Conducting Research to Answer Your Questions about Dual Enrollment

Melinda Mechur Karp
Dong Wook Jeong
Community College Research Center
Teachers College, Columbia University

Executive Summary

　　Dual enrollment, in which high school students take college courses, is an increasingly popular educational program. As states and local education authorities (LEAs) devote resources to this initiative, it is important to evaluate the effectiveness of these investments. Doing so presents unique challenges, however, because the data and analytic requirements of evaluating dual enrollment require coordination among individual high schools, LEAs, postsecondary institutions, and states. These challenges can be overcome, and the effort it takes to do so is well worth the financial and human investment. This summary is intended to help decision-makers understand why research is important and how policymakers can support research activities.

What is dual enrollment?
　　Dual enrollment programs are collaborative efforts between high schools and colleges through which high school students are permitted to enroll in college courses. Often credit earned in these courses also counts toward high school graduation requirements, an arrangement referred to as *dual credit.* In other cases, students earn only college credit. Under both arrangements, students are simultaneously enrolled in high school and college—thus, they are *dually enrolled* in the two institutions. The college credit earned through dual enrollment is recorded on the students' college transcripts, just as it would if they were regularly matriculated college students. During the 2002–2003 school year, 813,000 high school students enrolled in a college course (Kleiner and Lewis, 2005).

What do dual enrollment programs look like?
　　Dual enrollment programs range widely in their structure. Sometimes, these variations are regulated by state policies; other times, program structure is determined through cooperative agreements between high schools and colleges. Variation may occur along the following features, among others:
- Location: high school or college campus
- Instructors: college professors or high school teachers certified as college adjuncts
- Classroom mix: high school students only, or high school students and college students together
- Eligibility requirements
- Target population

What is the purpose of dual enrollment?
　　Traditionally, dual enrollment programs have been aimed at academically advanced students. The goal was to provide these students, who may have exhausted their high school's course offerings or needed academic challenge beyond what the high school could provide, with the chance to take classes that met their needs. In many places, this is still the primary reason for offering dual enrollment. Today, however, educators assume that dual enrollment participation can have additional benefits and should include a wider range of students, with the ultimate aim of improving student access to, success in, and completion of college.

Why data and research are important

We don't know much about dual enrollment's effectiveness

   Despite dual enrollment's popularity, not much is known about its effectiveness in meeting program goals, particularly those goals addressing college preparedness and completion. Two extensive reviews of the literature (Bailey and Karp, 2003; Lerner and Brand, 2006) found little evidence that dual enrollment contributed to students' college access or academic success. While a new report from the Community College Research Center (Karp, Calcagno, Hughes, Jeong, and Bailey, 2007) provides promising evidence that dual enrollment participation can lead to a range of positive outcomes for students of all backgrounds, it is only a start. At the state and local levels, studies from New York City (Michalowski, 2006; Skadberg, 2005) and Florida (Florida Department of Education, 2004) also show promising results, though those too cannot be seen as definitive evidence.

We are limited by the available data

   The dearth of research occurs, in large part, because appropriate data are not collected by programs, districts, or states. Few districts or states have been able to track students from high school into postsecondary education. As a result, analyses of dual enrollment either focus on high school completion and other short-term outcomes or examine long-term outcomes but cannot take into account differences between participants and non-participants before they entered the dual enrollment program. Studies on dual enrollment also do not generally use rigorous statistical methods, often because staff members lack capacity to conduct such analyses.

We are limited in our decision-making when we lack good information

   Although it might seem that investing in research is a low-priority item when compared to other education-funding needs, investing in research is important and can benefit states down the line. First, states and LEAs are devoting significant resources to the spread of dual enrollment programs. It is important to know whether the expenditure of these resources leads to the intended outcomes. Second, stakeholders assume that dual enrollment can address shortcomings in the current educational system; if it does not, then new reforms should be identified and implemented. Third, well-designed evaluations can help improve programs so that they effectively meet their goals.

How can policymakers encourage good dual enrollment research?

