IES Institute of Education Sciences

APPENDIXES
January 2021

Regional Educational Laboratory Southeast
At Florida State University

Exploring Teachers' Influence on Student Success in an Online Biology Course

See https://go.usa.gov/xAcUD for the full report.

Appendix A. Methods

This appendix provides details about the main models for analyzing segment completion rates, end-of-segment exam scores, and time to segment completion. Appendix B includes the findings of the alternative models and results for the sensitivity checks to determine whether the results change for different models.

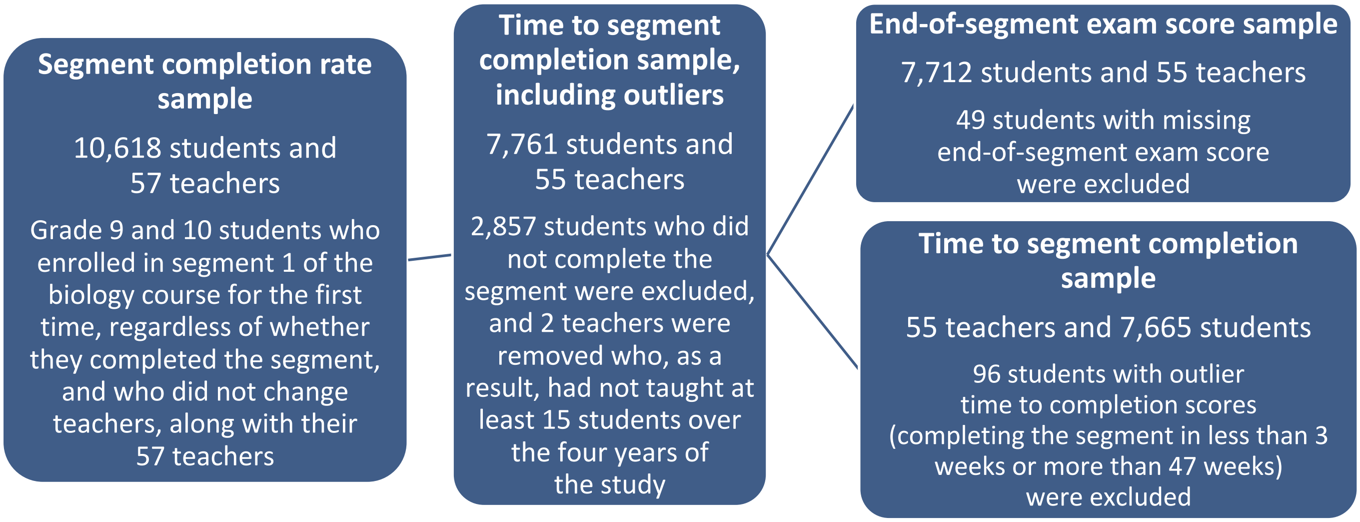
Study sample

The analysis focused on students in 2014/15–2017/18 who were taking biology for the first time, who enrolled in segment 1 of the two-segment online course, and who did not change teachers.[1] The study sample included four years of data (2014/15–2017/18) to reduce random error in the teacher-level estimates and to balance out any exceptionally positive or exceptionally negative scores for a teacher that might be observed in a single year by random error. The study focused on first-time students in segment 1 because including only first-time students would make the results more generalizable to "typical" online students. This criterion excluded students who retook courses in order to earn a higher grade or to recover credit after failing the course.

The analytical sample for research question 1 on variation in students' segment completion rates across teachers included the 10,618 students in grades 9 and 10 who enrolled in the biology segment 1 for the first time in 2014/15–2017/18 (regardless of whether they completed the segment) and did not change teachers, along with their 57 teachers (figure A1). The analytical sample for research question 2 on the end-of-segment exam score consisted of the 7,712 students who completed the segment and had nonmissing end-of-segment exam scores, along with their 55 teachers. The analytical sample for research question 3 on time to segment completion consisted of the 7,665 students who completed the segment in more than 3 weeks and less than 47 weeks, along with their 55 teachers. Based on Florida Virtual School guidance, students who completed the segment in less than 3 weeks or more than 47 weeks were excluded from the sample because, according to school staff, these cases could have been a result of data entry errors or other unusual circumstances. However, excluding these students did not affect the results of the analyses (see table B3 in appendix B). To ensure reliable estimates of teachers' potential influence on student outcomes, each analytical sample included only teachers who had at least 15 students in that sample combined across four years.

[1] Five percent of the students changed teachers in segment 1.

Figure A1. Structure of the analytical samples, 2014/15–2017/18



Note: Using multiple years of data reduces the random error of the teacher-level estimates and balances out exceptionally positive and exceptionally negative scores for individual teachers that might be observed in a single year, driven by random error. Each analytical sample included only teachers who had at least 15 students combined across four years. The same teachers remained in the end-of-segment exam score and time to segment completion samples.
Source: Authors' creation.

The study team also conducted secondary analyses of students who completed both segments 1 and 2 to determine whether the results might differ for students who spend a longer time in the course with the same teacher. These results are reported in appendix B after the findings of the main models and alternative models focusing on segment 1.

Main models for analyzing segment completion rates, end-of-segment exam scores, and time to completion

The study team used two-level hierarchical linear models that nested, or organized, students by teachers to examine teacher influence on students' segment completion rates, end-of-segment exam scores, and time to completion in Florida Virtual School online biology courses over 2014/15–2017/18. The analyses were conducted using HLM 7 software (Raudenbush et al., 2011).

