

Title: The effects of an online intervention program on math achievement in twelve southern Colorado school districts

Author(s): Dan Jorgensen, Becky Smith, and Wendy Dunaway
Colorado Department of Education

Date of Initial Publication: 01/12/2012

Abstract/Summary

The purpose of this report was to examine the impact of the Compass Learning Odyssey program on math achievement in low-performing middle school math students in twelve southern Colorado school districts. Our findings indicate that intervention participants fail to statistically differ on overall CSAP growth compared to the matched control. However, the Compass Learning students performing at the unsatisfactory level on the 2009 CSAP exhibited growth percentiles that were greater than those exhibited by the control. In addition, a larger percentage of intervention students that were at the unsatisfactory level in 2009 moved to a higher proficiency level in the 2010 school year compared to the control group (i.e. 25% to 19% respectively). Finally, the math measures of academic progress assessment revealed Rasch Unit (RIT) growth rates for program participants that exceeded expectations based on national norms.

Subject/Keywords: Math Achievement, online learning, propensity scores, quasi-experimental study

Document Type: Paper

Document Archive Number: 0012cdejo2012

**The Effects of an Online Intervention Program on Math Achievement in
Twelve Southern Colorado School Districts**

Dan Jorgensen, Evaluation & Research Coordinator
Becky Smith, Targeted Support Manager
Wendy Dunaway, Turnaround & Pilot Program Manager

Support and Intervention Unit
Colorado Department of Education

January 12, 2012



--THIS PAGE INTENTIONALLY LEFT BLANK --

Executive Summary

The Colorado Department of Education supports the implementation of research-based programs to improve achievement outcomes for all Colorado students. An online intervention program (i.e. CompassLearning Odyssey) was implemented in twelve school districts within southern Colorado during the 2009-2010 school year. This pilot report examines the impact of program participation on math achievement in low-performing middle school math students. Colorado Student Assessment Program (CSAP) math scale scores, proficiency levels, median growth percentiles, and math measures of academic progress scores were examined to identify growth that may be attributed to program participation. The study utilized a propensity score matching technique to establish group equivalence to account for possible bias that occurs due to non-random assignment when making comparisons in achievement. A logistic regression analysis was used to develop a control group that was matched with the intervention participants on a number of salient demographic and achievement measures. Next, cases were matched using a nearest-neighbor without replacement matching technique. Our findings indicate that intervention participants fail to statistically differ on overall CSAP growth compared to the matched control. However, the CompassLearning students performing at the unsatisfactory level on the 2009 CSAP exhibited growth percentiles that were greater than those exhibited by the control. In addition, a larger percentage of intervention students that were at the unsatisfactory level in 2009 moved to a higher proficiency level in the 2010 school year compared to the control group (i.e. 25% to 19% respectively). Finally, the math measures of academic progress assessment revealed Rasch Unit (RIT) growth rates for program participants that exceeded expectations based on national norms.

The longitudinal change in scores revealed small to moderate growth effects that may be partially attributed to program participation. This study provides preliminary evidence that supports the effectiveness of the CompassLearning Odyssey program in addressing the needs of low-performing middle school math students. Future studies will examine the impact of fidelity of implementation on observed effects.

Corresponding author: Dan D. Jorgensen, Evaluation & Research Coordinator, Support & Intervention Unit, Colorado Department of Education, 201 East Colfax Avenue, Denver, Colorado 80203-1799, U.S.A.; e-mail: Jorgensen_d@cde.state.co.us

Introduction

The Colorado Department of Education supports the implementation of research-based educational interventions to improve achievement outcomes for all Colorado students¹. A number of programs have been initiated and piloted within some of the lowest-performing districts to help address pedagogical deficits that may contribute to poor achievement outcomes. The implementation of these programs is followed by appropriate analytics to ascertain programmatic effectiveness. The CompassLearning Odyssey Program is one such online intervention program that has been implemented within the state to foster improved math achievement. The program was implemented at the direction of the Colorado Department of Education to serve as a pilot math intervention in a number of Southern Colorado school districts that weren't receiving supplemental instructional programmatic interventions at the time. This software-based curriculum "assesses each student's understanding of key objectives. Based on this assessment, a student is automatically prescribed a personalized learning path."² In addition, the developers claim, "The CompassLearning Odyssey suite of learning solutions includes lessons and activities that are based on current and confirmed research about the way today's students think and learn. Odyssey curricula for primary and secondary students facilitate Response to Intervention, differentiated and personalized instruction, and formative assessment that helps drive data-based decision making."²

The purpose of this technical report is to examine the one-year impact of the CompassLearning intervention program on math performance within twelve southern Colorado school districts. It was expected that students participating in the program would show academic growth rates that exceeded those of students that were non-participating but were demographically and academically similar prior to study initiation. In addition, it was believed that this supplemental intervention would move lower performing students to higher proficiency levels at a rate faster than that observed for non-participants. This result, if observed, would be expected to be due to the programs ability to provide appropriately targeted student remediation activities.