1. Develop a comprehensive state data system

   Studies of dual enrollment require data systems that follow individual students from high school to college. The most effective way to gather, store, and access this type of data is through a statewide data system that assigns each student his or her own identifying code or number and then collects transcript information for every year of his or her enrollment in school. Policymakers can encourage the creation of such a system in three ways:

- *Encourage collaboration among all entities involved in the creation of a data system.* One way to do this is to create a state-level task force that includes representatives from K-12 institutions, postsecondary institutions, and state-level

support staff charged with developing the data system and streamlining procedures for the entire state.

- *Provide funds for the establishment of the system and its continual upkeep.* This includes funding the hardware and software required and providing funding to hire staff members who are skilled in database management and data analysis. It also includes funding professional development for LEA and school staff, who will necessarily be involved in the data collection process and need to be kept up-to-date on state data requirements.

- *Ensure that data-related activities are integrated into staff members' job duties.* This means crafting job descriptions that allow for such activities. They should not be add-ons but aspects of the job that are used in hiring and evaluation decisions.

2. Construct the human infrastructure necessary for using the data system.

Having a good data system is pointless if states do not know how to make use of it. Policymakers are integral in ensuring that data are used to conduct well-designed research studies that can aid them in their decision-making.

- *Hire individuals who are schooled in research methods.* While many individuals have some familiarity with research, it is essential that states hire staff members who have expertise in actually conducting a wide range of research studies. Policymakers can ensure that such individuals are hired and supported, both by providing funding for these positions and stressing their importance.

- *Provide technical assistance to those responsible for collecting data.* This includes providing state-sponsored support to staff members in LEAs and individual schools who are participating in the state data system. Providing such support through the state sends the message that these activities are important, while also ensuring that data are collected accurately and consistently.

- *Set the tone.* Policymakers can set a tone where data-related professional development and use of the data system are expected.

3. Send the message that research is important.

Policymakers can support research by showing that they deem it valuable.

- *Make funding research activities a priority.*

- *Publicly support the research process.* By making public statements in support of research—even research that takes a long time to complete—policymakers underscore its necessity.

- *Use research to drive decisions.* The most important thing a policymaker can do to support research is to use research to make decisions. This includes working in conjunction with researchers to ensure that study designs are relevant for the current policy questions they are facing and also to ensure that they do not use the findings of studies for inappropriate purposes.

The important take-away message is that research is essential for understanding program implementation and outcomes. Decision-makers rely on this information to do their jobs. But they also have a part to play in making sure that the information exists in the first place. Without their support, good research cannot be conducted. This support

can be financial, in terms of earmarking funds for the development of data systems, research staff, and professional development. But it can also be philosophical, by sending the message to stakeholders and the public that research is indeed an important activity.

Dual enrollment, in which high school students take college courses, is an increasingly popular educational program. As states and local education authorities (LEAs) devote resources to this initiative, it is important that they evaluate the effectiveness of their investments. Doing so presents unique challenges, however, because the data and analytic requirements of evaluating dual enrollment require coordination among individual high schools, LEAs, postsecondary institutions, and states.

This paper is a tool for state and LEA personnel to use in their evaluation efforts. The executive summary was targeted at policymakers and sought to explain why research is important and how they can help support research activities. The full report is directed at those conducting research, though of course can also be valuable for those charged with policy decisions and program implementation. The paper provides a framework for thinking about and developing data systems that can be used to evaluate dual enrollment, as well as sample research activities to use as a guide for entities' future work. Throughout, we provide real-world examples of research that can serve as models for readers' own work.

I. Introduction

What is dual enrollment?

Dual enrollment programs are collaborative efforts between high schools and colleges through which high school students (usually juniors and seniors) are permitted to enroll in college courses. Often credit earned in these courses also counts toward high school graduation requirements, an arrangement referred to as *dual credit*. In other cases, students earn only college credit. Under both arrangements, students are simultaneously enrolled in high school and college—thus, they are *dually enrolled* in the two institutions.[1] The college credit earned through dual enrollment is recorded on the students' college transcripts, just as it would if they were regularly matriculated college students.