The general form for these models is provided in the following equations. Student-level and teacher-level outcomes are defined as follows:

$$\text{(A1)} \qquad (\textit{Student Outcome})_{ij} = \beta_{0j} + r_{ij} \qquad \text{(Student Level)}$$
$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad \text{(Teacher Level)}$$

where $(\textit{Student Outcome})_{ij}$ is the outcome value for student *i* of teacher *j*, r_{ij} is the random error associated with student *i* of teacher *j* and is independently normally distributed with mean 0 and variance σ^2, and u_{0j} is the random error associated with teacher *j* and is independently normally distributed with mean 0 and variance τ_{00}.

For the segment completion rate analyses, the outcome was a student's withdrawal from the segment (coded as 0) or completion of the segment (coded as 1), regardless of passing or failing grade at the time of completion or withdrawal. The second student outcome was the score on the end-of-segment exam in segment 1. For time to segment completion, the outcome was the number of weeks a student took to complete the segment. The Florida Virtual School administrative database captures time to completion as the number of days that a student took to complete segment 1 of the course. The weeks to complete outcome used in the main model is equal to the number of days divided by seven.

The aim of these analyses was to estimate the variation in three student outcome measures in the online biology course—segment completion rates, end-of-segment exam scores, and time to segment completion—and the extent to which these student outcomes might vary across teachers. In the hierarchical linear models, the overall variation in these outcomes was partitioned to separate the within-teacher variation (see box 1 in main report) from the between-teacher variation.[2] To that end the study calculated the intraclass correlation coefficient (ICC) for each model (equation A2). The ICC is the ratio of the between-teacher variance to the total variance. In this context, it can be interpreted as the proportion of variance in outcomes between teachers that is not due to random variation in the students assigned to a teacher:

(A2) *ICC* or *p* = (*Between-teacher variance*) / (*Between-teacher variance* + *Within-teacher variance*).

For sensitivity analyses, the study team also calculated ICCs for completion using a two-level logistic model to examine between-teacher variation (see appendix B). The study team also converted ICCs to *d*-type effect sizes for ease of interpretation by taking the square root of the ICCs (following Rowan et al., 2002).

References

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling.* Scientific Software International.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of Elementary Schools. *Teachers College Record, 104*(8), 1525–1567.

[2] One of the strengths of the models is that they consider that there would be some variation between teachers even if student observations were randomly grouped, ignoring which teachers they had. This adjustment results in smaller estimates of between-teacher variation in student outcomes compared to the observed between-teacher variation.

Appendix B. Supporting analyses

This appendix presents detailed results of the analyses underlying the findings in the main report, as well as supplemental analyses designed to test whether the findings depended on the construction of the outcome variables or the statistical models.

Primary analyses

Analyses examined between-teacher variance for segment completion rates, the end-of-segment exam scores, and time to segment completion to find out the extent to which outcomes were different with different teachers (see box 1 in the main report). Between-teacher variation accounted for 3.1 percent of total variation in student outcomes for segment completion rates, 0.7 percent of total variation for end-of-segment exam scores, and 4.9 percent of total variation for time to segment completion. Corresponding effect sizes were 0.18, 0.08, and 0.22 (table B1).

Table B1. Intraclass correlation coefficient values for analyses of segment completion rates, end-of-segment exam scores, and time to segment completion, main model, 2014/15–2017/18

Outcome measure	Standard deviation	Variance component	Intraclass correlation	*d*-type effect size
Segment completion rates				
Between-teacher variance	0.079	0.006	0.031	0.18
Within-teacher variance	0.436	0.190		
End-of-segment exam score				
Between-teacher variance	0.962	0.926	0.007	0.08
Within-teacher variance	11.343	128.667		
Time to segment completion				
Between-teacher variance	1.825	3.330	0.049	0.22
Within-teacher variance	8.024	64.382		

Note: The analytical sample consisted of 10,618 students and 57 teachers for segment completion rates, 7,712 students and 55 teachers for end-of-segment exam scores, and 7,665 students and 55 teachers for time to segment completion.
Source: Authors' analysis of data from Florida Virtual School.

Alternative models

The study team tested several alternative models using different versions of the outcome variables. As presented in the main report and table B1, segment completion rates were calculated using a linear probability model (or a binary linear model approach, as described by Goldstein et al., 2002). For sensitivity analyses the study team also calculated intraclass correlation coefficients (ICCs) using a two-level logistic model to examine between-teacher variation. The study team used the R statistical package ICCbin (Chakraborty & Hossain, 2018) to calculate the ICCs because there is no direct estimation of the residuals on level 1 (see equation A1 in appendix A) for the generalized nonlinear multilevel models. ICCbin is designed specifically for calculating ICC for multilevel models with a binary response outcome using several different methods. The study team used a simulation method (as described by Goldstein et al., 2002) to calculate the ICC. The ICC generated using this method for segment completion rates was 0.03, which is similar to the ICC calculated from a linear probability model.

The study team also estimated an alternative model for the time to segment completion variable. The main report treats time to segment completion as a continuous outcome. This appendix also presents results based on three category and binary versions of the time to segment completion variable (table B2). The initial time to completion categories were based on guidance from Florida Virtual School regarding what it considered to be an appropriate range of time to complete a segment. The study team provided Florida Virtual School with data showing the distribution of time to completion, and Florida Virtual School provided initial cutscores for three categories: slower than expected, as expected, and faster than expected times to completion. For sensitivity analyses the study team tested a model using these categories of time to completion as the outcome and another model using a binary outcome that was defined as completing a segment slower than Florida Virtual School guidelines (coded 1 if yes and 0 otherwise). These results (see table B2) are comparable to the results obtained in the initial analyses.

