A Study of Device Effects on the Smarter Balanced Assessments in Connecticut

February 2017

Connecticut State Department of Education

Contents

Introduction

Connecticut has administered the Smarter Balanced Assessment for two years. Over the past four years, the Connecticut State Department of Education (CSDE) has provided funding for school districts to bolster their technological infrastructure for online testing. In most school districts, four devices were used by students to access the Smarter Balanced Assessment: Chromebooks (Chrome), Macintosh desktops/laptops (Mac), Windows-based desktops/laptops, and iPads. In continuing its long standing commitment to deliver assessments of the highest quality, the CSDE decided that it is important to investigate whether or not the device students used for the Smarter Balanced tests influences their scores. Therefore, the purpose of this study was to investigate the effects of device usage on student performance.

The data used for this study was from Connecticut's 2016 Smarter Balanced operational administration of English Language Arts (ELA) and Mathematics assessments to students in grades 3 through 8. In total, 234,888 students in ELA and 234,286 in Mathematics are included in this study.

Descriptive Statistics

Across all grades, about 47% of students used Chromebooks, 44% used Windows, 7% used Macs, and 2% used iPads. Table 1 to Table 3 show the frequencies and proportion of device usage. Device usage was similar across grade, gender, and free/reduced price meal eligibility status.

Table 1: Device Usage across Grade: Count and Percent (Mathematics)

	Chrome		Mac		Windows		iPad		
Grade	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	**Total**
3	15,139	39.0%	3,276	8.4%	19,282	49.7%	1127	2.9%	38,824
4	17,014	44.4%	2,882	7.5%	17,363	45.3%	1090	2.8%	38,349
5	18,598	47.8%	3,004	7.7%	16,339	42.0%	959	2.5%	38,900
6	19,766	50.8%	2,471	6.3%	16,430	42.2%	252	0.6%	38,919
7	19,916	49.9%	2,520	6.3%	16,708	41.9%	763	1.9%	39,907
8	19,554	50.0%	2,577	6.6%	16,761	42.9%	229	0.6%	39,121

The vast majority of students used Chromebooks or Windows-based devices. In the lower elementary grades, more students used Windows while in the upper elementary/middle grades, more students used

Chromebooks. About 7 to 11% of students used Macintosh-based devices or iPads, and this declined in the upper grades.

Table 2: Device Usage by Gender: Count and Percent

		Chrome		Mac		Windows		iPad		
	Gender	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	Total
ELA	F	53,932	47.0%	8,321	7.3%	50,344	43.9%	2,139	1.9%	114,736
	M	56,181	46.8%	8,380	7.0%	53,366	44.4%	2,225	1.9%	120,152
Math	F	53,777	47.0%	8,342	7.3%	50,063	43.8%	2,183	1.9%	114,365
	M	56,210	47.0%	8,388	7.0%	52,820	44.2%	2,237	1.9%	119,655

The distribution of device usage was very similar across male and female students, and similar to the overall distribution.

Table 3: Device Usage by Free/Reduced Price Meal Eligibility Status: Count and Percent

		Chrome		Mac		Windows		iPad		
	Eligible for F/R Meals	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	Total
ELA	No	67,241	47.1%	11,360	8.0%	60,506	42.4%	3,538	2.5%	142,645
	Yes	42,872	46.5%	5,341	5.8%	43,204	46.8%	826	0.9%	92,243
Math	No	67,146	47.2%	11,374	8.0%	60,204	42.3%	3,565	2.5%	142,289
	Yes	42,841	46.7%	5,356	5.8%	42,679	46.5%	855	0.9%	91,731

The device usage for students with different meal eligibility statuses follows a similar pattern to that of all students, except for iPads which were used more by students ineligible for free/reduced price meals.

Table 4 shows the average ELA and Mathematics scale scores by grade for students using the same device while table 5 expands those average scores based on free/reduced price meal eligibility. These descriptive data are the first step in examining whether or not the device a student used influences his/her score. Students who tested on an iPad had higher scores on average (except grade 8) than students who used other devices. When disaggregated by eligibility for free/reduced price meal, among students who used the same kind of devices, those eligible had lower scores than those ineligible.

