EDVANTIA®
Partners in education. Focused on results.

PECE program evaluation center AT EDVANTIA℠

Globaloria Replication Study: An Examination of the Relationships between Globaloria Participation and Student Achievement in Year 5 of the West Virginia Pilot Implementation

February 2013

Hsiang-Yeh Ho, Ph.D.
Jessica Gore, M.A.
Kristine Chadwick, Ph.D.

500 Virginia Street, East, Suite 500
Charleston, WV 25301
One Vantage Way, Suite C200
Nashville, TN 37228
800.624.9120 • edvantia.org

Founded in 1966 as the Appalachia Educational Laboratory, Edvantia is a not-for-profit corporation with primary offices in Nashville, Tennessee and Charleston, West Virginia. Edvantia delivers high quality program evaluation services and develops award-winning reports to provide clients with timely information to improve their programs and document their successes. Edvantia staff members work collaboratively with clients to build their planning, data, and program evaluation capacity through just-in-time consulting and training. Edvantia's evaluation services are grounded in the latest research methodologies and evaluation best practices.

For information about Edvantia research, products, or services, contact

EDVANTIA®

500 Virginia Street East, Suite 500, Charleston, WV 25301 • 304.347.0400 • fax 304.347.0487
One Vantage Way, Suite C-200, Nashville, TN 37228 • 615.565.0101 • fax 615.565.0112
info@edvantia.org • www.edvantia.org

This publication was developed under a contract with the World Wide Workshop Foundation. However, its contents do not necessarily represent the policy of the Foundation, and you should not assume endorsement by the Foundation.

Table of Contents

List of Tables

List of Figures

Executive Summary

This study examined the relationships between Globaloria participation and student achievement, measured by the math, science, reading, and social studies subtests of West Virginia's state standardized exam, the WESTEST2. Propensity score matching (PSM) methods were used to create a counterfactual group that was similar to the group of Globaloria students in various aspects of student- and school-level characteristics. Due to small sample sizes and different patterns of achievement scores across racial groups, Edvantia researchers conducted separate analyses for each racial group. First, a series of hierarchical linear modeling (HLM) analyses were conducted for White students for each of the four WESTEST2 subtests. Second, a series of one-way analyses of covariance (ANCOVAs) were conducted for Black students for each of the four WESTEST2 subtests. For White students, subgroup analyses were also conducted to examine whether the associations between Globaloria participation and student WESTEST2 scores differed by gender and income status across all four WESTEST2 subtests. For Black students, subgroup analyses were not possible due to the small sample size.

Overall, findings showed that the associations between Globaloria participation and students' WESTEST2 subtest scores differed by race. First, there were no significant effects of Globaloria participation on Black students' achievement; yet, this finding should be interpreted with caution due to the small sample size. For White students, the key findings are as follows:

- For **math**, Globaloria participation had a positive effect on high school students who were in schools struggling with low math proficiency.
- For **reading**, Globaloria participation had no significant effect.
- For **science**, Globaloria participation had a positive effect on high school students who came from low-income families (i.e., students receiving free or reduced-price meals).
- For **social studies**, Globaloria participation had a positive effect on boys in middle school.

These findings support the effectiveness of Globaloria participation on White students' achievement in math, science, and social studies. Particularly, Globaloria participation seemed to have more of an effect on students who needed more supports to be successful or who were in schools in need of support. These results are similar to, and therefore validate, findings reported in the Year 4 Globaloria study (see Ho, Gore, & Chadwick, 2012).[1] The presence of Globaloria seems to be particularly important for students in schools that are struggling to meet student achievement benchmarks and/or have limited resources. Similarly, the effect of Globaloria on science is stronger for high school students from economically disadvantaged family backgrounds. This finding suggests that a program like Globaloria may mitigate some of the negative effects of family poverty

[1] Key Year 4 findings were as follows: (1) Globaloria participation was positively associated with White students' math achievement; (2) Globaloria participation was positively associated with White students' reading achievement; (3) there was a significant correlation between Globaloria participation and student science outcomes within the schools struggling with math proficiency; and (4) Globaloria participation was not associated with White students' social studies achievement (see Ho et al., 2012 for more detail). Please note that in the Year 4 study, analyses were not conducted separately for middle and high school students.

on student achievement, especially in science. A recent study by Reynolds and Chiu (2013, accepted) found similar results related to Globaloria participation and the effects of family poverty on student achievement.

Nevertheless, because the effects of Globaloria participation seemed to be conditional on certain school-level contextual factors and student-level attributes, more studies are needed to understand how and why Globaloria works to support student achievement. In addition, Globaloria was not consistently predictive of all four subsets of achievement outcomes. Most importantly, the inconsistent findings of the main effect of Globaloria participation over the past few years suggest that student achievement outcomes associated with Globaloria participation may be influenced by factors not yet studied (e.g., student interest in science, technology, engineering, and mathematics [STEM] subjects, student perceptions of STEM education, and student career aspirations in STEM fields). Lastly, as addressed in the Year 4 report, researchers did not find significant effects of Globaloria participation on Black students' achievement. Yet, this finding should be interpreted with caution due to the fact that the sample sizes were small. Given the associations found among socioeconomic status, school achievement, and Globaloria participation among White students, the Globaloria program would benefit from studies of these factors among a more racially and ethnically diverse student population.

Based on these findings, researchers offer several recommendations:

- More studies are needed to understand processes of change associated with Globaloria participation and the influences of contextual variables in order to more fully explain how Globaloria works to support students from various backgrounds and school contexts.
- The World Wide Workshop Foundation (herein referred to as the Workshop) is encouraged to re-examine the logic model suggested by Edvantia in 2008 to identify "mediating" variables explaining the processes that link program participation with student outcomes. These are the key ingredients for future program scale-up.
- To better understand the impact of Globaloria on racial and ethnic minority students, Globaloria staff should reach out to these minority communities and encourage participation. Even in West Virginia, the participation levels by minority students are lower than what would be representative of the state's minority student population. In addition, the Workshop program staff should consider replicating the current study's methodology using a more racially and ethnically diverse sample of students, likely from Globaloria implementation sites other than West Virginia, which is a rather homogenously White state.

For additional recommendations, see the *Globaloria Replication Study: An Examination of the Relationships between Globaloria Participation and Student Achievement in Year 4 of the West Virginia Pilot Implementation* (Ho et al., 2012).

Introduction

The World Wide Workshop Foundation (herein referred to as the Workshop) contracted with Edvantia to conduct a replication of the *Globaloria Replication Study: An Examination of the Relationships between Globaloria Participation and Student Achievement in Year 4 of the West Virginia Pilot Implementation* (Ho, Gore, & Chadwick, 2012). The goal of the study was to further explore the findings of the *Globaloria Pilot Study: The Relationship of Globaloria Participation and Student Achievement* (Chadwick & Gore, 2010) and the *Globaloria Replication Study: Examining the Robustness of Relationships between Globaloria Participation and Student Achievement* (Chadwick & Gore, 2011). The current study improves upon the prior studies in two key ways: (1) the sample size was larger, and (2) a modified, more rigorous quasi-experimental methodology was used that allowed researchers to take into consideration within- and between-school variations in student achievement. This report used the Year 5 data from the Globaloria West Virginia Pilot Implementation (heretofore referred to as Globaloria-WV). The main purpose of this study seeks to provide empirical evidence to support the effectiveness of Globaloria-WV on student achievement, as measured by West Virginia's standardized achievement test, the WESTEST2. The WESTEST2 provided a valid measure of the effects of the Globaloria program because the WESTEST2 was designed to measure 21st century skills, as represented by the West Virginia content standards and objectives (CSOs), while Globaloria is designed to increase them.

The World Wide Workshop Foundation

The Workshop, founded in 2004, is a non-profit organization that invents social media and digital technology applications to help youth and educators participate as leaders in the global knowledge economy. The Workshop partners with "forward-thinking leaders, corporations, school systems, universities, foundations, and research centers worldwide to enrich existing formal and non-formal education with the latest technology and innovative learning opportunities." The mission of the Workshop is to "develop applications for learning with technology that combine game mechanics and social networking to empower youth to be inventors and leaders in the global knowledge economy." The Workshop's programs are designed to "transform education by connecting youth to learning, community engagement, and economic development through game production" (World Wide Workshop Foundation, 2012).

Globaloria: Underlying Theory and Structure

The Workshop created Globaloria, a social network for learning web-based game design and simulation production to address the two digital divides encountered by poor and underserved communities in the United States and worldwide. The first digital divide is defined by issues with access to high-speed Internet. The second divide, coined as "digital literacy," is the ability to create, not just consume, digital media (Knestis, 2008). The Globaloria program is based on the constructivist educational philosophy (Harel, 1991; World Wide Workshop Foundation, 2008). According to Harel (1991), the constructionist approach to learning operates from the view that "building knowledge structures ('in the head') goes especially well when the subject is engaged in building material structures ('in the world')." Through this approach, "children learn how to learn,

and they learn how to think about thinking" (Harel, 1991). This is accomplished through "publicly shared, long-term projects that are complex, computational, immersive, and innovative" in which students learn by doing (World Wide Workshop Foundation, 2008). Research has shown that constructivist programs result in deeper forms of learning, cognitive integration, and improved approaches to learning (Rogers, Pertosino, Huebner, & Hacsi, 2000).

The goal of the Globaloria program is to create technology-based educational opportunities through a series of virtual learning networks (World Wide Workshop Foundation, 2008, 2009). That is, Globaloria allows participants to interact with games, puzzles, and creative tools, while also thinking as game developers (World Wide Workshop Foundation, 2009; 2012). When participants create their own games, they are encouraged to share their work on the shared wiki platform (World Wide Workshop Foundation, 2008). Through construction, interaction, and play, the Globaloria program empowers youth to be productive, successful 21st century citizens, by fostering the following six essential contemporary learning abilities (CLAs) inherent in digital literacy (Reynolds & Harel, 2009; Harel, Oliver, & Sullivan, 2010):

- Invention, progression, and completion of an original project: program an educational game, wiki, or simulation
- Project-based learning in Web 2.0[2] environments and processing complex project management (programmable wiki systems)
- Producing, programming, publishing, and distributing interactive, purposeful digital media
- Social learning, participation, and exchange
- Information-based learning, search, and exploration
- Thoughtful surfing of websites and web applications

Globaloria is the platform and curriculum that takes students through a rich learning process, as depicted in Figure 1. A complete review of this framework is available on the Workshop's website (http://www.worldwideworkshop.org/reports). A brief review is also available in Chadwick and Gore (2011).

Globaloria is a "rigorous turn-key instructional solution for [science, technology, engineering, and mathematics] STEM learning," comprising a year-long academic curriculum, game design and programming tutorials, game-content resources, and virtual support systems for educators and youth (http://www.worldwideworkshop.org/). Working independently and in small teams, students drive the design process, taking an original idea to final game product. No prior web design or programming skills are needed. Learning by doing, students are educated in both technical and computational skills and in content knowledge that prepares them for college-level studies, as well as for digital citizenship and careers in the global knowledge economy (World Wide Workshop Foundation, 2012). It is hypothesized that through participation in the Globaloria program, students will develop 21st century skills in digital literacy and social media while gaining a deeper understanding of curricular areas, such as science, mathematics, health, and global issues (Edvantia, 2008), which is essential for success in the 21st century (Harel et al., 2010).

[2] A Web 2.0 environment is a shift in the way users think about technology from consumers to creators.

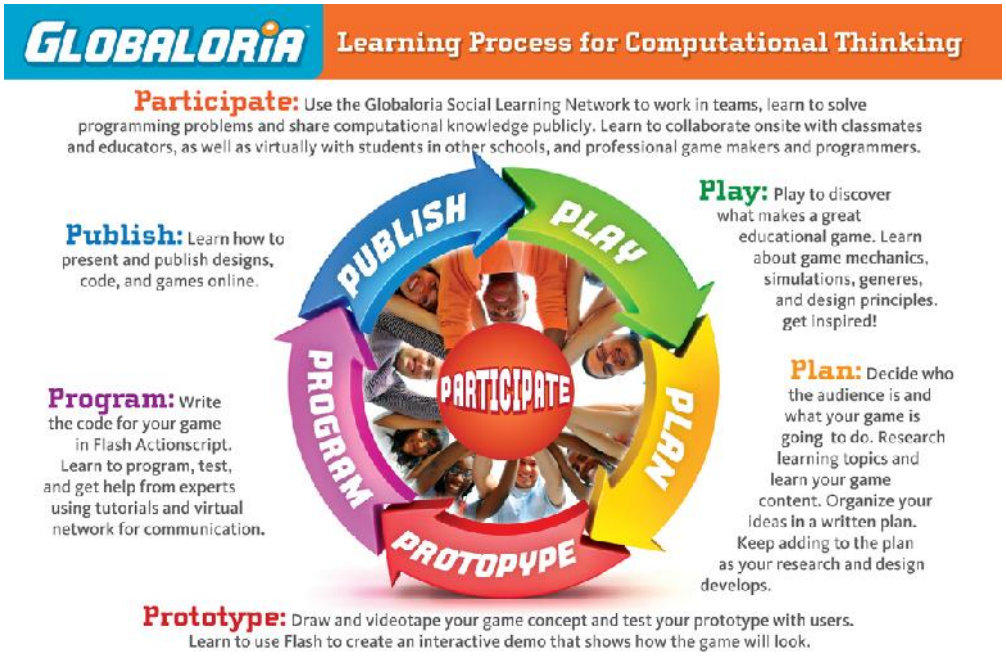


Figure 1. The Globaloria platform and its curricular units guide students and educators through a multidimensional blended learning process.

