Assessing the Knowledge, Skills, and Abilities of ELs, Selected SwDs, and Controls on Challenging High School Science Content: Results from Randomized Trials of ONPAR and Technology-Enhanced Traditional End-of-Course Biology and Chemistry Tests

Rebecca Kopriva, University of Wisconsin
Phoebe Winter, Pacific Metrics
Robert Triscari, Florida Gulf Coast University
Therese G. Carr, University of Wisconsin
Cathy Cameron, Center for Applied Linguistics
David Gabel, Center for Applied Linguistics

2950 students took part in the winter and spring 2012 administrations of the federally funded project's experimental study in Biology or Chemistry in a southern state. Four forms in each subject (2 long forms and 2 short forms) were randomly administered in each subject, and each of the forms included a subtest of ONPAR tasks and a subtest of technology-enhanced traditional tasks. In total, 2578 students received a non-0 score on at least one of the subtests, and most of these students completed both subtests with a non-0 score. The records with a subtest score of 0 were excluded from the analyses to address non-motivation. Test level analyses were completed on the long forms only, for those students with complete data sets, including test scores, an independent teacher rating of student Biology or Science ability in the classroom, and other demographic information. Students participating included Focal students (students with learning disabilities (LD), other students with disabilities (Other SwD), and English learners (EL)), and the control group of native English speaking, non-IEP students. In all, 1239 control students, 88 LD, 124 Other SwD, and 27 ELs participated in Biology. 872 control students, 8 LD, 12 Other SwD, and 12 ELs participated in the Chemistry test. Table 1 illustrates the various types of students included in Other SwD, and Table 2 shows the breakout of ELs by English proficiency level.

Table 1: Breakdown of Students with Other Disabilities

Subject	Biology	Chemistry
Group	**N**	**N**
Physical or Attention Disability	28	3
Multiple Disabilities	3	0
Visual Impairment	1	0
Hearing Impairment	0	0
Emotional Disability	9	0
Speech or Language Impairment	43	4
Autism	9	0
Traumatic Brain Injury	2	0
Qualified Individual with disabilities under 504	27	5
Intellectual	2	0

Disabilities		
Total	***124***	***12***

Table 2: English Learner Students by ACCESS for ELs Literacy Score

Test Type	Biology	Chemistry
ACCESS Score	**N**	**N**
3	2	0
4	6	8
5	8	0
6	14	4
Total	***30***	***12***

Tests were comprised of ONPAR and technology-enhanced traditional tasks developed for this study. Current teachers and state personnel from the southern state, with project staff, selected particular standards and objectives from which to build the ONPAR tasks, and subsequently developed the interactive ONPAR tasks aligned with these standards.

ONPAR tasks use computer-interactive novel presentation and response formats designed to communicate challenging content to and from students. Accommodation tools in the ONPAR tasks included non-construct relevant text roll-overs of words or phrases with pop-up language, static visual, animation, or symbol glosses or highlighted elements on the computer screen. Oral reading of the text questions or statements in the tasks is available in L1 or English when the speaker button is pushed. Animations and other visuals used in the presentation of the items and/or in the response sections of the tasks are used to convey contextual meaning instead of or in addition to text. Tasks are designed to provide redundancy of semiotic representations within a task to allow for the measurement of more challenging content by circumventing a certain amount of linguistic complexity or by deliberately building in non-construct relevant simultaneous supports.

Once the ONPAR tasks were completed, an independent contractor, Pacific Metrics, completed a version of each ONPAR task using selected technology-enhanced task techniques that anticipate the types of enhanced items that may be used in new online assessment tasks being developed for the assessment consortia. The technology-enhanced traditional tasks measured the same standards and were developed to measure, as closely as possible, the same topics as their ONPAR counterparts, at the same cognitive complexity level, and within the same context. The technology-enhanced traditional tasks in this project used traditional multiple choice and short and more lengthy constructed response item types, along with non-typical traditional item types such as multiple answer multiple choice and multiple selection. Enhanced components included using the keyboard for numerical responses and completion of constructed response, and enhanced visuals also used in the ONPAR tasks. The line was drawn at drag and drop, and other types of response spaces utilized only in the ONPAR tasks. To mirror the structure of the ONPAR tasks, many of the technology-enhanced traditional tasks included multiple items per task.

