A new report compares the performance of Florida Virtual School (FLVS) students with students in traditional brick-and-mortar schools and concludes the FLVS students perform about the same or somewhat better on state tests and at a lower cost. The report claims to be the first empirical study of K-12 student performance in virtual education. This is not correct, and the report in fact confirms the findings and repeats the methodological flaws and limitations of previous research. The report’s findings fail to account for the potential bias of student selectivity in the FLVS sample, the potential impact of regression effects, differential mortality in the two groups, and the fact that the virtual environment is simply a delivery medium. Given the limitations of research such as this new study, researchers have moved beyond simply investigating whether one medium is better than the other and begun—and need to continue—investigating under what conditions K-12 online and blended learning can be effectively designed, delivered, and supported.
Kevin Welner  
*Project Director*

William Mathis  
*Managing Director*

Erik Gunn  
*Managing Editor*

---

**National Education Policy Center**

School of Education, University of Colorado  
Boulder, CO 80309-0249  
Telephone: (802) 383-0058

Email: NEPC@colorado.edu  
http://nepc.colorado.edu

**Publishing Director:** Alex Molnar

---

This is one of a series of Think Twice think tank reviews made possible in part by funding from the Great Lakes Center for Education Research and Practice. It is also available at http://greatlakescenter.org.

This material is provided free of cost to NEPC's readers, who may make non-commercial use of the material as long as NEPC and its author(s) are credited as the source. For inquiries about commercial use, please contact NEPC at nepc@colorado.edu.
I. Introduction

K-12 online learning (also referred to as virtual schooling or cyber schooling) has been in existence since at least 1991. One of the earliest supplemental online learning options for K-12 students was the Florida Virtual School (FLVS), which began operating during the 1996-97 school year. Each year more states open up their educational markets to an increasing number of online options, and Florida currently has one of the most open markets for K-12 online learning providers. While the practice of K-12 online learning is more than two decades old, the empirical research into K-12 online learning is still developing. However, one of the few areas studied extensively has been the comparison of student performance in the K-12 online learning environment with that of students enrolled in brick-and-mortar settings.

This framework was the state of the practice of and research into K-12 online learning in the United States and the State of Florida at the time the Harvard Kennedy School’s “Program on Education Policy and Governance Working Paper” released the report reviewed here, Virtual Schooling and Student Learning: Evidence from the Florida Virtual School. The purpose of this report was to examine two potential goals of virtual education:

First, virtual education can increase access to education by enabling students to take courses that are not offered in their local school or that they cannot attend due to enrollment constraints or scheduling conflicts. Second, virtual education might improve the quality of education through personalization, competition resulting from increased choice among providers, and other channels. Even if virtual schools are no better than traditional schools, they may offer opportunities to increase productivity in education by operating at a lower cost (p. 3).

II. Findings and Conclusions of the Report

In the introduction, after setting out these potential goals, the report indicated “virtual schools meet the first goal, almost by definition, in that they provide a variety of courses
that students can take from anywhere and at any time” (p. 3). This is the extent of their exploration of this first goal. Later, in the conclusion, the report indicated that “in 2008-09 (the most recent year of our linked data), at least 1,384 AP [Advanced Placement] courses (916 unique students) were taken by students enrolled in high schools where those courses were not offered” (p. 13). There was no additional examination of course offerings in small, rural, or inner city schools beyond this statement. There was no exploration as to why these AP courses were not offered by these schools (e.g., low student interest/enrollment, inability to hire a qualified teacher, classroom space issues, cheaper option for AP offerings, etc.). The reason why the school did not offer the particular AP course could help explain whether virtual education did provide access to a course that would not have otherwise been offered in the student’s brick-and-mortar school.

As these two statements are the only references made to this first goal in the report, I will focus this review primarily on the report’s coverage of the second potential goal of virtual education: improved quality.