Table 4: Average Scale Score by Device Users and Grade

		3	4	5	6	7	8
ELA	**Overall**	**2,438**	**2,480**	**2,517**	**2,536**	**2,559**	**2,574**
	Chrome	2,438	2,483	2,520	2,538	2,562	2,577
	Mac	2,448	2,490	2,527	2,551	2,565	2,584
	Windows	2,435	2,474	2,511	2,531	2,553	2,569
	iPad	2,472	2,510	2,559	2,573	2,592	2,564
Math	**Overall**	**2,438**	**2,478**	**2,501**	**2,521**	**2,538**	**2,551**
	Chrome	2,437	2,478	2,501	2,523	2,539	2,554
	Mac	2,449	2,490	2,508	2,533	2,547	2,558
	Windows	2,436	2,475	2,498	2,516	2,534	2,548
	iPad	2,471	2,502	2,543	2,560	2,584	2,542

Table 5: Average Score by Lunch Status and Device usage

	Device	Free/Reduced Price Meal Eligibile	3	4	5	6	7	8
ELA	**Overall**		**2,438**	**2,480**	**2,517**	**2,536**	**2,559**	**2,574**
	Chrome	No	2,472	2,518	2,553	2,571	2,593	2,607
		Yes	2,392	2,432	2,467	2,484	2,509	2,525
	Mac	No	2,468	2,513	2,547	2,567	2,582	2,602
		Yes	2,410	2,447	2,489	2,510	2,522	2,542
	Windows	No	2,468	2,512	2,547	2,565	2,587	2,602
		Yes	2,392	2,426	2,462	2,481	2,502	2,517
	iPad	No	2,486	2,529	2,570	2,588	2,604	2,589
		Yes	2,416	2,442	2,507	2,459	2,523	2,494
Math	**Overall**		**2,438**	**2,478**	**2,501**	**2,521**	**2,538**	**2,551**
	Chrome	No	2,468	2,510	2,535	2,559	2,576	2,592
		Yes	2,395	2,433	2,449	2,462	2,478	2,488
	Mac	No	2,468	2,510	2,529	2,549	2,566	2,579
		Yes	2,414	2,451	2,470	2,494	2,499	2,505
	Windows	No	2,466	2,508	2,533	2,556	2,573	2,590
		Yes	2,395	2,432	2,449	2,459	2,475	2,483
	iPad	No	2,486	2,519	2,557	2,576	2,598	2,564
		Yes	2,419	2,441	2,481	2,440	2,497	2,481

Further examination of the data revealed that about 80% of all iPad users came from the following five historically high performing districts (Table 6).

Table 6: Number of Students Tested on iPad by District and Grade

District	Grade					
	3	**4**	**5**	**6**	**7**	**8**
Glastonbury School District					509	
Greenwich School District	629	588	613			
Somers School District	98	120	103	107	108	88
Woodbridge School District	92	90	96	119		
Regional School District 04					39	38

Significance Test of Score Differences

To definitively determine whether or not the device used by a student for the Smarter Balanced test influences his/her score, it is insufficient to look only at the descriptive statistics. Therefore, the CSDE conducted a hierarchical model to parse out the impact of free/reduced price meal eligibility status as well as school district because it is quite possible that any observed differences in student scores by devices is confounded by a student's socioeconomic status or the school district that they attend.

The first model run was an unconditional model which showed that 30% of the variance in the Mathematics Score is explained by between-district variability. This indicates that a hierarchical model with students nested within school district is appropriate.

Models were run by grade to take a closer look at the effects of device usage after controlling for school district and free/reduced price meal eligibility status. Since only a few students used iPad in each grade, the focus of this study was on the comparison among Mac, Windows and Chrome users. In addition, we excluded approved special education programs and schools, which normally have very small enrollment numbers, from this analysis.

All models used Chromebook as the reference group since it had the highest percentage of usage. Thus, the estimates can be interpreted as the score differences between Chromebook users and users of another device. The reference group for free/reduced price meal eligibility status were the students who were ineligible.