Methods

Participants

The names of the participating districts and the number of CompassLearning participants from each district are identified in table one. Our sample included all students that were enrolled during the 2009-2010 school year within the participating school districts. All 2009-2010 CompassLearning students were identified as participants in lists provided to the Colorado Department of Education by the program administrators. For all of the districts, except for one, the program was administered only to 6th to 8th grade students with the sixth grade being the most common for program participants.

Table 1. Count of CompassLearning Students by District (2009-2010)

District	Level	District Membership (Oct 2009)	Compass Student Count
0640- Centennial R-1	MS	220	22
2810- Center 26 JT	MS	605	47
2010- Creede Consolidated 1	MS	101	9
2730- Del Norte C-7	HS	597	53
2800- Moffat 2	MS	223	10
2740- Monte Vista C-8	MS	1181	105
2790-Mountain Valley Re 1	MS	128	10
0560-Sanford 6j	MS	340	11
0110- Sangre de Cristo Re-22j	MS	323	3
2750-Sargent Re-33j	MS	484	35
0740- Sierra Grande R-30	MS	252	25
0580- South Conejos Re-10	MS	280	27

Note. Total count of compass students = 357; total district membership=4734. MS: 6th to 8th grade only; HS: 9th to 12th grades only.

The students who participated in the program were selected due to their partially proficient or unsatisfactory performance on the Colorado Student Assessment Program (CSAP) math assessment and/or they were performing at the below proficient level on the measures of academic progress math assessment. In addition, students tended to be free or reduced lunch eligible, Hispanic, and male (see table two).

Propensity Score Analysis Procedure

The propensity score is utilized when random assignment is not feasible and reflects a conditional probability of a unit being assigned to a particular study condition given observed covariates³. A logistic regression analysis was conducted with program participation as the identified binary outcome variable to derive propensity scores. The model predictors included seven key variables that were dummy coded for analysis including gender (male & non-male), race (Hispanic & non-Hispanic; white & non-white), free-and-reduced lunch (FRL & non-FRL), year in school, special education (SPED vs. non-SPED), language proficiency (NEP/LEP vs. non-ELL) and the 2008-2009 CSAP math proficiency level (i.e. excluding no scores). The obtained propensity score analysis was applied to match program participants with non-participants based on a nearest-neighbor without replacement matching technique (aka the Greedy procedure).³ Following group matching, the “fit” of the match was determined based on frequency distributions of the model predictor variables.

Analysis

CSAP growth percentile outcomes were examined between the participating and non-participating matched groups using the non-parametric Wilcoxon sign test. Also, analysis included an examination of growth percentiles by 2009 CSAP math proficiency levels to identify program effects that may be restricted to specific performance levels. The final CSAP analysis consists of a two year comparison of the percentage of students moving between proficiency levels from 2009 to 2010 for both the intervention and control groups.

Measures of academic progress math Rasch Unit (RIT) scores of intervention students were available and used to derive an annual change estimate to be compared to national norms. In effect, an average RIT score for program participants was calculated for the fall and spring. Also, the mean change in scores between test administrations was calculated. Next, an independent sample t-test

and Cohen’s d effect size measure were calculated to identify statistically meaningful increases. Finally, NWEA norms were compared to the average scores of program participants.

Results

Propensity Score Analysis Matching

A total of 357 students participated in the CompassLearning pilot intervention program during the 2009-2010 school year. The October count enrollment for all students within the participating districts was 4,734 in 2009. A sizable percentage of these students were enrolled in grades other than those tested by the Colorado Student Assessment Program (i.e. CSAP). In sum, a total of 2,266 program non-participants were available for matching and utilization within this research study. In effect, the larger non-treatment population (i.e. 86.7% of total cases) facilitated the creation of the matched control group due to the expanded comparison sample that was utilized for precise matching with individual program participants. An initial examination of matching variables revealed substantial differences between comparison groups that would greatly limit causal arguments given the large achievement differences existing between groups prior to matching (see table two). The logistic regression and propensity score matching procedure served to reduce between-group variance to improve confidence that observed differences in outcomes result from the impact of the intervention of interest.