On that question, the report indicates that virtual students “performed about the same or somewhat better on state tests once their pre-high-school characteristics [were] taken into account” (p. 1). Based on this finding, the report concludes

\[\ldots\] that fears of reductions in the quality of education are misplaced. We do not find any evidence of negative effects of virtual education on student learning, and a finding of equivalent quality, on average, between FLVS and non-FLVS courses may suggest a higher level of productivity in the FLVS courses (p. 14).

The report also explores the characteristics of students enrolled in FLVS courses and finds FLVS students were “less likely to be eligible for a free or reduced-price lunch \ldots less likely to be in special education programs, and \ldots more likely to be white” (p. 10). Further, the report indicates that FLVS students performed better on the grade 8 state math and reading test and had fewer absences from school.

Finally, the report makes the claim that it represents “the first estimates of the effect of taking virtual courses by comparing the achievement of students in two traditional high school courses (algebra and English) to the achievement of students enrolled in the same traditional schools but who took one or both of these courses online” (p. 3).

III. The Report’s Rationale for Its Findings and Conclusions

The fundamental rationale is that grade 10 state standardized test scores represent an adequate measure of educational quality for virtual and traditional schools. The report uses Florida Department of Education’s PK-20 Education Data Warehouse, which

http://nepc.colorado.edu/thinktank/review-virtual-schooling-and-student-learning
“contained information on all Florida students attending public schools in grades 3 to 10 from the 2000–01 through 2008–09 school years” (p. 7). The report uses data related to “the school each student attends and its location; student characteristics such as ethnicity, gender, special education classification, and free lunch status; and annual measures of absences and state math and reading test scores” (p. 7).

The comparison of student performance relied on the first-time test results from the grade 10 Florida Comprehensive Assessment Test (FCAT) in Algebra I (DOE number: 1200310) and English I (DOE number: 1001310). The report uses the first-time test results to exclude those students who may not have had success on the test on the first instance and chose to repeat the course and the assessment (which often occurs when a student fails a course in the classroom, and then repeats the course in a virtual format). Each student outcome was calculated using a student’s math score, reading score, and number of days absent. In both the treatment and the control groups, the outcomes were controlled for demographic factors and for previous student performance by using the student’s results from the grade 8 state test.

Finally, the report used the per-pupil funding at FLVS and non-FLVS schools for the period of 2008-09 to 2012-13 to determine “that FLVS was producing similar outcomes at a lower cost” (p. 14).

IV. The Report’s Use of Research Literature

The report almost completely ignores the wealth of literature on K–12 online learning. The report begins with a series of statistics about the current practice of K–12 online learning in the United States, relying primarily upon two of the Keeping Pace with K–12 Online and Blended Learning reports. This is the only reference to K–12 online learning research contained in the report (although there is also one reference to an advocacy book that devotes significant space to K–12 online learning and one reference to an opinion-based news article focused on virtual education by one of the report’s authors).

In the Introduction, the report does outline a small selection of the research into the effect on educational outcomes of computer use at home, computer use in the school, technology-based interventions, and educational software. In discussing this research, the report does indicate that while several students may have found correlations between these conditions and educational outcomes, “it is not clear if this relationship is causal” (p. 2). This statement ignores the dominant perspective in educational technology research in general that technology is simply a delivery tool or medium and that educational outcomes are affected by changes in the design, delivery and support of the instruction provided.

The report goes on to state that “there is no existing high-quality research on the impact of fully online high school courses on student achievement in the U.S.” (p. 3). In fact, there is a wealth of research into student performance in the virtual school environment. Among the studies ignored by this report are three meta-analyses that have been conducted into
K-12 distance education. Two of these meta-analyses found positive effect sizes in favor of the online student cohort, while one found a negative effect size for the online student cohort. These meta-analyses are consistent with other research studies in the field. In 2004-05, the North Central Regional Laboratory sponsored eight quantitative studies to answer questions about online learning and to promote the growth of effective programs and practices. Seven of the eight studies included measures of student academic performance. For example, Ferdig, DiPietro, and Papanastasiou compared student performance of online students with 410 face-to-face students in five courses in Wisconsin. The authors reported “no significant differences between overall results of combined face-to-face versus combined online achievement scores, when the final course grades and assessments were analyzed by content area, higher scores for the online students were realized.” Similarly, Kleiman, Carey, Bonifaz and Haistead examined student performance in an online algebra course in Louisiana and found that students in the “online Algebra I course demonstrated, on average, at least as much or somewhat more achievement gain than students in comparison classrooms.” In fact, the vast majority of studies comparing student performance between online students and a face-to-face control group have found that the online students have done as well as or slightly better than their brick-and-mortar counterparts.