Nevertheless, through participating in Globaloria, students are also involved in and exposed to a variety of academic subjects and social issue topics, such as science, math, language arts, architecture, art, climate change, ecology, water, community services, technology skills, peace, and more (Harel et al., 2010). Therefore, the potential impacts of Globaloria participation are likely to go beyond the scope of STEM education and learning (Chadwick & Gore, 2011; Reynolds & Chiu, 2012).

Globaloria-WV

In August 2007, the Workshop deployed Globaloria into classrooms throughout West Virginia. It was the first statewide implementation of a social learning network using game design (Chadwick & Gore, 2010, 2011; World Wide Workshop Foundation, 2009). Since then, *Globaloria-WV* has included thousands of students from middle schools, high schools, community colleges, and alternative education programs across the state. Educators implement Globaloria as an elective, a high school completer course, and as an integrated program within the school core curriculum, aligning the program with Global21 state standards and objectives in English/language arts, mathematics, science, and 21st century skills (World Wide Workshop Foundation, 2012). As a first-of-its-kind statewide network for learning, this model demonstrates the potential for Globaloria to be implemented on a large scale, integrated at multiple grade levels, across the public school system and higher education.

To evaluate the cognitive, behavioral, and affective impact of the Globaloria program, the Workshop researchers have used a variety of evaluation methods, including the use of surveys; tracking of activities and behaviors; case studies; evaluations of work products; evaluations of wiki participation; interviews, conference calls, and e-mail exchanges with educators; in-person visits;

and videos and transcriptions from site visits. Information from these data collection measures informed a comprehensive evaluation of the impact of Globaloria and the refinement of the program over time (World Wide Workshop Foundation, 2012). This section summarizes the key findings from previous evaluations and activities. For a complete list of prior studies and reports related to *Globaloria-WV*, please visit http://www.worldwideworkshop.org/reports.

Brief Research Review for *Globaloria-WV*

In December 2008, Edvantia researchers developed a research agenda based on the review of extant data (e.g., project descriptions, evaluation reports, and participant artifacts) to guide future study of the Globaloria education intervention (Knestis, 2008). The agenda included a logic model describing the associations between program activities and different phases of program outcomes (see Appendix A for the logic model). To date, multiple research and evaluation studies have been conducted to examine the various relationships between program components and program outcomes depicted in the logic model. Findings have been positive for the program. Specifically, during the second year of the pilot study, research indicated that middle school student attitudes toward Globaloria activities and self-reported knowledge increased, especially for the more constructionist, effortful, and difficult activities (Reynolds, Scialdone, & Harel, 2010). For the Year 3 evaluation, the goal was to expand the evaluation of *Globaloria-WV* to include not only the impact of Globaloria on CLAs and teacher professional development (Whitehouse, 2009), but also its effect on student achievement, performance, and the relationship among classroom ecology, game design pedagogy, and Globaloria learning (Harel et al., 2010). These evaluation studies provided preliminary evidence that participation in Globaloria may have positive effects on student performance and academic achievement. Specifically, researchers found that Globaloria students scored moderately higher on five out of six assessment measures (three unit tests, a semester final, course average, and course grade) (Harder, Stuart, & Anthony, 2010; World Wide Workshop Foundation, 2012).

In recent reports, Reynolds and Chiu (2012) extended their work to examine how school-level factors (e.g., educators' experience with the program, educator training, motivation and expertise, and teamwork models) might be contributing to variations in student outcomes. This research is currently underway, and findings will provide a valuable framework for predicting classes that may require additional support and guidance.

Purpose of this Study

Edvantia has worked with the Workshop for the past four years to explore the effectiveness of *Globaloria-WV*. In 2010, Edvantia conducted a pilot study of the effects of Globaloria participation on student achievement. That study provided preliminary data concerning the performance of students participating in Globaloria relative to the Global21 standards, as measured by the WESTEST2, in math, reading/language arts, science, and social studies. Findings indicated that participants in Globaloria performed significantly better on the 2009 WESTEST2 science and social studies subtests than those who did not participate in the program (Chadwick & Gore, 2010); however, this study was limited by its small sample size.

In 2011, Edvantia conducted a replication study to gather further support for the pilot findings. In the replication study, researchers increased the sample size and used a more robust matching procedure (propensity score matching [PSM]) than was possible for the pilot study. Researchers found that students who participated in the Globaloria program scored slightly higher than comparison students on all four subsections of the 2010 WESTEST2; however, a significant difference was found only in the WESTEST2 science domain after controlling for previous achievement. This finding suggested that Globaloria participation was positively related to student science performance, but was not related—or only loosely related—to other subjects tested on the WESTEST2 (Chadwick & Gore, 2011).

Although the Globaloria program has shown promise for improving digital literacy in West Virginia students, more confidence in the relationships between program participation and achievement outcomes would result if similar outcomes were found with a larger sample size. In Year 4, researchers expanded upon the previous two studies conducted by Chadwick and Gore (2010, 2011) in two major ways: (1) further increasing the sample size, and (2) using a modified, more rigorous quasi-experimental methodology that allows researchers to take into consideration within- and between-school variations in student achievement. Due to the complex sampling design (i.e., students nested within schools), two-level hierarchical linear modeling (HLM) was used to analyze the data where students (level 1) were nested within schools (level 2). The findings from this study suggested that Globaloria participation positively affects students' math and reading achievement, and to a lesser extent, science achievement, as measured by the WESTEST2. Additionally, it seems that school-level contextual factors may be important in determining the extent to which participation will affect science and reading achievement (Ho et al., 2012).

While it is difficult to compare the findings across the three studies (pilot study, replication study, and the Year 4 study using a HLM design), researchers compared the key findings to the extent possible. In the pilot study, Globaloria participants outperformed their comparison counterparts in science and social studies achievement; however, only the difference for science was upheld in the replication study. In the Year 4 study, Globaloria participants outperformed their comparison counterparts in mathematics and reading, which is different from the findings of the previous two studies. In addition, while there continued to be some influence of Globaloria participation on science achievement, the influence was observed only in schools with low overall mathematics achievement. It is likely that the change in findings is due to the increased sample size, improved matching procedures, and the robust HLM design that takes within- and between-school variations in student achievement into consideration (Ho et al., 2012).

In Year 5, the study seeks to expand the previous studies and provide additional data concerning the performance of students participating in *Globaloria-WV* relative to the 2012 WESTEST2 using a modified replication of the Year 4 study. See Appendix A for how this study relates to the previously developed logic model and research agenda (Knestis, 2008).

Methods

In Year 5, 50 schools and education institutions participated in *Globaloria-WV*. Of those, Edvantia evaluators included 38 traditional middle and high schools (10 middle schools and 28 high schools) in the replication study.[3] To create a counterfactual group, evaluators used PSM (Stuart & Rubin, 2007) methods to select a group of comparison students from similar schools who were comparable to participating students on key student-level variables prior to participation. In the field of education and evaluation, PSM provides advantages to increase the analytic power for causal inferences when random assignment is not feasible or ethical (Stuart, 2010).

Sample Selection

The PSM matching for this study involved two stages. The first stage was to select a group of comparison schools[4] that were similar to the Globaloria schools on four key school-level characteristics: (1) percent of students achieving reading proficiency on the 2011 WESTEST2, (2) percent of students achieving math proficiency on the 2011 WESTEST2, (3) school enrollment, and (4) percent of students qualifying for free or reduced-price meal (FRPM) status. FRPM status was used as a proxy for income (i.e., students who qualified for FRPM status were considered to be low-income students). These school-level characteristics are commonly used in studies incorporating PSM because they have proven over time to be stable and significant covariates of student achievement outcomes (Stuart, 2010). The matching process ensured that the selected comparison schools had similar demographic and achievement characteristics to participating schools before participation in Globaloria. At this stage, the matching was conducted separately for middle schools and high schools.

Once comparison schools were identified, the second stage was to select a group of comparison students from the identified comparison schools to match with participating students on key student-level attributes that are associated with student achievement outcomes, including grade level, gender, race, FRPM status, limited English proficiency (LEP), disability, and 2011 WESTEST2 scale scores (i.e., pretest). For this stage of matching, the Workshop provided Edvantia researchers with a list of 931 students for which project staff obtained parental consent for participation in the study.[5] Of the 931 Globaloria students, 75 were removed from the sample due to incorrect identification numbers, which prevented the researchers from accessing their WESTEST2 scores. Of the remaining students, 230 were removed from the sample due to missing all four subsets of WESTEST2 pretest or posttest scores. This process resulted in a sample of 626

[3] Nine participating schools were removed from this study because they are not traditional middle or high schools. Of the remaining schools, one high school and three middle schools were removed because data from these four schools were not available for the analysis.

[4] Evaluators obtained a complete list of public middle and high schools in West Virginia from the West Virginia Department of Education website (http://wveis.k12.wv.us/nclb/pub). At the request of the Workshop, all traditional middle and high schools were included. Colleges and alternative schools were excluded.

[5] All students who participate in the Globaloria program are required by the Workshop to return parental consent to participate in the program and all research activities related to the program; students who do not return consent cannot participate.

students before matching.[6] [7] Again, the matching was conducted separately for middle school students and high school students. Table 1 displays the total number of participating students remaining in the sample for further analyses by participating schools. Details regarding PSM logic, processes, and procedures are reported in Appendix B.

Table 1. Number of Remaining Globaloria Participating Students by School

School	Initial Participating Students	Final Sample
Middle Schools	***n***	***n***
Chapmanville Middle School	19	19
Eastern Greenbrier Middle School	32	31
Elkins Middle School	44	43
Kasson Elementary/Middle School	9	9
Lenore K-8 School	14	13
Logan Middle School	10	9
Man Middle School	11	11
Sandy River Middle School	50	48
South Charleston Middle School	16	16
Tygarts Valley Middle School	36	35
MIDDLE SCHOOL TOTAL	**241**	**234**
High Schools	***n***	***n***
Braxton County High School	17	12
Buffalo High School	11	10
Capital High School	22	19
Chapmanville Senior High School	6	5
Doddridge County High School	55	44
Elkins High School	17	8
Fayetteville High School	7	5
George Washington High School	26	18
Greenbrier East High School	30	24
Greenbrier West High School	10	2

[6] Attrition analyses examined the differences between samples with at least one of the four WESTEST2 subscale scores (n = 626) in a given year and samples with none of the four WESTEST2 subscale scores (n = 230) in a given year. Results showed that the samples with none of the four WESTEST2 subscale scores in a given year were more likely to be those who did not qualify for FRPM status and were high school students, particularly Grade 12, before participation. The interpretation of study findings should be made with caution as a high proportion of students who did not qualify for FRPM status were removed from the analysis due to missing WESTEST2 data.

[7] Of the remaining 626 participating students, seven students had at least one 2011 WESTEST2 subtest score missing. Because PSM requires full information to proceed with the matching, multiple imputations were conducted to impute missing values for the cases with only partial missing data (Hill, 2004).