Table B2. Intraclass correlation coefficient values for analyses of time to segment completion based on alternative models, 2014/15–2017/18

Model	Standard deviation	Variance component	Intraclass correlation
Three-category time to completion outcome (slower than expected, as expected, faster than expected)			
Between-teacher variance	0.13	0.02	0.036
Within-teacher variance	0.68	0.46	
Binary time to completion outcome (slower than expected = 1; as expected or faster than expected = 0)			
Between-teacher variance	0.05	0.00	0.021
Within-teacher variance	0.32	0.10	

Note: The analytical sample consisted of 7,665 students and 55 teachers.
Source: Authors' analysis of data from Florida Virtual School.

For the time to completion analysis the study team worked with Florida Virtual School to determine outliers (unusually fast or unusually slow). Unusually fast cases were identified as students who took less than three weeks to complete segment 1 of the course (n = 25). Unusually slow cases were students who took more than 47 weeks to complete segment 1 (n = 71). These cases were removed from the analyses whose results are presented in table B1. The study team also tested a model that included the outliers. For the main model and for the alternative model with three categories of time to completion, results of the analyses that included outliers (table B3) were similar to the results of the analyses that excluded outliers (see table B1). For the alternative models the findings were similar regardless of whether outliers were excluded (see table B2) or included (see table B3).

Table B3. Intraclass correlation coefficient values for analyses of time to completion with outliers included, main and alternative analysis models, 2014/15–2017/18

Model	Standard deviation	Variance component	Intraclass correlation
Main model (number of weeks to completion as outcome)			
Between-teacher variance	2.12	4.51	0.057
Within-teacher variance	8.62	74.32	
Alternative three category model (slower than expected, as expected, faster than expected time to completion as outcome)			
Between-teacher variance	0.13	0.02	0.036
Within-teacher variance	0.68	0.46	
Alternative binary category model (slower than expected = 1, as expected or faster than expected = 0)			
Between-teacher variance	0.09	0.01	0.038
Within-teacher variance	0.44	0.19	

Note: The 96 outlier cases removed from the main models were included in these analyses, so the sample consisted of 7,761 students and 55 teachers.
Source: Authors' analysis of data from Florida Virtual School.

End-of-segment exams were scored on a separate scale for students in honors and nonhonors classes. The study team created a common metric across honors and nonhonors classes by scaling the end-of-segment exam scores to percent of correct answers (0–100) separately for students in honors and nonhonors classes. The study team tested a model that controlled for honors/nonhonors status at level 1. The ICC generated using this model was 0.007 (table B4), which is the same as the ICC calculated for the main model (see table B1).

Table B4. Intraclass correlation coefficient values for analyses of end-of-segment exam scores, with honors or nonhonors class status included in the model, 2014/15–2017/18

Model	Standard deviation	Variance component	Intraclass correlation
End-of-segment exam score			
Between-teacher variance	0.96	0.92	0.007
Within-teacher variance	11.30	127.69	

Note: The analytical sample consisted of 7,712 students and 55 teachers.
Source: Authors' analysis of data from Florida Virtual School.

Conditional intraclass correlation coefficients. The study team also calculated conditional ICCs using the same equation as the one used to calculate the unconditional ICC (see equation A2 in appendix A). The difference is that the conditional ICC calculation used conditional variances. Conditional variance is the variance remaining once covariates are included in the model. The study team used student covariates to control for observed heterogeneity while calculating between-teacher variation. The goal of these analyses was to account for the potential differences in the characteristics of students assigned to Florida Virtual School teachers in the calculation of teacher effects.

Florida Virtual School uses an algorithm to assign students to teachers. This algorithm emphasizes maintaining equal teacher loads and ensuring that students are taken off the wait list and assigned to a teacher quickly. No student criteria such as demographic characteristics or academic achievement are considered in assigning students to teachers. The result is that students' ability level can be expected to be randomly distributed across teachers.

The study team examined between-teacher variation on the student characteristics using covariates available to the team. Two-level hierarchical linear models with student variables as the outcome were employed to examine the extent of between-teacher variation. Across all student-level variables, between-teacher variation was low, as indicated by the small ICC values (table B5). The percentage of between-teacher variation ranged from 0 percent (for gender) to 1.2 percent (for grade level). This observed low variation suggests that as a result of the Florida Virtual School's student assignment algorithm, other unobserved student characteristics are randomly distributed across teachers.

Table B5. Intraclass correlation coefficient values for analyses of student covariates, 2014/15–2017/18

Model	Standard deviation	Variance component	Intraclass correlation
Prior year state math test score as an outcome[a]			
Between-teacher variance	1.12	1.26	0.003
Within-teacher variance	20.39	415.78	
Prior year state English language arts test score as an outcome[a]			
Between-teacher variance	1.31	1.71	0.005
Within-teacher variance	19.08	363.86	
National school lunch program status as an outcome			
Between-teacher variance	0.03	0.00	0.004
Within-teacher variance	0.50	0.25	
Race/ethnicity as an outcome			
Between-teacher variance	0.03	0.00	0.005
Within-teacher variance	0.50	0.25	
Gifted program status as an outcome			
Between-teacher variance	0.02	0.00	0.006
Within-teacher variance	0.26	0.07	
Honors class status as an outcome			
Between-teacher variance	0.01	0.00	0.001
Within-teacher variance	0.46	0.21	
Grade level as an outcome			
Between-teacher variance	0.05	0.00	0.012
Within-teacher variance	0.50	0.25	
Gender as an outcome			
Between-teacher variance	0.01	0.00	0.000
Within-teacher variance	0.50	0.25	