Like other credit-based transition programs such as Advanced Placement (AP) and International Baccalaureate (IB), dual enrollment programs provide students with a challenging academic experience and the opportunity to earn college credit prior to high school graduation. Unlike these other programs, however, dual enrollment students take an actual college course with a college syllabus, often on a college campus, rather than a college-level course intended to be taken by high school students. Achievement in dual enrollment courses is measured by the final course grade, rather than the score received on an examination.

Dual enrollment programs vary widely in their structure and targeted students. They may be taught on a high school campus or at a college, by regular college professors or high school teachers certified as college adjuncts, and may include college and high school students or high school students only. They vary in their eligibility requirements. Sometimes, these variations are regulated by state policies; other times, program structure is determined through cooperative agreements between high schools and colleges.

---

[1] For simplicity's sake, we refer to all programs that allow high school students to enroll in college courses as *dual enrollment programs*. For a more nuanced description of the defining features of these programs, as well as other credit-based transition programs, see Bailey and Karp (2003).

Why might states or LEAs invest in dual enrollment programs?

There are many reasons for offering dual enrollment to students, and dual enrollment is a strategy that is used by states and LEAs to meet a variety of educational goals. Initially, dual enrollment programs were targeted toward academically advanced students. The goal was to provide these students, who may have exhausted their high school's course offerings or needed academic challenge beyond what the high school could provide, with the chance to take classes that met their needs. In many places, this is still the primary reason for offering dual enrollment.

Today, however, many educators believe that dual enrollment participation can have additional benefits, with the ultimate aim of improving student access to, success in, and completion of college. Educators argue that these goals can accrue to a wide range of students, not just those at the top of the academic spectrum. They include:

- Increasing the academic rigor of the high school curriculum and encouraging students to take challenging coursework;
- Providing additional opportunities to students who have exhausted their high school's course offerings;
- Helping low-achieving students meet high academic standards and graduate from high school;
- Lowering the need for remediation in postsecondary institutions;
- Increasing student aspirations, particularly for students who do not have college-going role models in their families or communities;
- Helping students acclimate to college life;
- Improving relationships and coordination between secondary and postsecondary education institutions;
- Lowering the cost of college for students and their families; and,
- Improving college graduation rates.

Not all programs are expected to provide all of these benefits, and not all of these potential benefits have been substantiated by research. But this list outlines many of the reasons why states and LEAs may choose to implement a dual enrollment program, and why they may opt to include a wide range of students in that program.

Why data and research are important

Despite dual enrollment's popularity, not much is known about its effectiveness in meeting program goals, particularly those goals addressing college preparedness and completion. Two extensive reviews of the literature (Bailey and Karp, 2003; Lerner and Brand, 2006) found little evidence that dual enrollment contributed to students' college access or academic success. While a new report from the Community College Research Center (Karp, Calcagno, Hughes, Jeong, and Bailey, 2007) provides promising evidence that dual enrollment participation can lead to a range of positive outcomes for students of all backgrounds, it is only a start. At the state and local levels, studies from New York City (Michalowski, 2006; Skadberg, 2005) and Florida (Florida Department of Education, 2004) also show promising results. But given the rapid expansion of dual enrollment, it is surprising that more research, from additional states and districts, is not available.

The dearth of research occurs, in large part, because appropriate data are not collected by programs, districts, or states. Conducting analyses on dual enrollment outcomes, as we will see, requires data that span both the secondary and postsecondary sectors. Studies of program outcomes need to account for students' characteristics and achievement prior to participating in dual enrollment, as well as what happens to them years after they finish the course.

These types of data have been rare until recently. Few districts or states have been able to track students from high school into postsecondary education. As a result, analyses of dual enrollment either focus on high school completion and other short-term outcomes or examine long-term outcomes but cannot control for differences between dual enrollment and non-dual enrollment students at the high school level.

Second, dual enrollment studies often do not use rigorous statistical methods to control for preexisting student characteristics, even when such data are available, often because program staff members lack the capacity to conduct such analyses. For example, many studies compare the college success rates of dual enrollment students to their non-dual enrollment peers, without examining possible reasons for success outside of dual enrollment participation. They do not take into account student achievement in high school or other factors that may influence college success independently of dual enrollment. This may be due to a lack of data, as previously discussed, or a lack of time and/or knowledge at the program or state level to undertake such analyses.