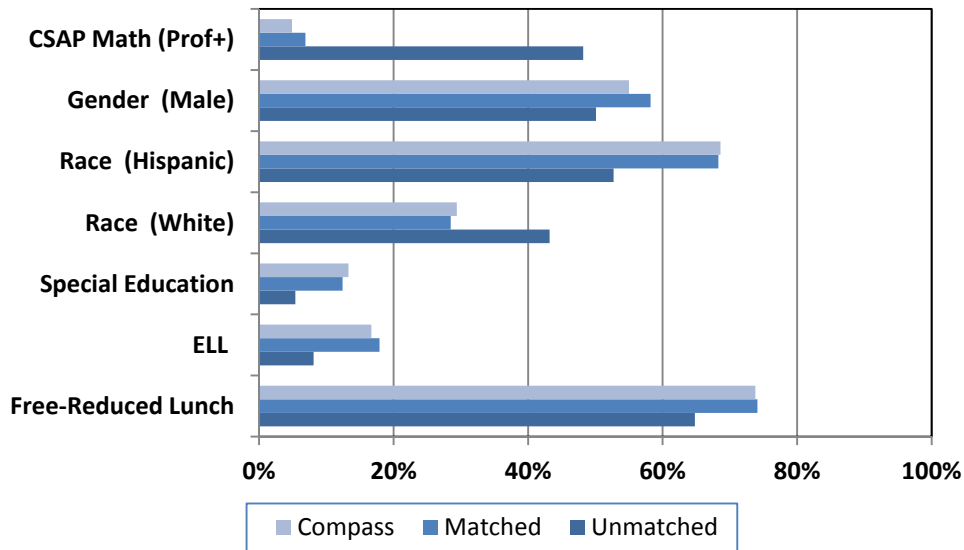
Table 2. Demographic Composition of Unmatched and Matched Groups

Matching Variable	Compass Participants	Non-Participants	
		Unmatched	Matched
FRL (Free/Reduced)	73.8%	64.8%	74.1%
ELL (NEP/LEP)	16.7%	8.1%	17.9%
SPED (Yes/No)	13.3%	5.4%	12.4%
Race (White)	29.4%	43.2%	28.5%
Race (Hispanic)	68.6%	52.7%	68.3%
Gender (Male)	55.0%	50.1%	58.2%
CSAP Math (Prof+)	4.9%	48.2%	6.9%
Grade in School (Mean)	6.83	6.52	6.55

Note. Values reflect 2009 CSAP data (i.e. prior to CompassLearning participation). CompassLearning participants n= 357; Non-participants (unmatched) n= 2,266.

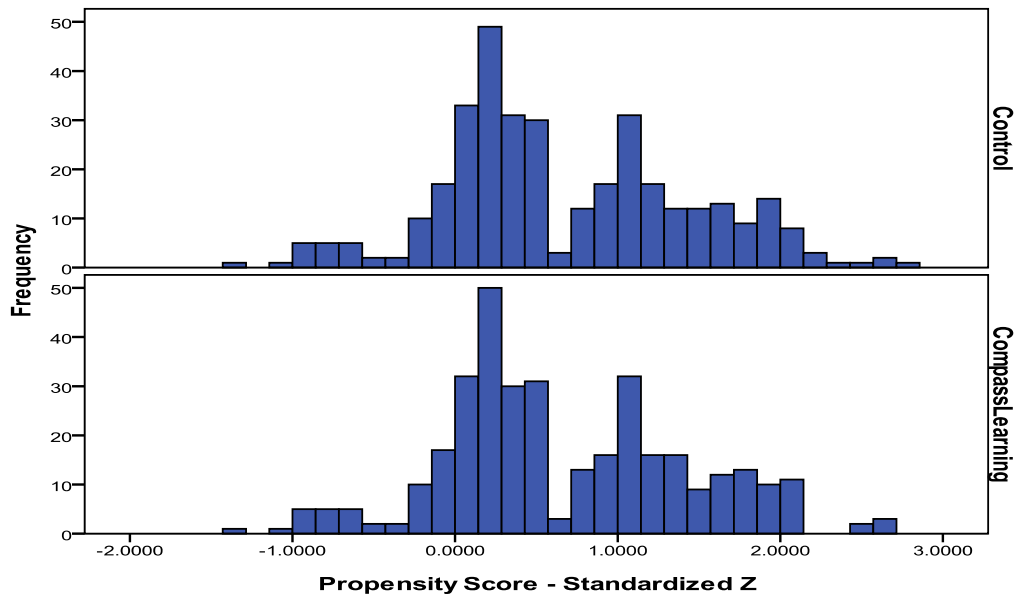
Following the application of the matching process the group differences were minimized thus increasing confidence in the results of treatment and control group comparisons (see figure one). It should be noted that substantial differences between the CompassLearning participants and unmatched non-participants existed prior to matching. The range of difference for each matching variable (prior to matching) went from -9.0% to 43.2%. Following matching, the range of difference was reduced to -0.9% to 3.2% for each variable thus indicating a high degree of matching success.