Test Level Results

Research Questions

Question 1: For each subject area, what is the relationship between overall student ability as defined by total score of teacher ability ratings, and total test scores of ONPAR or technology-enhanced traditional tasks?

Question 2*:* Controlling for overall student ability (as defined by teacher ratings) in each subject area, how do the test scores of each possible focal group compare on the two types of tests, and how do their scores compare to the control group?

Test level Reliability

Reliability of the total scores was computed by type for each subject. The Cronbach's Alpha reliability statistic findings for the ONPAR and technology-enhanced traditional total scores on the Biology long forms (comprised of 14 tasks each) were .67 for ONPAR, and .80 for the technology-enhanced traditional. The findings for the Chemistry subtests were .65 for ONPAR and .72 for technology-enhanced traditional. Given the novel nature of the ONPAR tasks it was expected that these reliability results would be lower than typically the case for traditional item types, but their showing seems to be reasonable and viable for our purposes.

Analyses

Two analyses were conducted to address both of these questions. First, regressions with the test scores as the independent variable and the independent teacher rating of ability as the dependent variable were undertaken for the overall sample and by focal and control groups for both subjects. Regressions were also performed in Biology for the LD and Other SwD samples. The focus here was how well the total teacher ability ratings correlate with and predict the variation in either the ONPAR or technology-enhanced traditional total test scores for the pertinent sample, and then to compare these relationships across types.

Second, the Analysis of Covariance (ANCOVA) with relevant contrasts was used to directly compare the test scores by type (ONPAR and technology-enhanced traditional) for the overall sample and for the samples disaggregated by group. Using the total score of teacher ability ratings to control for the differential influence of student achievement, the analyses within a sample asked if the students performed similarly on both types of tests. The assumption in the analyses within groups is that, for each of the focal groups, if one test type outperforms the other by more than the difference in the means for the control group, then the test type yielding the higher means may be more validly capturing the focal group students' knowledge, skills and abilities.

Regressions Regressions, for each subject per type, were performed. For all the test level regressions the dependent variable is total score ability in Chemistry or Biology from the teacher questionnaires, and the independent variable is either the ONPAR or TE traditional scores in each subject. Table 3 displays the regression results from Biology and Chemistry for the overall samples. Tables 4 and 5 display the Biology regression findings for omnibus-focal and control groups, and for the LD and Other SwD groups, respectively. Table 6 shows the omnibus-focal and control results for Chemistry.

Table 3: Biology and Chemistry Regression Results, Overall Sample

Overall Sample	Unstandardized	Standardized	

	Coefficients		Coefficients		
Biology	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	-18.762	11.949		-1.570	.117
ONPAR Scores	.140	.022	.296	6.308	.000
Intercept	5.907	7.277		.812	.417
TE Trad Scores	.098	.014	.318	7.007	.000
Chemistry	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	17.173	11.730		1.464	.144
ONPAR Scores	.082	.023	.220	3.639	.000
Intercept	-6.201	13.359		-.464	.643
TE Trad Scores	.121	.025	.285	4.855	.000

Table 4: Biology Regression Results, Omnibus-Focal and Control Groups

Group	Unstandardized Coefficients		Standardized Coefficients		
Omnibus-Focal	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	18.159	13.970		1.300	.196
ONPAR Scores	.063	.028	.206	2.264	.025
Intercept	40.374	11.490		3.514	.001
TE Trad Scores	.023	.025	.092	.933	.353
Control	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	47.744	5.417		8.814	.000
ONPAR Scores	.021	.011	.078	1.961	.050
Intercept	47.749	4.198		11.375	.000
TE Trad Scores	.023	.009	.106	2.651	.008

Table 5: Biology Regression Results, LD and Other SwD Groups

Group	Unstandardized Coefficients		Standardized Coefficients		
LD	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	43.823	25.615		1.711	.094
ONPAR Scores	.007	.051	.020	.132	.895
Intercept	34.563	18.932		1.826	.076
TE Trad Scores	.032	.041	.130	.799	.430
Other SwD	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	3.105	18.192		.171	.865
ONPAR Scores	.097	.037	.331	2.646	.011
Intercept	48.715	15.801		3.083	.003
TE Trad Scores	.006	.033	.024	.180	.858