This paper is one attempt to clarify what is needed to measure the effectiveness of dual enrollment. Such studies are important for many reasons. First, states and LEAs are devoting significant resources to the spread of dual enrollment programs. Educators, policymakers, and members of the public need to know whether the expenditure of these resources leads to the intended outcomes. Second, stakeholders assume that dual enrollment can address shortcomings in the current educational system; if students in dual enrollment programs do not attain expected goals, new reforms should be identified and implemented. Third, well-designed evaluations can help states, LEAs, and programs improve dual enrollment to more effectively meet their goals.

The current state of data systems

As noted, studies of dual enrollment require data systems that follow individual students from high school to college. The most effective way to gather, store, and access this type of data is through a state-wide data system that assigns each student his or her own identifying code or number. Such a system enables researchers to examine what happens to participating students no matter where in the state they end up after high school, while accounting for their pre-program characteristics. A statewide system, unlike single programs collecting data on their students, also allows for the creation of comparison groups, meaning that researchers can determine if there are any differences in outcomes between participants and non-participants.

As we have noted, however, this type of data system is quite rare. The Data Quality Campaign (www.dataqualitycampaign.org) has identified 10 elements of a good state data system. In 2007, only four states had data systems that included all 10 elements, and 14 states had 5 or fewer elements. With regard to the elements of most import to dual enrollment research, 21 states plus the District of Columbia have the capability to match student records between the K-12 and postsecondary sectors, and 16

states plus the District of Columbia have data systems that include transcript-level information such as courses completed and course grades. Even within these states, however, data systems may be missing elements essential for dual enrollment research, such as flags indicating dual enrollment courses or information on dual enrollment program features such as course location or instructor type. It is important to note that many states are making great progress in this area, and the number of states building their data capacity is rapidly increasing.

LEAs and states can still conduct research on dual enrollment without state data systems, of course. One way to do this is for programs or institutions to engage in data-sharing agreements, in which LEAs and postsecondary partners agree to share transcript information for their dual enrollment students, and then conduct outcomes analyses. This type of research can be useful, but it also is less likely to yield generalizable results or to include a comparison group. In addition, it is time-consuming and requires significant staff expertise to collect and analyze data at the institutional level. A simpler approach is for individual institutions to use the data they currently collect, such as participation rates or class grades, to examine their program practices. These studies will yield less rigorous results, however.

Outline of the paper

The remainder of the paper is organized as follows. In the next section, we discuss goal setting and the importance of conducting evaluations that are appropriate for the stated goals of a program. We examine the ways that different goals lead to different research questions, data sources, and analytic methods.

We then turn our attention to different types of analysis and discuss the data and methods used in each. First, we look at analyses focusing on program description, followed by analyses examining program quality, and finally, analyses examining program outcomes. This section serves as a "toolbox," providing a framework for developing your own data collection efforts and research studies. We follow this with a discussion of strategies for setting up a data system and conclude with a section identifying ways that researchers and policymakers can put research findings to use.

II. Setting goals to guide dual enrollment research and practice

In order to evaluate a program, policymakers and administrators need to be clear on what the program is meant to accomplish. Understanding the goals of a program helps clarify the research questions, data, and methods, and ensures that research activities are meeting their targets.[2] For a program like dual enrollment, which has a number of potential outcomes, clarifying program goals is particularly important. This section discusses the ways that research can be linked to program goals and provides guidance in determining what those goals are.

The importance of clarifying program goals

Clarifying program and policy goals is an essential task, especially for something as multifaceted as dual enrollment. Goals influence what stakeholders can expect a program to accomplish and the outcomes that should be sought in an evaluation. For

---

[2] Clarifying program goals is also essential for the program implementation process; dual enrollment programs and policies are likely to look different depending on their aims.

example, it is not reasonable to expect a program intended to provide enrichment for advanced students to lead to increased college preparedness among the broader high school student population. Nor is it reasonable to expect a program in which students pay their own tuition to decrease the cost of a college education. If policymakers and program staff are not clear about which goals the program is meant to address, research may assess outcomes that are not the intent—and thus show no impact.

