP subscales (coefficients ranged from .08 - .09). Overall, a positive relationship between NoRedInk performance and MAP growth was consistent across all models, even after controlling for demographic and baseline achievement variables¹. These findings suggest that NoRedInk performance is a robust predictor of end-of-year language improvement.

¹ Analyzing the results with a random effect for school yields a similar statistically significant positive correlation between topic mastery and MAP language growth, suggesting that the results are robust.

Table 3. Analysis results across different regression models (abbreviated Mx).

	M1: Overall score (growth)	M2: Overall score (growth)	M3: Overall score (growth)	M4: PODRR (growth)	M5: Grammar (growth)	M6: Mechanics (growth)	M7: Observed / projected growth	M8: Overall spring score
Independent variable	Coefficients (standard errors)							
Intercept	55.2*** (1.381)	60.4*** (1.432)	62.2*** (1.444)	82.9*** (1.698)	82.7*** (1.737)	77.7*** (1.755)	7.6*** (0.448)	62.2*** (1.444)
Overall MAP score fall 2015	-0.2*** (0.007)	-0.3*** (0.007)	-0.3*** (0.007)				-0.03*** (0.002)	0.7*** (0.007)
MAP Goal 1 (PODRR) fall 2015				-0.4*** (0.008)				
MAP Goal 2 (Grammar) fall 2015					-0.4*** (0.008)			
MAP Goal 3 (Mechanics) fall 2015						-0.4*** (0.008)		
Number of topics attempted	0.008*** (0.002)		-0.02*** (0.003)	-0.03*** (0.004)	-0.03*** (0.004)	-0.03*** (0.004)	-0.007*** (0.001)	-0.02*** (0.003)
Number of topics mastered		0.04*** (0.003)	0.06*** (0.004)	0.08*** (0.006)	0.08*** (0.006)	0.09*** (0.006)	0.02*** (0.001)	0.06*** (0.004)
Gender (Male as reference group)								
Female	0.4** (0.157)	0.3* (0.155)	0.4* (0.155)	0.5* (0.211)	0.6** (0.197)	0.9*** (0.198)	0.2*** (0.048)	0.4* (0.155)
Grade (grade 6 as reference group)								
Grade 7	1.0*** (0.186)	0.8*** (0.185)	0.7*** (0.184)	1.2*** (0.252)	0.6** (0.235)	1.0*** (0.235)	0.3*** (0.057)	0.7*** (0.184)
Grade 8	1.8*** (0.196)	1.7*** (0.194)	1.6*** (0.194)	2.0*** (0.264)	2.0*** (0.246)	2.2*** (0.246)	0.7*** (0.060)	1.6*** (0.194)

Ethnicity (White as reference group)								
Asian	1.9** (0.650)	1.8** (0.643)	1.6* (0.641)	2.0* (0.875)	1.4 (0.816)	3.1*** (0.817)	0.5* (0.199)	1.6* (0.641)
African American	-1.3*** (0.212)	-1.2*** (0.211)	-1.2*** (0.210)	-2.1*** (0.286)	-1.7*** (0.267)	-1.5*** (0.265)	-0.4*** (0.065)	-1.2*** (0.210)
Hispanic	-1.1*** (0.290)	-1.0*** (0.287)	-1.1*** (0.286)	-1.8*** (0.390)	-1.3*** (0.365)	-1.3*** (0.364)	-0.3*** (0.089)	-1.1*** (0.286)
Others	-0.9** (0.338)	-0.8* (0.335)	-0.8* (0.333)	-1.3** (0.456)	-1.4*** (0.425)	-0.3 (0.425)	-0.2* (0.103)	-0.8* (0.333)
Model statistics								
Adjusted R-square	0.159	0.176	0.183	0.227	0.222	0.197	0.054	0.717
Sample size	7,330	7,330	7,330	7,330	7,330	7,330	7,330	7,330

* $p < .05$, ** $p < .01$, *** $p < .001$

Discussion

The findings from this study suggest that NoRedInk performance (as measured by number of topics mastered) is a positive, significant predictor of language achievement (as measured by the NWEA MAP Language Usage assessment). These findings conceptually replicate the findings from Perez in a larger and more diverse sample (7,300 students from a large district, as opposed to a few hundred students within a single school), and with a performance predictor more relevant to learning (i.e., mastery of NoRedInk topics). The study provides correlational evidence that NoRedInk performance is related to student academic outcomes as measured by a widely used standardized language assessment. It also helps to validate the topic mastery algorithm developed by NoRedInk, as students who achieved mastery according to the algorithm tended to perform better on the MAP.

These findings suggest important hypotheses for schools and districts seeking to understand best practices when using NoRedInk. It is possible that structuring professional development and instructional goals around topic mastery benchmarks instead of usage benchmarks lead to stronger language growth. Using the site for mastery-driven practice as NoRedInk recommends lets teachers rely on the adaptive engine to progress students through a series of topics until they've achieved mastery in a given area. Future studies could investigate this possibility, as well as the relationship between more fine-grained metrics of authentic learning on the site and their effects on language growth.

Study Limitations

This study is correlational and did not involve a comparison group. For this reason, it is impossible to disentangle whether the observed effects are due to NoRedInk, or whether they are due to other external factors. Therefore, results from this study should not be used to draw causal inferences between NoRedInk usage and student achievement.

Conclusions

This study is exploratory in nature and represents a step towards understanding the relationship between use of NoRedInk and student achievement. Results are consistent with the hypothesis that performance on NoRedInk is linked to student language achievement. Because of the study's design, conclusions of causation cannot be drawn. In order to do so, a fruitful next step would be to conduct a two-arm trial with a comparison group of students who do not use NoRedInk.

References

Northwest Evaluation Association (2015). 2015 NWEA measures of academic progress normative data.

Retrieved from <https://www.nwea.org/content/uploads/2015/08/2015-MAP-Normative-Data-NOV15.pdf>

Perez, C. F. (2015, unpublished). NoRedInk and the NWEA MAP Language Usage Test: A case study at Shadow Ridge Middle school in Colorado. [https://s3-us-west-](https://s3-us-west-2.amazonaws.com/static.noredink.com/shadow_ridge_case_study.pdf)

[2.amazonaws.com/static.noredink.com/shadow_ridge_case_study.pdf](https://s3-us-west-2.amazonaws.com/static.noredink.com/shadow_ridge_case_study.pdf)