Figure 1. Demographic Comparison of Treatment and Unmatched/Matched Control Groups



An examination of matched propensity scores (see figure two) reveals that the nearest-neighbor matching procedure was successful in identifying appropriate counterparts for almost all program participants. The largest discrepancies were related to the most extreme propensity scores (i.e. greater than 2.0 standard deviations above the mean). However, these scores account for less than ten percent of all of the obtained values and appear to have had a minimal impact on matching success.

Figure 2. Distribution of Propensity Scores by Treatment and Matched Control Groups



CSAP Proficiency and Growth

Once the intervention and control groups were matched via propensity scores a series of outcome comparisons were conducted based on CSAP measures. Initial analysis involved a Wilcoxon’s signed-ranks test for significant differences in obtained CSAP math growth percentiles between groups. The Wilcoxon test is a nonparametric alternative to the correlated-samples t-test⁴. The key outcome variables, Wilcoxon value and obtained p-values are displayed in table three below. Our results fail to indicate a statistically significant difference in overall CSAP math growth percentiles between program participants and the control group. Additional analysis of growth percentiles by 2009 CSAP proficiency level (see table four) indicate that students at the unsatisfactory and proficient levels are impacted the most favorably by program participation. It should be noted that the number of participants that were proficient in 2009 was very low (i.e. less than twenty) so confidence in the impact on the program on proficient students is much more limited. In contrast, the intervention students who were partially proficient had growth percentiles that were lower than those displayed for the matched control group.

Table 3. Overall CSAP Math Growth Percentiles (2010): CompassLearning and Matched Control

Results	CompassLearning Participants		Matched Non-Participants		Wilcoxon-Rank Sum Test (W)	p-value
	CSAP Growth	n	CSAP Growth	n		
Overall (09-10)	50	333	54	346	112286.5	.715

Table 4. 2010 CSAP Math Growth Percentiles by 2009 Proficiency Levels

Group	Unsatisfactory	Part. Prof.	Proficient	Advanced
Compass	57.5	46.0	58.0	--
Matched Control	48.5	55.5	45.0	--

Note. Advanced category results aren't presented due to small n-size (i.e. less than two).

An additional analysis examines the change in CSAP math proficiency of participating students between years relative to the observed change in math proficiency of the control students (see tables five and six). The tables reveal that the control students had a greater overall increase in the percent of students increasing proficiency levels between-years. However, the program participants had a larger percentage of unsatisfactory students move into a higher proficiency level compared to the matched control group (i.e. 25% to 18.9% respectively).

Table 5. CSAP Proficiency Levels for Program Participants: Between-Years (2009 to 2010)

2009 CSAP Math PL	2009 Count	2010 CSAP Math PL				% Increase
		Unsatisfactory	Part. Prof.	Proficient	Advanced	
Unsatisfactory	116	87	27	2	0	25.0%
Partially Prof	203	48	138	17	0	8.4%
Proficient	16	1	4	9	1	6.3%
Advanced	1	0	0	0	1	na

Note. Total n (2009) =336. The overall percentage of Compass students that demonstrated an increase in proficiency between years was 13.9%.

Table 6. CSAP Proficiency Levels for Matched Controls: Between-Years (2009 to 2010)

2009 CSAP Math PL	2009 Count	2010 CSAP Math PL				% Increase
		Unsatisfactory	Part. Prof.	Proficient	Advanced	
Unsatisfactory	122	99	22	1	0	18.9%
Partially Prof	200	44	125	31	0	15.5%
Proficient	22	1	8	11	2	9.1%
Advanced	2	0	0	2	0	na

Note. Total n (2009) = 346. The overall percentage of non-participating students that demonstrated an increase in proficiency between years was 16.2%.