Consequently, understanding the goals of a state's or LEA's dual enrollment program is necessary prior to embarking on a research study of that program. The goals guide the questions to ask in an investigation, as illustrated in the text box below. Goals also guide the types of data that must be collected in order to conduct the analysis. For example, the data collected to examine the costs of a program (budgetary expenditures) are going to differ from data collected to examine student outcomes (transcript data).

Though this seems obvious, it is often not so in the real world. As we have pointed out, dual enrollment programs are increasingly seen as meeting an array of goals for a wide range of students. In some ways, there has been significant "mission creep" for dual enrollment—a program that used to be narrowly targeted toward one type of student for one purpose is now often seen as leading to multiple outcomes for many types of students. For example, in a 2002 issues briefing document, the American Association of State Colleges and Universities commented that dual enrollment programs can "serve a number of *different* purposes including providing more academically challenging course opportunities, expanding college access, and broadening high school curricular offerings" (p. 8).

This poses a serious challenge to policymakers and program evaluators. For policymakers, it means that regulations and budgeting processes must balance the competing demands of various program goals. It may also necessitate narrowing a program's scope, potentially alienating some stakeholders. For evaluators, mission creep may mean that it is not clear what, exactly, you are evaluating. Which outcomes can you realistically expect a given program to lead to? How do you know what to investigate if they do not know what a program is supposed to accomplish?

Thus, the first step for any state or LEA seeking to study their dual enrollment programs is to determine what they want the program to accomplish. We would hope that this is clear-cut and that program structures have been implemented that lead to the stated goals. But in the absence of such clarity (for example, in a state that is shifting from a restrictive to broad set of eligibility requirements), researchers need to determine the intended program goals before embarking on an evaluation project.

Program Goals Influence Research Questions and Data Required

Program Goal:
Provide enrichment to academically advanced students
Sample Research Questions:
Which types of courses are offered through dual enrollment?
Do dual enrollment students take more rigorous high school courses than non-dual enrollment students?
What are the academic backgrounds of dual enrollment students?
Possible Data to Collect:

Dual enrollment courses offered by institution
Dual enrollment courses taken by students
Number of dually-enrolled and non-dually-enrolled students completing a New Basics or other college-prep curriculum
Entering grade point average of dual enrollment students

Program Goal:
Open access to challenging courses to disadvantaged students
Sample Research Questions:
Who takes dual enrollment courses?
Which high schools offer dual enrollment courses?
Do student course-taking patterns change in low-income schools once dual enrollment is implemented?
Possible Data to Collect:
Demographics (i.e., race, socioeconomic status, parental education) of dual enrollment participants
Dual enrollment offerings by high school, along with demographic information about each school
Within a subset of low-income or otherwise disadvantaged high schools, information on student completion of college-prep or college-level courses, collected before dual enrollment is offered and after

Program Goal:
Increase college success
Sample Research Questions:
What are the college-going rates of dual enrollment participants, as compared to those who do not participate?
Do college-going and college graduation rates increase after a state implements a dual enrollment program?
What are the college outcomes (i.e., grades, persistence) of dual enrollment participants, as compared to those who do not participate?
Possible Data to Collect:
College enrollment for students who did and did not take dual enrollment courses
College-going rates for graduation cohorts before and after implementation of a state dual enrollment program
College transcript data, with indicators as to whether or not students took a dual enrollment course; ideally, this would also include high school transcript data in order to control for student pre-existing characteristics

Determining dual enrollment program goals
　　　　Ideally, goals will be clearly spelled out in state dual enrollment policies or individual program documents. When this isn't the case, researchers have a number of sources to help identify expected outcomes. The first is existing state policy and

regulations governing dual enrollment. (In locations without such policies, district and institutional regulations can serve the same purpose.)

Many state policies are explicit regarding program goals. Missouri's regulations, for example, state that dual credit is intended to "provide high-performing high school students an affordable opportunity to experience high-quality college-level courses." In other cases, policies and regulations implicitly state program goals. For example, states that have regulations limiting participation to high-achieving students are unlikely to have the goal of reducing remediation through dual enrollment. States that explicitly fund dual enrollment support services or developmental pathways, however, are likely to have the goal of improving broad access to and success in college through dual enrollment participation.