Table 6: Chemistry Regression Results, Omnibus-Focal and Control Groups

Group	Unstandardized Coefficients		Standardized Coefficients		
Omnibus-Focal	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	-107.762	151.185		-.713	.499
ONPAR Scores	.317	.295	.376	1.073	.319

Intercept	-134.284	43.509		-3.086	.027
TE Trad Scores	.356	.079	.897	4.536	.006
Control	**B**	**Std. Error**	**Beta**	**t**	**Sig.**
Intercept	19.139	11.656		1.642	.102
ONPAR Scores	.079	.022	.216	3.509	.001
Intercept	-4.629	13.630		-.340	.734
TE Trad Scores	.118	.026	.276	4.635	.000

The regression results for the overall sample by type for Biology and Chemistry have low moderate relationships for each subject (.296 ONPAR and .318 TE traditional for Biology, and .220 ONPAR and .285 technology-enhanced traditional for Bio) and seem to be very similar within subject. In Biology, the focal group correlation for ONPAR is higher than for technology-ehnanced traditional (.206 vs. .092), and the correlation for the TE traditional is not significantly different from 0. The control group correlations are very low and very similar (.078 for ONPAR and .106 for technology-enhanced traditional). By group in Biology, neither LD nor EL yielded correlations significantly different from 0, though EL relationships were high moderate but with very small n's (13 for each type). For the Other SwDs the ONPAR correlation was moderate (.331) compared to the technology-enhanced traditional, which was very low (.024). The Chemistry regression results are not surprising. The omnibus-focal findings are non-significant, which may or may not have been the case with a more robust n. The control group mirrored the overall findings.

ANCOVAs with Contrasts The ANCOVAs with contrasts, for each subject, overall and by some groups, were performed. The dependent variable in each case are the study scores for each task that the students took, and the independent variable is type (technology-enhanced traditional or ONPAR). The covariate is the total score teacher ability rating for each student, and all means and standard deviations illustrated in the tables are adjusted for the effect of the covariate. Table 7 illustrates the results of the Biology and Chemistry ANCOVAs for the overall sample. In both, difference by type (ONPAR or technology-enhanced traditional) is significant with the ONPAR scores significantly higher in Biology and the technology-enhanced traditional scores higher in Chemistry. Table 8 illustrates the adjusted means and standard errors of the ANCOVA main effects. As the tables show, differences in type (ONPAR or TE trad) are significant for both Biology and Chemistry. However, for Biology the ONPAR scores are higher while the technology-enhanced traditional scores are higher for Chemistry.

Table 7: Biology and Chemistry ANCOVA Results for Overall Sample

Biology					
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig**
Corrected Model	190363.941[a]	2	95181.971	29.392	.000
Intercept	18653275.471	1	18653275.471	5760.129	.000
Biology Ability Covariate	66155.495	1	66155.495	20.429	.000
TYPE	128850.192	1	128850.192	39.789	.000
Error	4776556.730	1475	3238.344		
Total	366348016.000	1478			
Corrected Total	4966920.671	1477			
Chemistry					
Source	**Type III Sum of**	**df**	**Mean Square**	**F**	**Sig**

	Squares				
Corrected Model	101577.287[a]	2	50788.643	8.288	.000
Intercept	7282753.891	1	7282753.891	1188.497	.000
Chemistry Ability Covariate	35023.690	1	35023.690	5.716	.017
TYPE	67261.762	1	67261.762	10.977	.001
Error	5521057.173	901	6127.699		
Total	207227104.000	904			
Corrected Total	5622634.460	903			

a. R Squared = .038 (Adjusted R Squared = .037) (Biology)
a. R Squared = .018 (Adjusted R Squared = .016) (Chemistry)

Table 8: Adjusted Means and Standard Errors of ANCOVA Main Effects

Biology			
Type	**N**	**Mean**	**Standard Error**
ONPAR	747	503.54	51.41
Traditional	731	485.21	62.72
Total	1478	494.48	57.99
Chemistry			
Type	**N**	**Mean**	**Standard Error**
ONPAR	464	463.89	79.54
Traditional	440	481.05	77.35
Total	904	472.24	78.91

Biology ANCOVA Findings Table 9 displays the results of the Biology ANCOVAs with focal/control as an independent variable in addition to type. Table 10 illustrates the adjusted means and standard errors for the main effects and interaction for this ANCOVA, and Table 11 shows the contrasts between type within focal and control, and between focal and control within type.