Note: National school lunch program status, race/ethnicity (non-Hispanic White vs. all), gifted program status, honors class status, grade level (grade 9 vs. grade 10), and gender (female vs. male) are binary variables. For these variables, linear probability models were used to calculate intraclass correlations (ICCs). Logistic models were also run, and ICCs were calculated using a latent variable approach (Goldstein et al., 2002). ICCs from the logistic models were as follows: 0.004 for national school lunch program status, 0.003 for race/ethnicity, 0.027 for gifted program status, 0.001 for honors class status, 0.015 for grade level, and 0.000 for gender. These were the same as or comparable to the ICCs from the linear probability models presented in this table. Prior year math test scores were available for 1,600 students (15 percent of the sample), and prior year English language arts test scores were available for 2,348 students (22 percent of the sample). As a result, the math model calculated ICCs using data for 42 teachers, and the English language arts model calculated ICCs using data for 43 teachers. The gifted program status model calculated ICCs using data for 56 teachers, as data were not available for 1 teacher. All other models calculated ICCs with data for 57 teachers using the expectation maximization (EM) estimation method, although data were not available for some students. Student-level sample sizes were 9,701 for the national school lunch program model, 10,613 for the race/ethnicity model, and 3,589 for the gifted program status model. For the grade level, honors class status, and gender models, the student-level sample size was 10,618 students. The EM estimation method used in hierarchical linear modeling treats the model parameters as missing values to be estimated. Although some student data were missing in some statistical models, available data contribute to the estimation of the model's parameters by implying probable values for missing values, borrowing information from the existing data at successive iterations until differences between successive iterations were trivial (Peugh & Enders, 2004).

a. Four years of data from 2014/15 through 2017/18 were used in the analyses except for the prior year math and English language arts test scores models, which used data from 2016/17 and 2017/18.

Source: Authors' analysis of data from Florida Virtual School

Since the between-teacher variation for the student variables did not identify any covariates to include as controls, the study team applied a two-step procedure as a second strategy to form models to calculate conditional ICCs. In the first step student covariates were entered as an independent variable in the main models (see equation 1 in appendix A) at level 1, and separate models were estimated in turn that included each student covariate. In the second step the significant predictors (using $\alpha = .05$ as a cutoff) from each of the separate models run in the first step were combined in a model. When a student covariate was no longer significant in the combined model at the $\alpha = .05$ level, it was dropped from the analysis. Final models were used to calculate conditional ICCs (table B6).

Table B6. Regression results and intraclass correlation coefficient values for analyses of segment completion rates, end-of-segment exam scores, and time to segment completion, 2014/15–2017/18

Segment completion rate model		
Fixed effect	Coefficient (standard error)	*p*
Intercept	0.40 (0.02)	***
Students' previous Florida Virtual School record[a]		
New student	0.24 (0.02)	***
Previous student, never assigned to a teacher	0.22 (0.03)	***
Previous student, successful completer	0.40 (0.01)	***
Academic year[d]		
2015/16	0.05 (0.01)	***
2016/17	0.03 (0.01)	*
2017/18	0.03 (0.01)	*
National school lunch program status[a]	–0.12 (0.01)	***
Honor class status	0.07 (0.01)	***
Grade 9 (vs. grade 10)	0.03 (0.01)	***
Non-Hispanic White (vs. all other)[a]	0.02 (0.01)	*
Random effect	Variance (standard deviation)	Conditional intraclass correlation
Between-teacher variance[b]	0.00 (0.05)	0.014
Within-teacher variance	0.16 (0.40)	

End-of-segment exam score model		
Fixed effect	Coefficient (standard error)	*p*
Intercept	77.68 (0.37)	***
Gifted program status[b]	1.81 (0.83)	*
National school lunch program status[b]	–0.97 (0.44)	*
Honor class status[c]	1.70 (0.46)	***
Random effect	Variance (standard deviation)	Conditional intraclass correlation
Between-teacher variance	0.07 (0.27)	0.001
Within-teacher variance	117.35 (10.83)	

Time to segment completion model		
Fixed effect	Coefficient (standard error)	*p*
Intercept	20.23 (0.43)	***
Students' previous Florida Virtual School record[a]		
New student	–2.99 (0.44)	***
Previous student, never assigned to a teacher	–2.58 (0.69)	***
Previous student, successful completer	–1.51 (0.33)	**
Academic year[d]		
2015/16	–0.41 (0.29)	
2016/17	1.64 (0.32)	***
2017/18	2.86 (0.32)	***
Random effects	Variance (standard deviation)	Conditional intraclass correlation
Between-teacher variance	1.66 (2.75)	0.042
Within-teacher variance	7.49 (63.00)	

* $p < .05$; ** $p < .01$; *** $p < .001$.

