Technical Research Report

Connections Academy Study 2020

Comparing state test achievement of Connections Academy students who remain enrolled for three or more years with state test achievement of traditional, brick-and-mortar students

Authors: Guido Gatti, MA**,** Anne Pier Salverda, PhD**,**
Clinton L. Johns, PhD**,** Muireann Kelly, PhD

Pearson Global Product Organization Efficacy & Learning Research Impact Evaluation – May 2021

Table of contents

Pearson

Executive summary

Virtual schools are increasingly sought out by students and their families as alternatives to local brick-and-mortar (B&M) schools. These students/families may face uncertainty about current public education, feel underserved or otherwise unsatisfied with their local school, or have circumstances that make it difficult for them to attend a traditional B&M school.

Virtual schools in the US have shown consistent growth since their origination in the late 1990s: as of 2019, 39 states have either virtual or blended schools. Given the forced transition from in-person instruction to online learning brought on by the Covid-19 pandemic, increases in virtual school attendance will likely continue at an accelerated pace. We would like to state upfront that this report, and the results contained therein, is for school years preceding the Covid-19 pandemic.

An online public education can provide a student-centered approach that is personalized and flexible. Customizing the learning experience for students by providing them with a flexible, portable education has the potential to benefit all students, but especially those unique students who feel online schools are a better fit. Online schools may therefore prove to be a logical investment of education funds if they can provide quality education.

Currently, rigorous peer-reviewed research on the effectiveness of virtual schools and the students who attend them is sparse. This is unfortunate, as it allows a few highly publicized studies to disproportionately influence the perceptions held by key stakeholders, such as parents and state legislators.

Research to date has not typically…

1. looked specifically at students attending Connections Academy schools across different states,
2. placed an emphasis on controlling for student mobility, or
3. controlled for local factors affecting educational access and quality.

No studies that we are aware of have done all three of these, where mobility is defined as any time a student changes schools for reasons other than grade promotion. While no such studies have come to our attention, we are interested in any that meet these criteria for our future reporting.

Overview of Connections Academy

Pearson Online & Blended Learning's Connections Academy schools are public schools that provide a tuition-free, online, personalized, full-time education to K-12 students. There are 42 Connections Academy schools in the United States, most of which serve students statewide, with students attending from their homes. All schools are state funded and regulated, and almost all are accredited. Schools are expected to consistently and reliably employ the same core instructional model for math and reading.

Connections Academy schools use the same "standard" public school curriculum and instructional model that covers the individual states' standards. The curriculum and standards are the same as those forming the basis for the annual state tests' content. In addition to course assignments and assessments, the core model also includes recommendations to teachers for synchronous live lessons and Response-to-Intervention tiered assignment and progress monitoring.

Connections Academy students are provided access to computer software, interactive learning technologies, assessment and reporting tools, digital curriculum materials, credit recovery options, multimedia curriculum tools and games (e.g., discussion boards, and online student groups such as book and robotics clubs), and social events (e.g., field trips). Students and their learning coaches (a parent/guardian or other legally designated caring adult) also receive support from teachers and other school educators working in the virtual school environment. Connections Academy teachers are certified to teach in their state and receive ongoing professional development on best practices in education and online learning.

In 2014, Connections Academy brought together a group of department leaders to formally review relevant literature and research. Through this exercise, the team formed a set of core beliefs about four major program elements that have been shown to have a significant impact on student learning. These four elements are: practice, feedback, student engagement/motivation, and intervention. Further, Connections Academy fosters a learner-friendly environment for its students via three mechanisms: building an environment of support, enhancing and improving access, and personalized learning.

Pearson

Intended learner outcomes

This study addresses the **Standard of Achievement outcome**: *Student academic performance meets or exceeds the performance of their peers.*

Research aims and research questions

Connections Academy schools aim to create active, independent learners. It is hoped that these learners will start to see achievement gains from the online education model. Based on previous research, we hypothesize that Connections Academy students will become effective learners after attending for three or more full consecutive school years and, as such, should close any initial achievement gaps and come to achieve comparably to non-mobile students attending traditional brick-and-mortar (B&M) schools.

It should be noted that online school-students in general tend to be highly mobile, and Connections Academy students are no exception. During the timeframe of this research, only 39% of Connections Academy students remained consistently enrolled for three school years. This fact makes it challenging to measure the effectiveness of Connections Academy over time. Relatedly, there are challenges with building an appropriate comparison group to Connections Academy students as the reasons for mobility vary and are difficult to measure and account for.

Connections Academy, however, believes it is imperative to assess student achievement in fair and meaningful ways. To that end, this study will build on recently published research and compare the state test scores of returning (for three school years) Connections Academy students from two southern states to the state test scores of closely matched, non-mobile students from the same grade who are attending what would have been their local B&M school districts.

By matching Connections Academy students to B&M students by location, and ensuring that location remains consistent, this research controls for many unmeasured would-be factors at the district, grade and community levels affecting access to education and achievement. This design allows us to provide evidence about the effectiveness of Connections Academy students who maintain consistent enrollment.

We matched each stable Connections Academy student to all similar stable B&M students (i.e., ***one-to-many*** matching) based on: state; school year; first-year school district; grade level; and first-year state test scores; as well as economically disadvantaged/low income status; English learner status; special education status; gender; and race/ethnicity. Separate matched groups were created for math (Connections Academy n = 341, B&M n = 12,086) and reading (Connections Academy n = 341, B&M n = 12,098). Stable B&M students attended their local district schools for four full consecutive years. Stable Connections Academy students attended Connections Academy for three full consecutive years after first attending the same B&M district (as their matching B&M students) for a full school year.

Main research question. Do students who previously attended traditional B&M schools, and who enroll for three or more consecutive years at the same Connections Academy school, score as well on standardized state tests of reading and math as demographically similar non-mobile students from those same local B&M school districts, and does the result differ for the two states currently under study?

Key findings

These findings come after matching students in elementary grade levels (3rd, 4th, 5th), following them for three consistently non-mobile school years, and then comparing them in middle grades (6th, 7th, 8th) on math and reading state test scores[1].

Attending Connections Academy was as effective as attending brick-and-mortar (B&M) schools for non-mobile middle-school students' reading state test scores. This was true for both states.

Compared to non-mobile middle school students in brick and mortar schools, attending Connections Academy resulted in lower math state test scores. On average, the performance difference was greater for State Two students (-0.50 standard deviations) than it was for State One students (-0.29 standard deviations).

The Connections Academy students from the two states under study in this report performed at the same level in reading as B&M students. This contrasts with previous findings on virtual school students included in the CREDO study (Woodworth et al., 2015), which examined data from 18 states, and the study that examined data specifically for the state of Indiana (Fitzpatrick et al., 2020). These two studies found that virtual school students in general performed more poorly in reading than matched B&M students.

Our results for math are similar to those reported in the Indiana study, which found a math gap of -0.50 standard deviations. In this study, the math gap for students who had attended Connections Academy for three school years, relative to matched B&M students, was similar for State Two (-0.50 standard deviations) but smaller for State One (-0.29 standard deviations).

Recommendations

The results suggest that students who remain enrolled in Connections Academy for three years achieved a similar reading level to their counterparts remaining in their local schools. The results also suggest the Connections Academy students achieved a lower level of performance in math compared to their counterparts who remained in their local schools.

The finding that attending Connections Academy had a negative impact on math performance is in line with much recent research on virtual schools. However, although this study compares student achievement between Connections Academy and B&M students, it tells us nothing about *why* observed differences exist or how virtual education may be improved. We also ignored implementation fidelity to the Connections Academy core instructional model in this study. However, it seems reasonable to hypothesize that schools implementing the core model with greater fidelity may have different results than those that do not. These are important areas for future research.

[1] A small number of students were in 6th grade in their first year (and repeated a grade across 4 years, e.g. 6-6-7-8). Likewise, a small number of students were in 5th grade in their final year (and repeated a grade across 4 years, e.g. 3-3-4-5).

A limitation of this study is its generalizability to both the general population of Connections Academy students and the general population of all students.

First, it should be noted that only two Connections Academy schools were included.
It is reasonable to assume that:

- across the Connections Academy schools, there is variation in students' math and reading state test performance (for example, our results demonstrate a difference in math scores between the two Connections Academy schools in this study)
- other Connections schools may perform better or worse, compared with both the Connections Academy schools in this study and with B&M schools.

Second, we did not include students in the entire kindergarten through 12th-grade continuum that Connections Academy schools serve.

Third, as we only included Connections Academy students who attended for three consecutive years, and these students have high mobility, the Connections Academy students remaining in the final analytic sample may not be representative of the broader Connections Academy student population. Further, Connections Academy students may continue to stay enrolled for different reasons than non-mobile B&M students (i.e., the majority). Many factors can influence the choice to attend, and then stay with, Connections Academy. We should limit our efficacy statements to families and students like the 'choosers' and 'stayers' in this study.

Fourth, there are aspects of the data related to student mobility and non-standard grade progression (i.e., skipping or repeating grades) that readers should consider. In the former case, although we accounted for student mobility at the *district* level, we were unable to account for student mobility at the individual school level. This is primarily due to characteristics of the sample population in which the majority of students' data necessarily included the transition from elementary to middle schools. In the latter case, the distribution of students with non-standard grade progression suggested that exclusion might both reduce statistical power and introduce selection bias.

Finally, our approach to matching B&M and Connections Academy students in order to define the analytic samples admits the possibility of bias by using demographic information from the full four years of these students' data to derive a single value for each student, rather than relying solely on first-year 'baseline' values to guide the matching process.

Introduction

Virtual schools are increasingly sought out by students and their families as alternatives to local brick-and-mortar (B&M) schools. These students/families may face uncertainty about current public education, feel underserved or are otherwise unsatisfied by their local school, or have circumstances that make it difficult for them to attend a traditional B&M school. Virtual schools in the US have shown consistent growth since their origination in the late 1990s.

As of 2019, 39 states have either virtual or blended schools. In 2017–18, 501 full-time virtual schools enrolled 297,712 students, and 300 blended schools enrolled 132,960 students. Enrollments in virtual schools increased by more than 2,000 students between 2016–17 and 2017–18, and enrollments in blended learning schools increased by over 16,000 during this same time period (Molnar et al., 2019). The forced transition from in-person instruction to online learning brought on by the Covid-19 pandemic may drive these numbers up at an accelerated pace. We would like to state upfront that this report and the results contained within are for school years prior to the Covid-19 pandemic.

The focus of this research is on students attending Pearson's Connections Academy schools. Connections Academy schools provide tuition-free, full-time, online public education to K-12 students that is both individualized and interactive. Connections Academy schools aim to create active independent learners. It is hoped that these learners will start to see achievement gains from the online education model. Based on previous research (e.g., Gatti, 2018), we hypothesize that Connections Academy students will become effective learners after attending for three or more full, consecutive school years and, as such, should close any initial achievement gaps, and come to achieve comparably to non-mobile students attending traditional B&M schools.

It should be noted that the general population of online school students is highly mobile, and Connections Academy students are no different. During the timeframe of this research, only 39% of Connections Academy students remained consistently enrolled for three school years. This makes it challenging to measure the effectiveness of Connections Academy for its students over time. Relatedly, there are challenges with building an appropriate comparison group to Connections Academy students as the reasons for mobility vary and are difficult to measure and account for. In addition, a review of the literature suggests that mobility puts students at greater risk for low reading and math test scores (Rumberger, 2015).

Connections Academy schools, however, believe it is imperative to assess their students' achievement in fair and meaningful ways. To that end, this study will build on recently published research and compare the state test scores of returning (for three school years) Connections Academy students from two southern states, to the state test scores of similar closely matched non-mobile students attending the same local school districts from which the Connections Academy students transferred.

By matching Connections Academy students to B&M students based on location - and ensuring that location remains consistent - this research controls for many unmeasured would-be factors at the district, grade and community levels affecting student access to education and achievement. Thus, the question is no longer whether or not Connections Academy students differ from the average of B&M students statewide, but whether or not their performance differs from other students from their neighborhood.

Background

Rigorous, peer-reviewed research on the effectiveness of virtual schools and the students who attend them is sparse. This state of affairs is unfortunate, as it allows a few highly publicized studies to disproportionately influence the perceptions held by key stakeholders, such as parents and state legislators.

Research to date has not typically…

1. looked specifically at students attending Connections Academy schools across different states,
2. placed an emphasis on controlling for student mobility, or
3. controlled for local factors affecting educational access and quality.

(Note: *mobility is defined as any time a student changes schools for reasons other than grade promotion.)*

Some rigorous studies have specifically compared the state test scores for Connections Academy students in grades 3–8 and their counterparts attending B&M schools. The results can be succinctly summarized as:

- Georgia Connections Academy students made significantly lower-than-expected gains in elementary reading, elementary math, middle-school math, and high-school geometry, but significantly higher-than-expected gains in 9th grade literature and high-school American literature (Sass, 2019).

- Connections Academy cohorts from 21 schools nationwide showed no statistical difference in math and reading compared to cohorts attending B&M schools (Gatti, 2018).

- California Connections Academy students made significantly higher-than-expected gains on the 10th grade reading state assessment, while they made gains as expected on the 8th-grade reading state assessment (Ford & Rice, 2015).

Other recent studies have examined state test scores for all available virtual school students (e.g., Fitzpatrick et al., 2020; Woodworth et al., 2015). These studies' analytic samples and reported results combined Connections Academy students with students enrolled in other virtual schools.

Woodworth and colleagues (henceforth the CREDO study) examined students from 18 states (including Washington, D.C.), while Fitzpatrick and colleagues focused only on students in Indiana. The CREDO study reported that B&M students outperform virtual school students by four percentile points in reading and 10 percentile points in math (after a single year's attendance).

The Indiana study reported considerably larger differences, with brick-and-mortar (B&M) students outperforming virtual school students by as much as 19 percentile points in math after attending for three years, and 13 percentile points in reading after three years.

Despite these methodological differences, all these studies compare Connections Academy students to B&M students and include controls to try and establish a fair comparison. Table 1 shows the important design features of each study and how they compare to this study.

Table 1. Relevant studies on virtual school academic performance

Study design features	CREDO	CA (2015), OH (2016), GA (2019), IN (2020)	Gatti 2018	This study
Connections Academy specific	No	No, Yes, Yes, No	Yes	Yes
Most recent data	No	No, No, Yes, Yes	Yes	Yes
Compares to B&M	Yes	Yes	Yes	Yes
Compares matched groups	Yes	No, No, No, Yes	Yes	Yes
Multiple school years	Yes	Yes	Yes	Yes
Gr 3–8 state test scores	Yes	Yes	Yes	Yes
Multiple states	Yes	No	Yes	Yes
Statewide comparison	Yes	Yes, Yes, Yes, No	Yes	Yes
Local or district comparison	No	No, No, No, Yes	No	Yes
Accounts for mobility	No	No	Yes	Yes
Compares non-mobile students	No	No, Partially, No, Partially	No	Yes
Accounts for student risk factors (other than mobility)	Yes	No, Yes, No, Yes	Yes	Yes
Sub-populations (e.g., at-risk)	No	No	No	Yes

(CREDO) Woodworth, J. L., Raymond, M. E., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W., & Van Donge, C. (2015). Online charter school study. Stanford, CA: Center for Research on Education Outcomes. Retrieved from: https://credo.stanford.edu/sites/g/files/sbiybj6481/f/online_charter_study_final.pdf

Ford, R., & Rice, K. (2015). Value-Added Results for Public Virtual Schools in California. Educational Technology & Society, 18(4), 412–423.

Ahn, J. (2016). Enrollment and Achievement in Ohio's Virtual Charter Schools. Fordham Institute.

Sass, T. R. (2019). The performance of state charter schools in Georgia, 2017/2018. Atlanta, GA: Andrew Young School of Policy Studies, Georgia State University.

Fitzpatrick et al. (2020). Virtual Illusion: Comparing Student Achievement and Teacher and Classroom Characteristics in Online and Brick-and-Mortar Charter Schools. Educational Researcher, 49(3), 161-175.

Gatti, G. (2018). A Comparison Study of Connections Academy Schools to Matched Brick and Mortar and Virtual Schools, Examining the Types of Students Who Attend K-12 Virtual School and the Effects on Performance of a Highly Mobile Student Body.

Description of Connections Academy

Pearson Online & Blended Learning's (OBL) Connections Academy schools are public schools that provide a tuition-free, online, personalized, full-time education to K-12 students. Connections Academy schools serve a variety of students, which includes those looking to be challenged, trying to catch up, with health concerns, with accessibility issues, being bullied, or otherwise unsatisfied with local brick-and-mortar (B&M) offerings.

Families hear about Connections Academy through the usual television advertisements and social media channels, but also via word of mouth and recommendations from B&M school officials. The section "How Connections Academy encourages students to stay", discusses reasons for joining Connections Academy. The reader will find common themes around the ways in which 26,377 new Connections Academy families (from 2015 through 2019, attending 39 academies across 28 states) felt that their local B&M schools did not fulfil their expectations.

There are 42 Connections Academy schools, most of which serve students statewide, with students attending from their homes. All schools are state funded and regulated, and almost all are accredited. Schools are expected to consistently and reliably employ the same core instructional model for math and reading. Connections Academy schools use the same "standard" public school curriculum and instructional model that covers the individual states' standards. These curriculums and standards are the same as those forming the basis for the annual state tests' content.
In addition to course assignments and assessments, the core model also includes recommendations to teachers for synchronous live lessons, and Response-to-Intervention tiered assignment and progress monitoring,

OBL students are provided access to computer software, interactive learning technologies, assessment and reporting tools, digital curriculum materials, credit recovery options, multimedia curriculum tools and games (e.g., discussion boards, and online student groups such as book and robotics clubs), and social events (e.g., field trips). Students and their learning coaches (a parent/guardian or other legally designated caring adult), also receive support from teachers and other school educators working in the virtual school environment. Connections Academy teachers are certified and receive ongoing professional development on best practice in education and online learning.

How learning science informs product design

Connections Academy has always reviewed relevant research when designing and developing the many products and services that make up the Connections Academy school program. Typically, each department within Connections Academy focuses on research topics specifically related to their areas of responsibility, such as curriculum content, instructional design, technology, teacher effectiveness, parental involvement.

In 2014, the Connections Academy chief academic officer brought together a group of department leaders to engage in a close study and discussion of learning sciences (e.g., Dweck, 2006; Hess & Saxberg, 2014; ASCD, 2010, 2011, 2012a, 2012b, 2013a, 2013b, 2013c; Bransford, Brown, & Cocking, 1999; Shechtman, DeBarger, Dornsife, Rosier, Yarnall, 2013; iNACOL, 2015). The aim was to strengthen cross-departmental collaboration focused on improving student outcomes, and to ensure that ongoing improvements to Connections Academy's products and services reflected findings from the most current, relevant research. Along with a formal review of relevant literature and research, the group brought many decades of experience and more than a decade of accumulated data to the discussions.

Pearson

Through this exercise, the team of experts responsible for the Connections Academy program crystallized a set of beliefs about four major program elements that have been shown through research to have a significant impact on student learning. Connections Academy regularly refers to these four elements, along with additional research and feedback from users, to guide ongoing improvements to the Connections Academy school program. The four elements are: practice, feedback, student engagement/motivation, and intervention.

Practice

Studies comparing novices and experts show that one characteristic of experts is relatively effortless or automatic retrieval of relevant knowledge, as well as easy recognition of problem types (Bransford et al., 1999). This ability to recognize problems so that appropriate solutions can be applied, as well as the ability to easily retrieve knowledge, comes primarily from practice. However, not all practice produces equivalent learning outcomes. Research suggests that practice should be relevant (Eccles, 1983), deliberate (Ericsson, Krampe, & Tesch-Romer, 1993), and ongoing (Cepeda et al., 2009), and give students multiple opportunities to learn and demonstrate learning without negative consequences.

While opportunities for practice are embedded throughout Connections Academy school's online curriculum, the Curriculum Development team has performed extensive analysis using performance and student feedback to make informed decisions about how and where to enhance practice within courses. Working closely with the Multimedia, Standards, and Design team, the team has created opportunities for students to practice using next-generation assessment functionality, allowing for multiple attempts, hint boxes that provide guided and targeted feedback, and opportunities to monitor and reflect on their thinking.