All three of the ANCOVA F's are significant. As expected the main effect of omnibus-focal/control indicates the control group's means are higher than the focal group. The type main effect significantly favors the ONPAR tests versus the technology-enhanced traditional. The contrasts relating to the significant interaction show that there is no difference between omnibus-focal and control on the ONPAR test, but that there is a significant difference between the two groups on the technology-enhanced traditional. Further, within group contrasts indicate that there is a significant difference between types for the focal group, as well as for the control group (always favoring ONPAR), but that the omnibus-focal group difference in the means is 37 points while the control group difference is 16.

Table 9: Biology ANCOVA Results for Focal and Control Groups

Biology					
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig**
Corrected Model	232019.867[a]	4	58004.967	18.045	.000
Intercept	18543794.577	1	18543794.577	5768.866	.000
Biology Ability Covariate	49883.444	1	49883.444	15.518	.000
FOCAL/CONTROL	22491.962	1	22491.962	6.997	.008
TYPE	129379.370	1	129379.370	40.249	.000
FOCAL/CONTROL* TYPE	21093.777	1	21093.777	6.562	.011

Error	4734900.804	1473	3214.461		
Total	366348016.000	1478			
Corrected Total	4966920.671	1477			

a. R Squared = .047 (Adjusted R Squared = .044)

Table 10: Adjusted Means and Standard Errors of Biology ANCOVA Main Effects for Focal/Control Groups

Biology: Main Effect by Focal/Control			
Group	**Mean**		**Standard Error**
Focal	484.824[a]		3.88
Control	495.982[a]		1.61
Biology: Main Effect by Type			
Type	**Mean**		**Standard Error**
ONPAR	503.496[a]		2.87
Traditional	477.311[a]		3.00
Biology: Interaction of Focal/Control*Type			
Group	**Type**	**Mean**	**Standard Error**
Focal	ONPAR	503.202[a]	5.28
	Traditional	466.447[a]	5.57
Control	ONPAR	503.789[a]	2.26
	Traditional	488.175[a]	2.27

a. Covariates appearing in the model are evaluated at the following values: Biology Ability Covariate= 57.30.

Table 11: Biology Contrasts between Type in Focal/Control, and between Focal/Control within Type

Contrast	**DF**	**Contrast SS**	**Mean Square**	**F Value**	**P Value**
Focal ONPAR/ Focal Trad	1	75031.52405	75031.52405	23.34	<.0001
Focal ONPAR/ Control ONPAR	1	33.39320	33.39320	0.01	0.9188
Focal Trad/ Control Trad	1	41654.26669	41654.26669	12.96	0.0003

Table 12 shows the results of the Biology ANCOVAs by group (LD, Other SwD, EL, Control) and type. Table 13 illustrates the adjusted means and standard errors for the main effects and the group by type interaction. Table 14 displays the results of the various within and between group contrasts. The graph in Figure 1 illustrates the difference in the ONPAR and TE traditional means by group in Biology.

The tables indicated that there is no difference for the main effect of group, but there is by type and for the group by type interaction. The type difference shows that, overall, students scored significantly higher on the ONPAR test. In reviewing the contrast results from the interaction, within group findings indicate that differences between types for all the groups are significant with ONPAR favored in all cases. As noted above, there is a difference of 16 points between the type means for the control group. This difference is contrasted to the difference in means for the LD group (38 points), the Other SwD group (27 points) and EL group (81 points). Contrasts of each focal group to the control within type consistently found that, while there was a significant difference between each particular focal group (LD, Other SwD, EL) and the control group (with the control group's adjusted mean higher than individual focal groups), there was no difference between the control vs. all focal groups. This suggests the focal students are being able to show

their knowledge, skills in abilities in a way that is commensurate with their non-EL, non-IEP peers.