The final model was written as:

S$S=\beta\_0j+\beta\_1j*SES+ \beta\_2j*G1+\beta\_3j*G2+ \varepsilon$

   $\beta\_0j=\gamma\_00+ u\_0j$

   $\beta\_1j=\gamma\_01+ u\_1j$

   $\beta\_2j=\gamma\_02+ u\_2j$

   $\beta\_3j=\gamma\_03+ u\_3j$

Where SS = scale score, SES is measured by free/reduced price meal eligibility status, and device type is dummy coded as G1 and G2, where G1 = 1 if device = Mac, 0 otherwise; G2 = 1 if device = Windows, 0 otherwise. Chrome users were used as the reference group. So γ_02 is the grand mean difference between Mac and Chrome users, and γ_03 is the grand mean difference between Windows and Chrome users.

Effect size of the group differences were also computed:

$$ES1=\frac{\gamma\_02}{\sqrt{\varepsilon+\mu\_2j}}$$

$$ES2=\frac{\gamma\_03}{\sqrt{\varepsilon+\mu\_3j}}$$

where ES1 is the effect size of the difference between Mac and Chrome users, and ES2 is the effect size of the difference between Windows and Chrome Users. ε and μ_3j are the variance in scale scores within and between school districts, respectively.

Results

Table 8 and Table 9 demonstrate the estimates of difference in student scores in both ELA and Mathematics between different device users. The t-test value, significance indication in terms of p-value, and effect size are also reported. For instance, in grade 3 math, the students who used Mac had

on average 5.46 score points higher than students who used Chrome, but the difference is not significant.

Table 8: Estimates of Difference of ELA Score

Grade	Effect	Estimate	t Value	Significant?	Effect Size
3	Mac - Chrome	2.00	0.37	No	0.025
	Windows - Chrome	-4.56	-1.34	No	-0.055
4	Mac - Chrome	4.33	0.7	No	0.047
	Windows - Chrome	-4.62	-1.59	No	-0.052
5	Mac - Chrome	-4.87	-0.67	No	-0.047
	Windows - Chrome	-10.39	-3.05	Yes	-0.106
6	Mac - Chrome	-0.68	-0.09	No	-0.007
	Windows - Chrome	-4.87	-1.14	No	-0.05
7	Mac - Chrome	-12.44	-1.54	No	-0.135
	Windows - Chrome	-6.33	-1.62	No	-0.069
8	Mac - Chrome	1.57	0.23	No	0.017
	Windows - Chrome	-9.60	-2.35	No	-0.01

Table 9: Estimates of Difference of Math Score

Grade	Effect	Estimate	t Value	Significant?	Effect Size
3	Mac - Chrome	5.46	1.10	No	0.073
	Windows - Chrome	-2.88	-0.95	No	-0.038
4	Mac - Chrome	10.5	1.98	No	0.127
	Windows - Chrome	-1.55	-0.58	No	-0.019
5	Mac - Chrome	-3.37	-0.49	No	-0.036
	Windows - Chrome	-2.66	-0.89	No	-0.03
6	Mac - Chrome	-11.00	-1.29	No	-0.107
	Windows - Chrome	-7.19	-1.83	No	-0.072
7	Mac - Chrome	6.63	0.83	No	0.069
	Windows - Chrome	-0.39	0.08	No	-0.003
8	Mac - Chrome	3.99	0.53	No	0.039
	Windows - Chrome	-9.63	-2.02	No	-0.09

In both ELA and Mathematics, the effect sizes indicate that the device used has no effect on student performance. In ELA, the one area where p-value indicates a significant difference in mean ELA scores is between windows users and chrome users for grade 5. However, both groups had a large sample of students, which made the p-value more sensitive to the significance test. Furthermore, the effect size in Table 8 between windows users and chrome users for grade 5 is very small. Therefore, we can safely conclude that even in the one instance where significance is detected, device usage has no impact on students' ELA scores.

Conclusion

As expected, across all grades, students who are eligible for free/reduced price meals performed significantly lower than students who are not eligible. When controlled for lunch status and attending school district, the device used by the student has no significant impact on student performance. A future study will consider employing a DIF analysis framework using the same data to see if specific test items function differently across devices.