Feedback

In online educational settings, feedback is generally regarded as a key component of knowledge and skill acquisition (Azevedo & Bernard, 1995). However, both the content and timing of feedback can influence its effectiveness (Shute, 2008). A review of the literature suggests that feedback should:

- focus on the task, not the learner
- elaborate with information about the 'what, how, and why' of a given problem, rather than just the correctness of the answer
- be specific and clear
- be objective
- promote a focus on growth, improvement, and learning (Shute, 2008)

Connections Academy teachers receive professional learning on giving feedback that is goal-referenced, actionable, user-friendly, timely, relevant, and ongoing (Wiggins, 2012). Feedback is delivered through rubrics, WebMail Messages to follow up on conversations, grade book comments, and most importantly, phone calls with the student known as Curriculum Based Assessments (CBAs). The learning coach or student receives an alert that their teacher has commented or given feedback.

Student engagement and motivation

Educational research (e.g., Richardson, Abraham, & Bond, 2012) has identified a number of non-cognitive factors that impact student success, with motivation playing a critical role. Many aspects of motivation have been researched, such as different beliefs, attitudes, goals, and interests. For example, when a student is feeling motivated, it may be because she or he feels interested in a topic, is feeling challenged or confident, wants to improve his or her career prospects, or wants to outperform his or her peers, among many other reasons. From this rich research base, various theories and frameworks have been developed for understanding this wide array of motivations and how they affect important learner outcomes. Some of these have proven quite powerful in helping Connections Academy understand how motivation impacts student success.

One non-cognitive factor that research has found to have a large impact on learning is students' mindset. Research has shown that people tend to gravitate towards one of two mindsets when it comes to learning in a given domain. One of these is a "fixed" (or "entity") mindset, where a person believes that how good one is in the domain is largely innate, and not much can be done to change that. For example, someone who believes that they are just not good at math and never could be has a fixed mindset. The other is a "growth" (or "incremental") mindset, where a person believes that ability in the domain comes through practice and effort. Someone who feels like they can improve with enough effort exemplifies a growth mindset.

A growing body of research has found that there are benefits associated with adopting a growth mindset. Students with a growth mindset are more likely to adopt more learning-oriented goals, to persist longer (Diener & Dweck, 1978), to use better learning strategies and, ultimately, to achieve better grades (Yeager & Dweck, 2012). In addition, research has begun to document different interventions that have been shown to move students to adopt this more beneficial growth mindset. For example, programs that provide training on learning strategies, informed by cognitive neuroscientific findings about how the brain changes with learning, have been shown to lead to greater adoption of growth mindset (Blackwell, Trzesniewski, & Dweck, 2007). In addition, growth-mindset tutoring by respected peers has been found to be effective (Good, Aronson, & Inzlicht, 2003), particularly for students from traditionally disadvantaged populations. More subtle approaches, including providing encouraging messages that focus on the development of a skill, have also been found to be successful (Williams, Paunesku, Haley, & SohlDickstein, 2013).

Students tend to work harder, spend more time on learning and have more positive motivational experiences when they find the content they are learning to be personally relevant. However, attempts to add interesting elements to engage students can actually be detrimental to learning, and students who are already interested may find those features distracting (Durik & Harackiewicz, 2007). Rather than attempting to design a universally interesting experience, a more promising approach is to get students to reflect on why the content they are covering in class may be useful to them.

For example, Hulleman and Harackiewicz (2009) prompted students to write about how they would apply the class content to their lives (or the lives of their friends or family), or how it fits in with their future plans, and found semester-long benefits for interest and achievement, particularly among students with low expectations of success. Approaches to help make clear the possible utility of the content for the student can have an impact on students' interest and achievement.

Pearson

Another element that is important to acknowledge has to do with the "epistemological framing" of the learning environment. This refers to the different ways in which students can think about the knowledge they are learning in class. Research has identified specific ways in which learning environments can promote a sense for students that the information they are learning is likely to be used in broad ways, outside of the particular context in which it is learned (Engle, Nguyen, & Mendelson, 2011). Some approaches to help students frame things in a more expansive way include:

1. connecting material explicitly across time (i.e., how what they learned earlier connects to what they are learning now; or how what they are learning now will connect to what they will learn in the future)
2. being connected to more people, such as other groups of students in the course, students in other courses, or even outside participants in the endeavor, such as scientists or other related groups
3. making clear that the student is involved in generating their own explanations and ideas, rather than simply being a (passive) recipient of canonical information.

Connections Academy schools foster student motivation and engagement by incorporating these empirical findings throughout the learning experience. As mentioned above, Connections Academy teachers receive professional learning on providing instruction consistent with a growth mindset (e.g., including user-friendly, goal-specific feedback). These principles are also integral to both curriculum development and to the orientation provided to students and their learning coaches.

Intervention

Research shows that targeted interventions to struggling students help those students significantly improve (Burns, Appleton, & Stehouwer, 2005; Tran, Sanchez, Arellano, & Swanson, 2011). The key components of intervention are identifying students at risk, delivering targeted, effective interventions, and monitoring progress (Shinn, 2010). Effective interventions explicitly teach the specific skills students need (Fuchs, Fuchs, & Compton, 2012).

At Connections Academy schools, universal screening tools are administered three times during the year to identify students who need intervention. Interventions include synchronous LiveLesson® sessions to provide targeted instruction or extended learning opportunities to individuals or small groups of students, or access to research-based supplemental instructional programs.

Using a multi-tiered instructional model, teachers can move identified students across tiers as appropriate to adjust the frequency and intensity of the intervention, based on students' responsiveness and learning needs. Teachers use ongoing progress monitoring to ensure that students are receiving appropriate interventions and making expected progress. Documentation of the data points reviewed during progress monitoring provides evidence that interventions are appropriate and is used to inform considerations around special education referrals.

How Connections Academy encourages students to stay

This is a study about the impact of Connections Academy on achievement for students returning to for three or more years, and this section explains how Connections Academy fosters a learner-friendly environment through three mechanisms:

1. building an environment of support
2. enhancing and improving access
3. personalized learning

Building an environment of support

The key to reaching long-term intended outcomes is building and providing an environment of support for students to be engaged in their education in a way that positions them to succeed. This environment is encouraged with a collection of foundational tasks and attributes that promote student engagement and learning. As discussed above, the curriculum provides opportunities for students to engage in meaningful practice, receive specific, actionable feedback from teachers, and reflect upon their mindset and set course-specific goals to support motivation and engagement. This supportive environment aims to provide students with:

1. current, standards-aligned curriculum and instructional delivery that supports varied learning preferences and/or needs
2. multiple opportunities to practice and learn without fear of negative consequences
3. guidance from supportive, qualified teachers
4. abilities to monitor their own learning and learn at a pace that matches their needs
5. frameworks for setting personal and academic goals
6. feedback that is timely, actionable and specific, intended to support students in becoming engaged, self-directed learners.

Enhancing and improving access

Connections Academy improves access to education for students who might struggle in conventional settings, both academically and physically. Hallmarks of Connections Academy schools include:

- an award-winning curriculum delivered via Connexus® (link)
- a proprietary education management system (EMS)
- state-certified and specially trained teachers
- a personalized approach to learning (Personalized Performance Learning®)
- a supportive school community that includes the involvement of a learning coach.

The learning coach is usually the student's parent, guardian, or another appropriate adult the parent/guardian designates. Primarily learning from home, students work with teachers online and via phone, while the on-site learning coach supports and monitors students' progress.

Student socialization occurs online in synchronous classes and clubs, and in person at events like school-organized field trips. The combination of academic and social support is aimed at keeping student engagement and motivation high. Connections Academy students are held to the same state standards as their traditional brick-and-mortar (B&M) peers and are required to take state assessments.

Personalized Performance Learning®

Personalized Performance Learning® refers to a school experience and approach to learning that meets students' needs. This philosophy places the student at the center of the learning experience; supported by teachers, the curriculum, and the learning coach, and all connected by technology. This technology-supported curriculum provides ample opportunities for practice and, since it is digital, immediate *feedback*. Not only does this support student learning in and of itself, but student performance analytics can also be used to tailor instruction and provide a more personalized experience.

Students are supported by state-certified and specially trained teachers. At the beginning of the school year, and within the construct of fulfilling state standards throughout the year, teachers discuss with each student and their learning coach the student's academic strengths and areas of need. This is to define a personal learning plan, which includes goal setting and discussions focused on making learning relevant and meaningful to the student.

Throughout the school year, teachers use real-time data tools and reports in Connexus to systematically monitor student progress. They use this data and regular synchronous contacts with each student to adjust the pace and content of student's lessons and coursework. This data is also used to identify and implement any necessary interventions or enhancements, ensuring students receive the right degree of challenge or support.

Teachers regularly connect with their students through online classes (LiveLesson® sessions), by phone, and via communications tools embedded in Connexus. This allows teachers to provide feedback and interventions as needed, and helps students review progress against their goals. Teachers can use this time to encourage students to develop a positive, growth-oriented mindset and help them understand how their courses can be personally meaningful. They also connect students with one another in LiveLesson® sessions and discussions. Connections Academy teachers hold at least a bachelor's degree; 60% hold advanced degrees.

The standards-aligned curriculum is designed to meet the needs of diverse learners and offers an expansive catalog of courses, including core academics, electives, and advanced placement courses. Extracurricular clubs and activities are also offered. The teams responsible for the Connections Academy curriculum combine research-based proprietary content with instructional resources and teaching materials from publishers to create units, lessons and instructional activities. They also develop interactive, multimedia, online educational tools and resources with the aim of engaging students and further supporting their learning.

The curriculum provides opportunities for students to engage in meaningful *practice*, receive specific, actionable *feedback* from teachers, and reflect upon their *mindset* and set course-specific goals to support *motivation* and *engagement*. Intervention programs to supplement the curriculum and support struggling students are incorporated into curricular offerings.

Students' learning in the virtual environment is supported by learning coaches. This is usually a parent/guardian, although the learning coach can be another adult designated by the student's parent/guardian. Connections Academy requires the involvement of the learning coach at grade-appropriate levels, which allows parents to be closely involved in their students' education, while also encouraging students to become increasingly independent learners as they move into higher grades.

Pearson

Connections Academy's *Get Coaching!* Program is dedicated to supporting learning coaches; it is designed to help them understand their role and to provide them with tools and strategies to support their students. It also provides access to a community of fellow learning coaches. Within Connexus, learning coaches also have access to Family 411. This is the family resource center that provides learning coaches with links to recorded orientations, interactive tutorials, how-to guides and digital learning tips, such as information on how to encourage a positive student mindset and the value of productive struggle.

Virtual schools are necessarily dependent on technology. Most Connections Academy schools provide students with loaned computers and subsidies for internet connection. The central technology feature at Connections Academy is the proprietary Connexus® EMS technology platform. Students use Connexus to engage with lessons, connect with teachers and classmates, and access a virtual library and communications and planning tools. Connexus is vital to the Connections Academy teachers, who use it to conduct lessons and grade assessments, track students' progress, communicate with students and families, and adjust coursework and lessons in support of each student's learning plan.

Parents/learning coaches also have insight into students' work and performance via Connexus EMS. Tools within Connexus support students as they set goals, take action on feedback provided by teachers, and engage in intervention programs. These tools also help students to reflect upon their mindset and assess their confidence in their ability to complete their coursework. Technology use is scaled by grade level.

Intended implementation

Connections Academy students are expected to navigate the online coursework as intended, according to the direction from their teachers, and under the supervision of their learning coach. In this study, we did not track students' progression through their math or reading courses. We made sure that all students in the analytic sample attended or were enrolled for at least 150 days in Connections Academy or, in the case of brick-and-mortar students, in the same school district for at least three school years, and were state tested in both math and reading in all school years.

The intended implementation of Connections Academy is to provide students with a complete, quality school experience outside the traditional classroom, characterized by a personalized approach to learning. The program is implemented consistently across grade levels, with minor age- and grade-appropriate variations.

Elementary school level
Elementary school students learn foundational educational concepts in reading, writing and mathematics, and are taught study skills. Science, social studies, technology, art, and physical fitness round out the core curriculum, and students work with hands-on instructional resources, including virtual tools, kits, and workbooks. Connections Academy offers electives, activities, and clubs to encourage further exploration. For example, students can take world language courses, learn basic music concepts and conduct home experiments. A minimum of 30 hours per week is spent learning (or as mandated by school and/or state requirements), and about 15–30% of the school day is centered on interactive online coursework.

Pearson

Students are assigned one expert elementary teacher who works with each student individually and also with groups of students to support and guide students as they engage in their coursework. Teachers have regular synchronous and asynchronous contact with students and use LiveLesson® sessions to engage students in online classes and support. A school counselor is also available. Learning coaches are encouraged to provide a high level of oversight for elementary students, which is generally a commitment of about five hours per day. Learning coaches typically support students by setting a schedule with varied activities and breaks, assisting with lessons, monitoring student comprehension and grades, and communicating frequently with the teacher.

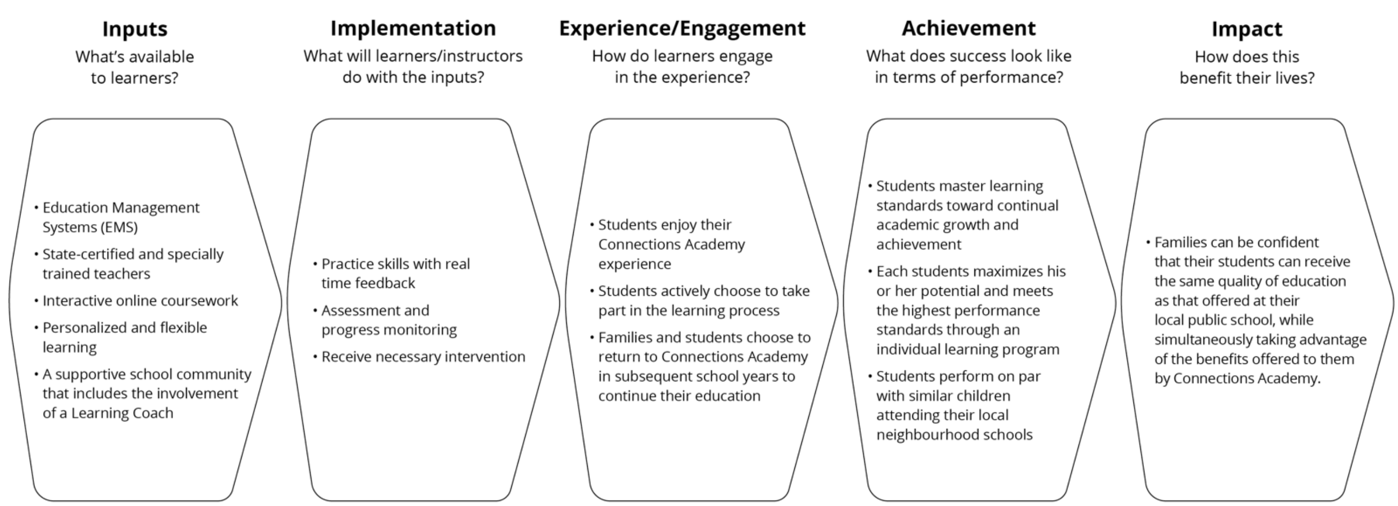
Middle school level

In middle school (6th–8th grades, or 7th–8th grades in some schools) Connections Academy aims to help students continue to develop their language, arts, math, and critical thinking skills through a blend of online and offline work. Electives provide students opportunities to learn new skills, find art in everyday life, and explore new technologies. Students can also join clubs to explore areas of interest; for example, students can learn about robotics or write for the school newspaper. When available in a school and approved by a counselor, gifted students can start earning high school credits early.

Connections Academy provides middle school students with a prescribed schedule, which requires a minimum of 30 hours per week, or as mandated by school and/or state requirements. Students work with teachers as needed to create their individual schedules. About 50–75% of the school day for middle schoolers is centered on interactive online courses. Connections Academy middle school students begin working directly with subject-specific teachers and a homeroom or advisory teacher who monitors and assists with all subjects. A school counselor is also available.

The role of the learning coach changes as the student becomes more independent and takes increased ownership of his or her learning. Connections Academy recommends that the learning coach spends about two to three hours a day overseeing learning. Activities may include supporting the transition to more independent learning, assisting with some lessons, monitoring student comprehension and grades, and communicating with teachers and referring the student to the teacher as needed.

The graphic below depicts the theory of change regarding the supports and interventions that will encourage students to stay at Connections Academy, and how retention should lead to parity in achievement over time.



This study

This study builds on recently independently evaluated, audited and published research conducted by Pearson (Gatti, 2018, pdf), that compared the state test achievement of Connections Academy and similar brick-and-mortar (B&M) student cohorts nationwide. The 2018 study first examined the information provided by the families of over 70,000 Connections Academy students to determine what types of students are most likely to seek out the Connections Academy virtual school option. Then, after accounting for relevant factors (including student mobility), the study combined multiple years of National Center for Education (NCES) data and 19 states' summative end-of-year assessment data to determine whether there was a statistically significant difference in math and reading proficiency between student cohorts in Connections Academy schools and cohorts in B&M schools.

Gatti (2018) revealed seven distinct profiles for students choosing a Connections Academy virtual school:

1. academically advanced students
2. academically struggling students
3. students experiencing health problems
4. new students who were experiencing bullying
5. returning students who originally enrolled with numerous challenges, including those captured in the previous clusters
6. students new to Connections Academy schools who were seeking flexibility and choice
7. returning Connections Academy students who continue to value flexibility and choice.

The study also confirmed that there are large differences in the mobility rates of Connections Academy schools compared to brick-and-mortar (B&M) schools in the same state. By way of demonstration, for 16 of the 19 Connections Academy schools, the mobility rate was above the 75th percentile, and in 10 states the Connections Academy had the highest or second highest mobility rate. Despite these differences in mobility, Gatti (2018) found that there was no statistically significant difference in reading and math achievement between student cohorts from Connections Academy schools and 3rd-8th grade student cohorts from B&M schools serving similar at-risk student populations.

Building on Gatti (2018), the goal of this study is to compare the state test scores of Connections Academy students, specifically *choosers-stayers*, to similar closely matched students of the same starting grade, attending what would be the Connections Academy students' local B&M district. Here, we use the designation of chooser-stayers to include those students who left their local B&M school to attend Connections Academy (i.e., choosers) and then attended Connections Academy the next three full consecutive school years (i.e., stayers). Students from two Connections Academy schools, each located in a southern state, are the focus of this report. Note that local B&M students who change schools yet remain within the same district for at least three consecutive years are considered non-mobile and eligible for matching.

The important contributions of this study are that, unlike previous studies:

1. student mobility is no longer an issue, as we compare location-stable students after receiving consistent uninterrupted instruction for multiple school years, and

2. factors at the district, grade (within a district) and community levels affecting student access and achievement are accounted for by our research design and statistical analytic approach.

This study also matches its comparison samples on students' observable characteristics (i.e., prior state test score, race/ethnicity, gender, English learner status, economically disadvantaged/low income, and special education).

By matching Connections Academy students to B&M students on location and ensuring that location remains consistent, the proposed research design controls many unmeasured would-be factors at the district, grade and community levels that affect student access and achievement. This method creates a rigorous, counterfactual comparison.

Comparing Connections Academy students to students attending their would-be local school districts permits a more pertinent comparison for the Connections Academy students' performance than does the more typical comparison of Connections Academy students to other students statewide. A statewide comparison essentially compares Connections Academy students to a predicted state-average student of the same gender, ethnicity, etc. In this study, the question is no longer whether or not Connections Academy students differ from the average of B&M students statewide, but whether or not their performance differs from other students from their neighborhood?

Research questions

This study addresses the **Standard of Achievement outcome**: *Student academic performance meets or exceeds the performance of their peers.*

Main research question: Do students who previously attended traditional B&M schools, and who enroll for three or more consecutive years at the same Connections Academy school, score as well on standardized state tests of reading and math as demographically similar, non-mobile students from those same local B&M school districts, and does the result differ for the two states currently under study?

Secondary research question: How do reading and math state test scores for these Connections Academy students compare to B&M students for each level of the following subpopulations? (Note: demographic information for each category provided by the Departments of Education for States one and two):

- metropolitan area (yes, no)
- race/ethnicity (White/Asian[2], yes or no)
- economically disadvantaged/low income (yes, no)
- special education (yes, no)
- English learner (yes, no)
- gender (i.e., male, yes/no)

[2] We dichotomized the race/ethnicity variable for purposes of coarsened matching. We opted for White/Asian vs non-White/Asian because inspection of state test scores showed that White and Asian students on average achieved higher test scores than non-White/Asian students.

To further support any main research findings, we decided *a priori* that if we did not find a statistically significant difference in achievement between Connections Academy students and B&M students, we would perform an additional analysis to test for the absence of a meaningful effect. This would be in the form of an *equivalence test*, which examines if the data provide evidence for an effect that is smaller than a chosen *smallest effect size of interest* (SESOI). Again, the size of the SESOI was determined ahead of time, prior to having seen any of the data or results for the present study.

We would like to note here that although the entire sample of Connections Academy students in a state attend a single Connections Academy school, this does not create an n=1 confound in the traditional sense. The What Works Clearinghouse (WWC) handbook provides guidance on this (Standards Handbook v4.0, p81): examples of circumstances that are not confounding factors include instances where “A school with unique organization and governance is compared to multiple comparison schools. When the intervention of interest is attending the school, the WWC does not consider this to be a confounding factor because the school and the intervention are the same.”

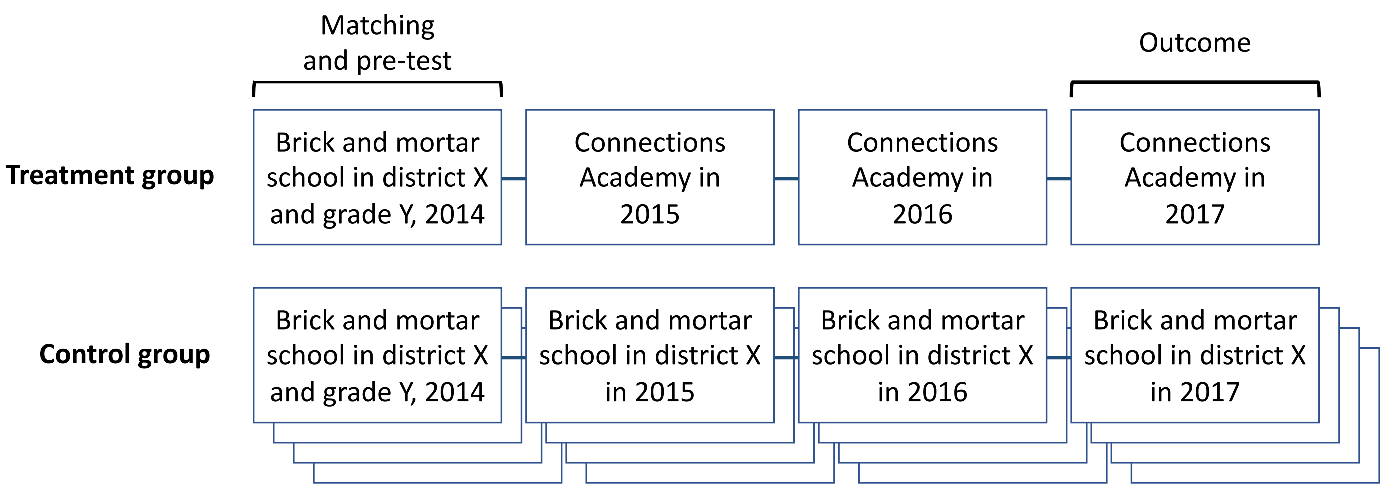
Table 2: Alignment between learner outcomes and research questions

Outcome category	Learner outcome	Main or secondary finding	Research question	Measure/ method of data collection	Type of efficacy statement (only for main finding)
Standard of achievement or level of competence	Student academic performance meets or exceeds the performance of their peers	Main	Do students who previously attended traditional B&M schools, and who enroll for three or more consecutive years at the same Connections Academy school, score as well on reading and math state tests, as similar non-mobile students from those same local B&M school districts, and does the result differ for the two states currently under study?	3rd through 8th grade math and reading state test scores	Causal
Standard of achievement or level of competence	Student academic performance meets or exceeds the performance of their peers	Secondary	How do Connections Academy students compare to B&M students for each level of the following subpopulations? ● Metropolitan area (yes, no) ● Race (White/Asian, yes or no) ● Low income (yes, no) ● Special education (yes, no) ● English learner (yes, no) ● Male (yes, no)	3rd through 8th grade math and reading state test scores	Causal

Method

This study compared groups for two southern states on 2016, 2017, 2018, or 2019 reading and math state test scores. Returning Connections Academy students were closely matched (four years prior, in 2013, 2014, 2015, or 2016, respectively) to returning brick-and-mortar (B&M) students who were in the same grade and attended a school in the same local traditional B&M district.

For illustrative purposes, the figure below represents hypothetical and simplified example data for the two study groups. The top row represents data for one particular Connections Academy student, and the bottom row represents data for four matching B&M students. (Note that a Connections Academy student can be associated with *any* number of matching B&M students.)



The main requirements for the students being compared are that they:

- attended either a B&M school followed by three years in a Connections Academy (treatment group), or attended (for four consecutive years) a non-charter B&M school in the same district as one of the Connections Academy students attended (control group)
- were continuously enrolled each of the four school years (150+ days in each year), and
- were state tested for both math and reading in each of the four consecutive years.

The Connections Academy schools under study use the same "standard" public school curriculum and instructional model that covers the individual states' standards and serves students statewide. The Connections Academy students attend virtually, while residing in all corners of each state. The curriculum and standards are the same as those forming the basis for the annual state tests' content. Only non-charter B&M school students will be used for comparison due to the variability among charter schools (e.g., in terms of student population, focus and/or curriculum), as well as to omit confounds from students who receive blended instruction.

To ensure a rigorous comparison between the Connections Academy and B&M student samples, students were matched on initial-year state test performance, economically disadvantaged/low income indicator, English learner status, special education, gender and race/ethnicity (i.e., coarsened to White/Asian or not). Connections Academy students were matched separately for the math analysis and the reading analysis. This means that a Connections Academy student...

1. will typically be matched to a different set of students from the same state in the math analysis than in the reading analysis
2. may not be matched in one or both analyses.

Initial-year state math and reading test performance was operationalized as the traditional population z-score (i.e., mean = 0; standard deviation = 1) of the students' scaled scores within each state, for a given year, for each grade level (and, in the case of one state, for each combination of test version and test language, because scaled scores for that state differed as a function of test version and test language). A match was defined as all otherwise similar B&M students in a range of $\pm$0.25 standard deviations from a Connections Academy student's initial-year state math or reading score.

Participants

We initially contacted nine states from different regions of the US in the hopes that some would share student-level data. Ultimately, our request for data was fulfilled by two southern states.

Appendix A, describes the data-cleaning process, including how the necessary exclusions reduce the analytic sample at each step. These exclusions were performed to ensure that students were continuously enrolled and state tested, and include the following basic criteria.

1. Connections Academy and B&M students state tested once for both math and reading for four consecutive years (required for non-mobility and statistical comparisons)
2. Connections Academy students enrolled for three or more consecutive school years (required for Connections Academy non-mobility)
3. B&M students enrolled in the same local school district for four or more consecutive school years (required for B&M non-mobility)
4. ***Note: Only non-charter B&M school students were included in the comparison group***
5. Connections Academy and B&M students enrolled between 150 and 180 days for each of the four school years (required for non-mobility)

After selecting for students that met our criteria, we performed a coarsened exact-matching procedure for math and reading separately. Note that initial-year B&M math or reading state test score was one of the variables used to match students.

After exclusions, there were 180 Connections Academy students from State One, and 218 from State Two. The final analytic sample for math consisted of:

- 341 Connections Academy students; 164 in State One and 177 in State Two
- 12,086 matching B&M students.

The final analytic sample for reading consisted of:

- 341 Connections Academy students; 157 in State One and 184 in State Two
- 12,098 matching B&M students.

Important demographic information is provided in Appendix B, for State One and in Appendix C, for State Two. The information reported in the appendices is for those students who were tested in both math and reading within a year. We consider this the largest population of relevant students.

Compared to students attending B&M schools, Connections Academy and other virtual school students are more often White in both states. In State One, the percentage of African American students is smaller in Connections Academy than in B&M schools. In State Two, the percentage of Hispanic students is smaller in Connections Academy than in B&M schools. In both states, B&M students tend to be more often economically disadvantaged/low income compared to Connections Academy students. Furthermore, Connections Academy has few students that are English learners compared to B&M schools.

Finally, an often overlooked and important difference is that many more virtual school students tend to be mobile, as the following numbers for State One show[3] (see also table B1 in Appendix B). Student mobility is defined as any time a student changes schools for reasons other than grade promotion. While 5.4–6.5% of traditional B&M school students in the state are mobile in a school year, a substantially higher percentage (13–26%) of Connections Academy or virtual school students are mobile.

The data also shows a difference between the percentage of students who are consistently enrolled and state testing for four years (i.e., 75% for B&M students and 39% for Connections Academy students). While the large majority of B&M students are non-mobile, mobility is much more common for Connections Academy students. This is significant because there is evidence that mobility is associated with negative impacts on students, including lower test scores and high school graduation rates (Rumberger, R. W., 2015 link).

To illustrate the effects of student mobility, Table 3 below shows an analysis of student data from our two states. We see that student mobility alone would be expected to lower students' state test scores by 5–7 percentile points. This expected drop in score comes in addition to the other risk factors and after adjusting for previous year state test score.

Table 3. The effects of student mobility in two southern states on 2016 and 2017 state test scores

State		Standardized test score	
		Math	ELA
One	If mobile	-0.154*** (-6.1%)	-0.121*** (-4.8%)
	Observations	522,183	521,103
Two	If mobile	-0.172*** (-7.0%)	-0.119*** (-5.4%)
	Observations	2,589,302	2,589,302

Parentheses indicate percentile rank decline in scaled state test score attributable solely to being mobile.

Groups of mobile and non-mobile students were matched and statistically compared to other students in the same state using doubly robust inversely weighted propensity score matching.

State One non-mobile students were enrolled in a single school for at least 150 days. State Two non-mobile students attended a single school for at least 150 days.

Students were matched on previous year state test score, gender, race/ethnicity (White, Black, Hispanic, Asian/Other), grade level 3rd–8th), eligible for free or reduced-price lunch, English learner, receiving special education, school year (2016 or 2017).

*** $p < .001$

[3] We do not report a corresponding table for State Two because for that state, the mobility variable provided by the Department of Education was not reliable for students attending virtual schools. 95% of their records had missing values for mobility, compared to 0.01% of the records for all other students.

Data collection

We initially contacted nine states from different regions of the US in the hopes that some would share student-level data. This report is concerned with the two states who were willing to share the necessary data. The process typically consisted of four steps:

1. opening a request for data
2. submitting an application
3. navigating a review and clarification process, and finally,
4. the data transfer.

Each state's Department of Education (DOE) de-identified and shared their students' data in strict accordance with their student records privacy regulations, and shared students' state test scores, demographic indicators (i.e., gender, economically disadvantaged/low income, English learner status, special education indicator, and race/ethnicity), and attendance (i.e., school, and/or district attended each school year).

State One provided data for 2012–13 through 2017–18 and State Two provided data for 2012–13 through 2018–19. State One did not provide data for the 2018–19 school year as it was not yet available at the time data collection efforts took place. In addition, we downloaded the most current 2012–13 through 2017–19 National Center Education Statistics (NCES) district data files for the two states. The NCES data provided the urban/non-urban distinction as well as identifying information for charter and other virtual schools (school ID, district ID).

Measures

This section describes state test scores, the outcome variable reported on for this research effort, as well as more detailed information on the student demographic variables used in the matching process (i.e., gender, English learner status, special education, race/ethnicity, economically disadvantaged/low income).

State test scores

The scaled test scores provided by the state were transformed to traditional z-scores for each grade level and school year. These standardized scores are computed by subtracting the mean scale score from each students' score and then dividing the result by the standard deviation. The mean and standard deviation were computed across all scores separately for each combination of year and grade for State One, and across each combination of year, grade, test version and test language for State Two. (Note that scaled scores differed as a function of test version and test language for State Two.) The standardizing population included all scale scores for students who tested only once within a particular year (State One), or once within a particular year at the earliest testing opportunity[4] (State Two), and had a testing record that included both a math and a reading score.

Standardizing scores within grade level and school year allows a student's relative position among the other state test scores to be compared across grade levels and school years. The comparisons in group means on standardized state test scores will therefore represent a valid and efficient estimate of the effect size. The figures in Appendices D (State One) and E (State Two) show the distributions for the standardized scores. The distributions vary considerably across grade levels, years and subjects. These distributions at times appear skewed, to spike or to be missing some scores at intervals on the continuum.

[4] Data for State One did not specify which testing opportunity, within a year, a test score was associated with.

We could not locate information on the internal consistency reliability of the state tests. However, the State Assessment Technical Advisory Committee guidelines call for test forms to show reliability indices of at least 0.85. Additional reliability evidence comes in the form of the correlations between subsequent years' state test scaled scores, see Table 4. The scores for each state are correlated as expected, with those scores closer in time more highly correlated. Also, the 2015 scores correlate less strongly than the 2016–2017 scores. This is expected, since the content standards changed after the 2014–15 school year. Pearson product moment correlation coefficients for scale scores across years range from 0.72 to 0.87.

Table 4. Correlations between subsequent years' state test scores

State	Subject Area	2015	2016	2017
One	Reading 2015	1.0000	-	-
	Reading 2016	0.7917	1.0000	-
	Reading 2017	0.7734	0.8653	1.0000
	Mathematics 2015	1.0000	-	-
	Mathematics 2016	0.7417	1.0000	-
	Mathematics 2017	0.7206	0.8492	1.0000
Two	Reading 2015	1.0000	-	-
	Reading 2016	0.8039	1.0000	-
	Reading 2017	0.7752	0.8074	1.0000
	Mathematics 2015	1.0000	-	-
	Mathematics 2016	0.8075	1.0000	-
	Mathematics 2017	0.7444	0.7941	1.0000

Student demographic information

The following list shows the state demographic data fields used for exclusion and matching and the values they can have.

1. Student number of days enrolled in school
    a. 1–180 (within each year)
2. Student grade level (within each year)
    a. 3–8
3. Student gender (within each year)
    a. Female, Male
4. Student race and ethnicity (within each year)
    a. **A** = Asian
    b. **B** = Black or African American
    c. **H** = Hispanic/Latino
    d. **W** = White
    e. **O** = Other
        i. American Indian or Alaska Native (I)
        ii. Native Hawaiian or Other Pacific Islander (P)
        iii. Two or more races within year (M)
        iv. No information provided (N/A)
    f. **M** = More than one race/ethnicity designation across years
5. Student received ESOL services (i.e., English is not student's first language)
    a. No, Yes (within each year)
6. Student received special education
    a. No, Yes (within each year)
7. Economic disadvantage/poverty/low income
    a. No, Yes (within each year)

Determining categories for race/ethnicity was complicated by the fact that the two states categorized race/ethnicity in different ways from each other *and* in different ways across school years. Briefly, State One used 10 different categories or combinations across the years 2013–2018, and State Two used five different categories or combinations (A, B, H, W, O). The categories and combinations for State Two are those provided in the list above and remained consistent across the years (i.e., 2013–2019).

To make State One's categories consistent and comparable, we merged the categories I, P, M, and N/A into a new category O (Other), and we split the category AP (Asian, Native Hawaiian or Other Pacific Islander) into either A or P, using information across years for the same student. This was possible because the AP designation was used in 2013–2015 and was split as A or P for 2016–2018. For instance, a student who was AP in 2015 and P in 2016 was designated P across both years (and thus categorized as O); a student who was AP in 2015 and A in 2016 was designated A across both years.

Pearson

Additionally, we added a category “More than one race/ethnicity designation across years”. For example, if a student’s family identified them as White for five years but then changed their designation to Hispanic/Latino for two years, this student would be categorized as M. (Note that for purposes of matching, where we categorized students as either White/Asian or not White/Asian, this student would be considered not White/Asian.)

Also, it should be noted that when race/ethnicity was missing in one or more years, but available in other years, we applied the rules to the information available to us. For example, if a student’s family identified them as White for five years but then their race/ethnicity designation was missing for two years, this student would be categorized as W.

Matching method

Matching is a nonparametric method of preprocessing data to control for some or all the potentially confounding influence of baseline factors by balancing or equating the treatment (here, Connections Academy students) and control group (here, B&M students). After preprocessing, various methods of analysis can be applied to estimate group differences. Effectively, matching attempts to identify groups of virtually similar students across the treatment and comparison groups where, ideally, the only substantive difference between groups is attending a Connections Academy. A research design such as this (i.e., a quasi-experimental design that closely matches students on prior outcome scores and other demographic variables) is typically considered to provide causal evidence.

In this study, we matched each Connections Academy student to all similar students (i.e., one-to-many matching) attending the same local traditional B&M school district that the Connections Academy student would have attended, at the same grade level and in the same school year in the initial year. Both groups were consistently enrolled each year (150+ days) and state tested in math and reading for four consecutive years. The matching process further incorporated initial year, grade, and initial-year state test scores, as well as demographic variables including economically disadvantaged/low income status, English learner status, special education, gender, and race/ethnicity. Separate matched groups were created for math and reading within each state.

Note that the one-to-many matching process includes replacement, and a particular B&M student can therefore match multiple Connections Academy students and occur more than once in the comparison group. Whenever this was the case, we did not make any adjustments. For State One, the percentage of matching B&M students who matched at least one other Connections Academy student was 7% for math and 8% for reading. For State Two, this percentage was 3% for math as well as for reading. Note that it was rare for a matching B&M student to match more than two Connections Academy students. This occurred only in State One, and for less than 1% of the matching students (each of whom matched three Connections Academy students).

The student matching profiles included the following indicators, for which exact matching was required:

- state of residence
- initial school year (B&M) school district
- initial school year (B&M) grade
  - ensures students are always matched at the same grade level
- initial school year (B&M)
  - ensures students are always matched in the same school year
- economically disadvantaged/low income indicator as defined by the state
  - False if "no" in all school years, or
  - True if "yes" in any of the school years
- English learner
  - False if proficient in all school years, or
  - True if not proficient in any of the school years
- special education indicator
  - False if "no" in all school years, or
  - True if "yes" in any of the school years
- most recent identified gender
- Male gender designation at last time student was tested
- Female gender designation at last time student was tested
- race/ethnicity defined as White/Asian vs non-White/Asian
  - False if not White or Asian in any of the school years, or
  - True if White or Asian in all the school years

Note that, as described above, we used relevant information across every student's full four school years to create, for each student, a single matching value for each of the following demographic variables: economically disadvantaged/low income, English learner status, special education, most recent identified gender and race/ethnicity (indicators 5–9 above).

We adopted this approach because it does not necessarily privilege the first year's information over subsequent years, which could themselves contain more accurate information (especially in light of, for example, changes in response options offered by a state). This approach has one drawback, in that it allows baseline matching to be informed by post-baseline information. We discuss this further in the Limitations section (pp. 43–45).

Caliper matching required:

- B&M initial-year's standardized scaled test score within a range of 0.25 standard deviations.

Note that the actual, "uncoarsened" values for indicators of race/ethnicity and the initial-year state test score were entered into the statistical models to further adjust for any remaining differences in the matched groups. The other indicators were matched exactly on reported levels/categories, but also entered into the statistical models. This is to adjust for any remaining imbalance between the matching variables across the matched pairs (where a *matched pair* consists of a Connections Academy student and their matching B&M students, whose data has been concatenated into a single matching profile as described below).

An imbalance may persist even with exact matching since the number of matches for each profile is not restricted but dictated by the population. This means that a student profile that is more prevalent for Connections Academy students (for instance, profiles including White students or a particular range of first-year test scores), but less prevalent among the B&M students in the matching district, will match to fewer B&M students. Across the matched groups, this can create an imbalance on those individual matching variables. The statistical models adjust for such imbalances.

It should be noted that this method of further adjustment in the statistical model...

- preserves the original exact matching
- does not preclude students from the analytic sample by using an unnecessarily strict matching protocol
- does not place an arbitrary restriction on the number of allowed matches
- provides for any necessary additional statistical controls
- is required by the What Works Clearinghouse Standards Handbook Version 4.0 (p14) where standardized group mean differences for each matching variable used in each analysis are between 0.05 and 0.25.

Students were always matched separately for the reading and math content area tests. This means that a Connections Academy student was typically matched to a different set of B&M students for reading than for math. The coarsening for race/ethnicity to White/Asian or not was determined after looking at the distributions of state test scores. White and Asian students had higher average scores than the other groups, and this was consistent across grade levels and years.

Coarsened Exact Matching (CEM) best describes our general matching procedure (see gking.harvard.edu/cem for more information on CEM). In CEM, the combination of specific levels across the matching variables defines a matching profile for each student, and the matching procedure exactly matches students on these profiles. All students who match each profile are included in the analytic sample. (Note that matching is performed with replacement, such that one B&M student can match multiple Connections Academy students.) The researchers choose both the matching variables and coarsening of some of those variables ahead of time, based upon what they know to be relevant factors and levels. These features combine to increase the generalizability and the efficiency of the resulting causal estimates. This is because...

- we always match exactly within a range, so there are no isolated overly influential cases
- the acceptable imbalance among the groups on each matching variable is set a priori
- the imbalance for one variable does not affect the imbalance for another variable
- no suitable matching student is left out of the comparison
- the number of students in each group is not set by the researchers, but is determined by the data
- the matching process maximizes statistical power to detect group differences
- groups are matched on actual values (not expected values, as with propensity scores), but can still match on an array of variables
- the matching process makes no underlying statistical assumptions
- the unmatched cases can provide information about how reasonable the matching plan and procedures were
- the method approximates a block or stratified randomized design, rather than a general randomized design (e.g., as with propensity score matching).

Pearson

Analysis method

For the main analyses, two series of statistical models, one for reading and one for math, will be used to test group differences on the fourth year (or third Connections Academy year, following an initial B&M year) state test scores, and the outcome of interest. The scores of non-mobile students attending Connections Academy will be compared to the scores of closely matched students attending B&M schools in the same initial district.

The first model in the series will include a treatment by state interaction to test if there is a difference in the treatment effect across the states. If this interaction is not statistically significant, a second model will be fit without the interaction to estimate the effect for both states combined, adjusted by the matching variables.

If the treatment by state interaction is statistically significant, two separate models will be run (without the interaction) separately for each state to estimate the state-specific effects adjusted by the remaining matching variables.

Subsequent, secondary follow-up analyses will apply the same statistical models to the following six dichotomous breakouts of student sub-populations within each state:

1. metropolitan area or non-metropolitan students
2. male or female
3. White-Asian or not
4. special education or not
5. low income or not
6. English learner or not.

All cases of student data that survive the exclusion criteria and matching process will be complete and included in the analytic sample. A hierarchical linear mixed-effects model with random intercepts for school district and matched group (within districts) will be employed to statistically test group mean differences. (Again, note that a matched pair is formed by one Connections Academy student and all their matching B&M students.) Models were fitted in the statistical software R, using the lme4 package.

The structural portion of the statistical model contains fixed effects for the main effect, matching variables' effects, the group by state interaction, and random effects for school district and matched pair (nested within school district).

The statistical model may be defined as...

y(smd) = b0 (intercept) +
b1 * x1 (Connections Academy v. B&M) +
b2 * x2 (State One v. State Two) +
B3 * **X**3 (Connections Academy v. B&M by State One v. State Two interaction effects) +
B4 * **X**4 (race/ethnicity orthogonal effects) +
b5 * x5 (male gender dichotomy) +
b6 * x6 (economically disadvantaged/low income dichotomy) +
b7 * x7 (special education dichotomy) +
b8 * x8 (English learner dichotomy) +
B9 * **X**9 (initial B&M year effects)
B10 * **X**10 (grade level in initial B&M year effects)
b11 * initial-year reading test score(smd) +
b12 * initial-year math test score(smd) +
G1 * **Z**1 school district(d) +
G2 * **Z**2 matched pair nested within school district (md) +
u1(d) +
u2(md) +
r(smd)

...where *y*(smd) = 2016, 2017, 2018, or 2019 population standardized (within grade and year) scaled state test score for student *s* in matched pair *m* in school district *d*; the **B**(1-12) are matrices of individual model fixed effects, with b2 and **B**4 through b12 serving as covariates to adjust the main effect (b1) for differences on state test score due to those variables; the **X** are matrices of orthogonal weights for the fixed effects; b11 and b12 are the parametric semi-partial correlations between initial B&M year reading test score and y and initial-year math test score and y; **G1** and **G2** are parameters for normally distributed random intercepts with variance t1 and t2; **Z1** and **Z2** are matrices of 1s and 0s determining group membership; and r(smd) is a normally distributed residual with variance v. Also note that all continuous and dichotomous fixed effects were mean-centered and the unexplained variance is the sum of the components t1, t2 and v.

Uncoarsened race/ethnicity (i.e., Asian = reference, White, Black, Hispanic, Other, identified with one unique value for each student using their race/ethnicity data across years; see p. 27 for details) and initial B&M year standardized test scores for reading and math, along with the other covariates, are entered into the statistical models to further adjust for any remaining imbalance in the study groups after matching.

Note that for the analysis on fourth-year reading state test scores, caliper matching is performed for first-year reading scores, but not first-year math scores; for the analysis on fourth-year math state test scores, caliper matching is performed for first-year math scores, but not first-year reading scores. However, in both cases, first-year math as well as first-year reading scores are used as covariates in the model. It should also be noted that in the secondary, follow-up analyses where the sample is broken out (e.g., by gender) the breakout variable is no longer included in the statistical model as a covariate.

The statistical tests performed for the main analyses will be adjusted using the Benjamini Hochberg (BH) procedure (Benjamini & Yekutieli, 2001) for controlling the family-wise false discovery rate (FWFDR) at the 5% level. The p-values for each test will be compared to the BH-adjusted critical p-values to determine statistical significance. The FDR is simply the proportion of erroneous significant tests among all tests found statistically significant for a family or grouping of tests. The BH adjustment controls the FDR at a specified level for each family of tests, limiting the probability of declaring comparisons statistically significant when there is no true effect (i.e., making type I errors). For this research, the families of tests are controlled at a FWFDR = 0.05. This means that, on average, we expect to only make a type I error in 1 in 20 families of tests.

The BH procedure is recommended by the WWC as suitable for most cases encountered in educational research. The WWC Procedures Handbook V4 quotes Benjamini and Yekutieli (2001, p1183) to make their point: "a modification of the original BH procedure could be made, although it is very often not needed, and yields too conservative a procedure".

The statistical tests performed for the primary matched groups analyses will be organized into two families. One for reading and one for math state test scores. Within each of these two families, the FWFDR will be controlled for a possible three significance tests, one for the interaction and then one for the combined sample, or two, one for each state. Three comparisons would have the following adjusted p-values in order from smallest to largest observed p-value:

- 0.0167 (lowest observed p-value must clear)
- 0.0333
- 0.05

To add extra support for the main research findings, if we do not find a statistically significant difference in achievement between Connections Academy students and B&M students, we will perform an additional analysis to test for the absence of a meaningful effect. An equivalence test will examine if the data provide evidence for an effect that is smaller than an a priori chosen smallest effect size of interest (SESOI).

Prior to carrying out the research and analyzing the data, we set a SESOI of 0.2 standard deviations (or a difference of 8 percentile points for a 50th percentile score). The equivalence test would actually consist of two directional tests. The first tests the null hypothesis that the effect is at least 0.2 in the direction favoring higher scores for Connections Academy students. (Note that null now refers to 0.2 whereas a traditional null hypothesis refers to 0.0.) The second tests the null hypothesis that the effect is at least 0.2 in the other direction favoring higher scores for B&M students. A conclusion of statistical equivalence is warranted when the larger of the two p-values is less than or equal to 0.05 (Lakens et. al., 2018; see also Magnusson, 2019, for an interactive visualization).

Here, we choose a SESOI of 0.2 standard deviations based on previous research and for practical reasons. We estimate that the smallest width for a 95% confidence band, given our total combined sample size, will be about 0.1 standard deviations. In addition, we feel a SESOI of 0.2 also strikes a balance with the average effect sizes found in the CREDO study (Woodworth et al., 2015), that is, 0.1 for reading and 0.25 for math. Note that choosing an unrealistically small SESOI will predetermine all the equivalence tests to be inconclusive; the tests would lack the statistical power to detect such differences.

Closeness of matched groups

This section demonstrates how closely we were able to match the Connections Academy students to the comparison B&M student group on the individual matching variables. Note that our data is structured into matched pairs, where each matched pair consists of one Connections Academy student and their matching B&M students.

In our main analysis, we incorporate this structure through the random intercept for matched pair in our hierarchical linear regression model. In this model, the weight of each B&M student's data is effectively inverse proportional to the total number of B&M students who match the same Connections Academy student.

In this section, in order to examine how closely matched each matching variable was between the treatment and control group, we therefore apply weighting. For instance, when considering matching by initial-year state test score, we take each Connections Academy student's test score and pair it with a virtual B&M student test score that is computed as the average test score of all the B&M students matching that Connections Academy student. This weighting counts each matched pair equally, regardless of the number of matched B&M students within that matched pair. Note that this is comparable to the weight that each student's data carries in the estimate for the treatment effect in our regression model.

Table 5 shows the reading and math group weighted mean differences for both states on the initial-year state test scores. All groups were closely matched on initial test score. In fact, there is only a 0.00003 to 0.004 standard deviation difference in the matched analytic samples when the scores are weighted equally by matched pair.

For completeness, Table 6 also reports unweighted mean differences, where each student's data has the same weight. Because there is variation in the number of B&M students who match each Connections Academy student, unweighted differences are typically less closely balanced. However, as explained above, the statistical model in our main analysis effectively controls for such imbalances. Table 6 shows that unweighted differences in initial-year state test scores ranged between 0.035 and 0.070, and that these were not statistically significant. Note that they are also well within the 0.25 range recommended by WWC guidelines (What Works Clearinghouse Standards Handbook Version 4.0, p14).

Appendix F (page 68) contains tables detailing the group counts, Cox Index effect sizes and associated p-value for each categorical matching variable, separately for the analytic sample for math and reading. Tables F1 and F3 show that there is no difference between matched groups for the four uncoarsened matching variables (i.e., economically disadvantaged/low income, special education, English learner, gender) and coarsened race/ethnicity (i.e., White/Asian) after weighting each matched group. This is as expected.

When we look at the group balance for uncoarsened race/ethnicity (i.e., each race/ethnicity category individually), Tables F2 and F4 show that there are some differences since the students were not matched on these specific categories and the number of matches for each profile is not restricted but dictated by the population. In State One, the offset is mainly between the B&M group having more matched Black/African American students and the Connections Academy group having more students who indicated multiple races. In State Two, the Connections Academy group has more White students and students who indicated multiple races, while the B&M group has more Hispanic/Latino and Black/African American students. Importantly, our regression model adjusts for these differences because it includes race/ethnicity as a covariate.

Table 5. Matched group weighted differences in initial-year population standardized state test

Subject	State	Group	Count	Mean	Std. deviation	Mean difference	t	P-value
Math	State 1	CA B&M	164 164	0.102 0.099	1.037 1.023	0.0037	0.766	0.445
	State 2	CA B&M	177 177	0.173 0.168	0.916 0.894	0.0043	0.855	0.394
Reading	State 1	CA B&M	157 157	0.197 0.201	0.995 0.982	-0.0042	-0.819	0.414
	State 2	CA B&M	184 184	0.262 0.262	0.787 0.781	-0.00003	-0.006	0.995

Note: Students were matched exactly on initial district and initial grade, and four uncoarsened demographic variables (i.e., gender, English learner status, special education, economically disadvantaged/low income). Students were also matched on two coarsened variables, race/ethnicity (i.e., White/Asian or non-White/Asian) and first-year test score (within 0.25 SD).

For this table, each Connections Academy (CA) student is paired with a virtual B&M student whose test score is computed as the average test score of all the B&M students matching that Connections Academy student. This weighting counts each matched pair equally, regardless of the number of matched B&M students.

Table 6. Matched group unweighted differences in initial-year population standardized state test scores

Subject	State	Group	Count	Mean	Std. deviation	Mean difference	HLM SE	P-value
Math	State 1	CA B&M	164 4706	0.102 0.137	1.037 0.741	-0.035	0.057	0.880
	State 2	CA B&M	177 7380	0.173 0.128	0.916 0.694	0.044	0.044	0.161
Reading	State 1	CA B&M	157 4468	0.197 0.257	0.995 0.752	-0.060	0.058	0.972
	State 2	CA B&M	184 7630	0.262 0.192	0.787 0.574	0.070	0.038	0.419

Note: Students were matched exactly on initial district and initial grade, and four uncoarsened demographic variables (i.e., gender, English learner status, special education, economically disadvantaged/low income). Students were also matched on two coarsened variables, race/ethnicity (i.e., White/Asian or non-White/Asian) and first-year test score (within 0.25 SD).

For this table, each Connections Academy student and each matched B&M student has the same weight. This means that matched pairs for Connections Academy students who have a greater number of matched B&M students count more toward the group mean difference.

SE indicates standard error for the group mean difference; these standard errors result from a hierarchical linear model with a random intercept for school district.

Results

This report looks at state test data from two southern states to address one main and one secondary research question.

Main research question: Do students who previously attended traditional B&M schools, and who enroll for three or more consecutive years at the same Connections Academy school, score as well on standardized state tests of reading and math as demographically similar non-mobile students from those same local B&M schools or districts, and does the result differ for the two states currently under study?

Secondary research question: How do reading and math state test scores for these Connections Academy students compare to B&M students for each level of the following subpopulations?

- metropolitan area (yes, no)
- race/ethnicity (White/Asian, yes or no)
- Disadvantaged/low-income (yes, no)
- Special education (yes, no)
- English learner (yes, no)
- Gender (i.e., male, yes/no)

To examine these research questions, we strictly matched students in elementary grade levels (3rd, 4th, 5th), followed them for three consecutive, consistently non-mobile school years, and then compared their state test scores for math and reading in middle grades (6th, 7th, 8th).

The final analytic sample for math consisted of...

- 341 Connections Academy students (164 in State One and 177 in State Two)
- 12,086 matched B&M students

and the final analytic sample for reading consisted of...

- 341 Connections Academy students (157 in State One and 184 in State Two)
- 12,098 matched B&M students.

We then employed the series of hierarchical linear models described in the Methods section. These models yield estimates of adjusted group mean differences, which correspond to the treatment effect from the statistical models (without interactions) where all other model effects are statistically controlled for. This difference can be viewed as a comparison (Connections Academy minus B&M) of the mean fourth-year test scores after further adjusting for the effects from all the matching variables that remain following our matching procedure.

To support the robustness of our findings, we performed a model specification check by computing the simple weighted group mean difference: the change from initial-year to fourth-year test score, a comparison that is often referred to as a difference in difference, value-added or comparison or pre-post score change.

Key findings

Attending Connections Academy was as effective as attending B&M schools for non-mobile middle-school students' reading state test scores. This was true for both states.

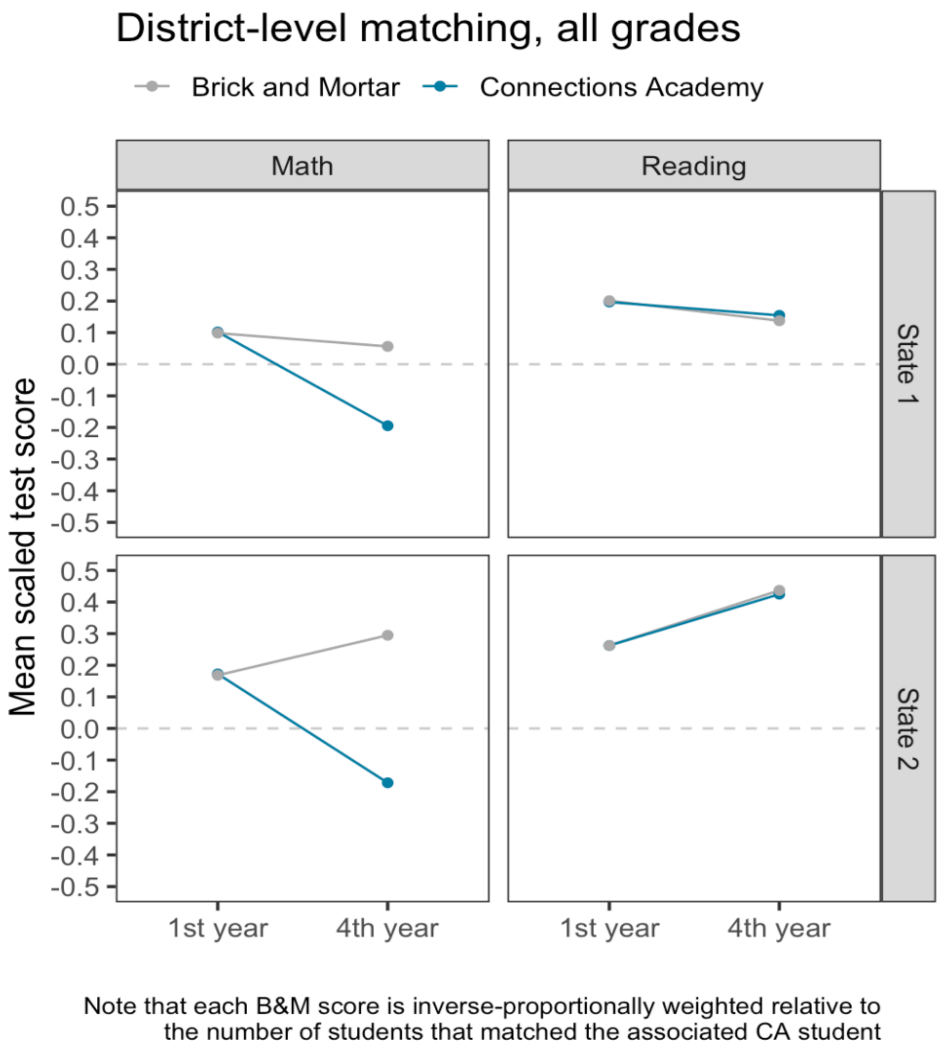
Compared to non-mobile middle school students in brick and mortar schools, attending Connections Academy resulted in lower math state test scores. On average, the performance difference was greater for State Two students (-0.50 standard deviations) than it was for State One students (-0.29 standard deviations).

We found that Connections Academy students perform equivalently to similar local B&M students in their reading state test scores. This was true for both states. We also found that Connections Academy students underperformed in math state test scores when compared to similar local B&M students. This difference was greater for State Two. Secondary analyses suggested that these results, for both reading and math, persist across several subpopulations (for instance, male vs. female students, or metropolitan vs. non-metropolitan students; see Appendix I, page 76).

Figure 1 below shows, separately for each state and for reading and math test scores, the change from initial-year test score to fourth-year test score for our analytic sample, following our matching procedure. We computed the mean test scores represented in this figure by inverse-proportionally weighting each B&M score relative to the total number of B&M students matching the same Connections Academy student. This gives each group of matching B&M students (with one such group per Connections Academy student) the same weight and yields a simple picture of how the means of the groups changed over the treatment period. Note that these mean test scores are not affected in any way by our statistical model.

The difference in fourth-year test scores between Connections Academy and B&M students in Figure 1 corresponds to the weighted group mean difference. Note that these differences correspond closely to our model estimates for the main effect of attending Connections Academy, which adjust for demographic variables and first-year math and reading state test scores. The weighted mean differences in math test scores for State One and State Two in Figure 1 are -0.25 and -0.47 respectively, while the corresponding estimates from our statistical model are -0.29 and -0.50.

Figure 1. Connections vs. brick-and-mortar group differences in population standardized state test scores



Matched group comparisons

Table 7 shows the main results for both the math and reading series of statistical models; math at the top and reading below. Complete model tables can be found in appendices G (for the math analysis) and H (for the reading analysis). Note that these models use the matched (and unweighted) sample consisting of Connections Academy students and their matching B&M students.
The outcome variable of each model is the students' population standardized state test score (see also subsection "State test scores" in the Method section).

The math portion of Table 7 has three rows of output, one each for the State One and State Two models and one for the model with data from both states and a treatment by state interaction term. The interaction term is statistically significant in the latter model (which was actually run first), suggesting that the difference in math state test scores between Connections Academy and B&M students is larger for State Two than it is for State One. We therefore analyzed data from each state separately. The B&M students in State One outperformed their Connections Academy peers by 0.29 standard deviations (or 11 percentile points, relative to the 50th percentile), which we consider a moderate difference. Likewise, the B&M students in State Two outperformed their Connections Academy counterparts by 0.50 standard deviation (or 19 percentile points, relative to the 50th percentile), which we consider a large difference.

The bottom portion of Table 7 shows the results for the reading sample. Here, the interaction was not statistically significant, so we combined the data from both states and analyzed them in the same model. The difference in reading state test scores between Connections Academy and B&M students for the two states' samples combined is 0.01 standard deviations (or less thanhalf of one percentile point, relative to the 50th percentile). This small difference is not statistically significant (or practically significant if a true difference of this size existed).

In addition, an equivalence test (see Lakens et al., 2018) demonstrated that the reading state test scores for the two groups are statistically equivalent. This can be determined most easily by noticing that the 95% confidence interval for the treatment effect, (-0.08, 0.06), falls within the range of our *a priori* determined smallest effect size of interest (SESOI) of 0.20 standard deviations. Note that the groups would still be considered equivalent if the SESOI had been set very low *a priori*, for instance, at 0.10. An equivalence test actually requires only the 90% confidence interval (CI), not the 95%, to fall within the range determined by the SESOI (here, [-0.20, 0.20]), but note that a 90% CI by definition includes the 95% CI. (Here, the 90% CI of the treatment effect was [-0.07, 0.04].)

Table 7. Connections Academy v. Brick-and-mortar group differences in population standardized state test scores

	4th year standardized math score			
Model effect	Treatment effect	SE	95% CI	p-value
CA v. B&M in State One	-0.29	0.05	-0.38 – -0.19	<0.001
CA v. B&M in State Two	-0.50	0.05	-0.60 – -0.40	<0.001
CA v. B&M by State (interaction)	-0.23	0.07	-0.37 – -0.09	<0.001

	4th year standardized reading score			
Model effect	Treatment effect	SE	95% CI	p-value
CA v. B&M, both States combined	-0.01	0.03	-0.08 – 0.06	0.72
CA v. B&M by State (interaction)	-0.05	0.07	-0.18 – 0.09	0.49

Treatment effect refers to the treatment effect from the statistical model (i.e., whether a student attended Connections Academy) described in the Methods section, where all other model effects are statistically controlled for. This difference can be viewed as a comparison (Connections Academy minus B&M) of the mean fourth-year test scores after the effects from all the matching variables that remain following our matching procedure are statistically controlled for.

SE refers to standard error for the group mean difference and ***95% CI*** is the traditional confidence interval around the group mean difference. The standard errors result from a statistical model with random intercepts for school district, and matched pair nested within school district.

Appendices G and H contain the complete statistical model tables for the math and reading analyses. Notice that the first table in each appendix presents the statistical model that includes the treatment by state interaction (which is labeled "Connections Academy by state"). The subsequent tables follow as necessary from the pending statistical significance of this interaction. These tables contain the estimates for the group mean differences (i.e., Connections Academy) in the form of simple main effects.

In a secondary analysis, we examined whether these main results differed across several different subgroups. This required fitting 72 separate statistical models: six subgroups by three models (one for subgroup A; one for subgroup B; one for both subgroups combined, including a subgroup interaction term) by two states by two subject areas. The results can be found in Appendix I (page 76).

Each model included the same covariates as the hierarchical linear model in our main analysis, plus a covariate for the relevant subgroup (e.g., gender). We did not evaluate results for models in which one or both of the subgroups contained less than 30 Connections Academy students because statistical analyses of such small sample sizes would not be informative; they would likely yield non-significant results due to insufficient statistical power. For instance, the combined model for special education subgroups is not reported because only 29 of the Connections Academy students in State One had a special education indicator.

Across both states and both subject areas, with one exception, we did not find a statistically significant difference in the size of the treatment effect between any of the subgroups. The one exception was that the negative impact of attending Connections Academy on math state test scores was larger for female students than for male students[5]. For all subgroup analyses (which were performed on subgroups with at least 30 students), we found a negative impact of attending Connections Academy on state test scores for math but not for reading.

[5] This result should be interpreted with caution because it was the only test that came out statistically significant among a large number of tests. With our false-discovery rate of 5% (alpha = .05), if we only tested data sets with no true effect, we would expect on average one out of 20 tests to yield a statistically significant result.

Pearson

Discussion

Students who attend Connections Academy tend to be highly mobile. Mobility is associated with negative effects on academic performance (Rumberger, 2015), and this makes it hard to evaluate the academic performance of Connections Academy students relative to regular brick-and-mortar (B&M) students. This study therefore focused on Connections Academy students who were not mobile (for at least three consecutive years) and examined the impact of Connections Academy on their reading and math performance after three years in Connections Academy.

We compared reading and math state test scores between non-mobile students who spent three years in Connections Academy (following a year in a B&M school) and non-mobile students from the same initial year B&M school district who stayed in that same school district during those years. In our analyses, the students we compared started from the same place (i.e., district, school year and grade) and we statistically controlled for additional demographic factors that can impact academic performance.

Our results show that Connections Academy students' reading state test scores are statistically equivalent to those of similar local B&M students. We also found that Connections Academy students underperformed in math state test scores relative to similar B&M students. These results come after matching students in elementary grade levels (3rd, 4th, 5th), following them for three consecutive school years, and then comparing their state test scores in middle grades (6th, 7th, 8th).

These results suggest that students who stayed enrolled in Connections Academy for three or more years improved their reading skills to the same degree as students who stayed enrolled in a B&M school. However, the math skills of these students did not improve to the same extent as those of students who stayed enrolled in a B&M school. This pattern of results was found across several demographically defined subpopulations.

Fitzpatrick et al. (2020), who also analyzed virtual-school students as a specific treatment group, found negative effects of attending a virtual school on math state test scores. The -0.50 standard deviations math gap that we found for State Two is identical in size to the math gap reported in Fitzpatrick et al. (2020). However, the math gap we found for State One (-0.29 standard deviations) is smaller in size than that found for State Two or in Fitzpatrick et al. (2020).

Notably, the Connections Academy students from the two states under study in this report performed equivalently in reading compared to B&M students. This contrasts with studies on virtual-school students by Woodworth et al. (2015), which included data from 18 states, and Fitzpatrick et al. (2020), which focused on data from the state of Indiana. These studies found that virtual-school students in general (a group that includes Connections Academy students) performed worse than matched B&M students in reading. In this study, we found that the state reading test performance of Connections Academy students was statistically equivalent (testing against a minimal standardized effect size of interest [SESOI] of 0.20 standard deviations) to that of matched B&M students.

To summarize, whereas prior studies have found negative impacts for virtual-school attendance on both reading and math performance, this study finds negative impacts for Connections Academy attendance only on math performance.

Limitations of the study

First, a limitation of this study is its generalizability to both the general population of Connections Academy students, and the general population of all students. It should be noted that only two Connections Academy schools were included. It is reasonable to assume there is variation across the Connections Academy schools in the students' math and reading state test performance (for example, our results demonstrate a difference in math scores between State One and State Two), and other Connections Academy schools may perform better or worse, compared with each other and with B&M schools.

Second, recall that we examined student performance after four school years and that the state testing programs are limited to grades three through eight. This necessitated matching students when they were in elementary grades (3rd, 4th, 5th) and comparing their performance when they were in middle grades (6th, 7th, 8th). Our results may, therefore, be specific to this subpopulation and not hold for the entire kindergarten through 12th-grade continuum that Connections Academy schools serve.

Third, since we only included three-year stable Connections Academy students, and because Connections Academy students have high mobility, the students remaining in the final analytic sample are not representative of the broader Connections Academy student population. Moreover, students may continue to stay enrolled in Connections Academy for different reasons than B&M enrollment-stable students. For example, families may choose to keep their students enrolled in Connections Academy because they feel the online instructional model works particularly well for them. In contrast, B&M families may keep their students enrolled in their local districts due to the lack of knowledge of alternatives or the inability to take advantage of such alternatives. As a consequence, students for whom Connections Academy's online instructional model works particularly well (and who tend to perform better than other Connections Academy students) may have been overrepresented in our analytic sample.

We know from analyzing responses to annual family surveys that many Connections Academy families feel pushed to find an alternative to their local schools. Connections Academy likely also appeals to more technology-ready families. We also know that a major initial constraint for choosing Connections Academy is the availability and training of the learning coach, typically a caregiver who facilitates their student's learning. To summarize, many factors can influence the choice to attend Connections Academy, and the ability to stay enrolled. We should limit our efficacy statements to families and students who are both 'choosers' and 'stayers'.

Fourth, there are two aspects of the data used in our analyses that require some discussion.

Mobility (see p. 22): because both states' data do not permit us to account for B&M students who changed schools *within* the same district, we have accounted for district-level (rather than school-level) mobility amongst B&M students. This means that the comparisons reported here and elsewhere are less conservative than they could be; for example, if we had been able to match non-mobile Connections Academy students to B&M students who had also remained in a single institution, rather than just in a single district. However, as noted in the second limitation above, the nature of the grade 3–8 sample means that the majority of students in the sample population switched institutions within a district due to the transition from elementary to secondary education.

Unfortunately, these changes are not distinguishable from school changes related to other factors of potential interest, such as availability of school resources; local changes of address that make one school more convenient than another; bullying or other interpersonal causes; and so on.

Although we recognize that this kind of sub-district mobility in the B&M population may have some impact, it is important to note that this limitation is common to all published studies that compare virtual-school students to those attending traditional B&M schools of which we are aware. Indeed, accounting for district-level mobility is a methodological advancement over the current state of the science.

Non-standard grade progression (see p. 23): although students were matched on their initial grade level, a subset of students who skipped or repeated grades during the treatment period had final grade levels that differed. Specifically, 2.6% of matching B&M students do not have a sequential grade progression, and 5.1% of Connections Academy students do not have a sequential grade progression. Although we standardized students' test scores by converting them to z-scores so as to promote comparability, this does not completely address the issue of out-of-level testing. It is possible that a student who is held back or promoted would have earned a different z-score if they had taken the test form associated with their first-year cohort. We nonetheless included these students in our matched analysis because these proportions suggest that the distribution of such students in the analytic sample differs between groups, and their exclusion could thus introduce selection bias to the analysis while also reducing statistical power.

In both the case of mobility and non-standard grade progression, our analytic decisions were based not just on the statistical practicalities of the data, but also on our assessment that the behaviors noted reflect variance inherent to the student experience. As we have noted, we consider that school changes based on the transition from elementary to middle grades – again, the majority driver of sub-district school changes in these data – are typical of the B&M educational experience in this sample's grade range, rather than being an indicator of the kind of mobility for which we would like to account. Similarly, repeating or skipping a grade is not uncommon, and is itself an indicator of academic achievement. In both cases, we feel the group analyses should include these natural sources of variance, which the reader can consider along with the other evidence presented.

Finally, we wish to address an aspect of our analytic approach that bears further consideration: the method by which we matched B&M students with corresponding Connections Academy students (see the “Matching method” section, pp. 26–28). As we describe, each student's demographic variables used in the matching process were derived from all four years of data, rather than according to the first ("baseline") year in a given four-year period.

Our goal in this was practical – we hoped to achieve consistency in each student's data within a given four-year window. There are many potential reasons for such inconsistencies in the demographic variables – including, but of course not limited to, differences in how the states collected these data – and it is beyond the scope of this study to speculate about these. However, one consequence of this approach is that, to the extent that group membership may influence a demographic variable, there is an increased risk for bias in the subsequent analyses. For example, if attending Connections Academy affects a student's designation as an English learner, then using a "post-treatment" (i.e., post-baseline) English learner designation to determine matching students is potentially problematic. That is, if students at B&M schools are less (or more) likely to be classified as an English learner than students attending Connections Academy, this could bias any effect we might otherwise attribute to attending Connections Academy in an unknown direction.

In order to assess the possibility of bias due to our matching process, we considered the absolute value of the difference between the baseline and the final values of the demographic covariates for each student group, for each standardized assessment; that is, abs((Treatment Baseline – Final) – (Control Baseline – Final)). For example, for the gender covariate, the first-year baseline for math rate for B&M (Control) students included in this study was 45.6% male, while the baseline for Connections Academy (Treatment) students was 45.7% male, and the final rate for the sample after including post-treatment information was 45.7%. Accordingly, the differential of the change between the two groups is abs((45.7 – 45.7) – (45.6 – 45.7)) = 0.1%, suggesting that the use of the post-treatment information is unlikely to have introduced significant bias.

Similarly, the differentials of two other demographic variables – English learner status, and special education status – were consistently small (< 1%), and therefore not potentially problematic. The differentials for the remaining covariates – race (white/Asian or not), 5.5%; and economically disadvantaged/low-income status, 5.7% – raise the possibility that our matching approach could have introduced troublesome bias on the basis of these variables. The pattern was similar for the reading sample, with gender, English learner status, and special Education status (< 1%), and race (5.5%) and economically disadvantaged/low-income status (5.2%) raising concern.

However, having said that, we note that the nature of the race and income covariates renders this concern less worrisome. It is difficult to see, for example, how attending Connections Academy could influence a student's self-reported race/ethnicity information, not least because state reporting standards are uniform for all schools (B&M and virtual). Similarly, it is challenging to imagine that participation in Connections Academy could cause a student's likelihood of needing economic need-based assistance to increase or decrease. Thus, although readers should consider this study's results in light of the fact that our method of matching students does not rely exclusively on pre-treatment "baseline" data, we believe it is unlikely that it creates a significant source of bias.

Implications of findings for product implementation and further research
The mixed results seen in this study are in line with much of the research on virtual schools already published. Although this study compares student achievement between Connections Academy and B&M students, it tells us nothing about why observed differences arise or exist, or how virtual education may be improved.

This study did not examine or take into account the implementation fidelity of the Connections Academy core instructional model. It is reasonable to assume that schools implementing the core model with greater fidelity may have better results than those that do not, and vice versa. These would all be important areas for future research.

Pearson

References

Ahn, J. (2016). Enrollment and Achievement in Ohio's Virtual Charter Schools. Fordham Institute.

Connections Academy. (2016, March 23). Retrieved from https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/efficacy-and-research/reports/Connections-Research-Report.pdf/

ASCD (2010). Giving Students Meaningful Work. Educational Leadership, (68)1. ASCD (2011).

Effective Grading Practices. *Educational Leadership*, (69)3.

ASCD (2012a). College, Careers, Citizenship, (69)7.

ASCD (2012b). Feedback for Learning. *Educational Leadership*, (70)1.

ASCD (2013a). Faces of Poverty. *Educational Leadership*, (70)8.

ASCD (2013b). Resilience and Learning. *Educational Leadership*, (71)1.

ASCD (2013c). Getting Students to Mastery. *Educational Leadership*, (71)4.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, 13(2), 111–127.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4), 1165-1188.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: *A longitudinal study and an intervention. Child development*, 78(1), 246–263.

Bransford, J., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C: National Academy Press.

Burns, M. K., Appleton, J. J., & Stehouwer, J. D. (2005). Meta-analytic review of responsiveness-to-intervention research: Examining field-based and research-implemented models. *Journal of Psychoeducational Assessment*, 23(4), 381–394.

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental psychology*, 56(4), 236–246.

Diener, C. I., & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 36(5), 451–462.

Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology*, 99(3), 597.

Dweck, C. S. (2006). Mindset: *The new psychology of success*. New York: Random House.

Eccles, J. S. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), Achievement and achievement motives: Psychological and sociological approaches (pp. 75-146). San Francisco, CA: Freeman.

Engle, R. A., Nguyen, P. D., & Mendelson, A. (2011). The influence of framing on transfer: *Initial evidence from a tutoring experiment. Instructional Science*, 39(5), 603–628.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3), 363.

Evergreen Education Group (2017). Why do students choose blended and online schools? The “end of average” requires personalized learning environments. Retrieved May 11, 2020, from https://www.gettingsmart.com/2017/01/why-students-choose-online-blended-schools/

Fitzpatrick et. al., (2020). Virtual Illusion: Comparing Student Achievement and Teacher and Classroom Characteristics in Online and Brick-and-Mortar Charter Schools. *Educational Researcher,* 49(3), 161-175.

Ford, R., & Rice, K. (2015). Value-Added Results for Public Virtual Schools in California. *Educational Technology & Society*, 18(4), 412–423.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2012). The early prevention of mathematics difficulty: Its power and limitations. *Journal of Learning Disabilities*, 0022219412442167.

Gatti, G. (2018). A Comparison Study of Connections Academy Schools to Matched Brick and Mortar and Virtual Schools, Examining the Types of Students Who Attend K-12 Virtual School and the Effects on Performance of a Highly Mobile Student Body.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents’ standardized test performance: *An intervention to reduce the effects of stereotype threat. Applied Developmental Psychology*, 24, 645–662.

Green, E. & Goldstein, D. (2019, October 30). Reading Scores on National Exam Decline in Half the States. The New York Times. Retrieved from https://www.nytimes.com

Hess, F. M., & Saxberg, B. (2014). *Breakthrough leadership in the digital age: Using learning science to reboot schooling*.

Hulleman, C. S., & Harackiewicz, J. M. (2009). *Promoting Interest and Performance in High School Science Classes.*

Science, 326(5958), 1410–1412. http://doi.org/10.1126/science 1177067.

iNACOL (2015). How to Start an Online Learning Program. Retrieved March 17, 2016, from https://www.onlineprogramhowto.org/teachers/role-of-teacher/

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259-269.

Lueken, M. et. al. (2015). Value-added in a virtual learning environment: an evaluation of a virtual charter school. *Journal of Online Learning Research*, 1(3), 305-335.

Magnusson, K. (2019). Equivalence, non-inferiority and superiority testing an interactive visualization. Retrieved from https://rpsychologist.com/d3/equivalence/

Molnar, A., Miron, G., Elgeberi, N., Barbour, M.K., Huerta, L., Shafer, S.R., Rice, J.K. (2019). Virtual Schools in the U.S. 2019. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/virtual-schools-annual-2019

Sass, T. R. (2019). The performance of state charter schools in Georgia, 2017/2018. Atlanta, GA: Andrew Young School of Policy Studies, Georgia State University.

Shechtman, N., DeBarger, A.H., Dornsife, C., Rosier, S., Yarnall, L. (2013). Promoting Grit, Tenacity, and Perseverance: Critical Factors for Success in the 21st Century. U.S. Department of Education: Office of Educational Technology.

Shinn, M. R. (2010). Building a scientifically based data system for progress monitoring and universal screening across three tiers, including RTI using curriculum-based measurement. *Interventions for achievement and behavior problems in a three-tier model including RTI*, 259–292.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153–189.

Tran, L., Sanchez, T., Arellano, B., & Swanson, H. L. (2011). A meta-analysis of the RTI literature for children at risk for reading disabilities. *Journal of Learning Disabilities*, 44(3), 283–295.

Williams, J. J., Paunesku, D., Haley, B., & Sohl-Dickstein, J. (2013). Measurably Increasing Motivation in MOOCs. In: MOOCshop Workshop, International Conference on Artificial Intelligence in Education, Memphis, TN.

Woodworth, J. L., Raymond, M. E., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W., & Van Donge, C. (2015). Online charter school study. Stanford, CA: Center for Research on Education Outcomes. Retrieved from: https://credo.stanford.edu/sites/g/files/sbiybj6481/f/online_charter_study_final.pdf

Pearson

Appendix A. Data-cleaning process

State One

We start with all 4,951,013 shared records. The records are for students from pre-K (up to three years before kindergarten) to 12th grade. Students are followed across years, can have multiple records per year, and each record may include scores for one or more state tests.

Across these records, there were 15,275 unique students with at least one year in which they were enrolled in Connections Academy for some days. Of these students, 11,336 were enrolled for some days in one or more schools other than Connections Academy (e.g., brick-and-mortar [B&M], other virtual, non-virtual charter) in at least one other year.

There were 1,198,727 unique students with at least one year in which they were enrolled for some days in one or more schools other than Connections Academy.

(There were 52,559 unique students with at least one year in which their school type could not be determined; that is, all records listed NA for days enrolled.)

This table summarizes the stepwise process of filtering the student data to include only records that meet the criteria of our study.

	Filtering step	Records excluded	Cumulative exclusions	Remaining
1	Import original data	0	0	4,951,013
2	Exclude grades < 3 and > 8	2,733,508	2,733,508	2,217,505
3	Exclude charter schools by school ID	68,328	2,801,836	2,149,177
4	Exclude charter schools by name	582	2,802,418	2,148,595
5	Exclude mobile students, < 150 days enrolled	280,077	3,082,495	1,868,518
6	Exclude students, > 180 days enrolled	13,957	3,096,452	1,854,561
7	Exclude students tested in multiple grade/year	220	3,096,672	1,854,341
8	Exclude records without both relevant tests	28,692	3,125,364	1,825,649
9	Exclude students w/ multiple records in a year	590	3,125,954	1,825,059
10	Exclude students with < 4 tested years	810,090	3,936,044	1,014,969
11	Exclude students with < 4 consecutive tested years	45,161	3,981,205	969,808

After row 10 in tabled exclusions, 1,014,969 records remain, 3,381 records belong to Connections Academy students, and 1,011,588 to B&M students.

Across these records, there are 216,398 unique students with at least one B&M year and 1,573 unique students with at least one Connections Academy year.

After row 11 in tabled exclusions, 969,808 records remain; 3,018 records belong to Connections Academy students, and 966,790 to B&M students. Across these records, there are 205,561 unique students with at least one B&M year and 1,352 unique students with at least one Connections Academy year.

	Filtering step *(B&M = brick-and-mortar; CA = Connections Academy)*	**Records excluded**	**Cumulative exclusions**	**Remaining**
12	School type across years for a student contains a valid B&M group or CA group sequence	3,691	3,984,896	966,117
	School type sequence contains ...	Action	Number of unique students	Number of records
	B&M CA CA CA **Note that this sequence can be contained in a sequence that is longer than 4 years, such as the 5-year sequence CA B&M CA CA CA.* *Thus, the number of records can be more than the number of unique students x 4 (years).*	Keep	180	915
	B&M B&M B&M B&M **Note that a 4-year BM sequence can be contained within a 5-year or 6-year B&M sequence. For instance, the 5-year sequence B&M B&M B&M B&M B&M contains two 4-year B&M sequences.* *Thus, the number of records can be more than the number of unique students x 4 (years).*	Keep	204,808	965,202
	Other sequence, with at least one CA	Exclude	825	3,691
		Total remaining	204,988	966,117
	Filtering step *(B&M = brick-and-mortar; CA = Connections Academy)*	Records excluded	Cumulative exclusions	Remaining
13	School type across years for a student contains a 4-year sequence within the same district for B&M group	54,559	4,039,455	911,588
	School type sequence contains ...	Action	Number of unique students	Number of records
	B&M CA CA CA	Keep	180	915
	a B&M B&M B&M B&M sequence in the same district	Keep	194,337	910,643
	no B&M B&M B&M B&M sequence in the same district	Exclude	10,471	54,559
		Total remaining	194,517	911,588

Pearson

State Two Department of Education data

	Filtering step	Records excluded	Cumulative exclusions	Records remaining
1	(Import original data)	0	0	24,664,645
2	Exclude duplicate records	11,533	11,533	24,653,112
3	Exclude ID "273146089"*	157,214		24,495,898
4	Exclude records that don't have a math or reading test score	9,588,212		14,907,686
5	Exclude within year all records for students who did not test on math and reading that year	951,155		13,956,531

** Note that the Department of Education (DOE) explained that this ID was used to group together all students whom they could not successfully match within their system.*

	Determine valid reading test scores			
	Filtering step	**Records excluded**	**Cumulative exclusions**	**Records remaining**
1	Select all records with a reading test score (including redacted ones)	554,889		13,401,642
2	Exclude records which were not the earliest testing opportunity within that year	926,836		12,474,806
3	Keep only records for which students were not mobile	490,630		11,984,176
4	Exclude records if student had multiple test scores for earliest testing opportunity	26,751		11,957,425
5	Keep only records if student also had a valid reading record	4,001		11,953,424 *

	Determine valid math test scores			
	Filtering step	**Records excluded**	**Cumulative exclusions**	**Records remaining**
1	Select all records with a math test score (including redacted ones)	837,065		13,119,466
2	Exclude records which were not the earliest testing opportunity within that year	641,953		12,477,513
3	Keep only records for which students was non-mobile	491,007		11,986,506
4	Exclude records if student had multiple test scores for earliest testing opportunity	25,426		11,961,080
5	Keep only records if student also had a valid reading record	7,656		11,953,424 *

** Note that this number matches that at the bottom row of the table above.*

	Determine valid B&M potential matching records			
	Filtering step	**Records excluded**	**Cumulative exclusions**	**Records remaining**
1	Keep only valid math and reading test scores, as obtained from all previous filtering steps	0	0	11,953,424
2	Exclude charter schools, virtual schools, and CA	458,000		11,495,424
3	Exclude instances where student did not test for math and reading in the same district	36		11,495,388
4	Exclude records associated with districts in which student did not test for 4 or more consecutive years This leaves, for each student, only data associated with the district in which they tested in reading and math for 4 or more consecutive years (if any such district exists, for the student)	6,201,605		5,293,783

** Note that this number matches that at the bottom row of the table below.*

State Two Connections Academy data

	Determine valid CA reading test scores (derived from CA test score records)				
	Filtering step	**Records excluded**	**Cumulative exclusions**	**Records remaining**	**Number of students remaining**
1	Import original data, math and reading	0	0	15,260	7,947
2	Exclude records which do not have a reading score	2,696		12,564	7,674
3	Exclude records which were not the earliest testing opportunity within that year	855		11,709	7,609
4	Exclude, by year, students with multiple test scores within that year	12		11,697	7,605

	Determine valid CA math test scores (derived from CA test score records)				
	Filtering step	**Records excluded**	**Cumulative exclusions**	**Records remaining**	**Number of students remaining**
1	Import original data, math and reading	0	0	15,260	7,947
2	Exclude records which do not have a math score	2,523		12,737	7,307
3	Exclude records which were not the earliest testing opportunity within that year	1,678		11,059	7,248
4	Exclude, by year, students with multiple test scores within that year	10		11,049	7,246

Determine valid CA math and reading test scores (derived from CA data)				
Merge reading score and math score records	**Valid reading score**	**Valid math score**	**Valid math and reading score**	**Number of students remaining**
	11,697	11,049	10,965	7,193

Filtering step	Records excluded	Cumulative exclusions	Records remaining	Number of students remaining
Valid math and reading score	0	0	10,965	7,193
Exclude records which cannot be linked to demographic information (provided by CA)	101		10,864	7,130
Keep only records for the initial 3 years for students who tested in reading and math for 3 or more consecutive years	8,233		2,631	877
Exclude records by student if their testing scores cannot be linked to TCH demographic information (by year)	27		2,604	868
Exclude records by student if the student was mobile in one or more of the 3 years (*mobility score derived from CA demographic data)	228		2,376	792

Linking Connections Academy data to State Two Department of Education data

Data processing step	Students excluded	Cumulative number of students excluded	Number of students remaining
Valid math and reading score in CA data	0	0	792
Keep students whose CA test scores can be linked to DOE data (by way of "score profile matching"; see below)	163		629
Exclude students who did not test at the earliest testing opportunity within one or more years	0		629
Exclude students who did not have valid test scores in a B&M school in the year prior to their first year in CA (because they were not in grades 3–8, or in a different state)	392		237
Exclude students whose test score(s) in their B&M year were ***redacted***	19		218

Redacted scores in Department of Education data from State Two
In the data that State Two's Department of Education (DOE) shared with us, some test scores were redacted. An employee of the agency explained this as follows:

> *"To comply with FERPA, [the Family Educational Rights and Privacy Act] the agency masks test results in the [...] student-level data. The results for each subject are masked (set to blank) if fewer than five students in a unit of masking are tested (have a score code of S). The unit of masking for each subject is a Concatenated Masking Variable (CMVAR) created by concatenating the masking variables specified in the request.*
>
> *"For example, the unit of masking for reading is created by concatenating the district, all the demographic fields, test version, test language and test accommodations. We count the number of students tested (have score code of S) in each unit of masking (CMVAR) for each subject and set the results to blank if fewer than five students are tested."*

Records associated with redacted test scores therefore contain all demographic variables and some test information (test version and test language), but not the actual test score. Across all test scores that were shared by the DOE, 21% of reading scores and 22% of math scores were redacted. However, across the subset of test scores that were associated with Connections Academy, 39% of reading scores and 43% of math scores were redacted. Note that the percentage of redacted scores is almost twice as high for Connections Academy students compared to the average student in the state.

The proportion of Connections Academy students for whom all scores associated with Connections Academy were redacted was 33% for math and 31% for reading. In order to ameliorate this missing test scores issue, we obtained all test scores directly from State Two's Connections Academy and tried to connect those to the data obtained from DOE using a matching procedure that we call "score profile matching".

Score profile matching procedure

For each remaining Connections Academy student, we used their Connections Academy data to construct a "score profile" for each year (N = 792; see table above). This profile included the year they tested, the grade level they tested at and their testing scores for both reading and math (including the total score as well as test subscores, 3–5 per test).

For example, a student who tested in 2016 at 5th-grade level could have the following Connections Academy score profile:

```
year = 2016
grade = 5
reading_category1 = 6
reading_category2 = 11
reading_category3 = 13
reading_raw = 30

math_category1 = 5
math_category2 = 9
math_category3 = 5
math_category4 = 3
math_category5 = NA
math_raw = 22
```

(Note that there is no score for the fifth math category because there were only four math category scores for 5th-grade students in 2016.)

Next, we constructed score profiles for students in the DOE data who attended Connections Academy. We then tried to link up unique score profiles across the two data sources, in order to obtain each Connections Academy student's unique DOE ID.

First, we excluded data (for one specific year) for each of five students whose profile for that year is not unique in the population of Connections Academy profiles for that year. (Note that this leaves three years of Connections Academy data for each of those five students.) We then tried to match the remaining score profiles. We successfully matched the score profile for at least one year for 629 out of the 792 remaining Connections Academy students. Thus, score profile matching retrieved the unique DOE ID for 79.4% of the Connections Academy students.

Using this DOE ID, we were then able to retrieve DOE demographic information for each of the three years that the Connections Academy student tested. Through score profile matching, we were able to analyze data for a substantially larger group (i.e., 80% instead of only 40%) of Connections Academy students than we would have if we only worked with DOE data (in which about 40% of test scores are redacted for Connections Academy students).

Pearson

Mobility data for State Two's Connections Academy students
DOE's mobility variable is categorical and based on how many days a student was enrolled in a particular school district in a particular year. A student is coded mobile in a school district if the student was enrolled in that district for less than 150 school days that school year. Unfortunately, mobility data was missing for 86% of the Connections Academy students' records with test scores in the DOE data and therefore not reliable.

Because DOE was unable to explain these missing data, we instead relied on mobility data obtained from State Two's Connections Academy. Unfortunately, the data that Connections Academy shared with us did not list the total number of days that a student was enrolled in Connections Academy in a particular year. However, Connections Academy did share each student's exact enrollment and withdrawal dates for each school year.

Using these dates, we were able to compute the number of days a student was enrolled in Connections Academy for each school year as the number of days that elapsed between their enrollment and withdrawal date minus weekends, holidays, and hurricane days for that specific school year (as obtained from a large independent school district's academic calendar that we were able to retrieve online for each school year).

Appendix B. Statewide student demographic information for State One

Table B1. Statewide mobility for 3rd through 8th-grade students with state test scores for both reading and math by school type by year for State One

Demographic	Year	Connections Academy	Brick-and-mortar	Other virtual schools
% Mobile	2013	24.2	6.5	23.3
	2014	13.9	5.9	25.8
	2015	15.9	5.9	21.8
	2016	15.6	5.5	21.2
	2017	13.3	5.4	22.8
	2018	16.1	5.5	22.8

These results are for the population of students who tested for both math and reading within a year. A student was counted as mobile if they were not enrolled in any one school within that year for at least 150 days.

Figure B1. Statewide race/ethnicity for 3rd through 8th-grade students with state test scores for both reading and math by school type by year for State One

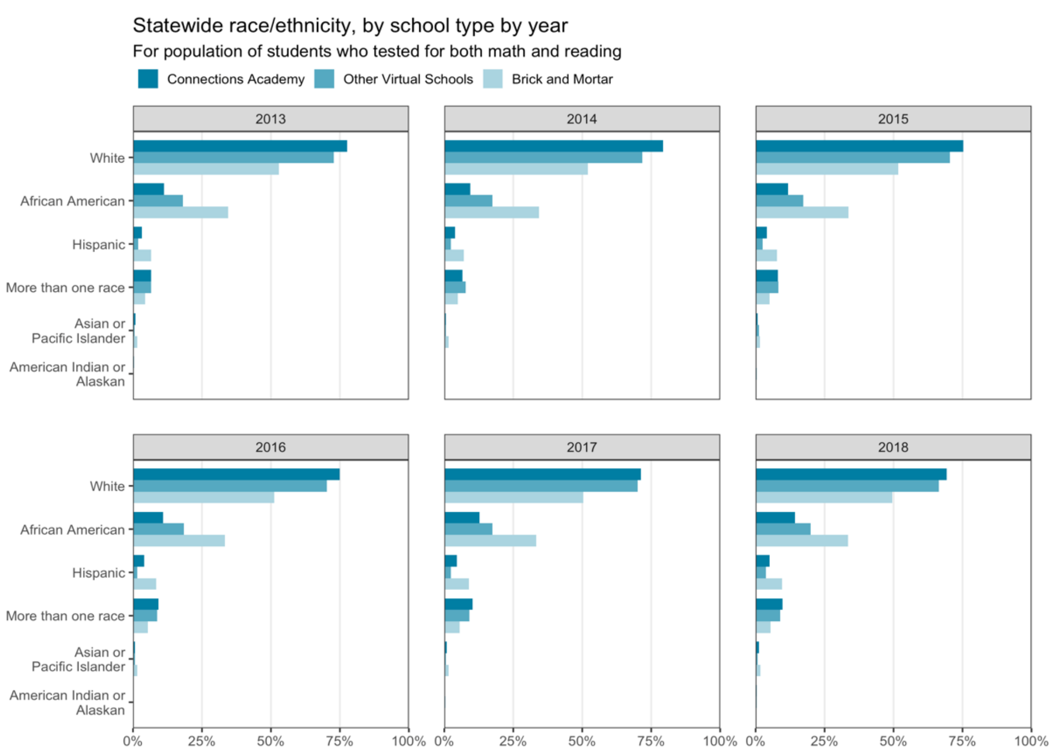
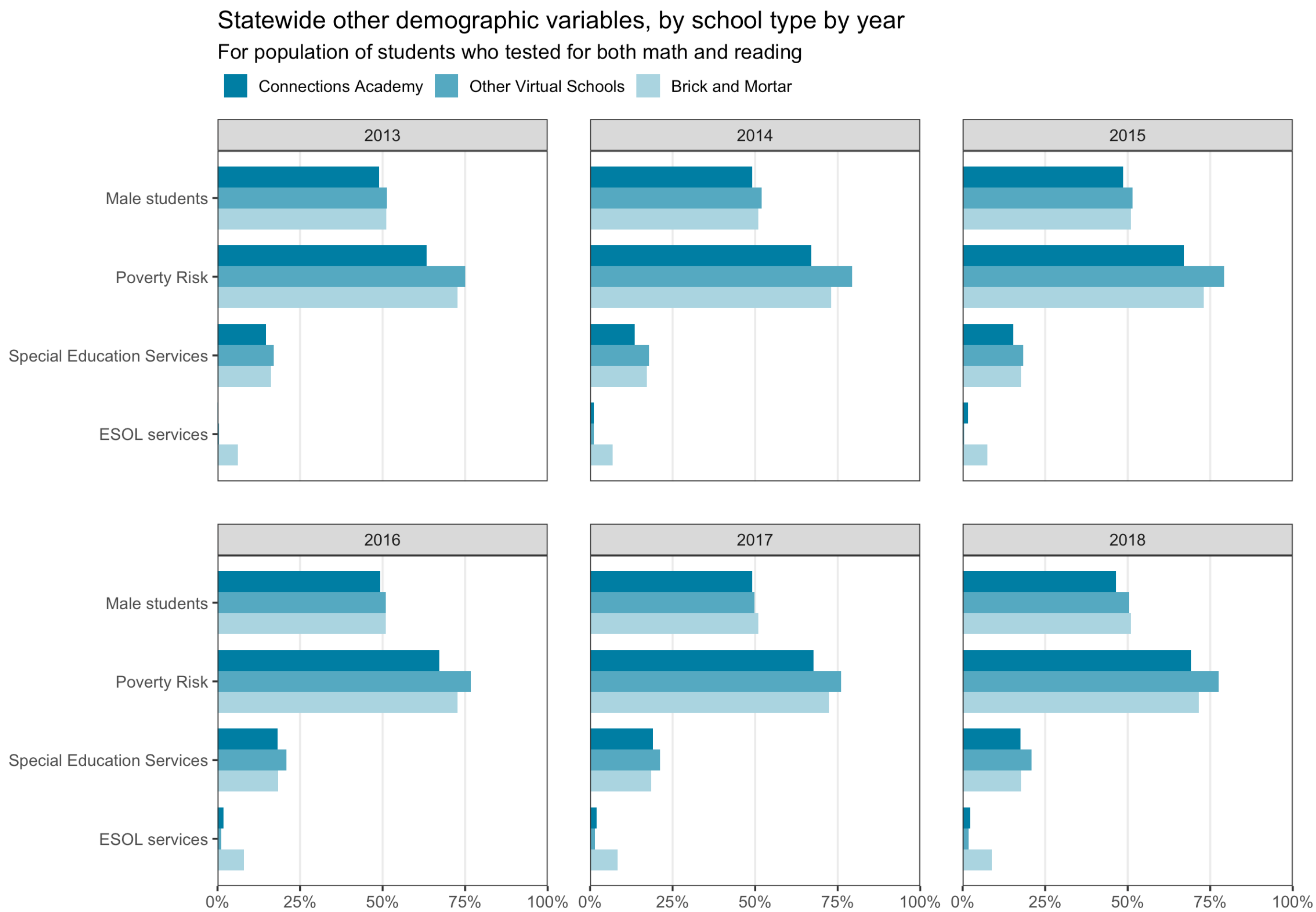


Figure B2. Statewide other demographic information for 3rd through 8th-grade students with state test scores for both reading and math by school type by year for State One



Appendix C. Statewide student demographic information for State Two

Figure C1. Statewide race/ethnicity for 3rd through 8th-grade students with state test scores for both reading and math by school type by year for State Two

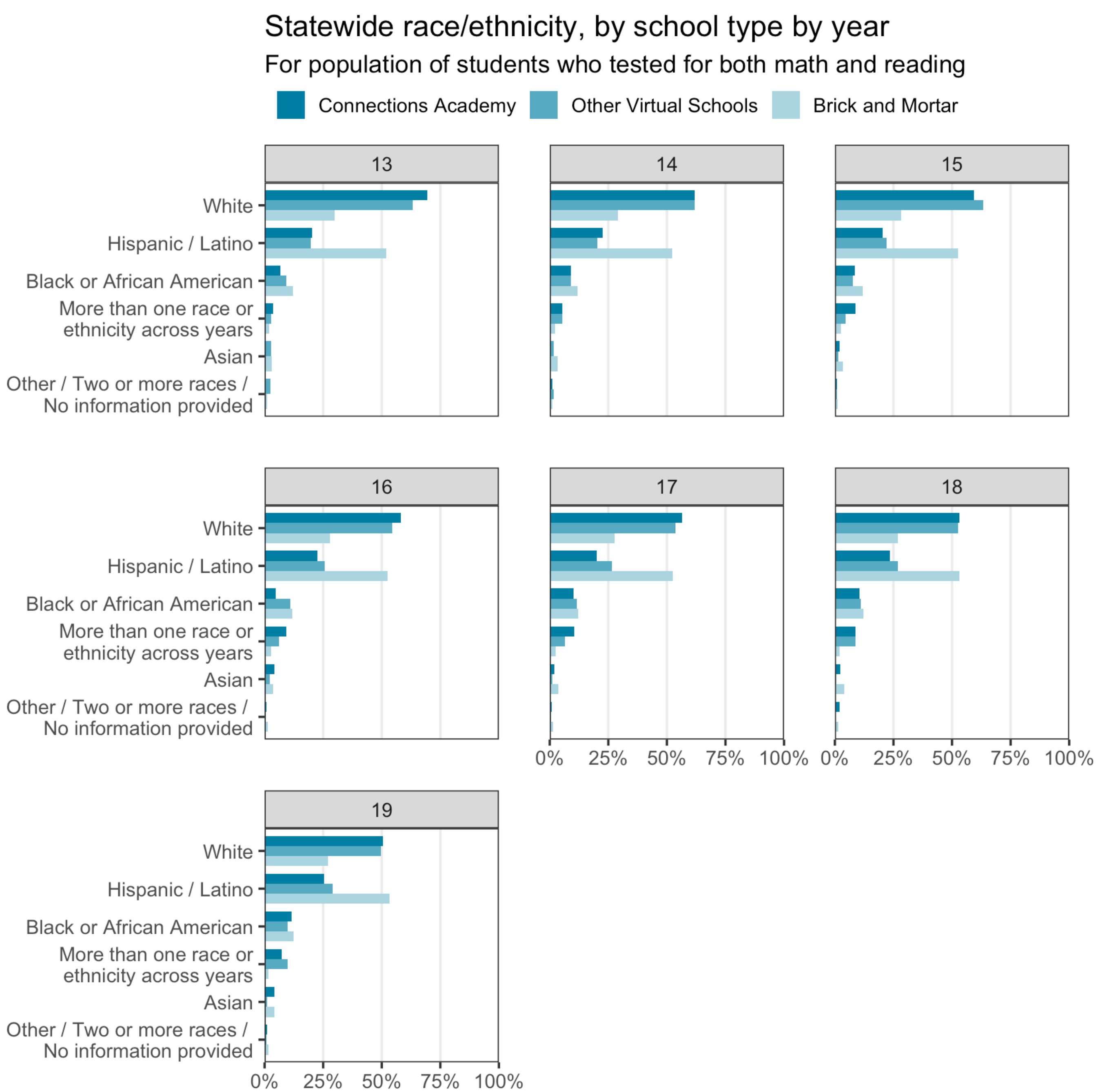
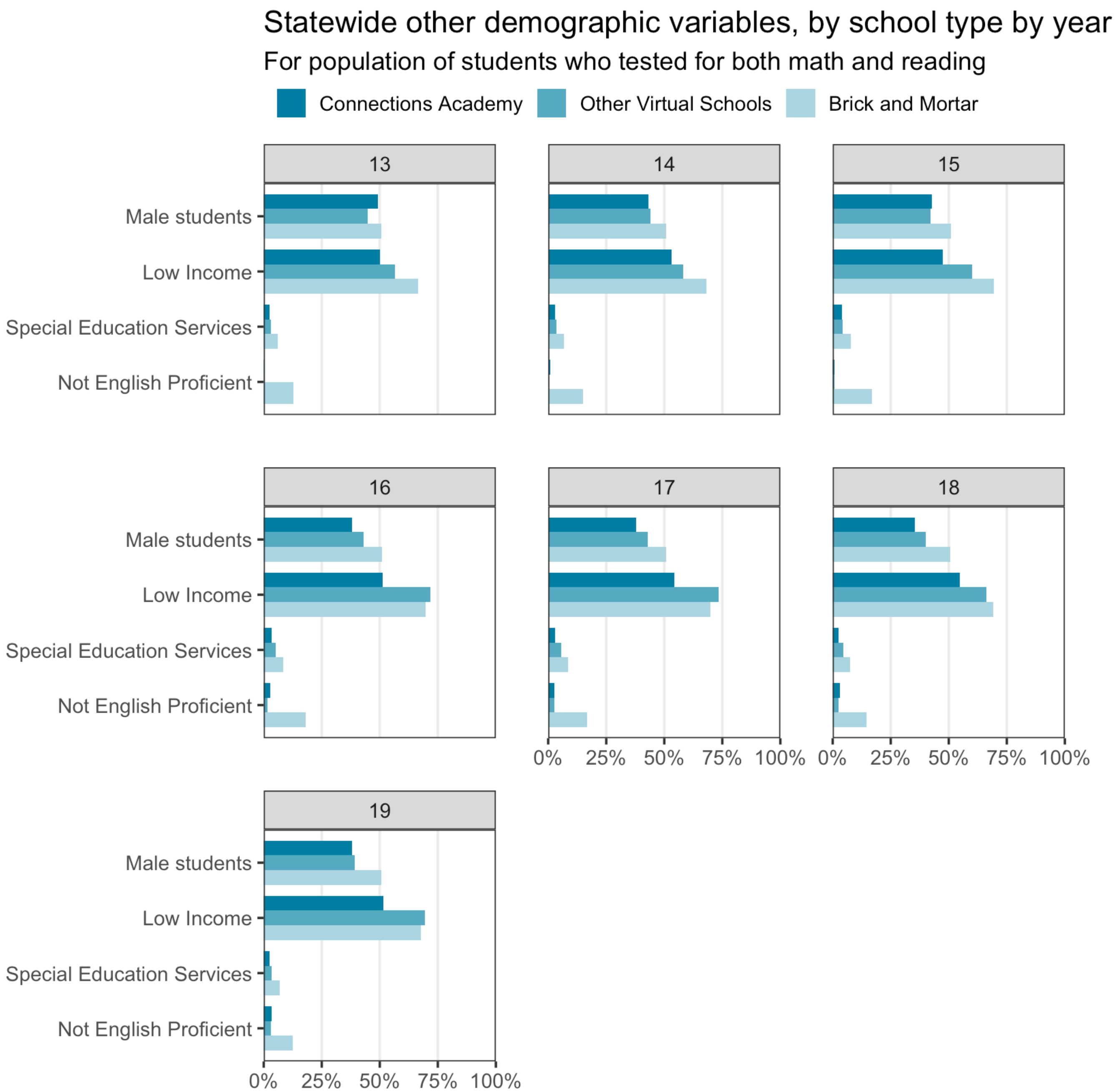
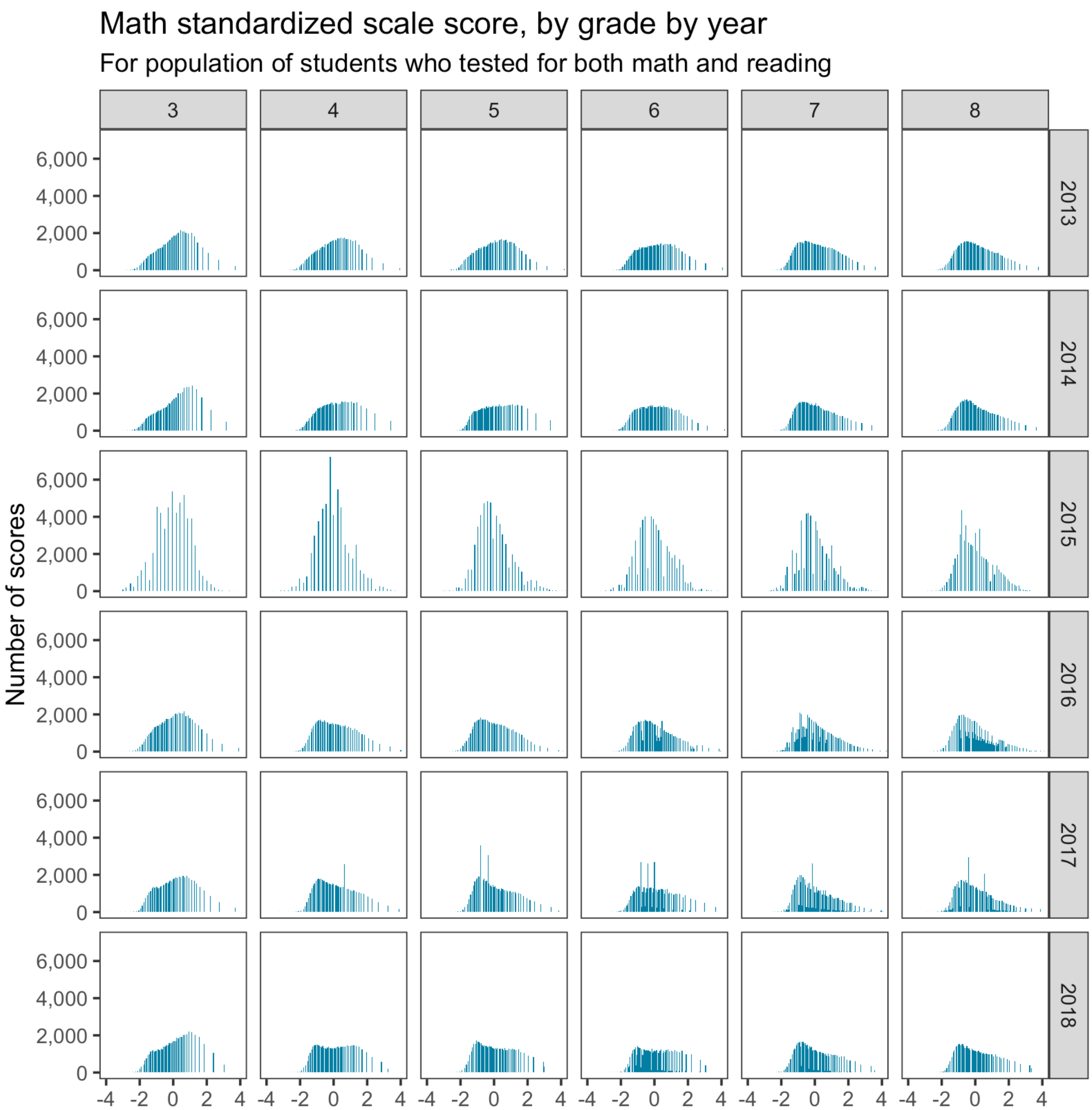


Figure C2. Statewide other demographic information for 3rd through 8th-grade students with state test scores for both reading and math by school type by year for State Two

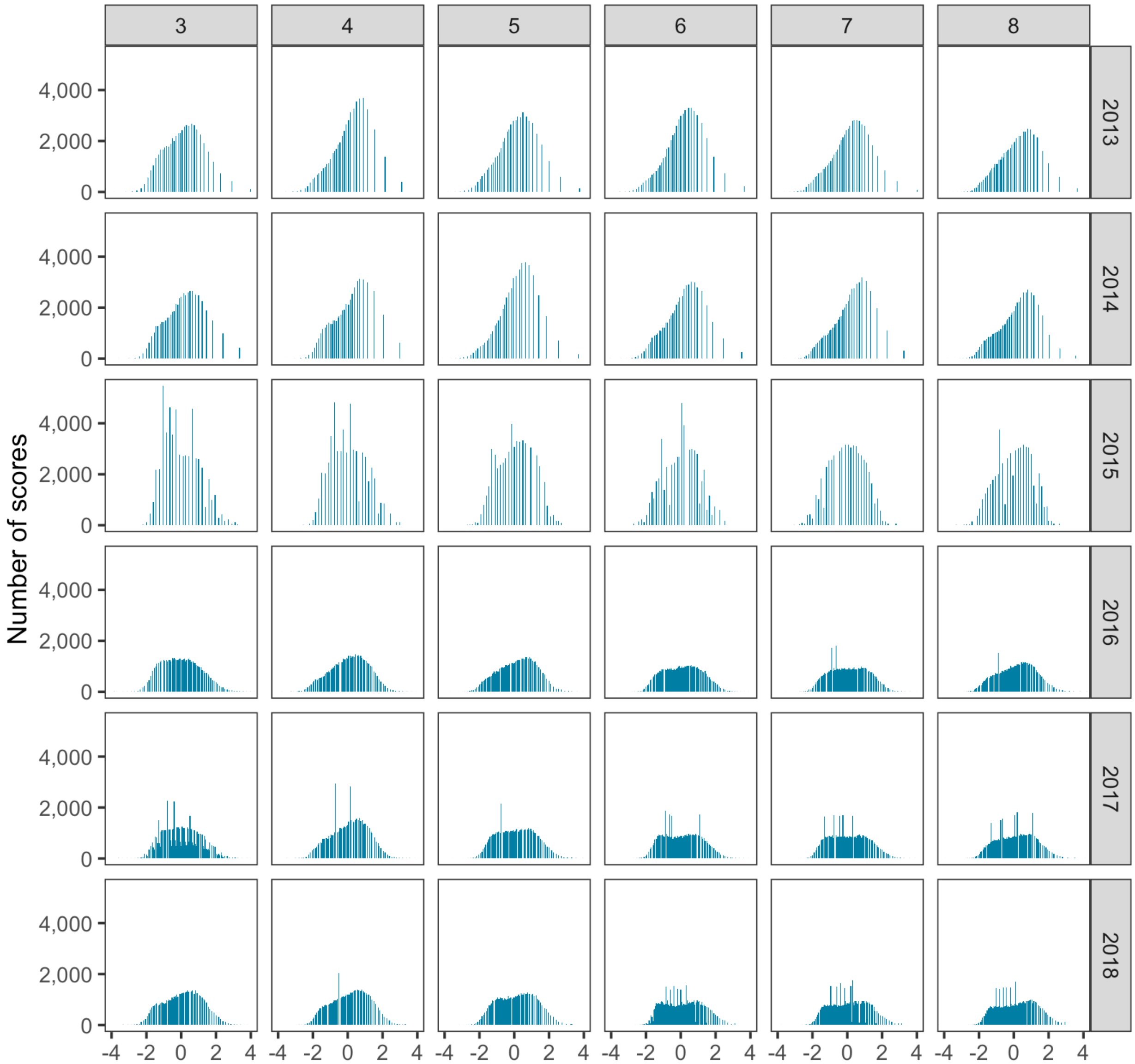


Appendix D. Standardized state test score distributions for State One

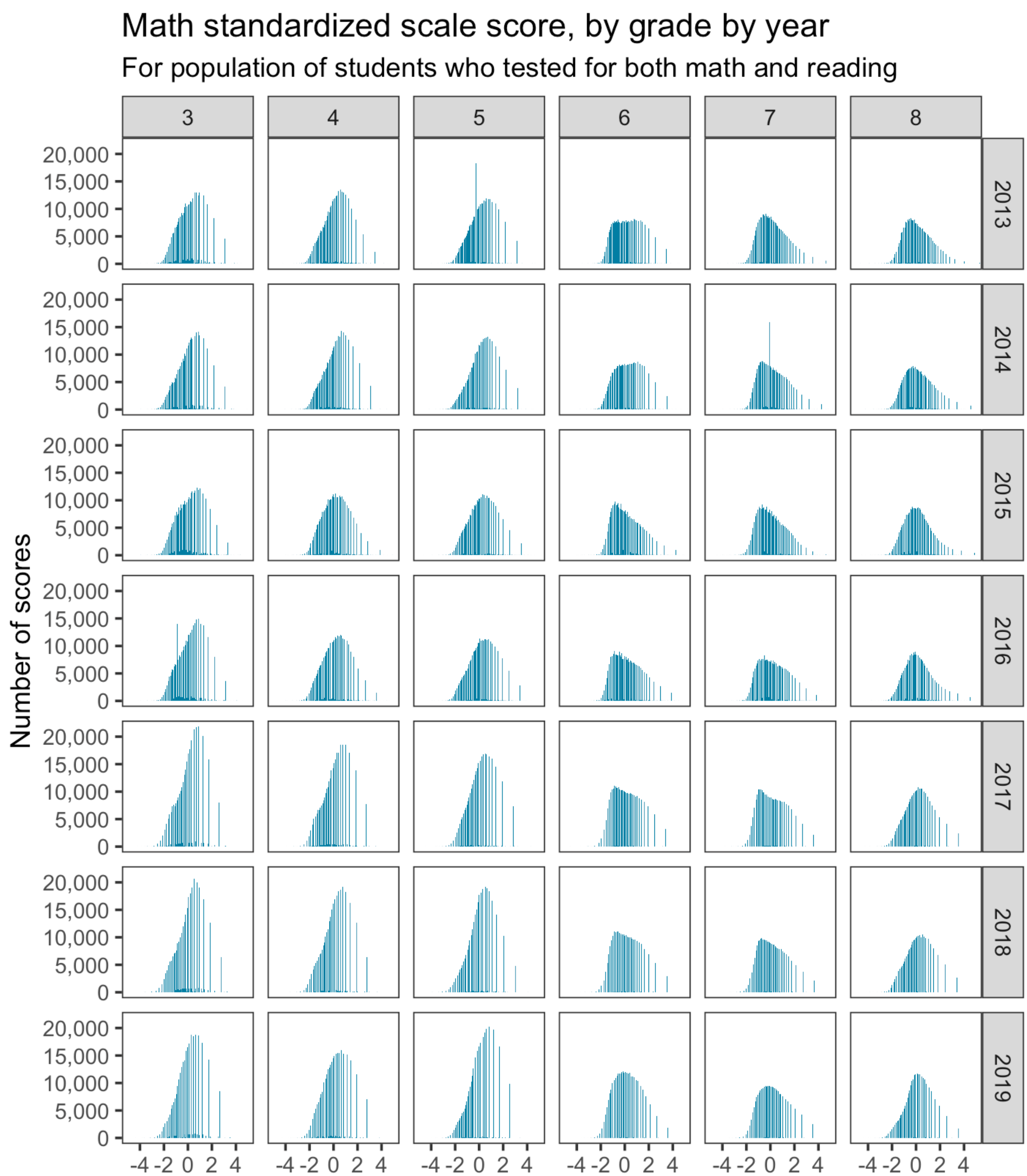


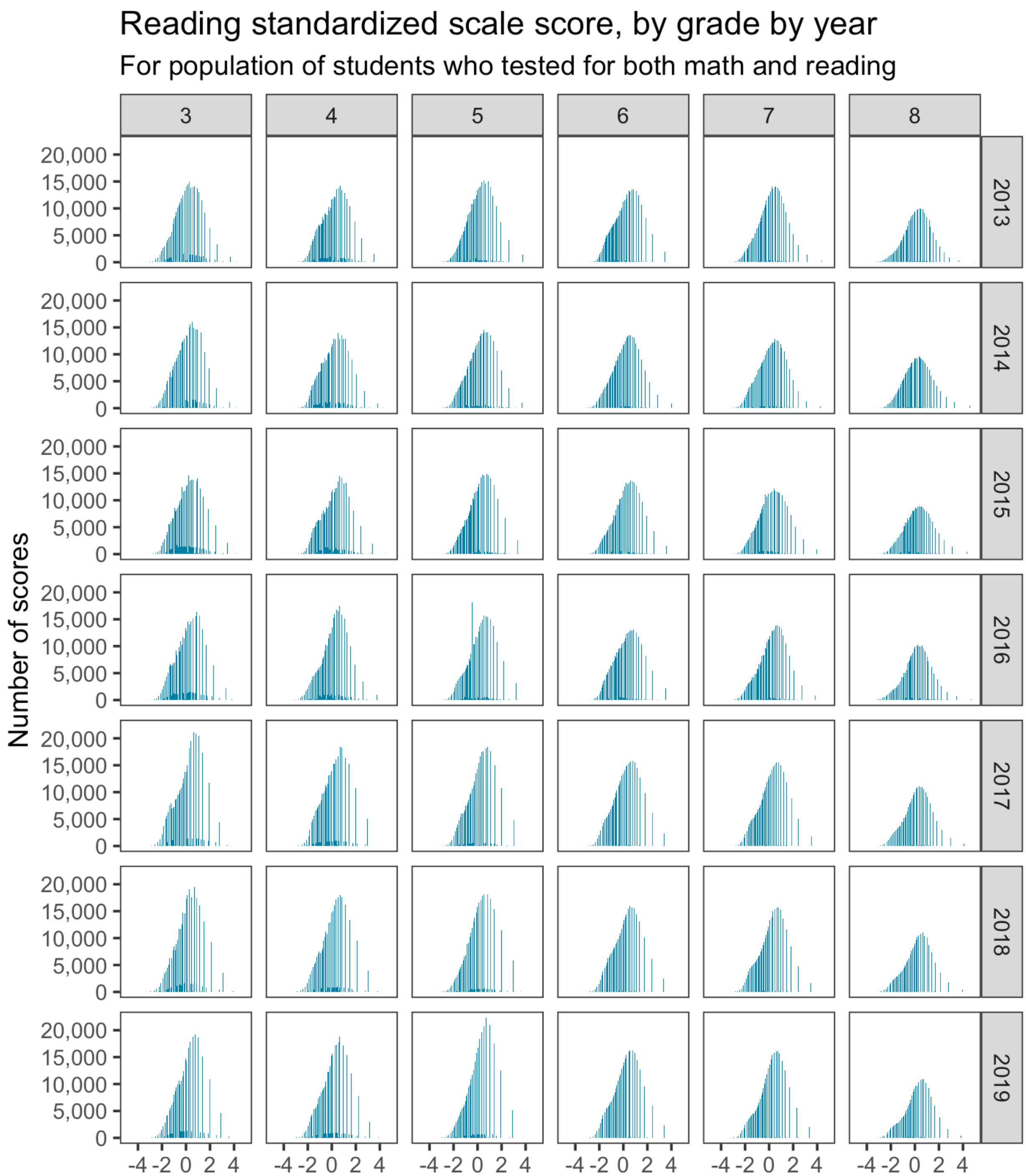
Reading standardized scale score, by grade by year

For population of students who tested for both math and reading



Appendix E. Standardized state test score distributions for State Two




Reading standardized scale score, by grade by year
For population of students who tested for both math and reading
3
4
5
6
7
8
2013
2014
2015
2016
2017
2018
2019
Number of scores
20,000
15,000
10,000
5,000
0
-4 -2 0 2 4

Appendix F. Matched group balance tables

Table F1. Weighted matched math sample balance

Variable	State	CA % (count)	Brick-and-mortar % (count)	Difference	Effect size (dCox)	p-value
Male	State 1	45.7% (75)	45.7% (1838)	0%	N/A	N/A
	State 2	41.8% (74)	41.8% (2680)	0%	N/A	N/A
Special education	State 1	17.7% (29)	17.7% (329)	0%	N/A	N/A
	State 2	1.7% (3)	1.7% (34)	0%	N/A	N/A
Low income	State 1	76.2% (125)	76.2% (3461)	0%	N/A	N/A
	State 2	49.2% (87)	49.2% (4329)	0%	N/A	N/A
English learner	State 1	2.4% (4)	2.4% (53)	0%	N/A	N/A
	State 2	2.9% (5)	2.9% (132)	0%	N/A	N/A
White/Asian	State 1	75.6% (124)	75.6% (3687)	0%	N/A	N/A
	State 2	51.4% (91)	51.4% (2648)	0%	N/A	N/A

Each Connections Academy student is paired with a virtual brick-and-mortar (B&M) student whose test score is computed as the average test score of all the B&M students matching that Connections Academy student. This weighting counts each matched pair equally, regardless of the number of matched B&M students.