Table 12: Results of the Biology ANCOVAs by Group (LD, Other SwD, EL, Control) and Type

Biology					
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig**
Corrected Model	247022.201[a]	8	30877.775	9.610	.000
Intercept	15497391.637	1	15497391.637	4823.339	.000
Biology Ability Covariate	49742.152	1	49742.152	15.482	.000
GROUP	23831.499	3	7943.833	2.472	.060
TYPE	99591.673	1	99591.673	30.996	.000
GROUP*TYPE	34845.716	3	11615.239	3.615	.013
Error	4719898.471	1469	3213.001		
Total	366348016.000	1478			
Corrected Total	4966920.671	1477			
Total	ONPAR	747	503.54	51.41	
	Traditional	731	485.21	62.72	

a. R Squared = .050 (Adjusted R Squared = .045)

Table 13: Adjusted Means and Standard Errors of Biology ANCOVA Results by Group and Type

Biology: Main Effect by Group			
Type	**Mean**		**Standard Error**
LD	484.096[a]		6.24
Other SwD	484.193[a]		5.33
EL	488.211[a]		11.93
Control	495.982[a]		1.61
Biology: Main Effect by Type			
Type	**Mean**		**Standard Error**
ONPAR	508.327[a]		4.88
Traditional	467.914[a]		5.41
Biology: Interaction of Group*Type			
Group	**Type**	**Mean**	**Standard Error**
LD	ONPAR	503.163[a]	8.43
	Traditional	465.028[a]	9.11
Other SwD	ONPAR	497.600[a]	7.41
	Traditional	470.787[a]	7.60
EL	ONPAR	528.756[a]	15.73
	Traditional	447.665[a]	17.93
Control	ONPAR	503.790[a]	2.26
	Traditional	488.175[a]	2.27

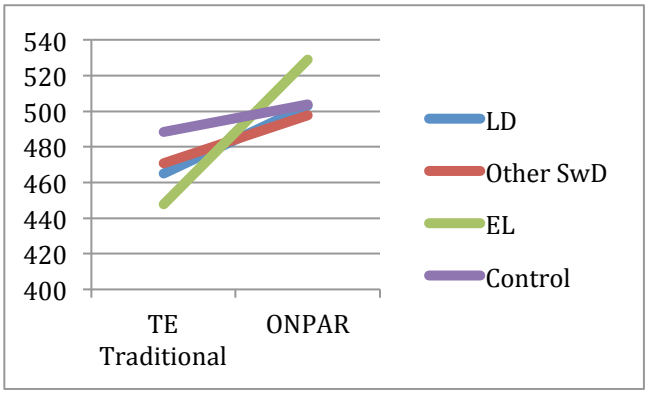
a. Covariates appearing in the model are evaluated at the following values: Biology Ability Covariate= 57.30.

Table 14: Results of Biology Within and Between Group Contrasts

Contrast	DF	Contrast SS	Mean Square	F Value	P Value
LD ONPAR/ LD Traditional	1	30682.03478	30682.03478	9.55	0.0020
Other SwD ONPAR/ Other SwD Traditional	1	20653.97574	20653.97574	6.43	0.0113
EL ONPAR/	1	37165.58159	37165.58159	11.57	0.0007

EL Traditional					
Control/ONPAR Control/Traditional	1	76473.17638	76473.17638	23.80	<.0001

Figure 1: Biology Adjusted Means by Type for Focal Groups and Control Group



Chemistry ANCOVA Findings Table 15 shows the results from the Chemistry ANCOVAs with focal/control and type as the independent variables, Table 16 illustrates the adjusted means and standard errors for main effects and the interaction, and Table 17 displays the contrast findings within and between focal/control and type. As expected, with such low n's the Chemistry results for the within focal group contrast are not significant and most likely not indicative of those from a larger sample. Further, since the control group reflected 97% of all those who took the chemistry tests, it is not surprising that the contrasts reflect those of the overall sample. The graph in Figure 2 illustrates the difference in adjusted means from the ONPAR and technology-enhanced traditional by focal and control in Chemistry.