Measures of Academic Progress: Math Growth

An additional analysis was conducted to examine the impact of the CompassLearning program on math achievement. The analysis involved a comparison of the observed change in math RIT scores on the measures of academic progress assessment across a one-year interval for program participants. In addition, the observed change was compared to national norms for an indication of program effectiveness⁵. The math RIT scores were only available for program participants so comparison to a matched control sample wasn't possible. The result of a paired sample t-test indicates a statistically significant increase in RIT scores between the fall and spring test administration ($t(259)=-11.99, p<.001$); $Mean_{fall}=215.07, Mean_{spring}=221.41$; Cohen's $d=.411$). The mean growth rate was 6.24 with a range of -20 to 36 points. For the change distribution, 19.2% of the included students (i.e. 50/260) experienced a decline in RIT scores while 76.5% (i.e. 199/260) experienced an increase in RIT scores between the fall and spring test administration. The observed overall increase in RIT score of 6.24 for program participants exceeds the expected change based on the average national norms reported out by NWEA for 6th (+5.5), 7th (+4.2), and 8th grades (+3.4) thus suggesting a programmatic impact⁵. The obtained effect-size was small to moderate ($d=.411$) and may indicate that math MAP scores are meaningfully impacted by CompassLearning participation. Effect sizes (Cohen's d) are a measure of the magnitude of the difference between means. Typically, effect sizes of 0.2 are considered small, 0.5 medium, and 0.8 or greater are considered large^{6,7}

Discussion

The purpose of this study was to examine the impact of an online math intervention program (i.e. CompassLearning Odyssey) on student math achievement levels and growth levels in middle school students within twelve southern Colorado school districts. The findings of this study indicate that the CompassLearning program tends have the greatest impact on the performance of the

lowest performing students. The results indicate that both growth percentiles and improvements in proficiency rates tend to be greater for those students that were performing at the unsatisfactory level prior to program participation.

The findings of this study indicate that the implemented program may serve to address the needs of middle school student populations that struggle with math achievement. However, it must be recognized that the quasi-experimental design (i.e. matching based on propensity score analysis) does not unequivocally assure causal attributions due to unaccounted factors that may exist and impact the presented findings. Nonetheless, the results of the current study do provide preliminary evidence of programmatic impact and a future replication of this study should provide additional evidence of the effectiveness of program implementation. Also, a closer examination of fidelity of implementation may provide corroborating evidence of the mechanisms by which the program appears to better support lower performing students. It's likely that the programmatic effectiveness may be mitigated by poor fidelity of implementation. In effect, a lack of fidelity may reduce programmatic impacts. Similarly, differential use of the program between the districts (e.g. time allotted for CompassLearning activities, etc.) could also reduce program effectiveness. This study serves to support the goal of the Colorado Department of Education in supporting research based programs and practices throughout the state of Colorado.

Works Cited

- ¹Colorado Department of Education (2011). *FY 2012-13 Budget Request: Strategic Plan*, Colorado Department of Education, Denver, Colorado. Available online at:
<http://www.cde.state.co.us/cdecomm/download/StrategicPlanOSPSubmission.pdf>
- ²Unidentified author (2011). *In About CompassLearning*. Retrieved December 15, 2011, from
<http://www.compasslearning.com/>
- ³Guo, S. & Fraser, M. (2009). *Propensity score analysis: statistical methods and applications*. Sage Publications, Thousand Oaks, CA.
- ⁴Schweigert, W. (1994). *Research methods and statistics for Psychology*. Brooks/Cole Publishing Company, Belmont, California.
- ⁵Northwest Education Association (2008). *2008 Normative Data Report, version 2*, NWEA Lake Oswego, Oregon.
- ⁶Coe, R. (2000). *What is an 'Effect Size'? A guide for users*. Curriculum, Evaluation and Management Centre: <http://www.cemcentre.org/File/CEM%20Extra/EBE/ESguide.pdf>.
- ⁷Valentine, J.C. & Cooper, H. (2003). *Effect size substantive interpretation guidelines: Issues in the interpretation of effect sizes*. Washington, DC: What Works Clearinghouse.

--THIS PAGE INTENTIONALLY LEFT BLANK --



Colorado Department of Education
Unit of Support and Intervention
201 E. Colfax Ave.
Room 401
Denver, CO 80203
<http://www.cde.state.co.us/>