This report also ignores the wealth of literature that would have confirmed their finding that the “FLVS students [were] a more advantaged group” (p. 10). Initial evaluations of the early supplemental online learning program Virtual High School (now VHS Collaborative) found that the “vast majority of VHS students in their courses were planning to attend a four-year college” and that “VHS courses are predominantly designated as ‘honors,’ and students enrolled are mostly college bound.” Further, Haughey and Muirhead wrote that preferred characteristics of virtual school students included being highly motivated, self-directed, self-disciplined, independent learners who could read and write well and also had a strong interest in or ability with technology. Roblyer and Elbaum found that only students with a high need to control and structure their own learning are likely to choose distance formats freely. Finally, Clark and his colleagues found virtual school students were highly motivated, high achieving, self-directed, or liked to work independently. Much of this research is not new, and could have easily been incorporated into the report.

The report even ignores previous research conducted with the FLVS that addressed each of the areas it examined. As early as the 1999-2000 school year, there have been public evaluations of FLVS that have found high rate of participation and success in its AP courses, as well as a high student performance in all of its online courses. Further, Cavanaugh and her colleagues found FLVS students performed better on a non-mandatory assessment tool compared with students from the traditional classroom. They speculated that this was likely due to the fact that the virtual school students who voluntarily took the assessment may have been more academically motivated and naturally higher-achieving students. More recently, Johnston and Barbour conducted an analysis of the 2009, 2010, and 2011 AP exams to compare FLVS student performance against brick-and-mortar students in the state. In each year, the percentage of FLVS students scoring a 3 or higher
increased (i.e., 51%, 53%, and 55%), while the percentage of face-to-face students doing so actually decreased (i.e., 45%, 43%, and 43%). The performance gap between the two cohorts actually widened (i.e., 6%, 10% and 12% in favor of the FLVS cohort).

Finally, in a study comparable to the one presented in this report, the Florida Tax Watch Center for Educational Performance and Accountability conducted an analysis of the academic performance of students enrolled in the FLVS and cost efficiency of the FLVS program compared with face-to-face counterparts. The authors used “data provided by FLVS includ[ing] student profiles, enrollment, final grades, and AP exam scores,” as well as “student results from the . . . (FCAT), FLVS students’ final grades from traditional courses, grade point average data, and financial allocations” provided by the Florida Department of Education. The Florida Tax Watch analysis came to many of the same conclusions presented in the report reviewed here. For example, “minority populations at traditional schools were significantly higher than were evidenced at FLVS,” and “students at FLVS represented various levels of academic proficiency but reflected a more narrowed student population than traditional sites.” Further, FLVS students generally scored as well or better than their face-to-face counterparts:

- “FLVS students earned higher grades in their online courses than they had earned in courses in that same subject area in the traditional public school setting.”
- In both 2004-05 and 2005-06, “FLVS students outperformed their traditional school counterparts in nine out of ten subject areas.” (The exception both years was art/visual arts.)
- “FLVS students consistently outperformed their public school counterparts on both the Reading and Mathematics FCAT Sections.”
- “FLVS students consistently outperformed their public school counterparts on Advanced Placement Examinations.”

Like the current report, the Florida Tax Watch Center for Educational Performance and Accountability determined that the FLVS was able to achieve this high level of performance at a cheaper cost.