Table 15: Chemistry ANCOVA Results for Focal and Control Groups

	Chemistry				
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig**
Intercept	Hypothesis	5452743.409	1	5452743.409	1014.738
	Error	228149.176	42.458	5373.551[a]	
Chemistry Ability Covariate	Hypothesis	35158.604	1	35158.604	5.728
	Error	5517859.515	899	6137.775[b]	
TYPE	Hypothesis	8303.935	1	.	.
	Error	.	.[c]	.	
FOCAL/CONTROL	Hypothesis	3159.092	1	.	.
	Error	.	.[c]	.	
TYPE * FOCAL/CONTROL	Hypothesis	.244	1	.244	.000
	Error	5517859.515	899	6137.775[b]	

a. .257 MS(CONTRO_OTHER) + .743 MS(Error)
b. MS(Error)
c. Cannot compute the error degrees of freedom using Satterthwaite's method.

Table 16 Adjusted Means and Standard Errors of Chemistry ANCOVA Main Effects for Focal/Control Groups

Chemistry: Main Effect by Focal/Control		
Group	**Mean**	**Standard Error**

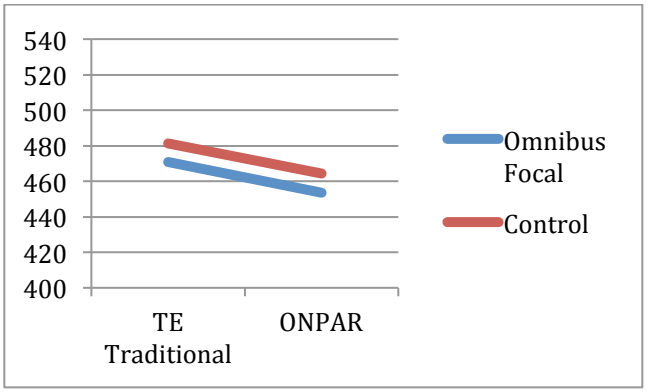
Focal	458.835[a]		9.97
Control	476.128[a]		11.03
Chemistry: Main Effect by Type			
Type	**Mean**		**Standard Error**
ONPAR	458.835[a]		9.97
Traditional	476.128[a]		11.03
Chemistry: Interaction of Focal/Control*Type			
Group	**Type**	**Mean**	**Standard Error**
Focal	ONPAR	453.456[a]	19.59
	Traditional	470.843[a]	21.73
Control	ONPAR	464.214[a]	3.70
	Traditional	481.414[a]	3.79

a. Covariates appearing in the model are evaluated at the following values: Chemistry Ability Covariate= 59.78.

Table 17 : Contrasts between Type in Focal/Control, and between Focal/Control within Type

Contrast	DF	Contrast SS	Mean Square	F Value	P Value
Focal ONPAR/ Focal Trad	1	2167.160848	2167.160848	.35	0.5525
Focal ONPAR/ Control ONPAR	1	1787.838781	1787.838781	.29	0.5895
Focal Trad/ Control Trad	1	1409.009040	1409.009040	.23	0.6320

Figure 2: Chemistry Adjusted Means by Type for Omnibus-Focal and Control Groups



V. Discussion

At the test level, the primary differences are seen in the interactions between group and test type in Biology in both the regression and ANCOVA results. Insufficient focal group information in chemistry makes regression results mute, and ANCOVA results hard to interpret. The main effect of test type in Biology found significantly higher scores for ONPAR for the entire sample while the technology-enhanced scores were significantly higher in Chemistry. While the focal group n's were so small in Chemistry, the test type findings clearly suggest the control group's preference for the traditional format. In Biology, it would seem that ONPAR tasks,

overall, yielded somewhat higher scores, regardless of group affiliation, though the differences in the focal groups were more pronounced.