Note: The analytical sample consisted of 9,698 students and 57 teachers for segment completion rate analyses, 2,471 students and 54 teachers for end-of-segment exam score analyses, and 7,665 students and 55 teachers for time to segment completion analyses. The expectation maximization (EM) estimation method used in hierarchical linear modeling treats the model parameters as missing values to be estimated. In the statistical models, although missing data for end-of-segment exam scores and completion were not imputed, available data contributed to the estimation of the model's parameters through implying probable values for missing values by borrowing information from the existing data (Peugh & Enders, 2004). Thus, the time to segment completion model calculated the intraclass correlation coefficient (ICC) using data from all teachers in the sample. For the end-of-segment exam score model, the ICC was calculated using data for 54 teachers, as data were not available for 1 teacher. A large amount of data was missing for the end-of-segment exam score model because of unavailable data on gifted program status. The same model was tested excluding the gifted program status covariate to calculate the conditional ICC using data for all 55 teachers. The analytical sample for that model included data for 6,934 students and 55 teachers. The conditional ICC was 0.006. Prior year math test scores were available for 1,600 students (15 percent of the sample), and prior year English language arts test scores were available for 2,348 students (22 percent of the sample). Thus, separate models were run, in which prior year math and English language arts scores were entered as student covariates. Across three study outcomes, prior year math and English language arts achievement were both statistically significant at the $p < .05$ level. In the model that controlled for students' prior year math test score, conditional ICCs were calculated using data for 42 teachers. In the model that controlled for students' prior year English language arts test scores, conditional ICCs were calculated using data for 43 teachers. For models using the prior year math score as a covariate, conditional ICCs were 0.027 for the segment completion rate analyses, 0.001 for the end-of-segment exam score analyses, and 0.042 for time to segment completion analyses. For models using the prior year English language arts score as a covariate, conditional ICCs were 0.017 for the segment completion rate analyses, 0.000 for the end-of-segment exam score analyses, and 0.040 for the time to segment completion analyses. Across all models, conditional ICCs were low.

a. Although all students in the sample were taking biology for the first time, it was possible that some students had previously taken another Florida Virtual School course. The students' previous Florida Virtual School record variable classified students as a "new student" (new to the Florida Virtual School system); "previous student, never assigned to a teacher" (in the Florida Virtual School system but never assigned to a teacher); "previous student, successful completer" (had previously taken and successfully completed another Florida Virtual School course); and "previous student, not yet completed successfully," the reference category (had previously taken but had not completed another Florida Virtual School courses).

b. National school lunch program status compared eligible and noneligible students; gifted program status compared students in the program with those not in the program; and non-Hispanic White students compared those students with Asian, American Indian or Alaska Native, Black, Hispanic, Native Hawaiian or other Pacific Islander, and multiracial students.

c. End-of-segment exam scores were scaled separately for students in honors and nonhonors classes. All scores were scaled to the percentage of correct answers (0–100) to create a common metric.

d. The reference group was 2014/15. Four years of data (2014/15–2017/18) were used in the analyses to reduce random error in the teacher-level estimates and to balance out any exceptionally positive or exceptionally negative scores for a teacher that might be observed in a single year by random error. The biology course was not substantively revised during this time period. Because multiple years of cross-sectional data were used, the study team examined whether a majority of teachers had a substantial portion of their students' data (more than 75 percent) from a single year (which could occur if teacher turnover was high); it found that 47 percent of the teacher sample had more than 75 percent of their student data from a single year.

Source: Authors' analysis of data from Florida Virtual School.

Analyses using segment 2 data. The study team was also interested in whether the results would change if students took the biology course from the same teacher for two segments. The longer duration of teacher influence for those students could result in greater between-teacher variation.

The study team used the same main models to answer the research questions (see equation A1 in appendix A), but this time using outcomes for segment 2 (table B7). For the analysis of time to segment completion, outliers were again dropped (the 120 students who completed the segment in less than 3 week and the 45 students who completed it in more than 34 weeks).

Table B7. Between-teacher and within-teacher variation in segment completion rates, end-of-segment exam scores, and time to segment completion, model using segment 2 data, 2014/15–2017/18

Outcome measure	Standard deviation	Variance component	Intraclass correlation
Segment completion rates			
Between-teacher variance	0.01	0.00	0.002
Within-teacher variance	0.20	0.04	
End-of-segment exam score			
Between-teacher variance	0.96	0.92	0.008
Within-teacher variance	10.45	109.27	
Time to segment completion			
Between-teacher variance	1.73	2.99	0.090
Within-teacher variance	5.49	30.09	

Note: There were 4,396 students for the segment completion rate analyses, 4,216 students for the end-of-segment exam score analyses, and 4,053 students for the time to segment completion analyses who had both segment 1 and segment 2 data with the same teacher. Thus, these models calculated intraclass correlations for 46 teachers. As a result, the findings for segment 2 are not directly comparable to the findings for segment 1.
Source: Authors' analysis of data from Florida Virtual School.

References

Chakraborty, H., & Hossain, A. (2018). R package to estimate intracluster correlation coefficient with confidence interval for binary data. *Computer Methods and Programs in Biomedicine*, *155*(1), 85–92. https://doi.org/10.1016/j.cmpb.2017.10.023.

Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding Statistics, 1*(4), 223–231. https://doi.org/10.1207/S15328031US0104_02.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research, 74*(4), 525–556. https://doi.org/10.3102/00346543074004525.

Appendix C. Descriptive analyses

This appendix provides descriptive information about each of the outcome measures examined in the study.

