

Fresno CA iACCESS Stage 2 Drama Integration Report: Control-Treatment Writing  
Sample Comparisons

**Executive Summary**

- Data sorting processes in the Data Collection revealed that name match between files and hand written name is improved from the previous dataset, but still an issue in some control schools.
- Writing prompts, while similar, appear to be not identical across all classrooms, therefore testing conditions were probably dissimilar and could account for differences in data analysis.
- *Differences in test conditions notwithstanding, treatment school students outperformed their control school counterparts on all qualitative rubric measures: Conventions, Organization, Development/Elaboration, and Empathy/Social Emotional Language.*
- Treatment school students wrote significantly *more* text (as determined by “word count” data) than control school students and the word count data, as a whole, is significantly correlated with average score of the qualitative measures (.64, p value .0001), but we don’t know if differences in test conditions were due to the apparent variation in testing prompts.
- Overall, without baseline data consistently collected from the students before the iACCESS project began, it is impossible to determine whether these statistically significant differences in control-treatment performance are due to the treatment or other factors.

**iACCESS Part 2 2015-16 Academic Year Report**

**Data Collection Methods Summary**

- Writing samples were collected from 241 Students in 3<sup>rd</sup> and 4<sup>th</sup> Grade, Treatment and Control Schools.
- The nature of the 3<sup>rd</sup> grade “Almond Farming” opinion piece (not dramatic writing or narrative) had no call for empathy or social emotional language and ratings were uniformly “1.0” across the entire grade, and thus are not included in the analyses below.
- The nature of the 4<sup>th</sup> grade “Dolores Huerta” opinion piece had occasions for including empathetic social emotional language and ratings were varied across the entire grade, and thus are included in the analyses below.
- Though topics were uniform across each grade level, format varied, with student worksheets having slightly different prompts and some writing on blank paper with no prompt present.

### Data Labeling Reliability

Name matching between the file name and the handwritten name was significantly more reliable on this data set than last year's. However, for the sake of future use of these data, the degree of match between the file and handwritten name was coded on the following scale:

*Name Matching Coding System:*

Complete Match (CM): No conflicting information
Probable Match (PM): Some matching information, some conflicting information
No match (NM): Names cannot be related between the list and the handwritten name

For example,

Student Name (File)	Student Name (Document)	Student Name Match?
George Garcia	George Garcia	CM
Bautista Tapia	Jasmin Batista	PM
Grace Crawford	Lauren Campbell	NM

Calculating the distribution of name matching codes for each cohort revealed that as was the case with the previous data set, control schools are still significantly less reliable than treatment schools in terms of data labeling, as shown in the table below.

Grade	Control			Treatment		
	NM	PM	CM	NM	PM	CM
3 <sup>rd</sup>	7%	33%	60%	6%	4%	90%
4 <sup>th</sup>	12%	39%	49%	5%	0%	95%

Other, more specific data collection issues included:

Data Collection Issue (Noted in the Database)	Description	N Students (Out of 241 Total)	Other Notes
<i>Dated 2015</i>	Student text indicated that it was from a previous scoring period and about a different subject; not scored	3	All from Addams school
<i>No Text</i>	Student file was blank; not scored	6	Half of these students came from one control school: Vang Pao
<i>Out of Focus Image</i>	Student writing unreadable due to image quality; not scored	4	
<i>Unfinished Text</i>	Student text appears to be unfinished; not scored	7	All in control schools

<i>Illegible Handwriting</i>	Student writing unreadable due to handwriting; not scored	2	
<i>Incomplete Text</i>	Student writing sample appears to be incomplete resulting in a lowered score	7 (at least)	Concentrated in Olmos school (Grade 4, Treatment), which had the by far the highest word counts. Many of the samples appeared to be the first of multiple pages

### Control-Treatment Comparisons

The following measures were analyzed through rubrics established January 2016 (Appendix A) for potential control-treatment differences in:

- 1) Word Count (quantitative)
- 2) Conventions (qualitative)
- 3) Organization (qualitative)
- 4) Development/Elaboration (qualitative)
- 5) Empathy/Social Emotional Language (qualitative)
- 6) Combined Average Score (all qualitative measures)

#### 1) *Word Count C-T Comparisons*

On average, treatment school students wrote significantly more text than control school students, with the differences most pronounced in the 3<sup>rd</sup> Grade.

Grade	Control Mean Word Count (Range)	Treatment Mean Word Count (Range)	Difference	ANOVA
3 <sup>rd</sup> Grade	44 (Range: 15-143)	75 (Range: 28-137)	+ 31	F Ratio = 40.81 p value = .0001
4 <sup>th</sup> Grade	60 (Range: 10-143)	82 (Range: 31-133)	+ 22	F Ratio = 11.50 p value = .001

#### *Individual School Differences*

Word count for individual schools also varied widely, raising questions about whether the student writing task is being proctored under identical conditions. Some of the essay prompts were visible on the images, confirming that the writing tasks, though similar in nature, were not identical.

The chart below, ordered from highest to lowest average word count by each school, reaffirms that both grade 3 and grade 4 treatment schools are noticeably more inclined to write more, implying that a higher degree of detail and elaboration from the individual students as a whole may be linked to the average student word count score.