The test level regressions in Biology and Chemistry asked how well the ONPAR and technology-enhanced traditional tests relate to and predict student's knowledge, skills and abilities as operationalized as the total teacher rating scores. The findings for the total sample show that the relationships between the teacher rating of abilities and the ONPAR or the technology-enhanced traditional tests are similar for Biology (R's are .296 for ONPAR and .318 for TE traditional), although the relationship with the technology-enhanced traditional in Chemistry (R = .285) is a little higher than it is for ONPAR (.220). These low moderate relationships are reasonable and typical for how well salient variables like test scores characteristically relate to other indicators of ability. In general it seems that neither technology-enhanced approach in and of itself is better at measuring the Biology and Chemistry abilities of students. The relationship is different for the Biology omnibus-focal group, however. Here, the ONPAR R is .206 versus .092 (and not significantly different from 0) for the technology-enhanced traditional, suggesting that the former approach could be more useful for measuring the Biology skills and knowledge of these students. Since the n's per test type in the Chemistry omnibus focal group were so small, any interpretation of group differences, even at the focal and control groups level, would seem to be premature.

In controlling for student variation in Biology or Chemistry ability on the topics measured by the two tests, the test level ANCOVAs evaluated the difference in scores between the two types (ONPAR or TE traditional), for the total sample and for the various groups. The findings from the task contrasts focused on why individual tasks might yield higher or lower scores. Because the TE traditional tasks were developed from the ONPAR tasks by an independent contractor to measure, as closely as possible, the same content and cognitive complexity but using a different approach, the interpretation suggests that, overall, higher total scores could be a better representation of what these students know as opposed to the lower scores. This could be for reasons of access, different approach of the item types used in each test, familiarity, or engagement, to name a few.

Findings indicate that differences in type for the total sample were significant for both Biology and Chemistry. However, for Biology the ONPAR scores are higher while the TE traditional scores are higher for Chemistry. The ANCOVA main effects and interaction tests in Biology for the omnibus-focal and control groups were all significant. As expected the main effect of focal/control indicates the control group's means are higher than the focal group, and that the ONPAR scores were significantly higher than the TE traditional. The interaction within-group contrasts indicate that there is a significant difference between types for the focal group, as well as for the control group (always favoring ONPAR), but that the focal group difference in the adjusted means is 37 points while the control group difference is 16. The between group contrasts show that there is no difference between focal and control on the ONPAR test, but that there is a significant difference between the two groups on the technology-enhanced traditional.

Unpacking the omnibus-focal group, within-contrasts explaining the significant interaction between type and group found that all groups (LD, Other SwD, EL) scored significantly higher on the ONPAR test as opposed to the TE traditional, with differences of 38 points (LD), 27 points (Other SwD) and 81 points (EL). Between group contrasts consistently found no difference on

the ONPAR tests between control and each of the LD, Other, and EL groups, but significant differences between control and each of the focal groups on the TE traditional test. These findings are interesting for two reasons. First, the differential boost the ONPAR test appears to provide to the omnibus and disaggregated focal groups is far larger than the boost for control (16 points), suggesting that this approach is allowing the focal groups to better display their knowledge, skills, and abilities then they could do on the TE traditional.

Second, controlling for Biology ability, no differences between each of the focal groups and the control on the ONPAR test implies that this approach may "level the playing field" for all students including students at various levels of ability as operationalized by the teacher rating measure. This same finding has been previously seen in a randomized trial of ONPAR and traditional elementary and middle school science tests for ELs of all levels of English proficiency (Kopriva, Gabel, and Cameron, 2009), adding to the defensibility of the interpretation.

Of interest, the high school English learners who performed significantly higher on the Biology ONPAR test is somewhat different than the group who performed the most differently in the 2009 elementary and middle school study. In the earlier study, the low English proficient students were the most effected by the difference in test type, while the high ELs performed similarly to the control group in all aspects. However, here, it was primarily the high English proficient that seemed to have benefited from the novel presentation and response types. This may be due to the additional complexity of the material at the high school level with possibly its concurrently more complex language demands that even the highest ELs are still mastering. This finding needs further research.

In Chemistry, the sample size of the focal group was so small, the ANCOVA test could not be computed properly. The control group's findings were essentially the same as those for the overall group, and here, the control group favored the TE traditional test. The traditional approach seemed better suited than the ONPAR for measuring the types of knowledge and skills the Chemistry test covered.