School	Initial Participating Students	Final Sample
Hurricane High School	76	47
Lincoln High School	17	7
Logan Senior High School	8	7
Man Senior High School	12	12
North Marion High School	19	11
Oak Glen High School	39	17
Oak Hill High School	22	16
Philip Barbour High School Complex	46	26
Preston High School	39	30
Riverside High School	57	3
South Harrison High School	13	9
Spring Valley High School	64	17
Tygarts Valley High School	18	17
University High School	14	6
Webster County High School	14	4
Weir High School	9	2
Wheeling Park High School	17	12
Woodrow Wilson High School	5	2
HIGH SCHOOL TOTAL	**690**	**392**
TOTAL PARTICIPATING STUDENTS	**931**	**626**

Balance Diagnosis

After the matching was completed, balance diagnostics were conducted to check the "goodness" of the matches. It was expected that the selected comparison schools and students would be similar to the participating schools and students on all covariates that were used for the matching (Rubin, 2001; Stuart, 2010). As shown in Figures 2 and 3, an examination of the distribution of propensity score distributions was first conducted to assess common support via a graphic diagnostic; then, three numerical balance measures were used to check covariate balance at the student level (Rubin, 2001):

- The ratio of the variances of the propensity score in the two groups must be close to 1.0. Rubin (2001) suggests that the variance ratios should be between 0.5 and 2.0.
- The difference in the means of the propensity scores in the two groups being compared must be small. Rubin (2001) suggests that the standardized differences of means should be less than 0.25.
- The percent of balance improvement, where the larger percentage of balance improvement indicates better results of the PSM.

As shown in Figure 2, the jitter plots suggest that the selected comparison schools and students have similar distributions of propensity scores. As shown in Table 2, the ratio of the variances of the propensity score ranged between 0.90 and 1.11, well within the range suggested by Rubin (2001). The analyses of standard mean differences suggest that the matching procedures have significantly minimized the group mean differences at both the school- and student-levels. Most importantly, after the matching, the majority of the covariates had a standardized mean difference smaller than 0.1, which is much smaller than the value of 0.25 suggested by Rubin (2001). The percent of balance improvement for the school-level variables ranged from 7% to 92% across middle schools and high schools, and student-level variables ranged from 27% to 100%. Taken together, these diagnostic criteria suggest that participating and comparison schools were similar in the key school-level covariates before Globaloria participation. Additionally, the comparison students selected from these identified comparison schools were very similar to the participating students before Globaloria participation in all aspects of the key covariates that are associated with student achievement.

Sample Characteristics

Table 3 shows the final samples' student- and school-level attributes by school type (i.e., middle schools versus high schools) after the matching. Overall, the majority of student-level attributes differed significantly by school type. School poverty level was significantly higher among middle schools than high schools.[8] These preliminary findings suggest that the participating middle schools and high schools differed in many characteristics that are associated with student achievement outcomes; therefore they should be analyzed separately.

Analytic Plan

Of specific interest in this evaluation was the relationship between students' participation status (whether or not they participated in the Globaloria program) and WESTEST2 outcomes. Due to the complex sampling design (i.e., students nested within schools), two-level HLM was used to analyze the data where students (level 1) were nested within schools (level 2). In the real world, students from the same schools are more similar than students from different schools. The analysis using HLM provides the advantages of taking within- and between-school variations in student achievement into account (Raudenbush & Bryk, 2002).

[8] Nationally, fewer high school students are reported eligible for the FRPM program compared to younger children (http://nces.ed.gov/programs/coe/indicator_pcp.asp).

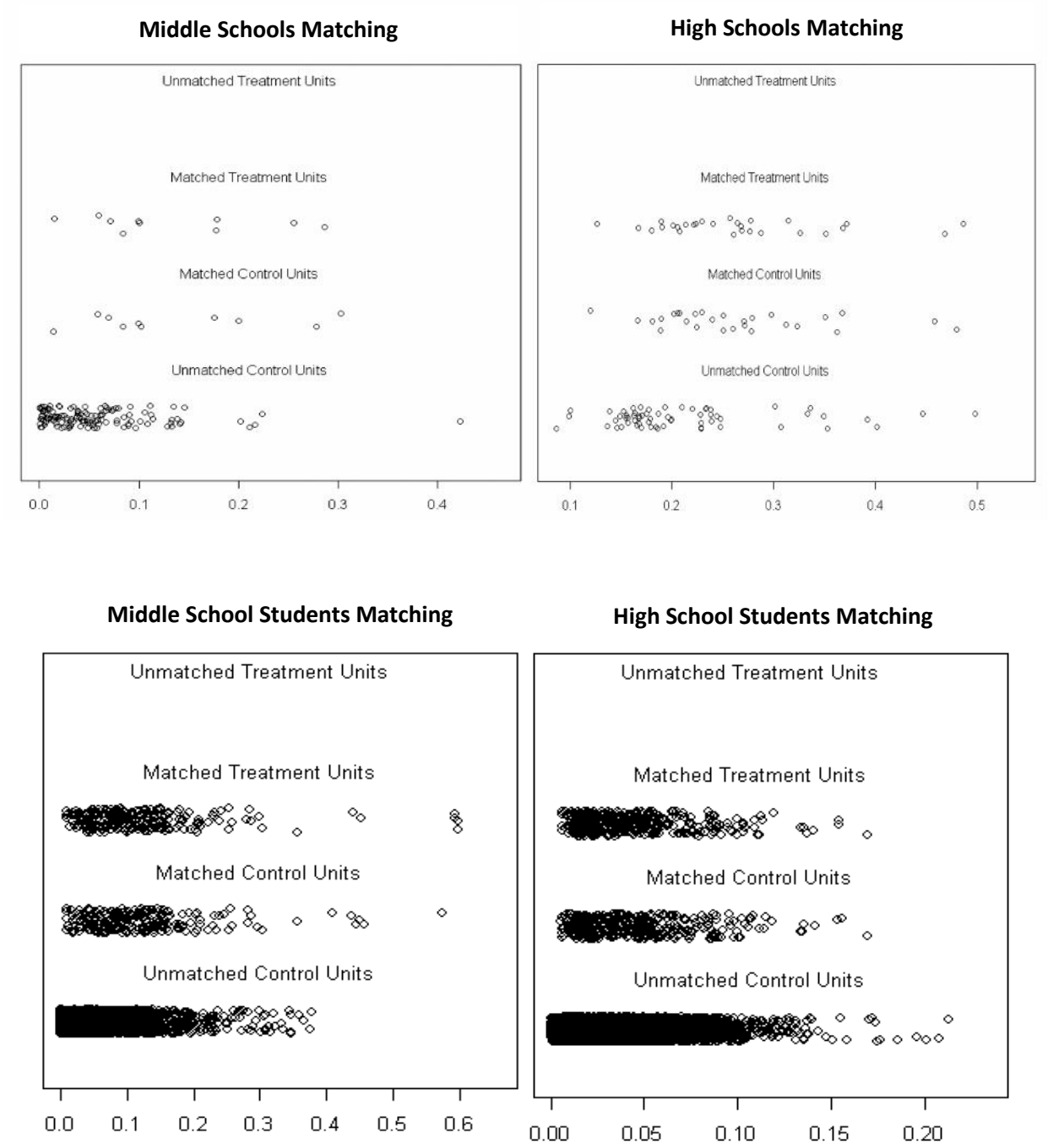


Figure 2. Graphical diagnosis of PSM results at the school-level and student-level.

Table 2. Balance Diagnosis Before and After the Matching

Variables	Participants		Comparison				Balance Diagnosis			
			Before		After		Propensity Score Variance Ratio	Standard Mean Differences		% Balance Improvement
	M	*SD*	*M*	*SD*	*M*	*SD*		Before	After	
School-Level										
Middle Schools										
Propensity score	0.13	0.09	0.06	0.06	0.14	0.10	0.90	0.82	-0.06	92.11
School Enrollment[a]	4.70	1.95	3.92	1.93	4.90	2.60		0.40	-0.10	74.24
% Reading Proficiency	45.96	7.50	49.17	9.58	47.25	9.90		-0.43	-0.17	59.68
% Math Proficiency	40.52	8.37	42.41	9.53	41.07	8.38		-0.23	-0.07	70.90
% Low-income Students	58.40	8.59	54.34	12.8	57.67	7.93		0.47	0.09	82.02
High Schools										
Propensity score	0.27	0.08	0.23	0.08	0.26	0.08	1.00	0.44	0.02	0.27
School Enrollment[a]	7.36	3.06	5.96	3.36	7.14	3.06		0.46	0.07	7.36
% Reading Proficiency	45.36	9.14	44.59	9.54	44.63	7.34		0.08	0.08	45.37
% Math Proficiency	40.57	11.09	40.08	10.11	39.57	9.02		0.05	0.09	40.57
% Low-income Students	45.12	9.62	47.12	12.26	44.29	11.69		-0.21	0.09	45.12
Student-Level[b]										
Middle Schools										
Propensity Scores (PS)	0.12	0.10	0.05	0.06	0.12	0.09	1.11	0.69	0.02	96.73
WESTEST2 2011 Math	653.31	41.66	609.05	59.48	653.33	40.80		1.06	-0.00	99.94
WESTEST2 2011 Reading	496.58	30.84	459.03	53.59	498.14	36.25		1.22	-0.05	95.85
WESTEST2 2011 Science	592.17	25.54	559.69	49.75	590.91	31.20		1.27	0.05	96.12
WESTEST2 2011 Social Studies	419.15	27.16	392.99	42.57	419.11	30.32		0.96	0.00	99.86
Grade Level	7.30	0.65	6.98	0.82	7.32	0.71		0.50	-0.02	96.06
Race[c]	0.05	0.29	0.13	0.42	0.06	0.29		-0.28	0.00	100.00

Variables	Participants		Comparison				Balance Diagnosis			
			Before		After		Propensity Score Variance Ratio	Standard Mean Differences		% Balance Improvement
	M	*SD*	*M*	*SD*	*M*	*SD*		Before	After	
Gender (Male)	0.57	0.50	0.52	0.50	0.57	0.50		0.10	0.00	100.00
FRPM (Low-SES)	0.44	0.50	0.59	0.49	0.43	0.50		-0.31	0.03	91.84
Disability[d]	0.02	0.15	0.14	0.34	0.03	0.16		-0.80	-0.03	96.30
High Schools										
Propensity Scores (PS)	0.04	0.03	0.03	0.02	0.04	0.03	1.00	0.58	0.00	99.37
WESTEST2 2011 Math	660.59	50.32	641.62	54.90	660.51	48.13		0.37	-0.01	97.53
WESTEST2 2011 Reading	491.55	53.61	478.81	55.79	493.09	53.13		0.24	0.03	87.85
WESTEST2 2011 Science	621.99	52.44	604.23	51.60	620.50	51.79		0.34	0.03	91.59
WESTEST2 2011 Social Studies	413.80	33.78	403.70	35.06	412.67	32.05		0.30	0.03	88.80
Grade Level	10.17	0.76	9.95	0.82	10.16	0.79		0.29	0.01	95.32
Race[c]	0.06	0.27	0.09	0.34	0.07	0.32		-0.06	-0.03	55.64
Gender (Male)	0.74	0.44	0.51	0.50	0.75	0.43		0.52	-0.04	92.25
FRPM (Low-SES)	0.40	0.49	0.45	0.50	0.44	0.50		-0.09	-0.07	26.76
Disability[d]	0.10	0.30	0.12	0.33	0.09	0.29		-0.08	0.03	55.71

[a] When conducting PSM, school enrollment was recoded as an ordinal scale to improve the matching result.

[b] LEP status was removed from student-level matching because only one participating student had LEP; adding LEP in the PSM models caused calculation issues.

[c] Race is a dichotomous variable: 0 = White, 1 = Black, 2 = Other.

[d] Disability is a binary variable: 0 = non-disability, 1 = disability.

Table 3. Student and School Characteristics by School Type

Variable	Middle Schools			High Schools			Independent Sample *t*-tests *p* value
	n	*M*	*SD*	*n*	*M*	*SD*	
Student Level							
Gender	470	0.57	0.50	786	0.74	0.44	<0.001
FRPM	470	0.43	0.50	786	0.42	0.49	0.644
LEP	470	0.00	0.00	786	0.01	0.07	0.045
Disability	470	0.02	0.15	786	0.09	0.29	<0.001
Race	470	0.06	0.28	786	0.06	0.29	0.565
WESTEST2 Math 2011	470	653.32	41.18	786	660.29	49.21	0.007
WESTEST2 Reading 2011	470	497.36	33.63	786	492.32	53.35	0.040
WESTEST2 Science 2011	470	591.54	28.49	786	621.24	52.08	<0.001
WESTEST2 Social Studies 2011	470	419.13	28.75	786	413.23	32.91	0.001
WESTEST2 Math 2012	470	658.62	40.01	782	666.52	50.60	0.002
WESTEST2 Reading 2012	470	497.81	31.62	783	495.34	53.25	0.302
WESTEST2 Science 2012	470	599.97	30.60	672	636.05	55.37	<0.001
WESTEST2 Social Studies 2012	470	419.60	28.30	782	412.30	36.29	<0.001
School Level							
% FRPM students	20	58.04	8.06	56	44.71	10.61	<0.001
% passing/meeting reading proficiency	20	46.61	8.58	56	44.99	8.22	0.458
% passing/meeting math proficiency	20	40.79	8.16	56	40.07	10.03	0.773

Note. Gender, FRPM, LEP, and disability are binary variables (Gender: 0 = female, 1 = male; FRPM: 0 = student without free/reduced-price meals, 1 = students receiving free or reduced-priced meals; LEP: 0 = without LEP, 1 = with LEP; disability: 0 = without disability, 1 = with disability). Race is a categorical variable (0 = White; 1 = Black; 2 = Other)

The preliminary examination of student demographic composition revealed that minority students (i.e., Black and others) comprised less than 5% of the sample. A series of one-way analyses of covariance (ANCOVAs) suggested that Black students and students from other races/ethnicities were significantly different in achievement outcomes from White students (i.e., Black students on average had significant lower 2012 WESTEST2 scores while students from other racial/ethnic backgrounds had higher 2012 WESTEST2 scores as compared to White students). These patterns suggest that it is more appropriate to examine the associations between participation and student achievement for each racial group independently. For White students, HLM was used given a large enough sample size. For Black students, HLM was not appropriate due to small sample size (n =45); therefore, a series of ANCOVAs were conducted to examine the main effect of participation on students' 2012 WESTEST2 scores by controlling for 2011 WESTEST2 scores. Results showed that participating and comparison Black students were not statistically

different on 2012 WESTEST2 subtests (i.e., mathematics, science, reading, and social studies).[9] Given the lack of significant findings, Black students will not be discussed further in subsequent sections. Similarly, students from other races/ethnicities had such a small sample size (n = 16) that it was not appropriate to perform any type of statistical analysis on the group; therefore, this group will not be discussed further.

To perform the HLM analyses, student-level demographic information (i.e., gender and FRPM status) and pretest scores (i.e., 2011 WESTEST2 scale scores) were entered into the models to adjust for within-school variations in 2012 WESTEST2 achievement outcomes. Since students from the same school would have the same participation status (i.e., participating schools versus non-participating schools), student participation was entered into the models as school-level predictors. Additionally, although not the main focus of this evaluation, three school-level contextual characteristics (i.e., percentage of low-income students, percentage of students scoring proficient and higher on the 2011 WESTEST2 math subscale, and percentage of students scoring proficient and higher on the 2011 WESTEST2 reading subscale) were entered into the model to adjust for between-school variations in student 2012 WESTEST2 outcomes.

The relationships among student participation status and school-level contextual factors on student outcomes can be quite complex; when there is no significant relationship between participation and student achievement, knowledge about the school-level characteristics may help researchers and program staff understand what may be influencing the outcomes (Lee, 2000). For program staff, this type of information is critical for the purpose of planning and programing. As such, studying school-level contextual factors is a key concern for future research. For this purpose, interaction terms of student participation status and school-level contextual factors were also added into the HLM model to examine whether the relationships between Globaloria participation and student 2012 WESTEST2 outcomes were moderated by school-level contextual factors. By including these interaction terms (i.e., participation status by school-level contextual factors) into the model, researchers may be able to better understand whether school-level contextual factors (e.g., school performance measured by percentage of students scoring proficient or higher) play a role in either "strengthening" or "weakening" the associations between Globaloria participation and student outcomes. For example, researchers examined whether Globaloria participation had a different impact on student achievement if the students were in schools with a high percentage of FRPM-eligible students. The three school-level contextual factors are (1) percentage of low-income students, (2) percentage of students scoring proficient and higher on the 2011 WESTEST2 math subscale, and (3) percentage of students scoring proficient and higher on the 2011 WESTEST2 reading subscale. It should be noted that, in this particular study, adding these school-level contextual factors in the model was primarily exploratory, and the factors selected for inclusion were based on information that was readily available in the dataset. For future studies, Globaloria program staff are encouraged to collect some "proximal" school-level or teacher-level factors that may help explain how Globaloria works to support student achievement.

[9] This finding should be interpreted with caution given that the sample sizes were small.

Results

The main purpose of this report is to document the relationships between Globaloria participation and student achievement in the four subject domains assessed by the WESTTEST2 (i.e., math, science, reading, and social studies). Due to insufficient sample sizes of racial/ethnic groups other than White students (e.g., Asian and Black student populations), only the results of HLM analyses for White students from middle schools and high schools are reported. Details with regard to model building and technical reports of HLM analyses are included in Appendix D.

In summary, student pretest scores (2011 WESTEST2 scores) are the strongest predictors of student achievement outcomes. However, student gender and FRPM status, as well as school-level contextual factors (i.e., school poverty and school achievement data), also contribute to the variations of student achievement outcomes. A snapshot of the overall findings is presented in Table 4; complete statistical reports are presented in Appendix D.

After controlling for the covariates, the main effects of Globaloria participation on student achievement were not statistically significant across all WESTEST2 subtests; however, the effects of Globaloria participation were moderated by school-level contextual factors, as well as student-level characteristics. These findings suggested that the effects of Globaloria were conditional on certain student-level attributes and school-level contextual characteristics. Since there were significant interaction effects, the interpretation of the fixed effect of Globaloria participation on student outcomes is less meaningful. Instead, it is necessary to conduct follow-up analyses to estimate the effects of simple intercepts and simple slopes to get a better understanding of how Globaloria participation effects may be different across subgroups and contexts. Graphs of interaction effects are also presented in each subsection for readers to visually understand these various interaction effects.

Overall, the final HLM models explained between 48% and 55% of within-school variance in student achievement outcomes, and 85% to 100% of between-school variance in student achievement outcomes.[10] More details on the key findings of each subset of WESTEST2 outcomes are presented.

[10] The majority of the between-school variation in student achievement were explained by both the level-2 predictors (i.e., school poverty and school math and reading proficiency) as well as level-1 predictors (i.e., grand mean centered level-1 pretest scores).

Table 4. Effects of Globaloria Participation on Student 2012 WESTEST2 Outcomes

Parameter Estimate	Math				Reading				Science				Social Studies			
	Middle		High		Middle		High		Middle		High		Middle		High	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Within-School Effects																
Intercept	656.19***	3.05	666.03***	5.46	500.11***	2.85	506.94***	3.63	598.72***	2.36	635.70***	5.59	421.82***	1.64	410.10***	2.63
Gender	7.11[a]	3.74	7.51[a]	4.14	-7.51**	2.86	-12.48*	4.92	3.76	2.47	4.84	5.76	-4.00[a]	2.18	1.89	2.76
FRPM	-1.89	3.29	-10.50**	3.89	-1.52	1.85	-7.99*	4.03	-2.48	2.15	-15.27**	5.26	-4.87**	1.76	-2.89	3.09
Math 2011	0.52***	0.07	0.34***	0.04	0.09***	0.02	0.15**	0.04	0.17***	0.05	0.33***	0.06	0.10*	0.04	0.08**	0.03
Reading 2011	0.10	0.07	0.12**	0.04	0.32***	0.05	0.44***	0.07	0.13**	0.04	0.00	0.07	0.15***	0.03	0.12**	0.04
Science 2011	0.29***	0.07	0.24***	0.05	0.24***	0.05	0.05	0.03	0.20***	0.05	0.22***	0.05	0.17*	0.07	0.07***	0.02
Social Studies 2011	0.05	0.07	0.12[a]	0.07	0.16**	0.06	0.33***	0.08	0.34***	0.06	0.50***	0.10	0.34***	0.06	0.50***	0.04
Between-School Effects																
Participation	-0.82	3.96	3.92	6.69	3.51	3.47	-0.81	4.51	0.92	3.49	7.35	6.98	-5.22[a]	2.47	0.83	3.42
SchPoverty	0.17	0.10	0.39	0.33	-0.21	0.16	0.48[a]	0.25	0.18	0.19	0.10	0.41	0.09	0.14	0.07	0.21
SchReadProficiency	0.18	0.14	-0.46	0.44	0.60**	0.19	0.75	0.58	0.45[a]	0.22	-0.13	0.47	0.11	0.12	-0.08	0.30
SchMathProficiency	0.09	0.18	1.04[a]	0.59	-0.74**	0.23	-0.26	0.44	-0.21	0.37	-0.05	0.59	0.24	0.23	0.18	0.28
Interactions Effects																
Participation x Gender Slope	-2.49	4.80	-8.43	5.54	-4.21	4.55	2.88	6.21	-1.82	3.52	-9.18	7.89	14.24***	3.59	2.00	3.72
Participation x FRPM Slope	0.52	4.53	7.60	5.71	2.60	3.53	8.45	6.15	-2.14	2.94	16.94*	8.21	2.65	2.98	5.61	3.94
Participation x SchPoverty	-0.06	0.23	-0.54	0.40	0.34	0.36	-0.74	0.53	-0.10	0.22	-1.07[a]	0.62	0.03	0.25	-0.15	0.31
Participation x SchReadProficiency	-0.15	0.30	0.71	0.49	-0.09	0.45	-0.13	0.66	-0.68[a]	0.34	-1.06	0.72	-0.08	0.34	0.05	0.38
Participation x SchMathProficiency	-0.18	0.27	-1.38*	0.63	0.43	0.34	-0.77	0.77	0.13	0.39	-0.01	0.67	-0.04	0.32	-0.20	0.35
Variance Explained																
Within-School Variance Explained	0.55		0.49		0.49		0.55		0.48		0.49		0.48		0.54	
Between-School Variance Explained	1.00		0.85		1.00		0.90		1.00		0.88		0.96		0.95	

*** $p < 0.001$
** $p < 0.01$
* $p < 0.05$
[a] $p < 0.10$

Mathematics

Although there were no significant correlations between overall Globaloria participation and student mathematics outcomes for middle and high schools, there was a statistically significant negative interaction between Globaloria participation and school-level math proficiency on student math outcomes (β = -1.38, SE = 0.63, p = 0.034) for high schools. The negative interaction effect suggests that the effect of Globaloria participation was stronger for students in schools with higher percentages of students struggling with math proficiency. When follow-up analyses were conducted, the results suggest that there was a positive correlation between Globaloria participation and student math outcomes within low math performing schools (i.e., schools with higher percentages of students who did not reach proficiency in math), $\chi^2(1) = 9.54$, $p = 0.002$. In contrast, there was no such correlation for higher math performing schools (i.e., schools with lower percentages of students failing to reach proficiency in math), $\chi^2(1) = 1.41$, $p = 0.23$. As shown in Figure 3, the visual presentation of the interaction effect suggests that the differences in math outcomes between the participants and non-participants was greater within schools where lower percentages of students were proficient in math; whereas there was almost no difference in math outcomes between these two groups within schools where high proportions of students reached proficient levels in math.

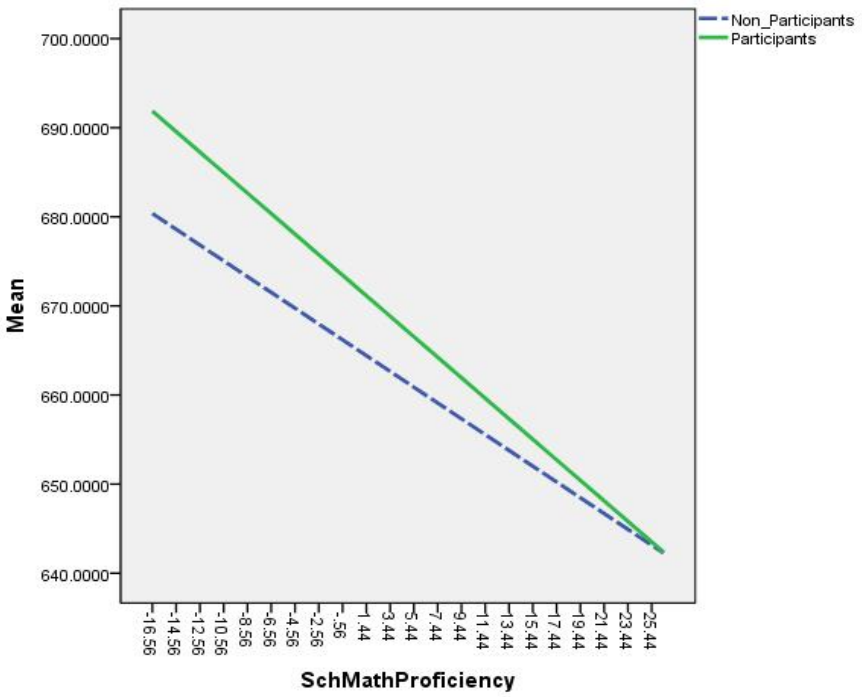


Figure 3. Interaction effect of Globaloria participation and school-level math proficiency status on student math achievement in high schools.

Overall, findings suggest that Globaloria participation had a positive effect on high school students from schools struggling with math proficiency. Additionally, this finding supports the idea that Globaloria might mitigate the effects of certain school-level contextual factors (i.e., school-level math proficiency status) on student math achievement.

Reading

Globaloria participation had no significant effects on student reading outcomes after controlling for student- and school-level covariates regardless of school type (i.e., high school versus middle school).

Science

The associations between overall Globaloria participation and student science outcomes were not statistically significant within middle schools and high schools. A significant interaction effect was found between Globaloria participation and FRPM status on student science outcomes (β = 16.94, SE = 8.24, p = 0.039) for high schools. The positive interaction effect suggests that the effect of Globaloria participation was stronger for students from low-income families (i.e., students enrolled in the FRPM program) than students who were not from low-income families (i.e., students not enrolled in the FRPM program). Follow-up analyses revealed that there were no significant correlations between student FRPM status and student science outcomes within the participating high schools, $\chi^2(1) = 0.09$, $p > 0.50$; however, there was a negative correlation between student FRPM status and student science outcomes within the non-participating high schools, $\chi^2(1) = 8.20$, $p < 0.001$.

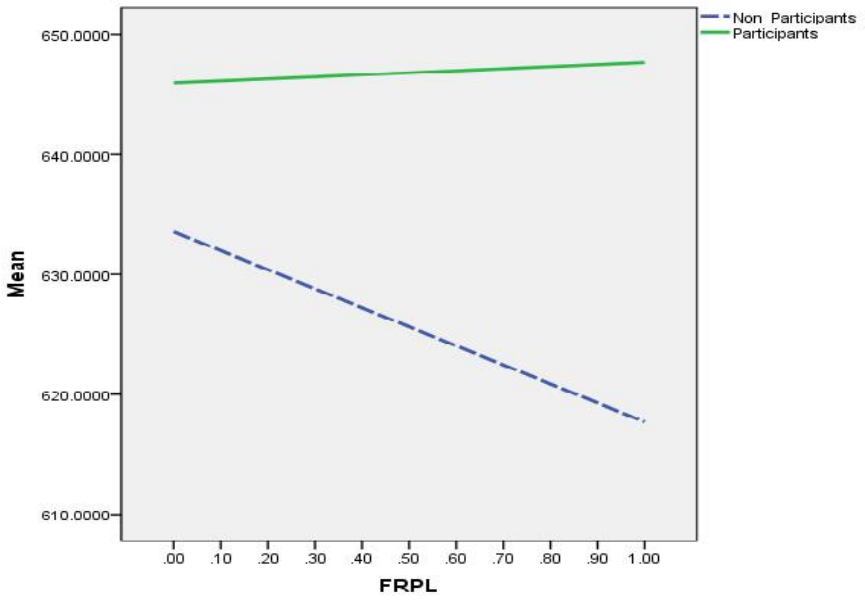


Figure 4. Interaction effect of Globaloria participation and student FRPM status on student science achievement in high schools.

Figure 4 shows that, regardless of FRPM status, Globaloria participants' science scores were approximately 647, whereas the scores of nonparticipants fluctuated by FRPM status (from higher than 630 among students who did not qualify for FRPM versus lower than 620 for students who did qualify for FRPM. This finding suggests that Globaloria participation had a positive effect on high school students' science outcomes for those who come from disadvantaged backgrounds (i.e., low-

income families). This implies that Globaloria participation might mitigate the negative effect of poverty on student science achievement.

Social Studies

The relationship between overall Globaloria participation and student social studies outcomes were not statistically significant within middle schools and high schools. A significant interaction effect was found between Globaloria participation and gender on middle school students' social studies achievement, $\beta = 14.24$, $SE = 3.59$, $p < 0.001$. The positive interaction effect suggests that the effect of Globaloria participation was stronger for males than for females. Follow-up analyses revealed no significant correlation between student FRPM enrollment and student social studies outcomes within the participating high schools, $\chi^2(1) = 0.09$, $p > 0.50$. However, a negative correlation was found between student FRPM enrollment and student social studies outcomes within the non-participating high schools, $\chi^2(1) = 8.20$, $p < 0.001$. As shown in Figure 5, male participants, on average, scored higher in social studies than male non-participants. Conversely, female participants, on average, scored slightly lower in social studies than female non-participants; however, this difference was not statistically significant ($p = 0.14$). This finding suggests that Globaloria participation had a positive effect on middle school male's social studies achievement.

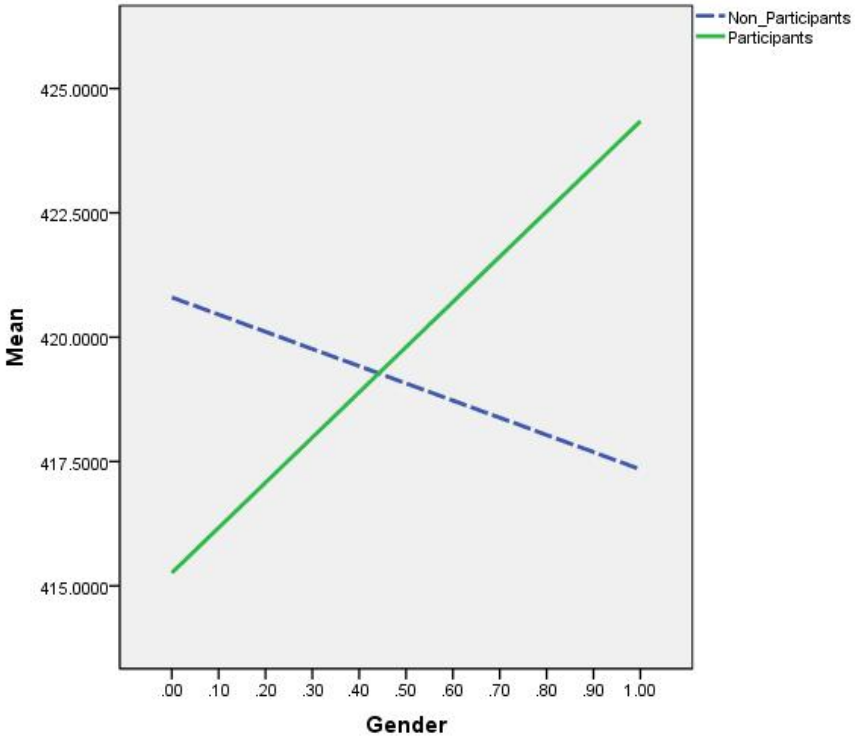


Figure 5. Interaction effect of Globaloria participation and gender on social studies achievement in middle schools.

Discussion and Recommendations

This study examined the relationships between Globaloria participation and student achievement, as measured by 2012 WESTEST2.[11] Edvantia researchers conducted a series of HLM analyses for White students for each of the four core subject areas, including math, science, reading, and social studies.[12] It is important to note that because of multiple analytical decisions, findings of this study are only generalizable to students included in the final sample (i.e., White students with at least one of the 2011 WESTEST2 scale scores). Overall, the key findings are as follows:

- For **mathematics**, Globaloria participation had a positive effect on high school students who were in schools struggling with low math proficiency.
- For **reading**, Globaloria participation had no significant effect.
- For **science**, Globaloria participation had a positive effect on high school students who come from low-income families (i.e., students receiving FRPM).
- For **social studies**, Globaloria had a positive effect on middle school males.

These findings support the effectiveness of Globaloria participation on White students' achievement in math, science, and social studies. Particularly, Globaloria participation seemed to have more of an effect on students who needed more supports to be successful or who were in schools where higher proportions of students needed more supports. These results are similar to, and therefore validate, findings reported in the Year 4 Globaloria study (see Ho et al., 2012).[13] The presence of Globaloria seems to be particularly important for students in schools that are struggling to meet student achievement benchmarks and/or have limited resources. Similarly, the effect of Globaloria on science is stronger for high school students from economically disadvantaged family backgrounds. This finding suggests that a program like Globaloria may mitigate some of the negative effects of family poverty on student achievement, especially in science. A recent study by Reynolds and Chiu (2013, accepted) found similar results related to Globaloria participation and the effects of family poverty on student achievement. Examining program contextual factors as well as the context in which the program is implemented adds valuable information for program staff to understand how and why the program works to support project outcomes (Lee, 2000). Similar to the findings of the Globaloria Year 4 study (Ho et al., 2012), the effects of Globaloria participation seemed to be conditional on certain school-level contextual factors or student-level attributes. More research is needed to further identify the type of contextual factors (e.g., teacher practice and peer interaction) that may moderate the effect of Globaloria on student outcomes. In addition,

[11] It is important to note that the WESTEST2 does not capture the full spectrum of the knowledge and skills that Globaloria may be able to influence, such as digital literacy and computational thinking skills.

[12] Black and students from other racial/ethnic backgrounds were excluded from HLM analyses (see Analytic Plan section for the reasoning).

[13] Key Year 4 findings were as follows: (1) Globaloria participation was positively associated with White students' math achievement; (2) Globaloria participation was positively associated with White students' reading achievement; (3) there was a significant correlation between Globaloria participation and student science outcomes within the schools struggling with math proficiency; and (4) Globaloria participation was not associated with White students' social studies achievement (see Ho et al., 2012 for more detail). Please note that in the Year 4 study, the analyses were not conducted separately for middle and high schools.

Globaloria was not consistently predictive of all four subsets of student achievement outcomes. Most importantly, the inconsistent findings of the main effect of Globaloria participation over the past few years suggests that student achievement outcomes associated with Globaloria participation may be influenced by factors/intermediate outcomes such as student interest in STEM subjects, students' perceptions of STEM education, and student career aspirations in STEM fields. Although these factors have been studied extensively in their own right by Reynolds and colleagues, researchers have not examined the mediating and moderating effects of these factors in relation to student achievement.[14] Globaloria program staff are strongly encouraged to identify and measure intermediate outcomes that affect the relationships between Globaloria participation and student academic outcomes; these are the key variables that help researchers and policymakers to understand how a program like Globaloria supports student achievement (i.e., the pathway links program participation to student academic success). Lastly, as addressed in the Year 4 report, researchers did not find a significant effect of Globaloria participation on Black students' achievement; therefore, they were excluded from the analyses.[15] This finding should be interpreted with caution due to the fact that the sample sizes were small. Given the associations found among socioeconomic status, school achievement, and Globaloria participation among White students, the Globaloria program would benefit from studies of these factors among a more racially and ethnically diverse student population.

Based on these findings, researchers offer several recommendations:

- More studies are needed to understand processes of change associated with Globaloria participation and the influences of contextual variables in order to fully explain how Globaloria works to support students from various backgrounds and school contexts.
- The Workshop is encouraged to re-examine the logic model suggested by Edvantia in 2008 to identify "mediating" variables explaining the processes that link program participation with student outcomes. These are the key ingredients for future program scale-up.
- To better understand the impact of Globaloria on racial and ethnic minority students, Globaloria staff should reach out to these minority communities and encourage participation. Even in West Virginia, the participation levels by minority students are lower than what would be representative of the state's minority student population. In addition, the Workshop program staff should consider replicating the current study's methodology using a more racially and ethnically diverse sample of students, likely from Globaloria implementation sites other than West Virginia, which is a rather homogenously White state.

[14] See http://www.worldwideworkshop.org/reports for research findings by Reynolds and colleagues related to contextual factors influencing Globaloria participation.

[15] While research including minority populations is not possible given the homogenous nature of West Virginia's population, research on the effects of Globaloria on minority students has been conducted by Laura Minnegerode using other implementation sites. See http://www.worldwideworkshop.org/reports for these reports.

For additional recommendations, see the *Globaloria Replication Study: An Examination of the Relationships between Globaloria Participation and Student Achievement in Year 4 of the West Virginia Pilot Implementation* (Ho et al., 2012).

Limitations

Readers should be aware of the limitations of this study. Although PSM provides an advantage to create a comparable counterfactual group by using covariates, it has its own limitations. A specific limitation is that similarity between the Globaloria group and comparison group can only be determined within the covariates used in the matching process. Researchers do not know whether these two groups were different in other aspects that were not measured. Additionally, because a subset of Globaloria participants was removed from PSM due to missing WESTEST2 data (see Sample Selection section), and the removed students had noted differences,[16] these findings should be interpreted with caution, as it does not reflect the full population of Globaloria participants in 2012.

The lack of self-selection indicators may minimize the generalizability of this study. For example, it is unclear who, why, and how students decide to participate in Globaloria. If students who were motivated to learn were more likely to participate and remain in the program, the effects of Globaloria were prone to be confounded with this self-selection factor. Also, self-selection factors are the key covariates that should be included when conducting PSM; however, these factors were not available for this study. Future study including these variables may increase the power of the design as well as results.

A third limitation is related to the fact that the findings of this study are only generalizable to White students with at least one set of pretest scores on the state standardized assessment (i.e. 2011 WESTEST2). Globaloria program staff are encouraged to recruit students from racially and ethnically diverse backgrounds to broaden stakeholders' understanding of Globaloria's impact.

A fourth limitation is that the study does not, nor was it designed to, examine the finer grained explanations for why Globaloria influences student achievement. This study was designed solely to examine participation and standardized test scores. Researchers did not have the information on the types of games created (i.e., science-focused, social studies-focused) and whether the game subject matter influenced the standardized achievement test scores. This information is available, and Edvantia researchers have recommended, and continue to recommend, that this be examined.

[16] Results showed that the samples with none of the four WESTEST2 subscale scores in a given year were more likely to be those who did not qualify for FRPM status and were high school students, particularly Grade 12, before participation.

References

Chadwick, K., & Gore, J. N. (2010). *Globaloria pilot study: The relationship of Globaloria participation and student achievement*. Charleston, WV: Edvantia, Inc.

Chadwick, K., & Gore, J. N. (2011). *Globaloria replication study: Examining the robustness of relationships between Globaloria participation and student achievement*. Charleston, WV: Edvantia, Inc.

Edvantia, Inc. (2008). *A chronicle of West Virginia's 21st century learning initiative*. Charleston, WV: Edvantia, Inc.

Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Thousand Oaks, CA: Sage.

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods, 15*(3), 234-249.

Harel, C. I. (1991). *Children designers: Interdisciplinary constructions for learning and knowing mathematics in a computer-rich school.* Santa Barbara, CA: Praeger.

Harel, C. I., Oliver, A., & Sullivan, S. (2010). *A blueprint for global learning and leadership in the 21st century. Transforming education through the Globaloria learning network (2009-2010 Annual Report)*. New York, NY: World Wide Workshop Foundation.

Hill, J. (2004). *Reducing bias in treatment effect estimation in observational studies suffering from missing data* (Unpublished manuscript). New York, NY: Columbia University Institute for Social and Economic Research and Policy (ISERP).

Ho, H., Gore, J., & Chadwick, K. (2012). *Globaloria replication study: An examination of the relationships between Globaloria participation and student achievement in Year 4 of the West Virginia pilot implementation* Charleston, WV: Edvantia, Inc.

Knestis, K. (2008). *Understanding Globaloria as the subject of research: An agenda for future study*. Charleston, WV: Edvantia, Inc.

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist*, *35*(2), 125-141.

Parsons, L. S. (2001). *Reducing bias in a propensity score matched-pair sample using greedy matching techniques*. A paper presented at the 26th annual meeting of the SAS Users Group International, Long Beach, CA.

Raudenbush, S. W., & Bryk, A. S. (Eds.). (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Reynolds, R., & Chiu, M. (2013, accepted). *How sustained engagement in game design and social media use among diverse students can mitigate effects of the digital divide.* A paper to be presented at the 2013 annual meeting of the American Education Research Association (AERA), San Francisco, CA.

Reynolds, R., & Chiu, M. (2012). *Contribution of motivational orientations to student outcomes in a discovery-based program of game design learning.* A paper presented at the International Conference of the Learning Sciences, Sydney, Australia.

Reynolds, R. & Harel, I. (2009). *Development of high school and community college students' Contemporary Learning Abilities in Globaloria.* A paper presented at the 2009 annual meeting of the American Educational Research Association (AERA), San Diego, CA.

Reynolds, R., Scialdone, M., & Harel, C. I. (2010). *Evidence of high school students' development of contemporary learning abilities in a game design program in rural West Virginia*. New York, NY: World Wide Workshop Foundation.

Rogers, P. J., Pertosino, A., Huebner, T. A., & Hacsi, T. A. (2000). *Program theory evaluation: Practice, promise, and problems*. San Francisco, CA: Jossey-Bass.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology, 2*, 169-188.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science, 25*, 1-21.

Stuart, E. A., & Rubin, D. B. (2007). Best practices in quasi-experimental designs: Matching methods for causal inference. In J. Osborne (Ed.), *Best Practices in Quantitative Social Science* (Vol. 11, pp. 155-176). Thousand Oaks, CA: Sage.

Whitehouse, P. (2009). *Networked teacher professional development: The case of Globaloria*. New York, NY: World Wide Workshop Foundation.

World Wide Workshop Foundation. (2008). *The case for the Globaloria network in West Virginia: Empowering West Virginia youth to create and collaborate online with a 21st-century game-making curriculum*. New York, NY: World Wide Workshop Foundation.

World Wide Workshop Foundation. (2009). *Globaloria: Empowering students and educators with 21st-century digital literacy through a game-making learning network*. New York, NY: World Wide Workshop Foundation.

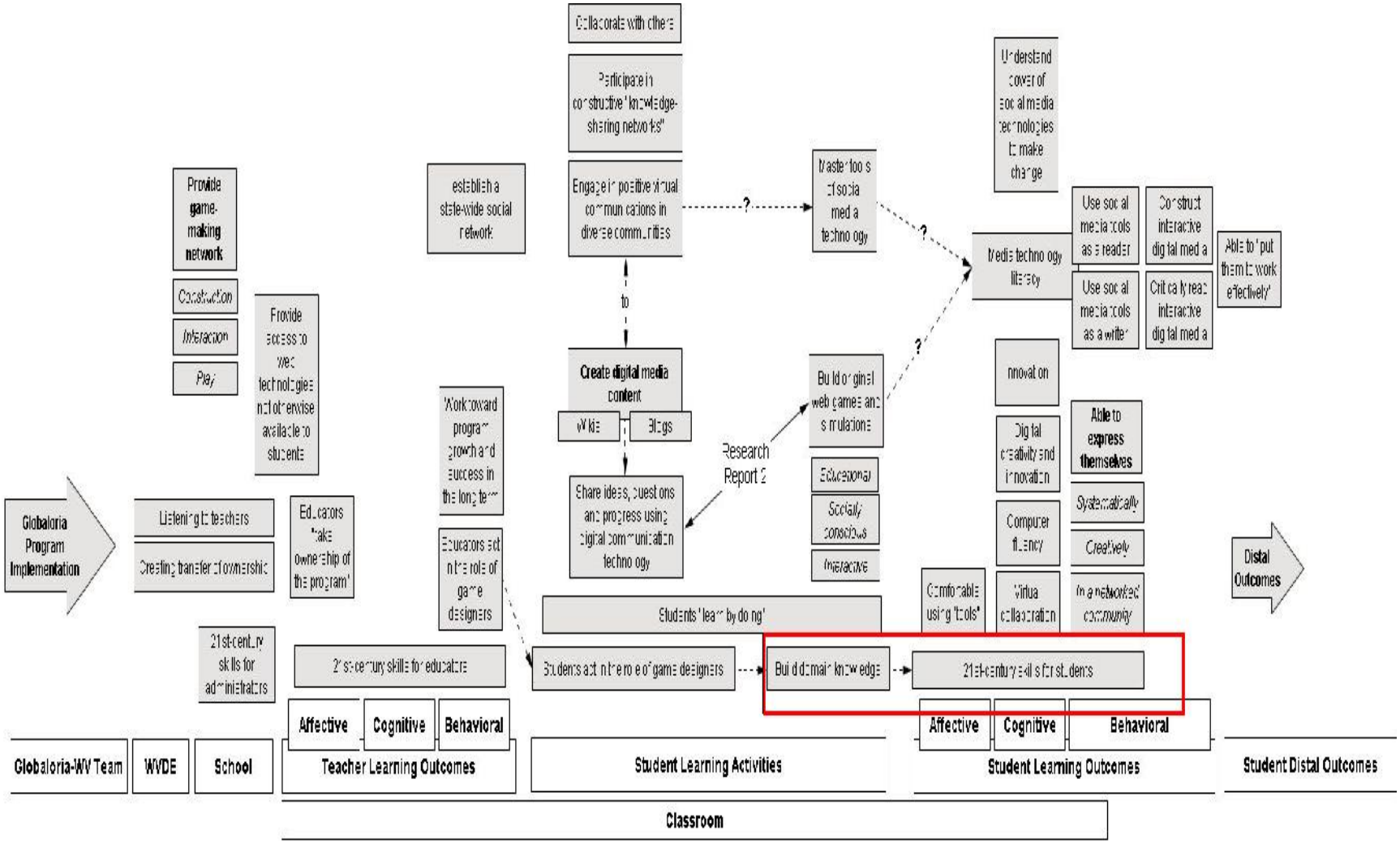
World Wide Workshop Foundation. (2012). *World Wide Workshop: Social media technology & learning*. New York, NY: World Wide Workshop Foundation.

Yeung, W. J., & Conley, D. (2008). Black-white achievement gap and family wealth. *Child Dev, 79*(2), 303-324.

Appendices

Appendix A
School-Level Globaloria Implementation and Learning Outcomes Logic Model

Appendix A: School-Level Globaloria Implementation and Learning Outcomes Logic Model



Appendix B
Propensity Score Matching Steps and Procedures

Appendix B: Propensity Score Matching Steps and Procedures

In the field of education and evaluation research, there is an increasing use of propensity score matching (PSM) methods to adjust a treatment effect based on the function of observed variables (i.e., covariates) in non-randomized observational studies (Stuart, 2010). PSM provides advantages in educational research where random assignment is not always feasible or ethical. In general, PSM consists of several analytic steps, including selecting covariates, estimating propensity score for matching, and diagnosing the matches (Stuart, 2010; Stuart & Rubin, 2007). Each step conducted for this report is described briefly.

Selection of Covariates

The logic behind PSM is that, in the absence of an experimental design, assignment to treatment is frequently nonrandom. Hence, units (e.g., schools, students, and teachers) receiving treatment and those without treatment may differ in many ways that may affect both participation (e.g., self-selection factors) and the outcome of interest (e.g., pretest scores or demographic attributes). To avoid the biased estimation of a treatment effect, PSM selects a comparison group that is similar to a treatment group based on observed covariates. Therefore, the main goal of PSM is to achieve balance on observed covariates through careful matching of the propensity score that is a function of observed covariates. In this sense, the selection of observed covariates is critical. In practice, three types of covariates commonly used in PSM are self-selection variables (e.g., individual preference or individual motivation), individual demographic information (e.g., gender, race, or income status), and individual pretest scores (e.g., pretest achievement data) (Stuart, 2012).

In this evaluation study, the matching involved two-stages. The first stage was to select a group of comparison schools matching the group of Globaloria participants. Four school-level covariates were used to select the matched comparison schools: (1) school enrollment; (2) percentage of low-income students (using free or reduced-price meal [FRPM] status as a proxy); (3) percentage of students achieving proficiency on the 2011 WESTEST2 reading subtest; and (4) percentage of students achieving proficiency on the 2011 WESTEST2 math subtest. The second stage was to select a group of comparison students from the identified comparison schools matching the group of Globaloria participating students. At this stage of matching, students' demographic information and pretest scores were used to estimate propensity scores for each individual, including grade level, gender, race, FRPM status, limited English proficiency (LEP) status, disability status, and the four 2011 WESTEST2 subtests scores. For both stages, matching was conducted separately by student grade level (i.e., high school versus middle school). This two-stage procedure ensures that the selected comparison schools were similar to the Globaloria participating schools, and students who were selected from the comparison schools were similar to the Globaloria participants prior to their participation in Globaloria.

Estimation of Propensity Scores

At each stage, matching was done using logistic regression to obtain a measure describing the "distance" between two individuals, so called propensity score. First, the logistic regression

with school-level data was conducted to select the matched comparison schools, followed by student-level matching. The school-level matching was conducted separately for middle schools and high schools using a one-on-one optimal matching algorithm. At the student-level, matching was conducted separately for middle schools and high schools using a greedy matching algorithm with a caliper of 0.15 to avoid the risk of bad matches (Guo & Fraser, 2010; Parsons, 2001). It should be noted that because PSM methods require full information to create propensity scores, multiple imputations were conducted at the student-level[17] to impute missing values before conducting the matching (Hill, 2004). For multiple imputations, student demographic data were used as predictors but were not imputed, whereas scores from the four subtests of the 2011 WESTEST2 (i.e., the pretest) were used as predictors and were imputed. A total of five sets of imputed data were obtained. The results of the imputed data were very similar in terms of means and standard deviations; hence, one set of imputed data from each imputed high school sample and middle school sample was randomly selected, and used for matching.