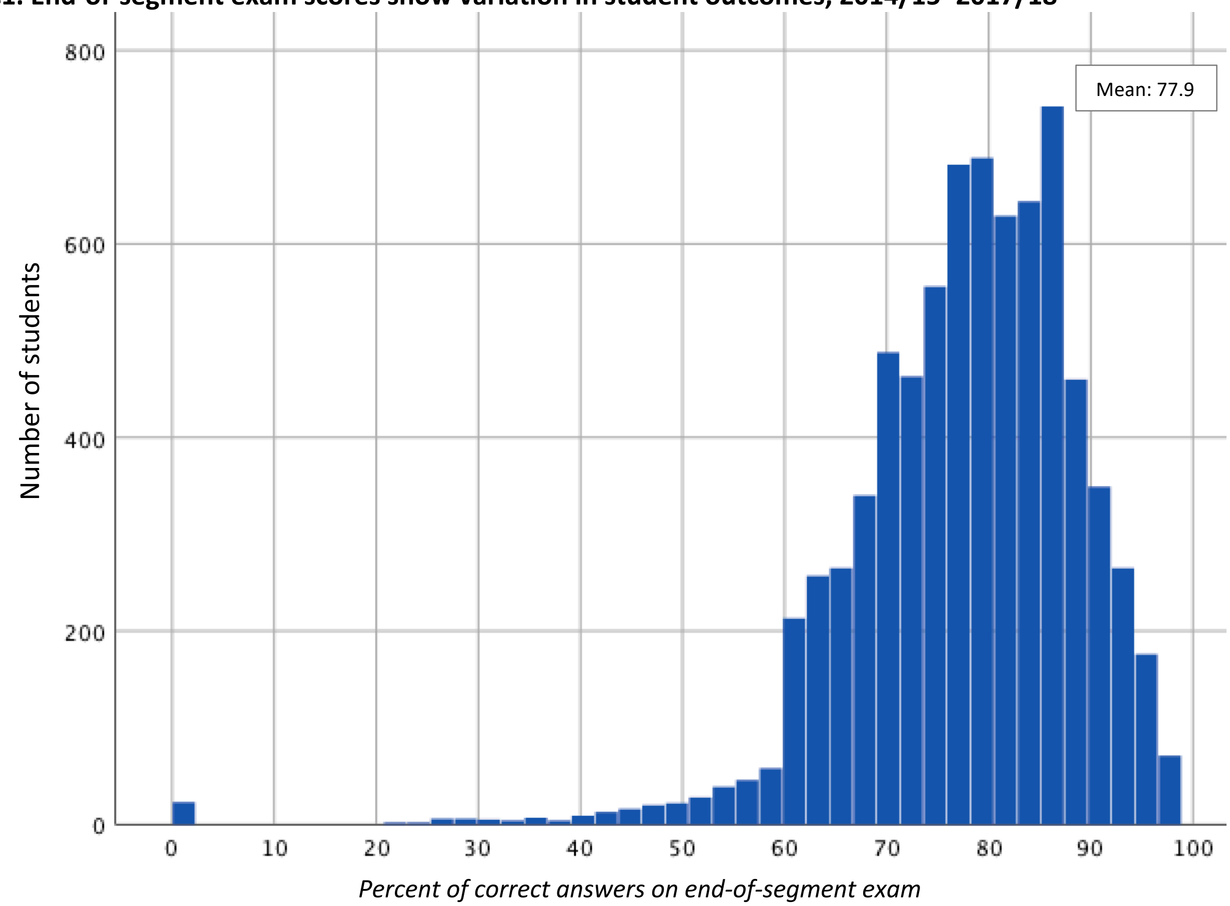
Segment completion rates

Segment completion is defined as a student having completed the course segment, regardless of the student's grade at the time of completion or withdrawal status.[1] In all, 73.4 percent of students completed all of the modules and took the end-of-segment exam.

End-of-segment exam scores

The average score on the exam at the end of segment 1 was 77.9, meaning that Florida Virtual School biology students answered nearly 78 percent of the exam questions correctly, on average. The distribution of scores shows substantial variation across students (figure C1).

Figure C1. End-of-segment exam scores show variation in student outcomes, 2014/15–2017/18



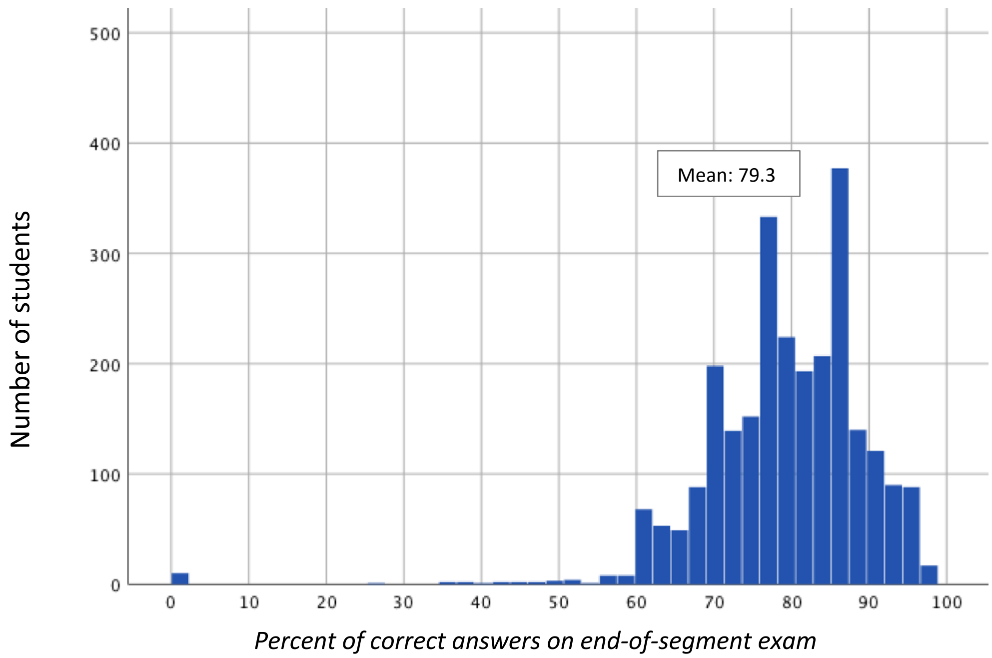
Note: The sample consisted of 7,712 students.
Source: Authors' analysis of Florida Virtual School data.