Simply put, the report reviewed here ignores all of the empirical research on K-12 online learning in general—and on the Florida Virtual School specifically—related to the characteristics of online students, student performance in the online environment, and the cost of online learning.
V. Review of the Report’s Methods

Many of the limitations to the methods used are outlined in the report itself:

First, although our study encompasses thousands of students at the largest state virtual school in the country, the last year of our linked data is 2008-09. It is possible that the quality of FLVS courses relative to the courses in traditional public schools has changed since then. Second, we are only able to compare student performance in two courses (Algebra I and English I), which make up just 8.7 percent of all FLVS course enrollments. Third, it is possible that our results are biased by unmeasured characteristics of students who choose to take these courses through FLVS versus at their local public school or by differences in teacher quality that lead students to take some courses through FLVS but not others. Finally, we are unable to measure any competitive effects that the availability of FLVS courses has on the quality of courses at traditional public schools (p. 14).

Further, like most virtual schools, the FLVS utilizes a trial period where students can drop a course without ever being recorded as officially registered. Trial periods, when a virtual program has one, can range from a single day or two to more than a hundred days. The FLVS has a 28-day trial period, and the report acknowledges that the data analysts “dropped from [the] analysis the course records for which the student decided not to continue past the trial period (about 23 percent of the original sample)” (p. 6). We know from other research that students who drop out tend to be those who do not possess the selective characteristics described earlier.

While the report acknowledges these limitations of their methods and sample, it does not dampen the authors’ enthusiasm for their results:

Despite these limitations, this analysis yields important new findings on virtual education, a topic that has generated much hype but little serious evidence. The results are mixed regarding the promise of technology to increase the quality of education through personalization (as of 2009), but they do strongly suggest that fears of reductions in the quality of education are misplaced. We do not find any evidence of negative effects of virtual education on student learning, and a finding of equivalent quality, on average, between FLVS and non-FLVS courses may suggest a higher level of productivity in the FLVS courses (p. 14).

Despite the acknowledged limitations, the report still makes definitive pronouncements about the success of virtual education—both in terms of student performance and cost effectiveness.

VI. Review of the Validity of the Findings and Conclusions

There are three main issues concerning the validity of the findings and conclusions in this report. The first issue is the inability to control for the reasons why FLVS students enrolled
in their virtual course. Many of these students chose virtual education because there were circumstances preventing them from being successful in that particular course in their brick-and-mortar school. Research has indicated a range of reasons for this decision, such as the course not being offered, a conflict in the student’s timetable, a conflict between the student and the face-to-face teacher, the student being bullied in school, specific learning disabilities or preferences, a lack of success in the past, or numerous other reasons or even a combination of several of those reasons. The issue arises when the report claims that improved educational outcomes are the result of the student being enrolled in a virtual environment, when they may simply be due to a lessening of the circumstances that caused the student to leave the traditional setting in the first place. For example, if a student being bullied in a brick-and-mortar school transfers to a cyber school, any improved performance may be completely divorced from the technology or delivery method, but rather could be attributable to the fact the student is no longer being bullied. While that is a benefit of virtual education, it wasn’t what the authors argued or were even researching.

Essentially, the control group and the treatment group are two non-randomly constituted groups, one of which is set to regress upward (i.e., from their poor motivation or attention or performance or whatever circumstance may have led them to select the virtual environment), and students in the control group, who are not on the same trajectory. In the language of experimental design, this source of internal invalidity is labeled a Regression-by-Selection Invalidity.

The second issue is the differential mortality in the two groups compared. The report indicates that almost a quarter of the potential FLVS sample dropped out during the trial period. We know that students who are retained and who persevere in the virtual environment are highly motivated, self-directed, high-achieving, independent learners. The report controls for this potential bias using standardized exam results from grade 8—a full two years prior to the data used in the study. The difference between the development of a 14-year-old and a 16-year-old—particularly in the areas of motivation, self-directedness, independence, and autonomy—can be significant. The report is unable to control for the development of these skills, and research suggests that this has the potential to skew the results in favor of the FLVS cohort.