Balance Diagnosis

Once the comparison schools and students were selected, the final step was to check for the covariate balance in the matched groups, which is defined as "the similarity of the empirical distributions of the full set of covariates in the matched participant and comparison groups" (Stuart, 2010, p. 11). For balance diagnosis, graphical diagnostics examining the distribution of propensity score distributions were first conducted to assess common support (see Figure 2 and 3 in the main report); then, three numerical balance measures were used to check covariate balances at student level (Rubin, 2001):

- The ratio of the variances of the propensity score in the two groups must be close to 1.0. Rubin (2001) suggests that the variance ratios should be between 0.5 and 2.0.
- The difference in the means of the propensity scores in the two groups being compared must be small. Rubin (2001) suggests that the standardized differences of means should be less than 0.25.
- The percent of balance improvement, where the larger percentage of balance improvement indicates better results of the PSM.

The results of balance diagnosis are reported in the Methods section; hence, they are not repeated here.

[17] There were no missing data at the school-level; therefore missing value imputation was not necessary.

Appendix C
Participating and Matched Comparison School Selection Criteria by Middle and High Schools

Appendix C: Participating and Matched Comparison School Selection Criteria by Middle and High Schools

Match	School	District	School Enrollment	% of Students Passing Reading Proficiency	% of Students Passing Math Proficiency	% of Low-Income Students
Participating Middle Schools						
1	Chapmanville Middle School	Logan	574	45.50	35.27	56.40
2	Eastern Greenbrier Middle School	Greenbrier	865	48.58	43.59	51.00
3	Elkins Middle School	Randolph	662	51.21	39.02	54.50
4	Kasson Elementary/Middle School	Barbour	200	39.53	46.51	45.00
5	Lenore K-8 School	Mingo	561	39.66	43.57	68.60
6	Logan Middle School	Logan	796	46.55	37.70	62.10
7	Man Middle School	Logan	478	46.95	31.30	61.30
8	Sandy River Middle School	McDowell	272	41.57	49.10	72.10
9	South Charleston Middle School	Kanawha	414	63.01	53.31	49.80
10	Tygarts Valley Middle/High School*	Randolph	478	37.06	25.80	63.20
PARTICIPATING MIDDLE SCHOOL STUDENT SAMPLE MEAN			**530.00**	**45.96**	**40.52**	**58.40**
Comparison Middle Schools						
1	B-U Middle School	Upshur	820	47.57	39.97	52.30
2	Collins Middle School	Fayette	720	33.19	29.46	67.60
3	Edison Junior High School	Wood	685	60.36	48.09	51.80
4	Herndon Consolidated Grade School	Wyoming	222	64.34	60.00	61.30
5	Independence Junior High	Raleigh	520	52.48	40.35	51.20
6	Keyser Primary/Middle School	Mineral	1150	45.48	37.59	52.40
7	Marlinton Middle School	Pocahontas	232	37.93	42.24	66.40
8	Martinsburg North Middle School	Berkeley	542	50.39	35.01	66.80
9	Walton Elementary/Middle School	Roane	386	40.34	35.19	61.40
10	West Preston Middle School	Preston	220	40.46	42.79	45.50
COMPARISON MIDDLE SCHOOL STUDENT SAMPLE MEAN			**549.70**	**47.25**	**41.07**	**57.67**

Match	School	District	School Enrollment	% of Students Passing Reading Proficiency	% of Students Passing Math Proficiency	% of Low-Income Students
Participating High Schools						
1	Braxton County High School	Braxton	658	37.94	39.06	48.60
2	Buffalo High School	Putnam	274	53.05	61.21	44.90
3	Capital High School	Kanawha	1236	48.47	34.57	48.20
4	Chapmanville Senior High School	Logan	712	39.29	29.12	48.90
5	Doddridge County High School	Doddridge	374	47.08	34.43	48.40
6	Elkins High School	Randolph	869	51.67	46.00	42.50
7	Fayetteville High School	Fayette	510	38.79	35.41	52.00
8	George Washington High School	Kanawha	1115	75.52	66.58	17.90
9	Greenbrier East High School	Greenbrier	1129	51.09	46.60	49.20
10	Greenbrier West High School	Greenbrier	430	43.32	31.27	60.90
11	Hurricane High School	Putnam	1142	61.41	63.98	24.70
12	Lincoln High School	Harrison	633	43.96	32.80	43.40
13	Logan Senior High School	Logan	826	34.98	29.20	41.90
14	Man Senior High School	Logan	392	41.63	29.18	51.50
15	North Marion High School	Marion	813	45.56	43.71	48.20
16	Oak Glen High School	Hancock	616	47.19	52.24	41.20
17	Oak Hill High School	Fayette	771	40.37	33.33	49.70
18	Philip Barbour High School Complex	Barbour	773	34.21	37.40	54.90
19	Preston High School	Preston	1336	33.81	36.04	42.40
20	Riverside High School	Kanawha	1244	35.82	27.36	49.30
21	South Harrison High School	Harrison	422	41.77	37.97	39.60
22	Spring Valley High School	Wayne	1041	49.46	39.60	34.40
23	Tygarts Valley Middle/High School*	Randolph	478	37.06	25.80	63.20
24	University High School	Monongalia	1281	54.47	54.69	36.10
25	Webster County High School	Webster	483	36.76	32.94	56.70
26	Weir High School	Hancock	619	45.53	48.27	37.50
27	Wheeling Park High School	Ohio	1690	51.65	44.25	42.30
28	Woodrow Wilson High School	Raleigh	1381	48.19	42.94	44.90

Match	School	District	School Enrollment	% of Students Passing Reading Proficiency	% of Students Passing Math Proficiency	% of Low-Income Students
PARTICIPATING HIGH SCHOOL STUDENT SAMPLE MEAN			**830.29**	**45.36**	**40.57**	**45.12**
Comparison High Schools						
1	Berkeley Springs High School	Morgan	746	48.09	39.01	47.50
2	Brooke High School	Brooke	1108	42.76	47.61	37.10
3	Buckhannon Upshur High School	Upshur	1140	42.83	34.82	41.70
4	Cabell Midland High School	Cabell	1836	58.81	48.65	32.60
5	Cameron High School	Marshall	326	44.59	41.46	46.60
6	Clay County High School	Clay	626	39.64	39.64	57.00
7	Frankfort High School	Mineral	535	45.40	51.96	28.80
8	Gilmer County High School	Gilmer	426	45.81	37.77	48.60
9	Hampshire Senior High School	Hampshire	1110	32.65	25.70	50.60
10	Hedgesville High School	Berkeley	1691	46.88	38.50	34.80
11	Herbert Hoover High School	Kanawha	739	48.97	39.29	37.50
12	Hundred High School**	Wetzel	115	44.30	26.58	47.80
13	Liberty High School	Harrison	627	46.23	33.97	49.10
14	Lincoln County High School	Lincoln	891	31.83	23.51	61.10
15	Montcalm High School	Mercer	340	34.58	28.19	69.70
16	Nicholas County High School	Nicholas	792	42.39	37.50	39.30
17	Nitro High School	Kanawha	755	64.36	59.10	28.90
18	Pikeview High School	Mercer	708	37.08	39.72	51.40
19	Point Pleasant High School	Mason	1214	45.13	40.55	54.20
20	Princeton Senior High School	Mercer	1075	47.45	46.20	46.80
21	Richwood High School	Nicholas	409	47.38	33.10	56.50
22	Ritchie County High School	Ritchie	449	38.48	39.45	39.90
23	Scott High School	Boone	648	42.21	41.02	42.90
24	South Charleston High School	Kanawha	1036	48.02	38.81	42.40
25	Valley High School	Fayette	533	39.62	31.70	59.30
26	Washington High School	Jefferson	1110	47.82	45.93	22.40
27	Wayne High School	Wayne	598	38.28	36.34	46.20

Match	School	District	School Enrollment	% of Students Passing Reading Proficiency	% of Students Passing Math Proficiency	% of Low-Income Students
28	Winfield High School	Putnam	823	57.96	61.92	19.40
COMPARISON HIGH SCHOOL STUDENT SAMPLE MEAN			**800.21**	**44.63**	**39.57**	**44.29**

* Tygarts Valley Middle and High School had exactly the same school-level characteristics. When conducting HLM, the same school-levels were used.

**Hundred High School was selected at school-level matching; yet, at student-level matching, none of the students from this school was selected.

Appendix D
Hierarchical Linear Modeling: Impact Model

Appendix D: Hierarchical Linear Modeling: Impact Model

This appendix presents the HLM models and technical report for each student's WESTEST2 outcome. Equations below show the definition of the parameters for level-1 and level-2 models, following Raudenbush and Bryk's (2002) notation.

Level-1 Model

$$y_{ij} = \beta_{0j} + \beta_{1j}*(Gender_{ij}) + \beta_{2j}*(FRPM_{ij}) + \beta_{3j}*(Math2011\_ganc_{ij}) + \beta_{4j}*(Reading2011\_ganc_{ij}) + \beta_{5j}*(Science2011\_ganc_{ij}) + \beta_{6j}*(SocialStudies2011\_ganc_{ij}) + \Upsilon_{ij.}$$

Notations:

y_{ij}	Dependent variable (student WESTEST2 score) for case *i* in school *j*
β_{0j}	The intercept (school mean WESTEST2) in school *j*
$\beta_{1j} - \beta_{6j}$	Regression coefficient (slope) in school *j*
$\Upsilon_{ij.}$	Level 1 residual for case *i* in school *j*

Predictors coding and interpretation:

$Gender_{ij}$	Male = 1; Female = 0
$FRPM_{ij}$	Students with free or reduce-priced meal status = 1; Students without free or reduced-priced meal status = 0
$Math2011\_grc_{ij}$	2011 WESTEST2 math (Grand mean centered)
$Reading2011\_grc_{ij}$	2011 WESTEST2 reading (Grand mean centered)
$Science2011\_grc_{ij}$	2011 WESTEST2 science (Grand mean centered)
$SocialStudies2011\_grc_{ij}$	2011 WESTEST2 social studies (Grand mean centered)

Level-2 Model

$$\beta_{0j} = \Upsilon_{00} + \Upsilon_{01}*(Participation_j) + \Upsilon_{02}*(SchPoverty\_ganc_j) + \Upsilon_{03}*(SchReadProficiency\_ganc_j) + \Upsilon_{04}*(SchMathProficiency\_ganc_j) + \Upsilon_{05}*(Participation\ X\ SchPoverty\_ganc_j) + \Upsilon_{06}*(Participation_j\ X\ SchReadProficiency\_ganc_j) + \Upsilon_{07}*(\ Participation_j\ X\ SchMathProficiency\_ganc_j) + \upsilon_{0j}$$

$$\beta_{1j} = \Upsilon_{10} + \Upsilon_{11}*(Participation_j) + \upsilon_{1j}$$

$$\beta_{2j} = \Upsilon_{20} + \Upsilon_{21}*(Participation_j) + \upsilon_{2j}$$

$$\beta_{3j} = \Upsilon_{30}$$

$$\beta_{4j} = \Upsilon_{40}$$

$$\beta_{5j} = \Upsilon_{50}$$

$$\beta_{6j} = \Upsilon_{60}$$

$$\beta_{7j} = \Upsilon_{70}$$

Notations:

Υ_{00}	The population intercept (grand mean of WESTEST2 score)
$\Upsilon_{01} - \Upsilon_{07}$	The population regression coefficient for the regression of the dependent variable on the level 2 predictor
$\Upsilon_{10} - \Upsilon_{70}$	The population regression coefficient (mean slope) for the regression of the dependent variable on the level 1 predictor

Υ_{11} & Υ_{21} The population regression coefficient for the interaction between the level 1 and level 2 variables in predicting the dependent variable (i.e., cross-level interaction terms).