Researchers should therefore engage in a policy analysis prior to embarking on dual enrollment projects. Systemically examining dual enrollment regulations will clarify the types of outcomes that stakeholders likely hope to see as a result of the program. Your analyses should include legislative language as well as regulatory codes; make sure to explore policies based in Departments of Secondary Education or State K-12 Departments and those based in the higher education sector.

A second arena through which to determine state or LEA goals for dual enrollment is through analysis of public discourse. Documents such as newspaper articles on dual enrollment, policymakers' testimony, or program recruitment materials provide insight into the goals that stakeholders have for the program. For example, if news stories about dual enrollment focus on the way that the program can save students and families money, then researchers can assume that such savings are an intended outcome.

Analyzing such discourse can be more complicated than analyzing policy documents. Documents are harder to find, for example, and they may be less clear in their perspective. However, they provide an important window into what the public expects dual enrollment programs to accomplish. And since policymakers are beholden to the public, these documents likely reflect the goals that policymakers (particularly the legislative members who fund the program) have, as well.

At times, you may find that policy language and public documents collide. This is particularly likely as dual enrollment programs continue to expand their reach. For example, a state whose policies support dual enrollment as an enrichment program may have stakeholders who increasingly see it as a means to broaden access to college for underrepresented groups. In these instances, there are multiple goals to examine. However, these are situations in which good research can provide a particularly useful service. Researchers can use both sets of goals to guide their research questions; most likely, the findings will indicate that not all of the goals are being met and will force stakeholders to reconsider the purpose of dual enrollment.

Take the hypothetical state discussed above, in which policies support dual enrollment as enrichment but the public sees it as a means for opening college access. Researchers could decide to investigate participation patterns throughout the state, in order to determine if dual enrollment programs enroll mainly high achievers (the stated policy goal) or a range of students including those typically disadvantaged in higher education (the emergent goal). The findings could then be used to encourage policy changes or a re-explanation of the program to the public, depending on the outcomes and policymakers' goals.

In sum, research should both reflect and test the state and local context. Prior to conducting an investigation of dual enrollment outcomes, researchers should be clear as to what the program is intended to accomplish. Research questions and data collection should examine the impacts intended for a particular program, rather than other possible outcomes that may not be valued by your state or LEA. In some cases, this means that you will spend time clarifying program goals prior to launching actual research activities.

III. Program Analysis—Ways to answer different types of questions

Not all program research uses the same data and methods. As we have seen, depending on the goals of a program, the questions research seeks to answer will vary. The necessary data and analytic techniques will therefore vary as well. Research can be seen as falling into one of three categories, depending on the types of questions it poses. This section describes these categories. For each, we describe the types of questions that can be answered, which dual enrollment goals can be evaluated, the types of data necessary to conduct the research, and the type of analytic strategies that are most useful. The questions and strategies we present are not, of course, exhaustive. Rather, this section can serve as a toolbox, providing a framework and guidance for developing your own data collection efforts and research studies.

A. Descriptive research: How big is our program, and who is in it?

The simplest type of research is that which describes a program and/or its participants. This type of research, not surprisingly called "descriptive research," provides basic information on who takes dual enrollment courses, which courses they take, where the courses are offered, and what grades the students earn. It provides stakeholders with a general idea of program participation and implementation throughout a state or district.

This type of research is most appropriate for answering "who" and "what" questions. For example:

- Who participates in dual enrollment, in terms of race/ethnicity or socioeconomic background?
- What are the participation rates in dual enrollment for a state or LEA?
- What types of high schools offer dual enrollment?
- What types of courses are offered through dual enrollment programs?
- What types of grades do students earn in dual enrollment courses?

Descriptive research, when conducted on a regular basis, can also provide information on changes in the program over time. It can show growth in participation rates, for example, or changes in the type of student participating. If a state or LEA decides to implement a policy modification, descriptive research conducted before and after this change can give stakeholders a sense as to whether or not the program changed as a result.