The third issue, as described earlier, is that the report is unable to attribute any gains made by FLVS students to the fact that they were engaged in virtual education. The online learning model utilized by the FLVS was simply a delivery medium. It would be like comparing a lecture that uses a chalkboard with a lecture that uses an electronic whiteboard. At the end of the day, students are still being lectured to with the aid of some visual displays. The report fails to consider whether the virtual environment changed how the instruction was designed, delivered, or supported. These are all factors that likely confounded the report’s findings and conclusions.

These issues with the validity of the findings and conclusions in the report call into question the findings and conclusions not only in this report, but in the vast majority of studies comparing student performance in the online environment with that in the face-to-face classroom.

http://nepc.colorado.edu/thinktank/review-virtual-schooling-and-student-learning
VII. Usefulness of the Report for Guidance of Policy and Practice

While not the first systematic investigation of student performance in virtual education (or even in the FLVS), this report does represent a reasonable attempt to examine how students perform in an online environment compared with their brick-and-mortar counterparts. The findings are consistent with a long history of empirical studies into this area. In fact, the findings in this report closely replicate the earlier study conducted by the Florida Tax Watch Center for Educational Performance and Accountability.

The problem with this line of inquiry, however, is that it continues to ignore the lessons learned from the broader field of educational technology research. Online learning is simply the delivery medium; gains or losses in student learning derive from how that face-to-face or online learning is designed, delivered, and supported. This line of inquiry also ignores the reality that online and blended learning continue to grow. Investigating whether one medium is better than the other serves little practical or academic purpose.

Within the field there has been a push for two separate, yet equally important lines of inquiry. The first line of inquiry is to explore under what conditions K-12 online and blended learning can be effective—that is, to explore what constitutes effective design, delivery and support of K-12 online and blended learning. The second line of inquiry focuses on researchers, in collaboration with individual K-12 online learning programs, identifying specific challenges that can be answered using a design-based research methodology. This approach has the potential to provide those individual K-12 online learning programs with data-driven solutions to real problems they are experiencing. Unfortunately, the report being reviewed here does not focus on either of these lines of inquiry.
Notes and References


11. The seminal piece in media comparison studies is:

In this article, Clark compares media or technology to a delivery truck and asserts that technology impacts learning as much as the truck impacts the nutritional value of the groceries that it carries.


Cavanaugh (2001) had an effect size of +0.147 and Means et al. (2009) had an effect size of +0.24 in favor of the online student cohort, while Cavanaugh et al. (2004) had an effect size of -0.028 for the online student cohort.


Smith, Clark, & Blomeyer (2005), 32.

Smith, Clark, & Blomeyer (2005), 39.


See http://thevhscollaborative.org/ for additional information on this program.

http://nepc.colorado.edu/thinktank/review-virtual-schooling-and-student-learning


This evaluation found that 79% of students earned an A or a B, and only 4% of students earned an F. The evaluation also found that 69% of students scored a 3 or better on the AP exam.


FLVS per student funding was $4,287 in 2003-2004, compared with $4,571 for face-to-face students.

FLVS per student funding was $4,496 in 2004-2005, compared with $5,238 for face-to-face students.

FLVS per student funding was $4,755 in 2005-2006, compared with $5,629 for face-to-face students.

FLVS per student funding was $5,243 in 2006-2007, compared with $6,291 for face-to-face students.

See notes 15-19.


DOCUMENT REVIEWED: Virtual Schooling and Student Learning: Evidence from the Florida Virtual School

AUTHORS: Matthew M. Chingos and Guido Schwerdt

PUBLISHER/THINK TANK: Program on Education Policy and Governance, Harvard Kennedy School of Government

DOCUMENT RELEASE DATE: September 2014

REVIEW DATE: October 28, 2014

REVIEWER: Michael K. Barbour, Sacred Heart University

E-MAIL ADDRESS: mkbarbour@gmail.com

PHONE NUMBER: (203) 997-6330

SUGGESTED CITATION: