Year 1 Technical Report:
Rigorous Tests of Student Outcomes in CTE Programs of Study

Marisa Castellano
Kirsten Sundell
Laura T. Overman

University of Louisville

January 15, 2010

National Research Center for Career and Technical Education
University of Louisville
Louisville KY 40292

Funding Information

Project Title:	National Research Center for Career and Technical Education
Grant Number:	VO51A070003
Act under Which Funds Administered:	Carl D. Perkins Career and Technical Education Act of 2006
Source of Grant:	Office of Vocational and Adult Education U.S. Department of Education Washington, D.C. 20202
Grantees:	University of Louisville National Research Center for Career and Technical Education 354 Education Building Louisville, KY 40292
Project Director:	James R. Stone, III
Percent of Total Grant Financed by Federal Money:	100%
Dollar Amount of Federal Funds for Grant:	$4,500,000
Disclaimer:	The work reported herein was supported under the National Research Center for Career and Technical Education, PR/Award (No. VO51A070003) as administered by the Office of Vocational and Adult Education, U.S. Department of Education. However, the contents do not necessarily represent the positions or policies of the Office of Vocational and Adult Education or the U.S. Department of Education and you should not assume endorsement by the Federal Government.
Discrimination:	Title VI of the Civil Rights Act of 1964 states: "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving federal financial assistance." Title IX of the Education Amendment of 1972 states: "No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving federal financial assistance." Therefore, the National Research Center for Career and Technical Education project, like every program or activity receiving financial assistance from the U.S. Department of Education, must be operated in compliance with these laws.

Rigorous Tests of Student Outcomes in CTE Programs of Study

The objective of this study is to evaluate the impact of programs of study (POS), a federally mandated education reform that was part of the Carl D. Perkins Career and Technical Education Improvement Act of 2006 (otherwise known as Perkins IV), the legislation that funds career and technical education (CTE) nationwide. This study seeks to measure the impact of POS on student academic and technical achievement. We are conducting this evaluation using two methodologies. First, we are conducting a set of randomized controlled trials (RCTs) in three sites in one large district. The second test of student outcomes uses rigorous quasi-experimental methods in another large district. Following Perkins law, the measures of effectiveness of POS include: (a) academic achievement, (b) technical skills achievement, (c) high school completion, (d) placement in postsecondary education, work, or the military, (e) program participation and completion by nontraditional students, and (f) program participation and completion by students from special populations as defined by the law. In order to better understand student outcomes, the study also describes the context and practices that produced them.

This study will help determine the overall impact of POS implemented under a variety of conditions in two different states (e.g., different student populations, different types of schools). The results will show the degree to which POS have any impact on student academic and technical skill outcomes in relation to the typical CTE programs and practices that exist in the control and comparison schools. Given the study design, we will be able to separately report differential impacts, if any, for some groups or under some conditions. The current report provides baseline information on the study design, site selection, and student samples.

Conceptual Framework

This study employs the same conceptual framework guiding the NRCCTE's other two field-based POS studies (Programs of Study Joint Technical Working Group, 2009).[1] As such, it acknowledges three important goals for all high school students: engagement in school, academic and technical achievement, and transition to post-high school endeavors, whether work or further education. This framework helped us to form the research questions and develop the instruments and measures described below; ultimately, it also allows us to judge the effectiveness of the participating POS.

POS change several aspects of the default high school experience. First, they represent an upgrade of the CTE experience. Popular among students but traditionally considered less academically rigorous than college preparatory tracks in high school, CTE has often been targeted at students not interested in college. POS infuse CTE with rigorous academics and seamless postsecondary connections, potentially leading to improved student outcomes in high school achievement, completion, and postsecondary matriculation.

Unlike pre-Perkins IV CTE programs, POS include mandatory combinations of new and familiar elements of CTE: non-duplicative sequencing of secondary and postsecondary curricula, rigorous content aligned with academic standards, opportunities for dual credit, and available credentials where possible. Standard CTE programs might offer several of these elements, but

[1] See http://136.165.122.102/UserFiles/File/Tech_Reports/POS_Joint_Technical_Report_Jan_2010.pdf

only POS offer the full set. POS also challenge college-preparatory students with hands-on, career-focused experiences and decrease the financial burden of postsecondary education by allowing students to earn college credit in high school. All of these factors may help convince students that their academic studies are relevant to their career goals.

Research Questions

The role of CTE in high school student engagement, achievement, and transition has been studied but many issues continue to be unclear. The recent changes to federal legislation on CTE and the new requirement of POS lead to the formation of our research question:

To what extent does participation in a POS lead to improved student outcomes as compared to outcomes of (1) a strand of control group students (who applied to be in the program but were not selected in a lottery process), or (2) a strand with a closely matched comparison group? Specifically, to what extent does POS participation increase student:

a. academic achievement?
b. technical skills achievement?
c. high school completion?
d. participation in and completion of POS by nontraditional (e.g., female) students and other special populations of students?
e. employability?
f. completion of coursework leading to college credits?

Study Design

As shown by research in the physical sciences and medicine, effectiveness can best be determined through true experiments with random assignment to treatment and control groups (Cook & Payne, 2002; Mosteller & Boruch, 2002). This design provides the highest assurance that any unobserved differences between the two groups are randomly distributed. As such, our first choice in study design was to conduct an RCT (Shadish, Cook, & Campbell, 2002). We were aware of the difficulties of creating random assignment conditions in educational contexts, however, and proposed considering a rigorous quasi-experimental design should we be unable to find sites amenable to random assignment of students into programs of study. In the end, we have a two-strand study: an RCT strand and a quasi-experimental strand.

Randomized Controlled Trial

The first strand of this study uses an RCT design (Shadish et al., 2002) with three simultaneous sites located in District 1. In this district, the unit of randomization is the school. Background characteristics (including prior achievement) are being collected on treatment and control groups in District 1 to serve as covariates in later analysis of covariance (ANCOVA) analyses as a way of minimizing the possibility that differences in outcomes reflect pre-existing differences in the groups.

Quasi-Experiment

The second strand of this study uses a quasi-experimental design in District 2. The sample for this strand consists of the 2008-2009 ninth-grade class that attends the participating wall-to-wall academy high school and a well-matched comparison group that was identified by the district research office using propensity score matching (PSM; Shadish et al., 2002).

In both strands, data on background characteristics were collected on the sample as well as the overall district for descriptive purposes. These data include gender, race/ethnic origin, prior achievement, a proxy for poverty/socioeconomic status (SES, such as eligibility for free or reduced-price lunch), limited English proficiency, and special education status (i.e., the presence of an individualized education program or IEP). These variables have been shown to be vital to understanding school effectiveness; further, they are being employed in statistical analyses as appropriate to explore questions related to student-level outcomes either as covariates and/or individual differences variables.

Measures

All of our outcomes measures are found in the systems data in our districts and are of practical interest to schools; these measures capture precisely those skills and content that POS are designed to address. We measure academic gains using systems data, including course grades and state test scores. We also measure technical skills achievement using systems data, which include course grades and program assessments (either locally or commercially developed). We collect work readiness assessments if these are available. Data on high school completion and dual credits earned are collected from systems records. In an ideal universe, we would also be able to administer comparable technical skills assessments across sites along with senior year academic achievement exams that incorporate the time in which students are exposed to the most advanced concepts in the POS. However, district leaders understandably limit the amount of research-initiated additional assessment, and we agreed not to add to their testing calendars.

If, as hypothesized, infusing CTE programs with rigorous academics and postsecondary connections increases student engagement, then we must include measures of student engagement. Student attendance and retention over the four years of high school are our primary measures of engagement. Our classroom observations also contain some measures of student engagement. In addition, we include questions on the student surveys that elicit student attitudes about school and their engagement to it.

Data on participation by nontraditional (e.g., female) students and other special populations (e.g., students with disabilities, special education students, economically disadvantaged students, and/or students with limited English proficiency) are being collected from Perkins reporting data. We use Perkins data because localities already collect them, saving considerable time and expense.

Threats to Validity

For both study strands, we have taken measures to reduce the potential influence of selection bias on membership in the treatment/intervention or control/comparison group. In District 1, this was accomplished by the random assignment process of the lottery, as described below. In District 2,

district personnel attempted to reduce selection bias at the outset through a sample identification procedure known as propensity score matching, also described below.

In both strands, the comparison groups were closely matched. Comparison students are in the same grade as those in the intervention groups, and all data will be collected over roughly the same time period, thus minimizing multiple-group threats to internal validity such as history and maturation. Multiple sites consisting of regular schools with typical teachers are being used to reduce the likelihood of systematic bias related to site-specific factors, as recommended by the quantity of evidence guidelines (U.S. Department of Education, Institute of Education Sciences, 2003).

The external validity or generalizability of the results will be strongest in the RCT strand because the sample was randomly assigned to the treatment. We are conducting these RCTs in three different school structures as described in an earlier section, extending the generalizability to a variety of school structures implementing POS.

We are cognizant that the addition of high academic standards to POS could conceivably drive out lower achieving students to the extent that such students do not even apply to the programs, affecting the generalizability of the findings. Like many RCT studies using a "naturally occurring" randomization technique (i.e., district-run lotteries), this study only includes students who are motivated enough to apply for these programs. We will address this potential limitation by comparing the students who applied to POS to students in the same district and grade who did not apply, and perhaps also to national samples of students such as the Educational Longitudinal Study of 2002 (ELS).

It should be noted that in both strands, we cannot control student attrition over the course of the study. What we can and will do is to continue to collect and analyze data on all sample members, even those who fail to complete the intervention, so as to maintain equivalence. Also, to the extent possible, we will follow up with and interview students who leave the groups in order to support an intention-to-treat analysis (Fisher et al., 1990).

Power

At this stage, the following power analysis is an estimate that uses the current sample sizes because there are corrections yet to be made for clustering—that is, for the fact that the students are nested in programs that are nested in schools, and therefore each student is not an independent observation (Cohen, 1988; Murphy & Myors, 2004). This fact makes a multilevel statistical analysis most appropriate; however, given that we have resources for only a small number of sites, a multilevel analysis would likely be extremely underpowered. We used the estimates below in order to determine that the sample sizes we had available to us were large enough to sufficiently minimize the probability of making a Type II error (i.e., failing to reject a false null hypothesis). We chose an effect size of $d = 0.2$. Using a Type I error rate of .05 and entering the number of students in the sample at each site yielded the following:

District 1
School #1: 509 treatment students, 661 control students: 96% statistical power

School #2: 481 treatment students, 873 control students: 97% statistical power
School #3: 274 treatment students, 527 control students: 85% statistical power

District 2
School #4: 376 intervention students, 752 control students: 95% statistical power

New statistical tests of power will be conducted to take the lack of statistical independence into account. This will eliminate the danger of resulting standard errors that are spuriously small (and hence, confidence intervals that are similarly too small). To address the problem of statistical dependence, we will employ the correction offered by Kish (1965).

Quantitative Data Analysis

Our primary statistical technique is the ANCOVA. All comparisons between treatment and control groups (as well as any interactions) have the effects of gender, race/ethnic origin, prior achievement, SES, limited English proficiency, and special education status statistically controlled. Cohen (1988) recommended employing one-tailed statistical tests when a negative effect for a treatment has functionally the same meaning as a null result. In this case, if POS are a harmful treatment (e.g., they result in lower academic achievement) then the practical implication is the same as a null result—we would not recommend further implementation of POS. We therefore conduct all statistical analyses using one-tailed tests.

We express dichotomous outcomes as odds ratios and all others as the standardized mean difference (Cohen's *d*) to help interpret the effect size of our results. Because our statistical analyses are based on ANCOVA, the standardized mean difference effect sizes are based on adjusted means (instead of unadjusted means) and unadjusted pooled posttest standard deviations. Odds ratios are similarly adjusted for baseline differences. In addition, we disaggregate results by special populations and other student background characteristics, including gender, ethnicity, SES, and limited English proficient, special education, and disability status.

Qualitative Data Analysis

To help us understand the process through which POS effect change, this study employs a qualitative component that measures mediating and moderating variables. It is insufficient to report student outcomes without a sense of the context and practices that produced them. As such, we are conducting interviews and classroom observations in both the treatment/intervention and the control/comparison conditions. Although we discuss the qualitative analyses separately from the quantitative, the qualitative findings will inform the quantitative outcomes and serve as another source of convergent evidence about the effectiveness of a CTE program of study.

As an example, a central part of POS is the connection between secondary and postsecondary education. Discussions of high school CTE program alignment to postsecondary programs often assume that all such programs are homogeneous. However, they differ in regard to the specific alignment activities, the contacts that take place for both students and faculty, and the level of

commitment to that alignment. Of particular interest is whether these cross-institution connections serve to lower the need for student remediation upon entrance to the college. This study will examine this relationship between the participating high schools and colleges at each site.

We are conducting interviews with teachers, students, and administrators in both the treatment/intervention and control/comparison conditions. These interviews focus on what participants perceive to be the most important differences between POS and the control/comparison condition, participants' sense of how those differences influence outcomes, and how participants perceive the usefulness of their school's programs as preparation for postsecondary education and work. The interviews also assist us in accounting for sources of variation in outcomes across these settings. We conduct classroom observations so that we may (1) discern program differences ourselves and (2) verify the presence and fidelity of treatment (O'Donnell, 2008).

Interviews are audio-recorded and summarized. The summaries are classified by type (e.g., CTE teachers, students). Following the case study methods of Yin (2009) and the grounded theory approach of Strauss and Corbin (1990), the content of the summaries are being coded and entered into a qualitative data analysis software package called HyperResearch© version 2.8. The classroom observation data are also being coded and entered into HyperResearch. Coding involves marking the instances of concepts or topics relevant to the study as they appear in the interview or observation data (e.g., project-based learning). Such open or general coding identifies concepts that can then be developed through axial coding, or grouping the data into topical categories that are more specific and descriptive (e.g., evidence of academic standards). These data can often be reduced to a series of matrices (Miles & Huberman, 1984) to aid in within- and cross-site analyses that will identify the mediating and moderating variables that can help explain the outcomes. Codes can be queried within HyperResearch and yield a report rich with information from various stakeholders, including students, as well as data from the classes they attended. Such an analysis process maximizes triangulation of as many data sources as possible.

The Site Selection Process

We began our search in several places, one of which was examining the websites of the National Charter School Clearinghouse and the National Alliance for Public Charter Schools organizations, looking for charter high schools with a career theme. We hypothesized that charter schools with popular programs would be likely to run lotteries for admission. An initial search generated several programs that were very promising, but upon closer inspection, we learned that many did not provide a curriculum that could be called CTE. Most offered small learning communities with a college preparatory focus, heavy on technology across the curriculum and overlaid with internships, project-based learning, and senior capstones—but nothing approximating a sequence of CTE courses.

Study Sites

We were successful in finding POS sites for study when we sought out districts that already used lotteries to distribute admissions opportunities for oversubscribed programs. Many districts have developed opportunities for students at all grade levels to receive innovative or themed instruction. In some cases, these opportunities are developed as a means of integrating schools so that students who live in various parts of the district and would otherwise attend relatively segregated schools instead attend a magnet school that attracts students from all parts of the district. Students are attracted by school themes, such as International Baccalaureate or career academies. Transportation is usually provided to students who live outside the normal school boundaries.

In other cases, the goal may not be integration per se, but simply to provide options that are outside the traditional high school experience. This may be in response to the notion of the "shopping mall high school" (Powell, Farrar, & Cohen, 1985), in which high schools have been described as attempting to cater to such a wide variety of students that they were unable to respond adequately to the needs of most. By providing schools of choice, students and teachers can be certain of the focus at any particular school and attend to the business of learning. The districts participating in this study chose to develop POS for some or all of these reasons.

District 1

District 1 is located in a large city in a Western state. It serves 300,000 students, over 60% of whom self-identify as ethnic minorities. Over 40% of the students are eligible for the federal free lunch program. Table 1 summarizes important characteristics of Districts 1 and 2, described below, compared to national averages.

Table 1
Selected Characteristics of Participating School Districts and Their Populations

Characteristics	*National Average 2006-2007*	*District 1 2007-2008*	*District 2 2008-2009*
District Characteristics			
Enrollment	3,424[1]	308,554	134,060
Minority students (%)	45.0[2a]	63.9[2b]	66.3[2b]
Free lunch eligible (%)	41.8	42.5	48.7
Number of teachers	220.7[3]	15928.8[4]	9363
Student/teacher ratio	15.4	19.0[4]	14.3
Number of schools in district	6.9	319.0[4]	172.0
Per pupil expenditure[5]	$10,400[6]	$6,913	$8,794
Population Characteristics			
Total population in district	16,702[7]	1,375,365	695,454
Population density of city per square mile[8]	79.6	4222.3	2232.3
Minority (%)[9]	31.9	39.8	38.9
All persons living below poverty line (%)[10]	12.4	10.8	9.2
Children under 18 living below poverty line (%)[11]	16.6	14.6	11.5
Youths ages 16 to 19 who are not enrolled in school (%)	20.2	31.4	21.4

Youth unemployment rate (16 to 19)	9.3	9.7	11.1
Unemployment rate	3.7	4.2	3.7
Households receiving public assistance (%)	3.4	2.4	2.3
Median household income	$41,994	$44,616	$50,579
Persons age 25 or over with a high school diploma (%)	80.4	79.5	86.2

Sources. National averages of characteristics were obtained from the website of SchoolDataDirect, an online service of the Council of Chief State School Officers' State Education Data Center, http://www.schooldatadirect.org/app/location/q/stid=1036196/llid=162/stllid=676/locid=1036195/catid=-1/secid=-1/compid=-1/site=pes.
Notes. District characteristics data were obtained from district websites or reports and the Common Core of Data website, http://nces.ed.gov/ccd/. Population characteristics were based on 2000 Census data obtained from the U.S. Census website, http://www.census.gov/.
[1]Total number of students divided by total number of districts.
[2a]Sum of number of Native American, African American, Latino, and Asian students divided by the total number of students with reported race/ethnicity data.
[2b]Percent of students enrolled in the school district who identified as not White (i.e., identified as African American, Asian, Latino, or Native American).
[3]Based on projected number of public school teachers for 2005-2006 obtained from the Digest of Education Statistics, Table 61, http://nces.ed.gov/programs/digest/d07/tables/dt07_061.asp, divided by total number of districts obtained from SchoolDataDirect.
[4]Based on 2006-2007 data.
[5]Per pupil expenditures may not be comparable because districts may vary in their definition of what to include in total expenditure per student.
[6]Based on 2005-2006 data.
[7]Total number in population from U.S. Census, divided by the number of districts in 1999-2000 from http://nces.ed.gov/pubs2001/100_largest/discussion.asp#tableA.
[8]Population density is based on the city in which each school district is located.
[9]These percentages were calculated by subtracting the percent White (one race only) who are not Latino from 100%.
[10]Percent of all individuals with a known poverty status who are living below the poverty line in the county in which the district resides.
[11]Percent of individuals under 18 living below the poverty line in the county in which the district resides.

District 1 has a much larger enrollment than the national average, as well as a larger percentage of minority students. District 1 has more teachers and more schools than the national average, yet its per-pupil expenditure is lower. In terms of the population in the area that the district serves, it has more people per square mile than the national average, and a higher percentage of people who identify as minorities, but fewer people living below the poverty line. District 1 has a high percentage of youths ages 16 to 19 who are not enrolled in school. Unemployment rates are higher than the national average, but median incomes for workers are higher.

The impetus behind District 1's magnet and POS high schools was to improve student achievement, promote diversity, and create an awareness of career opportunities relative to the POS (and the magnet themes). A district administrator attributed the development of these schools of choice to visionary leaders who saw the rapid growth in the district and decided to provide high-quality educational options in all parts of the district. Another administrator described the goal as "elite schools for all children" as opposed to "schools for elite children."

Perkins IV was not a driving force behind the creation of POS high schools, according to one principal. The POS were developed with the help of the Joint Technical Skills Committees made up of the district, the community college, the state university, and relevant business communities. District 1's POS schools offer the latest in CTE technology and curriculum (i.e., project-based learning, opportunities to earn academic and technical postsecondary credits in the form of Advanced Placement [AP] classes and Tech Prep, and working partnerships with industry). Programs provide students with the opportunity to learn all aspects of an industry (e.g., the construction management POS includes instruction in related software, labor issues, and the scheduling tasks that contractors normally perform).

These CTE programs meet the requirements of POS and have proven to be popular enough that the district holds a lottery to select and assign student applicants to their high schools of choice. Three of these high schools are participating in the experimental strand of this study. These three high schools differ from each other in structural ways that allow us to extend the generalizability of the findings beyond any one high school.

School #1, Navajo HS,[2] is housed in a brand-new, specially designed facility. Its CTE programs are housed in modular units that have been customized through input from the Joint Technical Skills Committees for that program area. Academic classrooms are also located within the CTE modules, facilitating collaboration. The CTE curriculum begins in the ninth grade. This new school was designed around "best practice" tenets of contemporary education (e.g., learning communities, integrated curriculum, project-based curriculum, and the expectation that students will earn postsecondary credits while in high school and subsequently attend college).

School #2, Sioux HS, is the district's former career center, a 40-year-old facility that has been renovated and modernized in some areas. The integration of academic and CTE content has not occurred to the extent that it has at Navajo, but Sioux has a strong focus on academic subjects—it is not the vocational center of yore. As the district began to upgrade its CTE offerings, Sioux improved its academic offerings in order to continue attracting students. It now offers many AP classes and boasts a strong pass rate on the state high-stakes exam. The CTE curriculum begins here in the sophomore or junior year, which is a somewhat more traditional timetable than that found at Navajo. Although not as modern as Navajo, Sioux retains a reputation of high performance, safety, and certification/licensure opportunities that makes it popular with both parents and students.

School #3, Apache HS, is a magnet high school of wall-to-wall academies; in previous years it offered high-tech CTE programs (i.e., computer science) as a complement to Sioux's more traditional programs before the district overhaul of CTE. Other schools offer the same information technology programs as Apache, but it boasts "a culture of technology in both the core and program classes." Apache maintains a reputation as a high-performing school. Like Navajo, the CTE curriculum begins in ninth grade.

POS are being implemented at all three of these very different sites. One school has a long history of delivering CTE, another has built a reputation of "academic distinction and technological excellence," and the third was just opened in the last few years. Each school delivers POS slightly differently: one through a project-based curriculum, another through an academic-technological focus, and the third through more traditional CTE infused with high-level academics. These differences reflect practices employed in many high schools implementing POS across the country, so the findings from these replications can be generalized beyond what could be done with any one high school.

Eighth-grade students applied to the POS schools and participated in a lottery in 2008 for the 2008-2009 entering freshman class. In the three participating POS schools in this district, two had around 1,800 eligible applicants each for about 500 seats, and the third school, a smaller school, had about 1,000 qualified applicants for about 300 seats.

[2] All school names are pseudonyms.

There are two tiers of student eligibility, depending on the specific POS a student is applying to (see Table 2). The following table depicts the number of POS each school contains and how many require rigorous versus standard eligibility requirements.

Table 2
District 1 POS Schools, by Number of POS and Eligibility Requirements

School	*Total Number of POS*	*POS Requiring Rigorous Eligibility*	*POS Requiring Standard Eligibility*
POS School #1 (Navajo)	9	3	6
POS School #2 (Sioux)	8	3	5
POS School #3 (Apache)	8	8	0

Student applicants are assigned points based on their attendance, citizenship, GPA, and standardized test scores. The total number of points determines whether a student has met the criteria for the program she or he applied to. However, to be eligible for the rigorous programs, the added criteria of student grades in academic courses are examined. There are three extra points possible for having taken advanced (i.e., honors or accelerated) English, math, or science courses in middle school.

District 2

District 2 is located in a large city in the Eastern United States. It serves almost 150,000 students, over 65% of whom self-identify as ethnic minorities. Almost half of the students in this district are eligible for the federal free lunch program. Table 1 summarizes important characteristics of the district. For example, District 2 is larger than the national average, and the percentage of minority and free lunch-eligible students enrolled in its schools is substantially larger. Although not as large as District 1, District 2 also has more teachers and more schools than the national average, and its per-pupil expenditure is lower than the national average. The population density per square mile in the District 2 catchment area is greater than the national average. Like District 1, District 2 has a higher percentage of people who identify as minorities than the national average, yet fewer people living below the poverty line. The percentage of youths ages 16 to 19 who are not enrolled in school in District 2 is similar to the national average. The overall unemployment rate is the same in District 2 as the national average, although the youth unemployment rate is higher. Workers in the District 2 area earn higher median incomes; there are more high school graduates than the national average.

District 2 has multiple magnet programs across all grade spans. At the high school level, a number of schools (or programs within schools) offer a range of magnets that are available to students through a lottery process. Applicants must demonstrate successful completion of core competency standards (in reading, math, and computers). Cherokee Academy is a state-of-the-art high school featuring three integrated technology academies combining rigorous academics with strong, well-articulated POS in engineering, medical sciences and biotechnology, and information technology. Cherokee is in a state with secondary/postsecondary statewide articulation agreements, assuring students a seamless transition and the opportunity to earn

college credit that can be applied to an associate degree or postsecondary certificate program. There are also local articulation agreements between the district and the local community college.

Cherokee was poised to be another experimental site for the study, but after submitting an application to conduct research in the district, we learned that there had been no lottery in 2008 when the target cohort of ninth graders applied. The district had increased the number of students allowed to enroll in the school; with no oversubscription, no lottery was necessary. Our negotiations with the district instead turned to designing a quasi-experimental strand and the details of selecting a well-matched comparison group.

We sought to retain District 2 because of the excellent POS we found at Cherokee that provided advanced technological training to obviously engaged students. As a result, we designed a quasi-experimental strand for this site, which provides a fourth look at POS in a different environment than the three sites in District 1. POS have a different developmental history in this district, and our goal is to capture that variability and how it affects outcomes of interest. In each case, we will compare the sample to the entire district and to national data, such as the Education Longitudinal Study of 2002 (ELS), in order to assess their generalizability.

Although we have done our best to ensure that the control or comparison conditions are "business as usual" and do not contain programs that are CTE POS, there are nonetheless many options for students in both of these districts. In District 1, there are two control groups. The first is being used for a dosage analysis, and is comprised of students who did not attend other magnet schools or programs. But many "home high schools" offer academy-like experiences. The second control group includes all students who applied to a POS school, which necessarily includes students in other school choice options. This group is being used for an "intent-to-treat" analysis. In District 2, the comparison group comes from four high schools that are most similar to the intervention school. These high schools likely also have programs that offer students more than "business-as-usual" experiences. There are few high schools, especially in large, innovative districts like these, that are not engaged in some sort of reform (Lee & Smith, 2001). This fact could raise the bar on this study, making it harder for the treatment and intervention conditions to have additional significant impact. We are minimizing this by paying attention to the programs offered at these control and comparison schools, and will report details on the control and comparison conditions at each site.

District 1 Baseline Student Sample

A lottery is held every spring in District 1 for oversubscribed high school choice options. Like many districts, certain lottery preferences are provided in order to minimize student travel and maximize student diversity across schools. District 1 reserves 25% of the available seats at any school for sibling preference, 25% for geographical preference, and 25% for feeder preference. The remainder of the seats available (and those not taken in the preference lotteries) are utilized in a general lottery.

Applicants may submit one application per school. Each application is reviewed for eligibility. Special education students receive individual reviews; both program area and special education

professionals determine eligibility. All qualified applicants are then placed into the general lottery. Students selected by the lottery are mailed a notification with a Letter of Intent, which must be signed by the student's parent/guardian and returned to the district lottery office within 10 working days of the mailing. Students not selected at this time are randomly ordered in an alternate pool. Failure to return the letter by the due date will result in the student losing the seat.

Students who apply and are selected to more than one school are required to choose only one school. This choice is made by signing and returning the Letter of Intent indicating the student's choice. Once this selection is made, that student's applications for other schools are no longer considered.

After all Letters of Intent have been received, if there are still available seats, an alternate lottery is held. Any applicant who was not selected in the initial lottery and did not accept a seat for another program is placed in the alternate lottery. As seats become available, students are contacted in the (randomly determined) order in which they appear in the pool. Once a student has accepted any seat, his or her name is removed from all other alternate pools.

Late applications are accepted and are defined as any application submitted after the deadline of the initial lottery. A late application process is only utilized for schools that have remaining seats available after the qualified alternate pool has been exhausted. As openings become available, the lottery office runs a computerized random lottery of late applicants.

Identification of Student Samples for District 1

One year after the lottery of Spring 2008, the records of all students who had applied to each of the three treatment schools were separated into two groups each, still separated by school: those who were attending the POS school and those who were not. These files were sent from the district lottery office to the district research office. There, the files were merged into one large treatment group and one large control group, in order to minimize the destinations for the student information to be provided. There were 1,264 students in the treatment dataset and 4,285 students in the control dataset. This section provides more detail on those datasets.

Treatment groups. The experimental sample was formed from a dataset that consisted of ninth-grade students enrolled in one of the three experimental schools. Three experimental datasets were formed, one for each treatment school. This resulted in 509 treatment students at Navajo, 481 students at Sioux, and 274 students at Apache.

There were four possible lottery outcomes listed for the treatment school in which the student was enrolled:

1. *Not Qualified* – The student did not meet eligibility requirements, appealed, and was entered into a later lottery.
2. *Alternate Pool* – The student met eligibility requirements, was entered into the lottery, was not selected but put in the alternate pool, and was selected through subsequent lotteries.

3. *Selected* – The student met eligibility requirements, was entered into the lottery, was selected, and enrolled in the school.
4. *Late Lottery* – The student turned in the application after the deadline, but because seats remained after the lottery, the student was able to enter a late lottery and was selected.

There were two ways for students who had been identified as *Not Qualified* to get into the treatment group. First, they could appeal the decision and be put into a later lottery if their appeal showed they were eligible. Second, students who were not qualified for the program they applied to could be placed into another program at that school for which they did qualify, if there were seats available after all lotteries had been exhausted. There was only a small percentage of students who were identified as *Not Qualified* at each school (3.9% at Navajo, 3.7% at Sioux, and 1.5% at Apache, respectively).

Because students could apply to more than one school, there were 445 cases of students who were in both district datasets (i.e., students who enrolled in one of the treatment schools but who had also applied to another of these schools but did not enroll). Of those 445, 351 met eligibility requirements. These students remained in both datasets; as a result, there are cases in which a student is in the control group for one or more treatment schools while also being in the sample for a different treatment school. Thus:

- School #1 (Navajo) had 126 crossover students in its control group (52 at Sioux; 74 at Apache)
- School #2 (Sioux) had 183 crossover students in its control group (96 at Navajo; 87 at Apache)
- School #3 (Apache) had 135 crossover students in its control group (77 at Navajo; 58 at Sioux)

Control groups. The control sample was derived from the second dataset which contained 2,853 usable records from students who had applied to at least one of the experimental schools. We developed a control group for each experimental school. To do this, we first eliminated from the dataset of potential control students all students who had applied to one or more schools and did not qualify for any of them. This reduced the original sample to 1,570 students. From that reduced sample, we then selected a control group for each experimental school by limiting the respective control group to students who had applied at least once to the experimental school of interest and had qualified at least once for the lottery. Because a student could apply to more than one of the experimental schools, they could be in more than one control group if they qualified for each school to which they applied. This process resulted in 661 control students for Navajo, 873 control students for Sioux, and 527 control students for Apache.

The file of students who had applied to one of the POS schools presented 4,285 students by their current school location. Of these, 1,132 were duplicate records (students who applied to more than one school or program) and were removed. There were 300 students for whom there was no record of their having enrolled in the ninth grade in this district for the Fall of 2008. These were students who had either formally or informally withdrawn from the district between the time they applied for the lottery and the subsequent fall. This brought the original sample to 2,853.

The control sample dataset provided information on up to four schools applied to and the outcome of each application. For each school applied to, one of three lottery outcomes was possible:

1. *Not Qualified* – The student did not meet eligibility requirements and thus was not entered into the lottery.
2. *Alternate Pool* – The student met eligibility requirements, was entered into the lottery, was not selected, and was placed in the alternate pool. From here it is impossible to tell if the student was ultimately selected and turned down a seat or was not selected.
3. *Selected* – The student met eligibility requirements, was entered into the lottery, was selected, but did not enroll in the school.

A control group for each treatment school was formed based on the outcomes for each student at each school to which the student applied.

Background Characteristics of District 1 Student Sample

A comparison of student background characteristics was conducted for each treatment school and its respective control school. Comparisons of gender, race, free lunch status (as a proxy for poverty level), Limited English Proficiency (LEP) status, Individual Education Plan (IEP) status, age, and eighth-grade GPA (both unweighted and weighted) were performed. School engagement was assessed using eighth-grade average daily attendance and the variable *Number of disciplinary occurrences*, which was collapsed into four categories: (1) none, (2) one to two occurrences, (3) three to 10 occurrences, and (4) more than 10 occurrences. Achievement was assessed using results of two tests taken in eighth grade. Scaled scores in reading, math, and science from the state's criterion referenced test (CRT) were compared as well as percentile scores in reading, language, and math from the norm-referenced ITBS (International Test of Basic Skills).

Crosstabulations and chi-square analyses were performed on gender, race, free lunch status, Limited English Proficiency status, IEP status, and number of disciplinary occurrences. *T* tests were performed to compare age, eighth-grade average daily attendance, eighth-grade GPA (unweighted and weighted), and achievement.

Who Was Lost to the Sample as a Result of the Lottery Eligibility Requirements

As in most district lotteries for oversubscribed programs, the programs in these POS high schools have eligibility requirements. Many students applied who were not qualified for these programs. This section describes the students who were lost to the sample as a result of failing to meet the eligibility requirements of the programs to which they applied. Tables are not provided for these descriptions but are available upon request.

Qualified was defined as (1) being in a treatment school or (2) being in the control dataset and having an outcome of selected or alternate pool for that school. *Not Qualified* was defined as being in the control dataset and having an outcome of not qualified for that school. If a student in

the control dataset applied to more than one program in a school and was labeled as not qualified for one but selected or alternate pool for the other, the student was counted as qualified.

The race/ethnicity and poverty level of a student were not criteria for eligibility for the lottery, but the district had regional preferences built in, so students from some minority groups and low-income students were differentially likely to enter the lottery for certain school choice options located in the various parts of the district. However, POS high schools were located throughout the district, so although a student might be less likely to be selected at one of these schools, they would be more likely to be selected at another.

For this reason as well as others, significantly fewer Black and Latino students were qualified to enter the lottery for Navajo and Apache, whereas significantly more White and Asian students were qualified for these two schools. At Sioux, Black students were equally likely to qualify for the lottery. Significantly fewer Latino students were qualified to enter the lottery for Sioux, whereas significantly more White and Asian students were qualified.

Similarly, gender was not a criterion for admission into any of the lottery-based high schools. However, there were significantly more males who were not qualified for the lotteries at all three schools than males who were qualified. Gender is often associated with middle school achievement, which was an eligibility requirement, and could explain the gender imbalance in the qualified pools of students for these three POS schools.

In this district, special education status is recorded in the system by the presence of an IEP. When IEP students apply to the school choice options, that information is used to identify the application for a second review. This review is done for each individual student; program area and special education professionals jointly determine eligibility. At Navajo and Apache, students with an IEP were significantly less likely to be qualified for the lottery. At Sioux, students with an IEP were equally likely to be qualified or not qualified for the lottery.

At Sioux and Apache, LEP students were less likely to be qualified for the lottery than students who were not classified as LEP. At Apache, younger students were more likely to be qualified for the lottery than older students.

Significantly higher numbers of qualified students had no disciplinary occurrences in the previous school year than the not qualified students for Navajo, Sioux, and Apache. An eligibility requirement for the lottery was eighth-grade citizenship grades but these were not available to us in a reliable fashion, so we report disciplinary occurrences instead.

Attendance and prior achievement were eligibility requirements for the lottery, so it is not surprising that at Navajo and Apache, qualified students had significantly higher attendance rates during the previous school year and significantly higher GPAs, CRT scores, and ITBS scores than those who did not qualify. At Sioux, attendance and GPA were higher among qualified than not qualified students. However, the CRT and ITBS scores are not significantly different between qualified and not qualified students.

As far as the entire baseline sample is concerned, the three separate comparisons all lose student applicants who share these characteristics: They are Latino, low-income, male, and special education students as well as students with at least one disciplinary occurrence, lower attendance rates, and lower eighth-grade GPAs. At Apache, the qualified and not qualified students differ significantly on every measure we collected. Because of these significant differences at baseline, we will make the appropriate adjustments with covariates in the final analyses.

As a final word on which students were lost to the sample as a result of the admissions procedures, we had wanted to compare students who were attending a POS high school with the students who had participated in the lottery and won a seat in one of the schools but either turned it down or did not return the Letter of Intent. Unfortunately, this information was not fully specified for all students in the control sample. For those students who were not initially selected and went into the alternate pool, we could not determine whether they were subsequently selected but did not enroll or if they were never selected. This omission from the lottery records made this important comparison impossible.

Comparisons of District 1 Treatment and Control Schools

School #1 (Navajo) Compared to Its Control Group

***Demographics*.** Table 3 shows that although both samples are predominantly female, the control group has significantly more female students (67.5%) than Navajo (56.2%). There is a significant difference in racial composition, with Navajo having significantly fewer Asian students (12.4% vs. 16.5%), significantly fewer Latino students (22.2% vs. 35.3%), and significantly more White students (53.0% vs. 35.3%). The control group has significantly more students receiving free lunch (31.1%) than Navajo (21.0%). Each school has a very small LEP population, ranging from 1.8% to 2.3%, and few students with an IEP (2% at Navajo vs. 3.2% in the control group).

Table 3
Baseline Characteristics of Student Sample, School #1 (Navajo)

Baseline Characteristics	*School #1 (Navajo) (N = 509) (%)*	*Control #1 (N = 661) (%)*	*Total (N = 1,170) (%)*
Demographics			
Gender			
Male	43.8**	32.2	37.3
Female	56.2	67.8	62.7
Race/Ethnicity			
Black	11.6	12.3	12.0
White	53.0**	35.6	43.2
Latino	22.2**	35.4	29.7
Asian	12.4*	16.5	14.7
Native American	0.8	0.3	0.5
Other/Multiracial	0.0	0.0	0.0
Poverty			

Receives free lunch	21.0***	31.3	26.8
Does not receive free lunch	79.0	68.7	73.2
Limited English Proficiency			
Yes	1.8	2.3	2.1
No	98.2	97.7	97.9
Has Individual Education Plan			
Yes	2.0	3.2	2.6
No	98.0	96.8	97.4
School Engagement			
# Disciplinary Occurrences			
None	74.1**	80.6	77.8
One to two	18.1	15.3	16.5
Three to ten	7.5	4.1	5.6
More than ten	0.4	0.0	0.2

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

There is no significant difference in age or eighth-grade average daily attendance (see Table 4). The majority of students in both groups are between the ages of 14 and 15. The average eighth-grade daily attendance in both groups was 97%. A significant difference in average eighth-grade GPA (both weighted and unweighted) was found favoring the control students. For both groups, eighth-grade GPA ranged from 3.4 to 3.5.

Table 4
Baseline Characteristics of Student Sample, School #1 (Navajo)

Baseline Characteristics	*N*	*School #1 (Navajo) Mean (SD)*	*N*	*Control #1 Mean (SD)*	*N*	*Total Mean (SD)*
Age	509	14.71 (0.51)	661	14.70 (0.52)	1170	14.70 (0.52)
Eighth-Grade Attendance	459	0.97 (0.03)	647	0.97 (0.03)	1106	0.97 (0.03)
Eighth-Grade GPA – not weighted	464	3.39 * (0.47)	650	3.45 (0.45)	1114	3.42 (0.46)
Eighth-Grade GPA –weighted	464	3.40 * (0.48)	650	3.47 (0.46)	1114	3.44 (0.47)

Note. * $p < .05$.

School engagement. A significant difference in the number of disciplinary occurrences was found, with 19.4% of treatment students having at least one occurrence compared to 23.9% of control students (see Table 3).

Achievement. According to district records, the majority of treatment and control students (87% for control vs. 84% for treatment) met or exceeded standards for proficiency in reading, math, and science on the state's eighth-grade CRT exam. A comparison of scaled scores revealed that

students from Navajo did not perform as well on the reading and math subtests as their control school counterparts (see Table 5). Comparisons of the eighth grade ITBS test showed that students from both groups performed between the 61st and 63rd percentile for reading, the 65th and 66th percentile for language, and 65th and 68th percentile for math. The difference in performance on the math subtest significantly favored the control group.

Table 5
Prior Achievement of Student Sample, School #1 (Navajo)

Achievement (Scaled Scores)	*N*	*School #1 (Navajo) Mean (SD)*	*N*	*Control #1 Mean (SD)*	*N*	*Total Mean (SD)*
CRT-Reading Scaled Score	458	347.71* (53.16)	650	355.36 (53.08)	1108	352.16 (53.22)
CRT-Math Scaled Score	458	369.44* (69.96)	650	379.08 (71.28)	1108	375.10 (70.87)
CRT-Science Scaled Score	458	370.39 (60.50)	650	369.47 (64.62)	1108	369.85 (62.92)
ITBS Reading Percentile	439	60.51 (25.04)	610	63.03 (24.15)	1049	61.98 (24.54)
ITBS Language Percentile	438	64.47 (23.94)	604	66.17 (24.58)	1042	65.45 (24.32)
ITBS Math Percentile	438	64.64* (23.64)	607	68.39 (23.97)	1045	66.82 (23.89)

Note. $^{*}p < .05$.

School #2 (Sioux) Compared to Its Control Group

Demographics. In this comparison, shown in Table 6, both samples are predominantly female but the control sample has significantly more female students (65.5%) than Sioux (55.5%). There is no significant difference in racial composition or percent receiving free lunch. Both schools are comprised of predominantly Latino or White students and students who do not receive free lunch. Each school has a small LEP population (3.5% to 2.7%) and few students (about 5%) with IEPs.

Table 6
Baseline Characteristics of Student Sample, School #2 (Sioux)

Baseline Characteristics	*School #2 (Sioux) (N = 481) (%)*	*Control #2 (N = 873) (%)*	*Total (N = 1,354) (%)*
Demographics			
Gender			

Male	44.5**	34.5	38.0
Female	55.5	65.5	62.0
Race/Ethnicity			
Black	11.0	11.2	11.2
White	28.5	27.5	27.8
Latino	50.9	46.5	48.1
Asian	8.5	14.0	12.0
Native American	1.0	0.8	0.9
Other/Multiracial	0.0	0.0	0.0
Poverty			
Receives free lunch	40.7	36.8	38.2
Does not receive free lunch	59.3	63.2	61.8
Limited English Proficiency			
Yes	3.5	2.7	3.0
No	96.5	97.3	97.0
Has Individual Education Plan			
Yes	5.0	5.3	5.2
No	95.0	94.7	94.8
School Engagement			
# Disciplinary Occurrences			
None	73.2*	81.0	78.2
One to two	20.6	14.5	16.7
Three to ten	6.0	4.5	5.0
More than ten	0.2	0.0	0.1

Note. * $p < .01$, ** $p < .001$.

There was no significant difference in age or eighth-grade average daily attendance (see Table 7). The majority of students in both groups were between the ages of 14 and 15. The average daily attendance in both groups was 97%. A significant difference in average eighth-grade GPA (both weighted and unweighted) was found favoring control students. For both groups, eighth-grade GPA ranged from 3.3 to 3.4.

Table 7
Baseline Characteristics of Student Sample, School #2 (Sioux)

Baseline Characteristics	*N*	*School #2 (Sioux) Mean (SD)*	*N*	*Control #2 Mean (SD)*	*N*	*Total Mean (SD)*
Age	481	14.70 (0.55)	873	14.71 (0.53)	1354	14.71 (0.54)
Eighth-Grade Attendance	471	0.97* (0.04)	865	0.97 (0.03)	1337	0.97 (0.03)
Eighth-Grade GPA – not weighted	473	3.16*** (0.47)	872	3.33 (0.45)	1345	3.28 (0.47)

Eighth-Grade GPA –weighted	473	3.17*** (0.48)	872	3.35 (0.46)	1345	3.29 (0.48)

Note. * $p < .05$. *** $p < .001$.

School engagement. A significant difference in the number of disciplinary occurrences was found, with 26.8% of treatment students having at least one occurrence compared to only 19 % of control students (see Table 6).

Achievement. According to district records, significantly more control students met or exceeded standards for proficiency in reading, math, and science on the state's eighth-grade CRT exam than treatment students (82% vs. 73% for reading, 82% vs. 71% for math, and 81% vs. 75% for science). A comparison of scaled scores revealed that students from Sioux did not perform as well on the reading, math, or science subtests as control students and the differences were significant (see Table 8). Comparisons of the eighth-grade ITBS test showed that students from both groups performed between the 51st and 57th percentile for reading, the 55th and 61st percentile for language, and 56th and 62nd percentile for math. The control students performed significantly better than students from Sioux on reading, language, and math.

Table 8
Prior Achievement of Student Sample, School #2 (Sioux)

Achievement (Scaled Scores)	*N*	*School #2 (Sioux) Mean (SD)*	*N*	*Control #2 Mean (SD)*	*N*	*Total Mean (SD)*
CRT-Reading Scaled Score	471	324.82*** (53.18)	871	340.14 (52.28)	1342	334.77 (53.08)
CRT-Math Scaled Score	471	335.48*** (76.09)	871	358.63 (76.21)	1342	350.38 (76.92)
CRT-Science Scaled Score	471	337.89*** (66.42)	869	354.02 (66.09)	1340	348.35 (66.63)
ITBS Reading Percentile	441	51.22*** (24.25)	823	56.82 (24.39)	1264	54.87 (24.48)
ITBS Language Percentile	442	54.6*** (24.69)	819	60.78 (25.34)	1261	58.62 (25.28)
ITBS Math Percentile	442	56.34*** (25.89)	815	62.33 (25.17)	1257	60.22 (25.57)

Note. *** $p < .001$.

School #3 (Apache) Compared to Its Control Group

Demographics. Table 9 shows that Apache has significantly more male students (61.7%) than its control group (46.5%). There is a significant difference in racial composition, with Apache having significantly more Asian students (29.2% vs.19.7%) and significantly fewer Latino students (20.8% vs. 31.3%). The control group has significantly more students who receive free

lunch (29.6%) than Apache (17.9%). Each group was similar in having a very small LEP population (less than 1%) as well as very few students with an IEP (1.5% or less).

Table 9
Baseline Characteristics of Student Sample, School #3 (Apache)

Baseline Characteristics	*School #3 (Apache) (N = 274) (%)*	*Control #3 (N = 527) (%)*	*Total (N = 801) (%)*
Demographics			
Gender			
Male	61.7**	46.5	51.7
Female	38.3	53.5	48.3
Race/Ethnicity			
Black	9.1	12.1	11.1
White	40.1	36.4	37.7
Latino	20.8*	31.3	27.7
Asian	29.2*	19.7	23.0
Native American	0.7	0.4	0.5
Other/Multiracial	0.0	0.0	0.0
Poverty			
Receives free lunch	17.9**	29.6	25.6
Does not receive free lunch	82.1	70.4	74.4
Limited English Proficiency			
Yes	0.7	0.6	0.6
No	99.3	99.4	99.4
Has Individual Education Plan			
Yes	1.5	1.3	1.4
No	98.5	98.7	98.6
School Engagement			
# Disciplinary Occurrences			
None	80.7	83.7	82.6
One to two	16.4	13.5	14.5
Three to ten	2.9	2.7	2.7
More than ten	0.0	0.2	0.1

Note. *p < .01, **p < .001.

Table 10 shows that there is no significant difference in age, eighth-grade average daily attendance, or eighth-grade GPA (weighted or unweighted). The majority of students in both groups are between the ages of 14 and 15. The average daily attendance in both groups was 98%. The average eighth-grade GPA for both groups ranged from 3.5 to 3.6.

Table 10
Baseline Characteristics of Student Sample, School #3 (Apache)

Baseline Characteristics	*N*	*School #3 (Apache) Mean (SD)*	*N*	*Control #3 Mean (SD)*	*N*	*Total Mean (SD)*
Age	274	14.67	527	14.71	801	14.69
		(0.55)		(0.52)		(0.53)
Eighth-Grade Attendance	248	0.98	517	0.98	765	0.98
		(0.03)		(0.03)		(0.03)
Eighth-Grade GPA – not weighted	249	3.54	519	3.59	768	3.57
		(0.42)		(0.36)		(0.38)
Eighth-Grade GPA –weighted	249	3.57	519	3.62	768	3.60
		(0.43)		(0.37)		(0.39)

School engagement. There was no significant difference in number of disciplinary occurrences with 19.3% of students having one or more occurrences at Apache compared to 16.3% of students in the control group (see Table 9).

Achievement. According to district records, the majority of treatment and control students (more than 90%) met or exceeded standards for proficiency in reading, math, and science on the state's eighth-grade CRT exam. A comparison of scaled scores revealed that students from Apache performed better on the math and science subtests than their control counterparts (see Table 11). Comparisons of the eighth-grade ITBS test showed that students from both groups performed between the 72nd and 74th percentile for reading, the 74th and 76th percentile for language, and 78th and 82nd percentile for math. The difference in performance on the math subtest significantly favored the treatment group.

Table 11
Prior Achievement of Student Sample, School #3 (Apache)

Achievement (Scaled Scores)	*N*	*School #3 (Apache) Mean (SD)*	*N*	*Control #3 Mean (SD)*	*N*	*Total Mean (SD)*
CRT-Reading Scaled Score	247	368.68	518	372.32	765	371.14
		(47.21)		(46.05)		(46.43)
CRT-Math Scaled Score	247	417.60*	518	407.89	765	411.02
		(54.65)		(61.26)		(60.25)
CRT-Science Scaled Score	247	406.16**	516	393.12	763	397.34
		(56.66)		(55.20)		(55.97)
ITBS Reading Percentile	229	73.74	489	72.21	718	72.70
		(20.10)		(19.56)		(19.73)
ITBS Language Percentile	228	75.85	486	74.34	714	74.82
		(19.61)		(19.50)		(19.53)
ITBS Math Percentile	229	82.03**	487	78.07	716	79.34
		(16.45)		(19.15)		(18.41)

Note. $^{*}p < .05$, $^{**}p < .01$.

District 2 Baseline Student Sample

The intervention group in District 2 consists of the students enrolled at Cherokee Academy at the start of their ninth-grade year. The comparison group was created by the District 2 research office after the prospect of an RCT design was eliminated.

District 2 Comparison Group Identification

Because students were not randomly assigned to treatment and control conditions, different sampling methods were employed. First, District 2 personnel ran a cluster analysis on the other high schools in the district, searching for four comparison schools that were most similar to Cherokee in terms of student demographic variables, mobility, and prior achievement.

Then propensity score matching (PSM) was employed to identify a comparison group. Propensity scores are the estimated probability that a subject is assigned to an intervention based on various pieces of information (Pasta, 2000). This predicted probability is obtained by conducting a logistic regression that predicts membership in the intervention group utilizing a vector of covariate predictors, in effect mitigating the effects of inherent differences among students in the different study conditions (Dehejia & Wahba, 2002; Foster, 2003; Shadish et al., 2002). Students with similar distributions across the covariates will have similar estimated propensity scores. Thus a student in the intervention condition can be matched with a comparison student possessing a similar propensity score, yielding reduced-bias estimates of effects during analysis (Rosenbaum & Rubin, 1985).

A matrix of intervention group variables was selected on which to match students from the comparison schools identified in the cluster analysis:

- Performance on standardized eighth-grade reading and math tests
- Gender
- Race/ethnicity
- Limited English Proficiency (LEP) status
- Special education classification
- Free lunch status (as a proxy for poverty)
- Eighth-grade retention
- Number of violent incidents during the school year in which the student was involved
- Number of nonviolent incidents during the school year in which the student was involved
- Number of unexcused absences
- Number of excused absences
- Number of out-of-school suspensions
- Number of in-school suspensions
- Mobility during eighth grade (the number of school changes)
- Performance on standardized eighth-grade reading and math tests

Once predicted probabilities were obtained, intervention students were matched with comparison students possessing exact probabilities when such were available, and similar probabilities with minimal difference when an exact match was not available. This matching procedure was conducted without replacement, so that a student identified as a comparison student for a particular intervention student was removed from the potential comparison pool before the next match for an intervention student was conducted. The propensity match procedure was carried out twice to provide two matched comparison students for every intervention student, to take into account attrition over the next three years of the study. A comprehensive file was created containing student background information from the eighth-grade year (2007-2008) of all intervention and comparison students.

Background Characteristics of District 2 Student Sample

Tables 12 and 13 show the results of the propensity score matching and the background characteristics of the sample for this strand of the study.

Table 12
Baseline Characteristics of District 2 Student Sample

Background Characteristics	*Cherokee (N=376) (%)*	*Comparison (N=752) (%)*	*Total (N=1228) (%)*
Student Demographic Characteristics			
Gender			
Male	52.1	49.9	50.6
Female	47.9	50.1	49.4
Race/Ethnicity			
Black	71.0	70.5	70.7
White	10.6	11.2	11.0
Latino	10.4	10.6	10.5
Asian	3.7	3.6	3.6
Native American	0.3	0.1	0.2
Other/Multiracial	4.0	4.0	4.0
Poverty			
Receives free lunch	68.4	66.4	67.0
Does not receive free lunch	31.6	33.6	33.0
Limited English Proficiency			
Limited English proficient	4.3	4.8	4.6
Not limited English proficient	95.7	95.2	95.4
Has an Individual Education Plan			
Yes	2.4	2.3	2.3
No	97.6	97.7	97.7
Repeated eighth grade in 2007-2008			
Retained	1.6	2.3	2.0
Not retained	98.4	97.7	98.0
Mobility in 2007-2008			

No school changes	85.4	98.5	88.1
One to three school changes	14.6	10.5	11.9
High School			
Intervention – Cherokee			376
Comparison 1			228
Comparison 2			148
Comparison 3			304
Comparison 4			72
Student Engagement in School			
Nonviolent incidents			
None	76.3	75.9	76.1
One to Two	17.0	17.0	17.0
Three to Ten	6.6	6.8	6.7
More than Ten	0.0	0.3	0.2
Violent incidents			
None	100.0	100.0	100.0
One to Two	0.0	0.0	0.0
Three to Ten	0.0	0.0	0.0
More than Ten	0.0	0.0	0.0
Unexcused absences			
None	22.1	22.5	22.3
One to Two	28.2	32.4	31.0
Three to Ten	44.9	40.7	42.1
More than Ten	4.8	4.4	4.5
Excused absences			
None	66.8*	61.0	62.9
One to Two	17.0	25.3	22.5
Three to Ten	15.2	12.6	13.5
More than Ten	1.1	1.1	1.1
Ever had an out-of-school suspension			
No	79.8	83.9	82.5
Yes	20.2	16.1	17.5
Ever had an in-school suspension			
No	85.6	85.1	85.3
Yes	14.4	14.9	14.7
Ever had an excused tardy			
No	43.4	38.0	39.8
Yes	56.6	62.0	60.2
Ever had an unexcused tardy			
No	13.8	18.4	16.8
Yes	86.2	81.6	83.2

Note. $*p < .05$.

Student engagement in school. On most variables, the intervention and comparison samples were very similar. There were no few violent for either group. About a quarter of each group had at least one nonviolent incident. Few had more than two incidents (7%). There was little

difference between groups on number of unexcused absences: A little under one third of each group had "one to two" unexcused absences. More students had "three to ten" unexcused absences (28% treatment compared to 32% for control). There were far fewer excused absences with about two thirds of the sample having no excused absences. However, the treatment and control groups did differ significantly on this variable, with the treatment group having 5.8% more excused absences. Eighty percent of treatment students and 84% of control students had never had an out-of-school suspension. About 85% of students in each group had never had an in-of-school suspension. More than half of each group had had an excused tardy and over 80% had had unexcused tardy.

Student achievement. Table 13 shows no significant differences in eighth-grade reading or math state test scores. According to district records, over 80% of students had performed at proficient or advanced level in math. A smaller percentage was proficient or advanced in reading, a little over 60%.

Table 13
Prior Achievement of District 2 Student Sample

Achievement (Scaled Scores)	*Cherokee Mean (SD)*	*Comparison Mean (SD)*	*Total Mean (SD)*
	(N=376)	*(N=752)*	*(N=1128)*
Eighth-Grade Reading	359.91 (7.10)	359.69 (7.49)	359.76 (7.36)
Eighth-Grade Math	363.22 (6.96)	362.90 (7.54)	363.00 (7.35)

Late transfers. Some students in the intervention group transferred to Cherokee Academy after the beginning of the school year. Of the 376 students at Cherokee, most (n = 370) were recorded as officially entering the school in July, August, or September 2008—prior to or just after the start of the school year. One student came to Cherokee in October 2008. Another came in November of 2008. Another came in January and two came in February of 2009. Finally, one came in May of 2009. These six students did not receive the same amount of the intervention as the other students in the sample and will be flagged accordingly in subsequent analyses.

Summary

In this first year of a longitudinal study, we have learned a good deal about POS and how they operate in some districts. We have learned that POS (along with International Baccalaureate and other high school choice programs) are popular enough across the nation that many districts have had to develop a way to fairly allocate these opportunities. In some districts, new sections of popular programs are added. In others, applicants are rank ordered and offered admission in that order. In still other districts, lotteries are held for such programs, and it was among these that we focused our efforts to find sites amenable to an experimental evaluation of POS.

We have also learned that districts were already upgrading CTE programs prior to the passage of Perkins IV in ways that in essence met the criteria for POS without calling them POS. Neither of the districts in this study refers to these programs as POS, although we will use this term

throughout this paper and in other work. Instead, these districts' POS are known as *academies*, a term perhaps more familiar to parents and students. These POS were developed based on best practices throughout the nation. CTE directors, like other professionals, are aware of developments in their field through participation in conferences and other communications. Best practices spread to districts just as they do to the federal level, where new mandates develop based on what states and localities find feasible and promising. By the time the Perkins IV legislation was passed, states and localities were already implementing what would legally become known as POS. We sought out such sites so that this study could examine well-functioning POS in practice as opposed to documenting efforts to develop and implement them.

Early observations about POS are presented jointly with those of other field-based POS studies from the National Research Center on Career and Technical Education (Programs of Study Joint Technical Working Group, 2009).[3]

[3] See http://136.165.122.102/UserFiles/File/Tech_Reports/POS_Joint_Technical_Report_Jan_2010.pdf

References

Carl D. Perkins Career and Technical Education Improvement Act of 2006. Pub L. No. 109-270. (2006).

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, T. D., & Payne, M. R. (2002). Objecting to the objections to using random assignment in educational research. In F. Mosteller & R. Boruch (Eds.), *Evidence matters: Randomized trials in education research* (pp. 150-178). Washington, DC: Brookings Institution Press.

Dehejia, R., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics, 84*, 151-161.

Fisher, L. D., Dixon, D. O., Herson, J., Frankowski, R. F., Hearron, M. S., & Peace, K. E. (1990). Intention to treat in clinical trials. In K. E. Peace (Ed.), *Statistical issues in drug research and development* (pp. 331-350). New York, NY: Marcel Dekker Inc.

Foster, E. M. (2003). Propensity score matching: An illustrative analysis of dose response. *Medical Care*, *41*, 1183-1192.

Kish, L. (1965). *Survey sampling*. New York, NY: Wiley.

Lee, V. E., & Smith, J. B. (2001). *Restructuring high schools for equity and excellence: What works*. New York, NY: Teachers College Press.

Miles, M., & Huberman, M. (1984). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Mosteller, F., & Boruch, R. (Eds.). (2002). *Evidence matters: Randomized trials in education research.* Washington, DC: Brookings Institution Press.

Murphy, K. R., & Myors, B. (2004). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (2nd ed.). Mahwah, NJ: Erlbaum.

O'Donnell, C. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33-84.

Pasta, D. (2000, April). *Using propensity scores to adjust for group differences: Examples comparing alternative surgical methods*. Paper presented at the 25th Annual SAS Users Group International Conference, Indianapolis, IN.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The shopping mall high school: Winners and losers in the educational marketplace*. Boston, MA: Houghton Mifflin.

Programs of Study Joint Technical Working Group. (2009, January). *Programs of study: Year 2 joint technical report*. Louisville, KY: National Research Center for Career and Technical Education, University of Louisville. Retrieved from http://136.165.122.102/UserFiles/File/Tech_Reports/POS_Joint_Technical_Report_Jan_2010.pdf.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*, 1, 33-38.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage Publications.

U.S. Department of Education, Institute of Education Sciences. (2003). *Identifying and*

implementing educational practices supported by rigorous evidence: A user friendly guide. Washington, DC: Coalition for Evidence-Based Policy. Retrieved from http://www.ed.gov/rschstat/research/pubs/rigorousevid/index.html.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Los Angeles, CA: Sage.