Word Count Rank Ordered by School Type

School	Grade	School Type C/T	Mean Word Count
Olmos*	4	T	106
Columbia	3	T	84
Storey	4	T	78
Vinland Tatum	4	T	78
Balderas	3	T	77
Vinland Plumb	4	T	75
Forkner	4	C	72
Aynesworth	3	T	71
Hidalgo Flores	4	C	64
Fremont	3	T	63
Hidalgo Gong-Chun	4	C	61
Addams	3	C	57
Mayfair	3	C	56
Mayfair	4	C	54
Hidalgo Gonzalez	3	C	45
Hidalgo Clayton	3	C	43
Vang Pao	3	C	40

\*This school had a very high proportion of incomplete text, suggesting that even though it has the highest average length, the true value would be even higher.

## 2) Adherence to ELA Writing Conventions

Our main finding is unambiguously clear: Treatment school students scored significantly higher on Conventions compared to control school students. That is, 3<sup>rd</sup> and 4<sup>th</sup> grade treatment school students both excelled in the grasp and execution of ELA elements such as spelling, grammar, capitalization, etc. Additionally, both cohorts showed improvement comparing 3<sup>rd</sup> to 4<sup>th</sup> grade, but the gap was wider in favor of the treatment schools.

Grade	Control	Treatment	Difference	ANOVA
3 <sup>rd</sup> Grade	1.81	2.15	+ .34	F Ratio = 4.15 p value = .04
4 <sup>th</sup> Grade	2.22	2.77	+ .55	F Ratio = 9.23 p value = .003

### 3) Quality of Writing *Organization*

Our main finding here is also unambiguously clear: treatment school students scored significantly higher on Organization compared to control school students in both grade levels. That is, 3<sup>rd</sup> and 4<sup>th</sup> grade treatment school students outperformed their control school counterparts in effective use of organizational devices, such as sentence structure, transitional strategies, coherence, etc.

One minor caveat that needs further explanation is that the 4<sup>th</sup> grade treatment and control students are not outperforming the 3<sup>rd</sup> grade students. This counterintuitive finding is likely due to the fact that these are entirely different student cohorts that we cannot be compared longitudinally.

Grade	Control	Treatment	Difference	ANOVA
3 <sup>rd</sup> Grade	1.74	2.34	+ .60	F Ratio = 44.60 p value = .0001
4 <sup>th</sup> Grade	1.75	2.16	+ .41	F Ratio = 23.47 p value = .0001

### 4) Degree of *Development/Elaboration* in Writing Samples

Similar to the previous qualitative measures of writing skills, it is unambiguously clear is that Treatment school students scored significantly higher on Development/Elaboration compared to control school students. That is, treatment school students included significantly more specific details and description in their writing.

Grade	Control	Treatment	Difference	ANOVA
3 <sup>rd</sup> Grade	1.67	2.11	+ .44	F Ratio = 49.82 p value = .0001
4 <sup>th</sup> Grade	1.68	2.25	+ .57	F Ratio = 46.87 p value = .0001

### 5) Degree of *Empathy/Social Emotional Language* in 4<sup>th</sup> Grade Writing Samples

As discussed previously, the nature of the 3<sup>rd</sup> Grade writing task prompted virtually no use of empathetic or social-emotional language, and ratings were uniformly “1.0” for all 3<sup>rd</sup> Grade Students.

While the average 4<sup>th</sup> Grade treatment school “Empathy” score was slightly higher than the control school average, this difference was not statistically significant, nor did either cohort’s scores rise to the level of consistent presence in their writing samples.

Grade	Control	Treatment	Difference	ANOVA
4 <sup>th</sup> Grade	1.54	1.62	+ .12	NS

**6) Summary Statistics: Combined Average C-T Comparisons and 3<sup>rd</sup>/4<sup>th</sup> Grade Comparisons**

Treatment school students’ combined average scores (combining conventions, organization, development/elaboration, and empathy/social emotional language) were superior to those of the control school students at a high level of statistical significance. Furthermore, these results show that the treatment schools scored categorically higher than the control school student cohorts in both grades. That is, whereas the control school scores were within the range of the lowest categorical response of skill development (“not evident/indiscernible”), the treatment school cohort scores were consistently within the second level of performance skill (evident though “uneven/inconsistent”).

Grade	Control	Treatment	Difference	ANOVA
3 <sup>rd</sup> Grade (not including “Empathy” score)	1.74	2.19	+ .45	F Ratio = 28.24 p value = .0001
4 <sup>th</sup> Grade	1.80	2.20	+ .40	F Ratio = 22.98 p value = .0001

As discussed earlier, the quantitative measure of “Word Count” was significantly higher in treatment schools. Correlation analysis revealed that word count is a significant positive predictor of the combined average score across all students ( $r=.64$ ,  $p$  value = .0001)

*Program effect exceeds grade-level effect*

Performing an analysis of variance (ANOVA) on the combined average scores revealed that, as a whole, the “Control-Treatment” variable was a statistically significant predictor of differences in combined average scores, while “Grade Level” was not, suggesting that the impact of the iACCESS program is statistically more powerful than the developmental effect of being enrolled in school for an additional year, as summarized in the table below.

Comparison	ANOVA
Grade Level to Combined Average Score	NS
Control-Treatment to Combined Average Score	F Ratio = 51.30 p value = .0001

**Final Conclusions**

On the surface, all of the indications from this stage of the iACCESS project are positive: treatment school students are outperforming control school students on all qualitative measures (Conventions, Organization, Development/Elaboration, Empathy/Social

Fresno iACCESS Stage 2 Drama Integration Report  
*Lawrence Scripp/Josh Gilbert CMAIE*

Emotional Language), and a quantification of writing proficiency (word count). Taken together these results suggest strongly that the effect of the iACCESS program exceeds the developmental expectations for grade. While these statistical findings are unambiguously clear, without any preprogram baseline data, the validity of these control-treatment school cohort differences cannot be determined.

*August 10, 2016*