Υ_{0j} Level 2 residual for school *j*

Predictors coding and interpretation

$Participation_j$	Participating schools = 1; Non-participating schools = 0
$SchPoverty\_ganc_j$	Percentage of students receiving free or reduced-price meals (grand mean centered) within school *j*
$SchReadProficiency\_ganc_j$	Percentage of student scoring proficient or high in reading (grand mean centered) in school *j*
$SchMathProficiency\_ganc_j$	Percentage of student scoring proficient or high in math (grand mean centered) in school *j*
$Participation_j$ X $SchPoverty\_ganc_j$	Interaction term of participation and school poverty status (grand mean centered) in school *j*
$Participation_j$ X $SchReadProficiency\_ganc_j$	Interaction term of participation and school reading proficiency status (grand mean centered) in school *j*
$Participation_j$ X $SchMathProficiency\_ganc_j$	Interaction term of participation and school math proficiency status (grand mean centered) in school *j*

Tables D1 through D4 show the results of the two-level HLM examining the effect of Globaloria on student WESTEST2 achievement outcomes. Model 1 is the null model (i.e., unconditional model) showing the unconditional level-1 and level-2 variance. Intra-class correlation coefficients (ICCs) were calculated using information obtained from the null models. The greater the ICCs, the stronger the clustering effect. Model 2 is the final model that is reported in the Results section of this report.

Table D1. HLM Results for Two-Level Model Examining the Effect of Globaloria on Student WESTEST2 Math Outcomes

Parameter Estimate	Middle Schools						High Schools					
	Null Model			Final Model			Null Model			Final Model		
Fixed Effects	***β***	***SE***	***p***	***β***	***SE***	***p***	***β***	***SE***	***p***	***β***	***SE***	***p***
Within-School Effects												
Intercept (r_{00})	657.22	3.53	< 0.001	656.19	3.05	0.000	667.83	2.40	<0.001	666.03	5.46	0.000
Gender (r_{10})				7.11	3.74	0.058				7.51	4.14	0.070
FRPM (r_{20})				-1.89	3.29	0.564				-10.50	3.89	0.007
Math 2011(r_{30})				0.52	0.07	0.000				0.34	0.04	0.000
Reading 2011 (r_{40})				0.10	0.07	0.147				0.12	0.04	0.001
Science 2011 (r_{50})				0.29	0.07	0.000				0.24	0.05	0.000
Social Studies 2011 (r_{60})				0.05	0.07	0.451				0.12	0.07	0.091
Between-School Effects												
Participation (r_{01})				-0.82	3.96	0.839				3.92	6.69	0.560
SchPoverty (r_{02})				0.17	0.10	0.133				0.39	0.33	0.250
SchReadProficiency (r_{03})				0.18	0.14	0.225				-0.46	0.44	0.294
SchMathProficiency (r_{04})				0.09	0.18	0.636				1.04	0.59	0.084
Interaction Effects												
Participation x Gender Slope (r_{11})				-2.49	4.80	0.604				-8.43	5.54	0.128
Participation x FRPM Slope (r_{21})				0.52	4.53	0.909				7.60	5.71	0.184
Participation x SchPoverty (r_{05})				-0.06	0.23	0.798				-0.54	0.40	0.178
Participation x SchReadProficiency (r_{06})				-0.15	0.30	0.611				0.71	0.49	0.159
Participation x SchMathProficiency (r_{07})				-0.18	0.27	0.513				-1.38	0.63	0.034

Parameter Estimate	Middle Schools				High Schools			
	Null Model		Final Model		Null Model		Final Model	
Random Effects	**Variance**	***p***	**Variance**	***p***	**Variance**	***p***	**Variance**	***p***
Between school variation (τ_{00})	161.57		0.072	>0.500	105.1	<0.001	15.47	0.04
Within school variation (σ_{00})	1488.44		675.074		2431.54		1234.68	
No. of parameters	3		18		3		18.00	
Deviance (FIML)	4587.21		4208.73		7914.17		7396.03	
Variance Explained								
Intra-Class Correlation Coefficient (ICC)	0.10				0.04			
Within school variation explained (level-1 R^2)	0.55				0.49			
Between school variation explained (level-2 R^2)	1.00				0.85			

Note. FIML (Full Information Maximum Likelihood) robust estimations are reported. The calculation of variance explained is based on Raudenbush & Bryk's (2002) version.

Table D2. HLM Results for Two-Level Model Examining the Effect of Globaloria on Student WESTEST2 Reading Outcomes

Parameter Estimate	Middle Schools						High Schools					
	Null Model			Final Model			Null Model			Final Model		
Fixed Effects	*β*	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*	*β*	*SE*	*p*
Within-School Effects												
Intercept (r_{00})	495.49	2.7	<0.001	500.11	2.85	0.000	496.08	3.04	<0.001	506.94	3.63	0.000
Gender (r_{10})				-7.51	2.86	0.009				-12.48	4.92	0.012
FRPM (r_{20})				-1.52	1.85	0.410				-7.99	4.03	0.047
Math 2011 (r_{30})				0.09	0.02	0.000				0.15	0.04	0.001
Reading 2011 (r_{40})				0.32	0.05	0.000				0.44	0.07	0.000
Science 2011 (r_{50})				0.24	0.05	0.000				0.05	0.03	0.157
Social Studies 2011 (r_{60})				0.16	0.06	0.008				0.33	0.08	0.000
Between-School Effects												
Participation (r_{01})				3.51	3.47	0.331				-0.81	4.51	0.859
SchPoverty (r_{02})				-0.21	0.16	0.219				0.48	0.25	0.062
SchReadProficiency (r_{03})				0.60	0.19	0.009				0.75	0.58	0.202
SchMathProficiency (r_{04})				-0.74	0.23	0.008				-0.26	0.44	0.560
Interaction Effects												
Participation x Gender Slope (r_{11})				-4.21	4.55	0.355				2.88	6.21	0.642
Participation x FRPM Slope (r_{21})				2.60	3.53	0.463				8.45	6.15	0.171
Participation x SchPoverty (r_{05})				0.34	0.36	0.365				-0.74	0.53	0.173
Participation x SchReadProficiency (r_{06})				-0.09	0.45	0.852				-0.13	0.66	0.838
Participation x SchMathProficiency (r_{07})				0.43	0.34	0.230				-0.77	0.77	0.327

Parameter Estimate	Middle Schools						High Schools					
	Null Model			Final Model			Null Model			Final Model		
Random Effects	**Variance**	***p***		**Variance**	***p***		**Variance**	***p***		**Variance**	***p***	
Between school variation (τ_{00})	86.65	<0.001		0.23	0.177		257.1	<0.001		26.26	0.01	
Within school variation (σ_{00})	922.26			474.37			2671.91			1191.59		
No. of parameters	3			18			3			18.00		
Deviance (FIML)	4370.00			4050.16			8002.5			7374.96		
Variance Explained												
Intra-Class Correlation Coefficient (ICC)	0.09						0.09					
Within school variation explained (level-1 R^2)	0.49						0.55					
Between school variation explained (level-2 R^2)	1.00						0.90					

Note. FIML robust estimations are reported. The calculation of variance explained is based on Raudenbush & Bryk's (2002) version.

Table D3. HLM Results for Two-Level Model Examining the Effect of Globaloria on Student WESTEST2 Science Outcomes

Parameter Estimate	Middle Schools						High Schools					
	Null Model			Final Model			Null Model			Final Model		
Fixed Effects	***B***	***SE***	***p***	***B***	***SE***	***p***	***B***	***SE***	***p***	***B***	***SE***	***p***
Within-School Effects												
Intercept (r_{00})	597.98	3.07	<0.001	598.72	2.36	0.000	636.65	2.69	<0.001	635.70	5.59	0.000
Gender (r_{10})				0.17	0.05	0.000				0.33	0.06	0.000
FRPM (r_{20})				3.76	2.47	0.128				4.84	5.76	0.401
Math 2011 (r_{30})				-2.48	2.15	0.250				-15.27	5.26	0.004
Reading 2011 (r_{40})				0.13	0.04	0.002				0.00	0.07	0.977
Science 2011 (r_{50})				0.20	0.05	0.000				0.22	0.05	0.000
Social Studies 2011 (r_{60})				0.34	0.06	0.000				0.50	0.10	0.000
Between-School Effects												
Participation (r_{01})				0.92	3.49	0.797				7.35	6.98	0.298
SchPoverty (r_{02})				0.18	0.19	0.364				0.10	0.41	0.810
SchReadProficiency (r_{03})				0.45	0.22	0.068				-0.13	0.47	0.786
SchMathProficiency (r_{04})				-0.21	0.37	0.572				-0.05	0.59	0.928
Interaction Effects												
Participation x Gender Slope (r_{11})				-1.82	3.52	0.605				-9.18	7.89	0.245
Participation x FRPM Slope (r_{21})				-2.14	2.94	0.467				16.94	8.21	0.039
Participation x SchPoverty (r_{05})				-0.10	0.22	0.660				-1.07	0.62	0.090
Participation x SchReadProficiency (r_{06})				-0.68	0.34	0.068				-1.06	0.72	0.147
Participation x SchMathProficiency (r_{07})				0.13	0.39	0.751				-0.01	0.67	0.994

Parameter Estimate	Middle Schools						High Schools					
	Null Model			Final Model			Null Model			Final Model		
Random Effects	**Variance**	***p***		**Variance**	***p***		**Variance**	***p***		**Variance**	***p***	
Between school variation (τ_{00})	137.26	<0.001		0.15	0.259		105.81	<0.001		13.05	0.063	
Within school variation (σ_{00})	821.21			424.03			2940.84			1513.21		
No. of parameters	3			18			3			18.00		
Deviance (FIML)	4325.55			3999.59			6902.25			6466.74		
Variance Explained												
Intra-Class Correlation Coefficient (ICC)	0.14						0.03					
Within school variation explained (level-1 R^2)	0.48						0.49					
Between school variation explained (level-2 R^2)	1.00						0.88					

Note. FIML robust estimations are reported. The calculation of variance explained is based on Raudenbush & Bryk's (2002) version.

Table D4. HLM Results for Two-Level Model Examining the Effect of Globaloria on Student WESTEST2 Social Studies Outcomes

Parameter Estimate	Middle Schools						High Schools					
	Null Model			Final Model			Null Model			Final Model		
Fixed Effects	***B***	***SE***	***p***	***B***	***SE***	***p***	***B***	***SE***	***p***	***B***	***SE***	***p***
Within-School Effects												
Intercept (r_{00})	418.35	2.76	<0.001	421.82	1.64	0.000	412.31	1.78	<0.001	410.10	2.63	0.000
Gender (r_{10})				-4.00	2.18	0.067				1.89	2.76	0.495
FRPM (r_{20})				-4.87	1.76	0.006				-2.89	3.09	0.349
Math 2011 (r_{30})				0.10	0.04	0.013				0.08	0.03	0.002
Reading 2011 (r_{40})				0.15	0.03	0.000				0.12	0.04	0.002
Science 2011 (r_{50})				0.17	0.07	0.016				0.07	0.02	0.000
Social Studies 2011 (r_{60})				0.34	0.06	0.000				0.50	0.04	0.000
Between-School Effects												
Participation (r_{01})				-5.22	2.47	0.056				0.83	3.42	0.809
SchPoverty (r_{02})				0.09	0.14	0.534				0.07	0.21	0.737
SchReadProficiency (r_{03})				0.11	0.12	0.363				-0.08	0.30	0.782
SchMathProficiency (r_{04})				0.24	0.23	0.302				0.18	0.28	0.528
Interaction Effects												
Participation x Gender Slope (r_{11})				14.24	3.59	0.000				2.00	3.72	0.591
Participation x FRPM Slope (r_{21})				2.65	2.98	0.374				5.61	3.94	0.156
Participation x SchPoverty (r_{05})				0.03	0.25	0.915				-0.15	0.31	0.618
Participation x SchReadProficiency (r_{06})				-0.08	0.34	0.809				0.05	0.38	0.896
Participation x SchMathProficiency (r_{07})				-0.04	0.32	0.893				-0.20	0.35	0.562

Parameter Estimate	Middle Schools						High Schools					
	Null Model			Final Model			Null Model			Final Model		
Random Effects	**Variance**	***p***		**Variance**	***p***		**Variance**	***p***		**Variance**	***p***	
Between school variation (τ_{00})	109.25	<0.001		4.27	0.01		62.74			3.05	0.01	
Within school variation (σ_{00})	704.27			365.67			1260.99			575.10		
No. of parameters	3			18			3			18.00		
Deviance (FIML)	4255.34			3937.30			7417.79			6813.45		
Variance Explained												
Intra-Class Correlation Coefficient (ICC)	0.13						0.05					
Within school variation explained (level-1 R^2)	0.48						0.54					
Between school variation explained (level-2 R^2)	0.96						0.95					

Note. FIML robust estimations are reported. The calculation of variance explained is based on Raudenbush & Bryk's (2002) version.