[1] The final grade for each segment is based on the total points for all assignments in the course: weekly assignments associated with individual modules, module exams, phone quizzes, and the end-of-segment exam, which is 20 percent of the final grade.

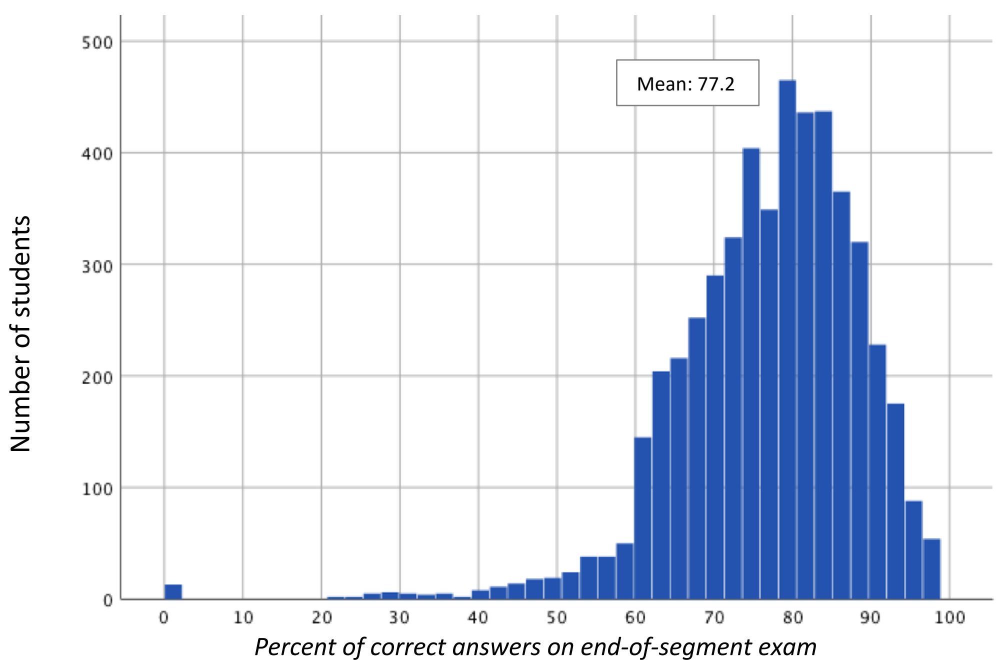
End-of-segment exam scores were scaled separately for students in honors and nonhonors classes. So, the maximum total number of points differed for honors and nonhonors students. All scores were scaled to the percentage of correct answers (0–100) to create a common metric. The distribution of end-of-segment exam scores for students in honors and nonhonors classes show considerable variation (figure C2).

Figure C2. Distribution of end-of-segment exam scores show variation in student outcomes for students in honors classes and those in nonhonors classes, 2014/15–2017/18