Table F2. Unweighted matched math sample balance for race/ethnicity

Variable	State	CA % (count)	Brick-and-mortar % (count)	Difference	Effect size (dCox)	p-value
Asian	State 1 State 2	0.6% (1) 1.7% (3)	1.1% (50) 5.7% (424)	-0.5% -4.1%	-0.339 -0.765	1.000 0.019
Black or African American	State 1 State 2	9.1% (15) 9.6% (17)	17.4% (817) 15.2% (1121)	-8.2% -5.6%	-0.446 -0.316	0.004 0.042
Hispanic/Latino	State 1 State 2	4.3% (7) 23.2% (41)	2.2% (104) 46.3% (3417)	2.1% -23.1%	0.412 -0.637	0.101 < 0.001
White	State 1 State 2	75.0% (123) 49.7% (88)	77.3% (3637) 30.1% (2224)	-2.3% -19.6%	-0.076 0.503	0.507 < 0.001
American Indian or Alaska Native / Native Hawaiian or Other Pacific Islander / Two or more races / No information provided	State 1 State 2	0.6% (1) 2.8% (5)	0.3% (12) 1.4% (106)	0.4% 1.4%	0.531 0.419	0.360 0.189
More than one race/ethnicity across years	State 1 State 2	10.4% (17) 13.0% (23)	1.8% (86) 1.2% (88)	8.5% 11.8%	1.107 1.525	< .001 < .001

Each Connections Academy student and matched brick-and-mortar (B&M) pair counts the same. This means that matched groups for Connections Academy students who have a greater number of matched B&M students more strongly influence the group mean difference.

Table F3. Weighted matched reading sample balance

Variable	State	CA % (count)	Brick-and-mortar % (count)	Difference	Effect size (dCox)	p-value
Male	State 1 State 2	46.5% (73) 42.9% (79)	46.5% (1875) 42.9% (2804)	0% 0%	N/A N/A	N/A N/A
Special education	State 1 State 2	19.1% (30) 1.6% (3)	19.1% (332) 1.6% (20)	0% 0%	N/A N/A	N/A N/A
Low income	State 1 State 2	75.8% (119) 50.0% (92)	75.8% (3093) 50.0% (4590)	0% 0%	N/A N/A	N/A N/A
English learner	State 1 State 2	1.9% (3) 2.7% (5)	1.9% (53) 2.7% (36)	0% 0%	N/A N/A	N/A N/A
White/Asian	State 1 State 2	75.2% (118) 51.6% (95)	75.2% (3543) 51.6% (2544)	0% 0%	N/A N/A	N/A N/A

Each Connections Academy student is paired with a virtual brick-and-mortar student whose test score is computed as the average test score of all the B&M students matching that Connections Academy student. This weighting counts each matched pair equally, regardless of the number of matched B&M students.

Table F4. Unweighted matched reading sample balance for race/ethnicity

Variable	State	CA % (count)	Brick-and-mortar % (count)	Difference	Effect size (dCox)	p-value
Asian	State 1 State 2	0% (0) 1.6% (3)	1.2% (52) 5.3% (407)	-1.2% -3.7%	N/A -0.742	N/A 0.019
Black or African American	State 1 State 2	9.6% (15) 9.8% (18)	17.0% (759) 17.8% (1359)	-7.4% -8.0%	-.401 -.603	0.012 0.001
Hispanic/Latino	State 1 State 2	4.5% (7) 23.9% (44)	2.1% (93) 46.0% (3507)	2.4% -22.1%	0.477 -0.603	0.083 < 0.001
White	State 1 State 2	75.2% (118) 50.0% (92)	78.1% (3491) 28.0% (2137)	-3.0% 22.0%	-0.101 0.572	0.378 < 0.001
American Indian or Alaska Native / Native Hawaiian or Other Pacific Islander / Two or more races / No information provided	State 1 State 2	0.6% (1) 2.7% (5)	0.2% (9) 1.6% (122)	0.4% 1.1%	0.700 0.328	0.292 0.227
More than one race/ethnicity across years	State 1 State 2	10.2% (16) 12.0% (22)	1.4% (64) 1.3% (98)	8.8% 10.7%	1.246 1.421	< 0.001 < 0.001

Each Connections Academy student and matched brick-and-mortar (B&M) pair counts the same. This means that matched groups for Connections Academy students who have a greater number of matched B&M students more strongly influence the group mean difference.

Appendix G. Math statistical model output

Table G1. Math, with state by group interaction

	4th-year standardized score			
Predictor	*Estimate*	SE	95% CI	p
(Intercept)	0.66	0.05	0.57 – 0.75	**<0.001**
Connections Academy	-0.42	0.04	-0.49 – -0.35	**<0.001**
Connections Academy by state (interaction)	-0.23	0.07	-0.37 – -0.09	**<0.001**
State Two vs State One	0.12	0.05	0.02 – 0.21	**0.01**
First-year math scaled score	0.52	0.01	0.49 – 0.55	**<0.001**
First-year reading scaled score	0.21	0.01	0.19 – 0.23	**<0.001**
First year: 2013	(ref)			
First year: 2014	-0.01	0.03	-0.07 – 0.05	0.67
First year: 2015	-0.00	0.03	-0.06 – 0.06	0.96
First year: 2016	-0.04	0.04	-0.12 – 0.03	0.28
First-year grade: 3	(ref)			
First-year grade: 4	-0.01	0.03	-0.06 – 0.05	0.86
First-year grade: 5	0.05	0.03	-0.00 – 0.10	0.05
First-year grade: 6	0.36	0.30	-0.24 – 0.95	0.24
Race: Asian	(ref)			
Race: Black	-0.56	0.04	-0.64 – -0.48	**<0.001**
Race: Hispanic	-0.54	0.04	-0.63 – -0.46	**<0.001**
Race: multiple across records	-0.48	0.06	-0.60 – -0.37	**<0.001**
Race: Other	-0.39	0.07	-0.53 – -0.26	**<0.001**
Race: White	-0.49	0.03	-0.56 – -0.43	**<0.001**
CEM gender male	-0.00	0.02	-0.05 – 0.04	0.92
CEM special education	-0.03	0.05	-0.12 – 0.06	0.45
CEM disadvantaged	-0.25	0.03	-0.30 – -0.19	**<0.001**
CEM English second language	0.17	0.09	0.01 – 0.34	**0.04**
Random effects				
σ^2	0.407			
$\tau_{00\ \text{matched\_pair\_id}}$	0.011			
$\tau_{00\ \text{first\_year\_district}}$	0.028			
$N_{\text{first\_year\_district}}$	122			
$N_{\text{matched\_pair\_id}}$	341			
Observations	12427 (164 CA students in State One, 177 CA students in State Two)			

Table G2. State One math

Predictor	4th-year standardized score Estimate	SE	95% CI	p
(Intercept)	0.47	0.09	0.29 – 0.65	**<0.001**
Connections Academy	-0.29	0.05	-0.38 – -0.19	**<0.001**
First year: 2013	(ref)			
First year: 2014	-0.06	0.04	-0.13 – 0.01	0.07
First year: 2015	-0.01	0.04	-0.08 – 0.06	0.82
First-year grade: 3	(ref)			
First-year grade: 4	-0.00	0.04	-0.08 – 0.08	0.96
First-year grade: 5	0.02	0.03	-0.04 – 0.09	0.52
First-year grade: 6	0.09	0.44	-0.77 – 0.94	0.84
First-year math scaled score	0.50	0.02	0.46 – 0.54	**<0.001**
First-year reading scaled score	0.27	0.01	0.24 – 0.29	**<0.001**
Race: Asian	(ref)			
Race: Black	-0.47	0.09	-0.65 – -0.29	**<0.001**
Race: Hispanic	-0.31	0.11	-0.53 – -0.09	**0.01**
Race: multiple across records	-0.36	0.11	-0.57 – -0.16	**<0.001**
Race: Other	-0.39	0.19	-0.76 – -0.03	**0.03**
Race: White	-0.34	0.08	-0.50 – -0.18	**<0.001**
CEM gender male	-0.00	0.03	-0.06 – 0.05	0.98
CEM special education	0.01	0.05	-0.08 – 0.11	0.76
CEM disadvantaged	-0.22	0.04	-0.30 – -0.15	**<0.001**
CEM English second language	0.10	0.12	-0.14 – 0.34	0.42
Random effects				
σ^2	0.343			
τ_{00} matched_pair_id	0.009			
τ_{00} first_year_district	0.006			
N first_year_district	43			
N matched_pair_id	164			
Observations	4870 (164 CA students)			

Pearson

Table G3. State Two math

	4th-year standardized score			
Predictor	*Estimate*	*SE*	*95% CI*	*p*
(Intercept)	0.68	0.06	0.55 – 0.81	**<0.001**
Connections Academy	-0.50	0.05	-0.60 – -0.40	**<0.001**
First-year math scaled score	0.53	0.02	0.49 – 0.57	**<0.001**
First-year reading scaled score	0.17	0.01	0.15 – 0.20	**<0.001**
First year: 2013	(ref)			
First year: 2014	0.07	0.05	-0.04 – 0.18	0.20
First year: 2015	0.00	0.05	-0.10 – 0.10	0.95
First year: 2016	-0.01	0.05	-0.11 – 0.08	0.77
First-year grade: 3	(ref)			
First-year grade: 4	-0.00	0.04	-0.09 – 0.08	0.98
First-year grade: 5	0.08	0.04	-0.00 – 0.16	0.06
First-year grade: 6	0.46	0.41	-0.33 – 1.26	0.25
Race: Asian	(ref)			
Race: Black	-0.52	0.06	-0.63 – -0.41	**<0.001**
Race: Hispanic	-0.52	0.05	-0.63 – -0.42	**<0.001**
Race: multiple across records	-0.47	0.08	-0.63 – -0.31	**<0.001**
Race: Other	-0.36	0.08	-0.52 – -0.20	**<0.001**
Race: White	-0.52	0.04	-0.59 – -0.45	**<0.001**
CEM gender male	0.01	0.03	-0.06 – 0.07	0.81
CEM special education	-0.04	0.14	-0.31 – 0.23	0.77
CEM disadvantaged	-0.28	0.04	-0.36 – -0.19	**<0.001**
CEM English second language	0.21	0.14	-0.07 – 0.49	0.15
Random effects				
σ^2	0.447			
$\tau_{00\ \text{matched\_pair\_id}}$	0.011			
$\tau_{00\ \text{first\_year\_district}}$	0.048			
$N_{\text{first\_year\_district}}$	79			
$N_{\text{matched\_pair\_id}}$	177			
Observations	7557 (177 CA students)			

Appendix H. Reading statistical model output

Table H1. Reading, with state by group interaction

	4th-year standardized score			
Predictor	*Estimate*	*SE*	*95% CI*	*p*
(Intercept)	0.43	0.04	0.35 – 0.50	**<0.001**
Connections Academy	-0.02	0.03	-0.09 – 0.05	0.64
State Two vs State One	0.18	0.03	0.12 – 0.23	**<0.001**
Connections Academy by state (interaction)	-0.05	0.07	-0.18 – 0.09	0.49
First year: 2013	(ref)			
First year: 2014	-0.01	0.03	-0.06 – 0.04	0.63
First year: 2015	-0.01	0.02	-0.06 – 0.04	0.77
First year: 2016	0.01	0.03	-0.05 – 0.07	0.70
First-year grade: 3	(ref)			
First-year grade: 4	0.02	0.02	-0.03 – 0.07	0.39
First-year grade: 5	0.20	0.02	0.16 – 0.24	**<0.001**
First-year grade: 6	0.04	0.24	-0.42 – 0.50	0.88
First-year math scaled score	0.22	0.01	0.20 – 0.24	**<0.001**
First-year reading scaled score	0.55	0.01	0.53 – 0.58	**<0.001**
Race: Asian	(ref)			
Race: Black	-0.23	0.04	-0.31 – -0.16	**<0.001**
Race: Hispanic	-0.19	0.04	-0.26 – -0.12	**<0.001**
Race: multiple across records	-0.18	0.06	-0.29 – -0.07	**<0.001**
Race: Other	-0.15	0.06	-0.27 – -0.02	**0.02**
Race: White	-0.18	0.03	-0.24 – -0.12	**<0.001**
CEM gender male	-0.15	0.02	-0.18 – -0.11	**<0.001**
CEM special education	-0.15	0.04	-0.23 – -0.07	**<0.001**
CEM disadvantaged	-0.17	0.02	-0.21 – -0.13	**<0.001**
CEM English second language	0.08	0.07	-0.07 – 0.22	0.30
Random effects				
σ^2	0.376			
$\tau_{00\ \text{matched\_pair\_id}}$	0.007			
$\tau_{00\ \text{first\_year\_district}}$	0.002			
$N_{\text{first\_year\_district}}$	126			
$N_{\text{matched\_pair\_id}}$	341			
Observations	12439 (157 State One CA students, 184 State Two CA students)			

Pearson

Table H2: Reading, no state by group interaction

	4th-year standardized score			
Predictor	*Estimate*	*SE*	*95% CI*	*p*
(Intercept)	0.43	0.04	0.35 – 0.50	**<0.001**
Connections Academy	-0.01	0.03	-0.08 – 0.06	0.72
State Two vs State One	0.17	0.03	0.12 – 0.23	**<0.001**
First year: 2013	(ref)			
First year: 2014	-0.01	0.03	-0.06 – 0.04	0.63
First year: 2015	-0.01	0.02	-0.05 – 0.04	0.77
First year: 2016	0.01	0.03	-0.05 – 0.07	0.70
First-year grade: 3	(ref)			
First-year grade: 4	0.02	0.02	-0.03 – 0.07	0.39
First-year grade: 5	0.20	0.02	0.16 – 0.24	**<0.001**
First-year grade: 6	0.03	0.24	-0.43 – 0.50	0.88
First-year math scaled score	0.22	0.01	0.20 – 0.24	**<0.001**
First-year reading scaled score	0.55	0.01	0.53 – 0.58	**<0.001**
Race: Asian	(ref)			
Race: Black	-0.23	0.04	-0.31 – -0.16	**<0.001**
Race: Hispanic	-0.19	0.04	-0.26 – -0.12	**<0.001**
Race: multiple across records	-0.18	0.06	-0.29 – -0.07	**<0.001**
Race: Other	-0.15	0.06	-0.27 – -0.02	**0.02**
Race: White	-0.18	0.03	-0.24 – -0.12	**<0.001**
CEM gender male	-0.15	0.02	-0.18 – -0.11	**<0.001**
CEM special education	-0.15	0.04	-0.23 – -0.07	**<0.001**
CEM disadvantaged	-0.17	0.02	-0.21 – -0.13	**<0.001**
CEM English second language	0.08	0.07	-0.07 – 0.22	0.30
Random effects				
σ^2	0.376			
$\tau_{00\ \text{matched\_pair\_id}}$	0.007			
$\tau_{00\ \text{first\_year\_district}}$	0.002			
$N_{\text{first\_year\_district}}$	126			
$N_{\text{matched\_pair\_id}}$	341			
Observations	12439 (157 State One CA students , 184 State Two CA students)			

Appendix I. Statistical tests for individual subpopulations

Table I1: Estimated effect of attending Connections Academy in State One on math state test scores

Demographic variable split	Difference between Level A and Level B	Level A	Level B
Gender (Male/Female)	**Statistically significant**	Male -0.19 (sig) (n =75)	Female -0.37 (sig) (n = 89)
Economically disadvantaged (Yes/No)	Not statistically significant	Yes -0.28 (sig) (n = 125)	No -0.32 (sig) (n = 39)
Race/ethnicity White/Asian (Yes/No)	Not statistically significant	Yes -0.29 (sig) (n = 124)	No -0.26 (sig) (n = 40)
Special education (Yes/No)	N/A	Yes N/A* (n = 29)	No -0.31 (sig) (n = 135)
English learner (Yes/No)	N/A	Yes N/A (n = 4)	No -0.28 (sig) (n = 160)
Metropolitan (Yes/No)	Not statistically significant	Yes -0.30 (sig) (n = 119)	No -0.23 (sig) (n = 45)

Note that (n = ...) indicates the number of Connections Academy students; **(sig)** indicates statistically significant mean difference; **ns** indicates statistically non-significant mean difference; and **N/A** indicates statistical model was not recommended for fewer than 30 Connections Academy students.

Table I2: Estimated effect of attending Connections Academy in State One on reading state test scores

Demographic variable split	Difference between Level A and Level B	Level A	Level B
Gender (Male/Female)	Not statistically significant	Male 0.08 (ns) (n = 73)	Female -0.04 (ns) (n = 84)
Economically disadvantaged (Yes/No)	Not statistically significant	Yes 0.04 (ns) (n = 119)	No -0.06 (ns) (n = 38)
Race/ethnicity White/Asian (Yes/No)	Not statistically significant	Yes -0.01 (ns) (n = 118)	No 0.13 (ns) (n =39)
Special education (Yes/No)	Not statistically significant	Yes 0.05 (ns) (n = 30)	No 0.01 (ns) (n =127)
English learner (Yes/No)	N/A	Yes N/A (n = 3)	No 0.01 (ns) (n = 154)
Metropolitan (Yes/No)	Not statistically significant	Yes 0.03 (ns) (n = 115)	No 0.00 (ns) (n = 42)

Note that (n = ...) indicates the number of Connections Academy students; **(sig)** indicates statistically significant mean difference; **ns** indicates not statistically significant mean difference; and **N/A** indicates statistical model was not recommended for data sets with fewer than 30 Connections Academy students.

Table I3: Estimated effect of attending Connections Academy in State Two on math state test scores

Demographic variable split	Difference between Level A and Level B	Level A	Level B
Gender (Male/Female)	Not statistically significant	Male -0.44 (sig) (n = 71)	Female -0.53 (sig) (n = 101)
Economically disadvantaged (Yes/No)	Not statistically significant	Yes -0.44 (sig) (n = 83)	No -0.49 (sig) (n = 89)
Race/ethnicity White/Asian (Yes/No)	Not statistically significant	Yes -0.50 (sig) (n = 87)	No -0.44 (sig) (n = 85)
Special education (Yes/No)	N/A	Yes N/A (n = 3)	No -0.48 (sig) (n = 169)
English learner (Yes/No)	N/A	Yes N/A (n = 5)	No -0.47 (sig) (n = 167)
Metropolitan (Yes/No)	N/A	Yes -0.45 (sig) (n = 157)	No N/A (n = 15)

Note that (n = ...) indicates the number of Connections Academy students; **(sig)** indicates statistically significant mean difference; **ns** indicates statistically non-significant mean difference; and **N/A** indicates statistical model was not recommended for fewer than 30 Connections Academy students.

Table I4: Estimated effect of attending Connections Academy in State Two on reading state test scores

Demographic variable split	Difference between Level A and Level B	Level A	Level B
Gender (Male/Female)	Not statistically significant	Male -0.11 (ns) (n = 74)	Female 0.05 (ns) (n = 101)
Economically disadvantaged (Yes/No)	Not statistically significant	Yes 0.08 (ns) (n = 85)	No -0.08 (ns) (n = 90)
Race/ethnicity White/Asian (Yes/No)	Not statistically significant	Yes 0.02 (ns) (n = 90)	No -0.03 () (n = 85)
Special education (Yes/No)	N/A	Yes N/A (n = 3)	No -0.02 (ns) (n = 172)
English learner (Yes/No)	N/A	Yes N/A* (n = 5)	No 0.00 (ns) (n = 170)
Metropolitan (Yes/No)	N/A	Yes -0.01 (ns) (n = 157)	No N/A (n =18)

Note that (n = ...) indicates the number of Connections Academy students; **(sig)** indicates statistically significant mean difference; **ns** indicates no statistically significant mean difference; and **N/A** indicates statistical model was not recommended for data sets with fewer than 30 Connections Academy students.