Honors



Nonhonors

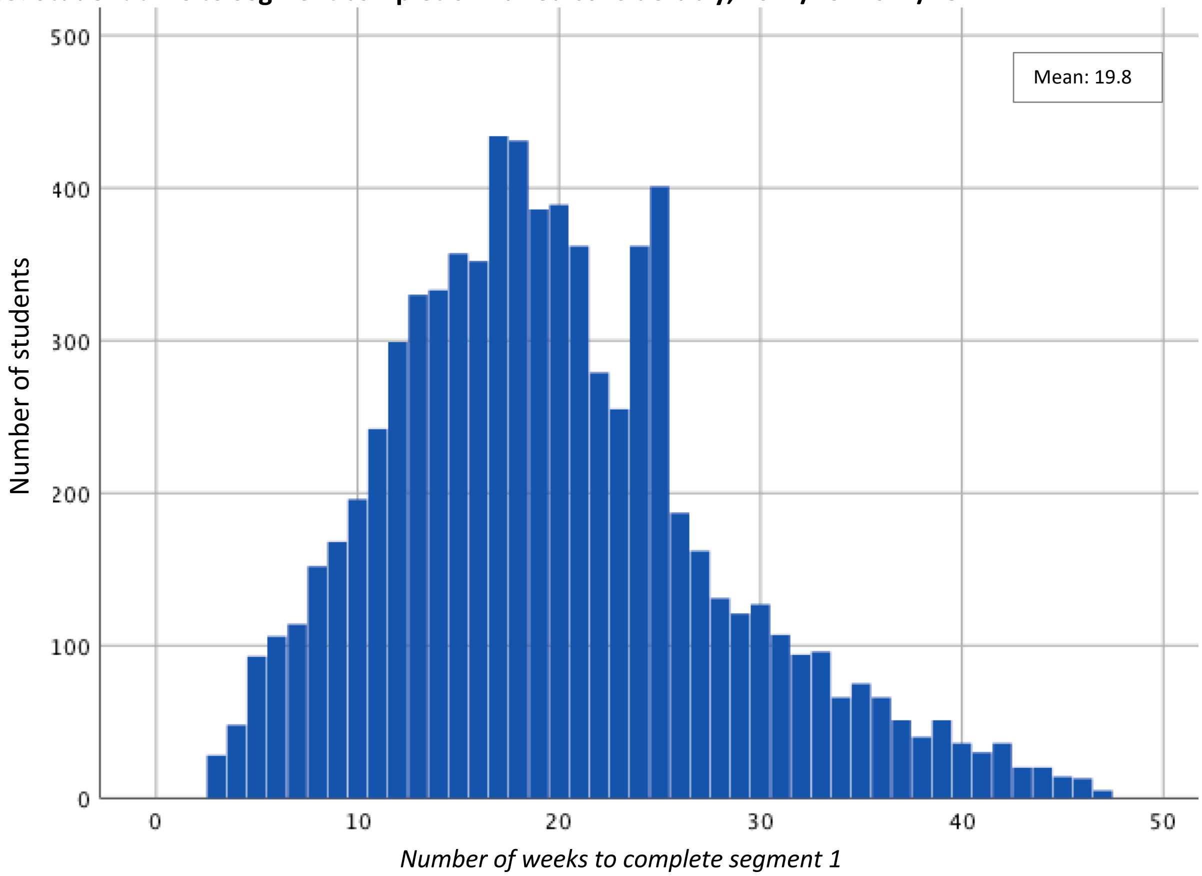


Note: The honors students sample consisted of 2,617 students. The nonhonors students sample consisted of 5,095 students.
Source: Authors' analysis of Florida Virtual School data.

Time to segment completion

The time in which students completed the segment varied considerably. Students took an average of 19.8 weeks to complete segment 1 of the biology course (figure C3).

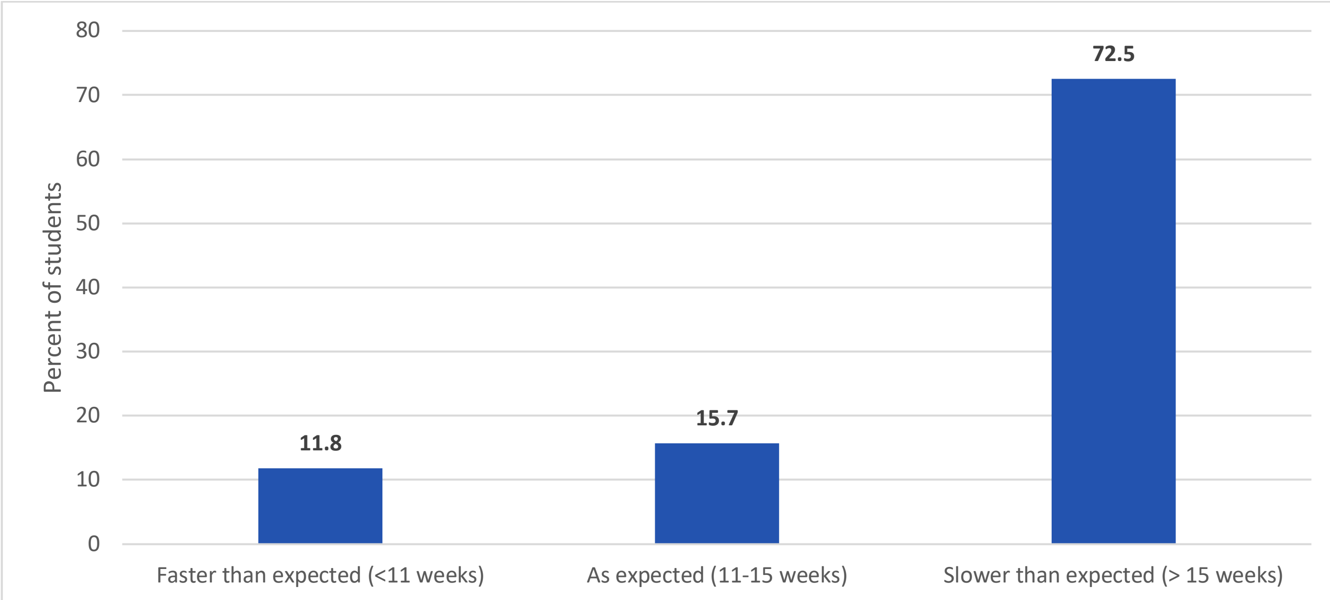
Figure C3. Student time to segment completion varied considerably, 2014/15–2017/18



Note: The sample consisted of 7,665 students.
Source: Authors' analysis of Florida Virtual School data.

Given the wide range in time to completion and based on what Florida Virtual School considered "normal" or expected progress in the course, students were classified into one of three categories: those who completed the segment faster than expected (lees than 11 weeks), those who completed it as expected (11–15 weeks), and those who completed it more slowly than expected (more than 15 weeks). Nearly 73 percent of students completed the segment more slowly than expected, 16 percent completed it as expected, and 12 percent completed it faster than expected (figure C4).

Figure C4. Most students completed segment 1 of the biology course more slowly than expected, in an average of 20 weeks, 2014/15–2017/18



Note: Expectations for completion were defined by Florida Virtual School for segment 1 of the biology course. The sample consisted of 7,665 students. It includes four years of data from 2014/15 to 2017/18.
Source: Authors' analysis of Florida Virtual School data.