One very important caveat for this type of research is that it cannot explain what causes the outcomes we see. Because descriptive research only describes a program (i.e., providing the frequency with which a type of student participates or the number of course sections offered), it cannot account for the many possible reasons for the findings. It can illustrate the state of a program, but it cannot tell us about causes.

For example, if descriptive research conducted over a number of years shows that the number of students of color participating in dual enrollment has changed, we cannot know why. It might be due to a policy modification, but it may also be due to a factor outside of the program's purview. Likewise, if dual enrollment students have higher grade point averages than other students, we cannot know if that is due to their participation in dual enrollment or some other factor. Thus, it is important to interpret the results of this type of research carefully, and not use them to draw causal inferences.

Still, descriptive research is useful for answering many of the questions facing policymakers. In many locations, such basic information is not well known. Understanding who participates in dual enrollment, if access to dual enrollment is more prevalent among certain types of students, and what types of courses are offered through the program goes a long way toward understanding how dual enrollment plays out in a state or district. It can also be used to help stakeholders understand why the program might be valuable, why it needs to be modified, or how it has grown.

The data needed for these types of analyses are those commonly collected by educational institutions. At the individual level, these include demographic information such as race/ethnicity, free/reduced lunch eligibility, and gender; achievement data such as grade point average and standardized test scores; and dual enrollment course grades. The key for using these data in program analyses is to make sure that institutions include an indicator that identifies which students participate in dual enrollment. Ideally, programs (particularly high school partners) should notate which courses on a students' transcript were dual enrollment courses, but a flag identifying dual enrollment participants is sufficient for these types of analyses in most cases.

At the institutional level, descriptive analyses rely on information that high schools and colleges already collect, but which they may not consider "data." This includes courses offered for dual enrollment, dual enrollment class sizes, and aggregate course grades in dual enrollment classes. Here, it is important that institutions organize this information and retain it in a way that allows for analysis, such as arranging dual enrollment course offerings in a spreadsheet or database.

As noted, all of the data necessary for descriptive studies are collected by institutions. Often, this is done as part of a state or district data system. For example, even in states that do not have fully-developed systems, LEAs are often required to provide an array of data to their state governing body. Similarly, colleges often collect such information as part of accountability or accrediting requirements. In these situations, the addition of a flag identifying dual enrollment students and/or courses is sufficient to provide researchers with the data they need to conduct descriptive analyses.

Where such data systems do not exist, LEAs or states can still conduct descriptive studies of dual enrollment. Most of the data required are already collected by programs, as part of course registration forms (i.e., demographic information) or program guides (i.e., course offerings). Researchers will need to take the time to convert this raw information into a usable database, but the data do exist. For example, if an LEA wishes to understand who takes dual enrollment courses, they may choose to analyze the demographics of dually-enrolled students. They would need to collect enrollment forms from all participating high schools and enter student demographics into a program like Excel or SPSS in order to conduct the analysis. They would not need to collect anything

else, presuming that demographic information is already collected as part of the registration process.

Analyses of descriptive data are relatively straightforward. They include calculating percentages of a whole (i.e., percent of participants who are eligible for free/reduced lunch; percent of students receiving a B or higher in their dual enrollment course); tabulating total rates of participation (i.e., tabulating the number of courses offered by each high school in an LEA); or creating cross-tabs for variables of interest. Again, when conducting these analyses, it is important to remember that while they can describe what is happening at the program, district, or state level, they do not provide information on why these things are occurring.

Sample Study 1: *What are statewide participation rates and do they vary by race or socioeconomic status?*
Answering this question provides a snapshot of student participation in dual enrollment programs across the state. The data elements required include student dual enrollment participation flags and demographic information (e.g. gender, race/ethnicity, socioeconomic status).

One way to conduct this analysis is to run a cross-tab of program participation and demographic variables, which will show the participation rates by each variable. You can use a program like SPSS to tabulate the number of students in dual enrollment by their various characteristics. A cross-tab analyzing participation by gender might look like this:

		Dual enrollment		Total
		Non-participants	DE students	
Gender	Males	800	160	960
	Females	840	200	1040
Total		1640	360	2000

Based on the table, we can calculate that, statewide, the participation rate is 18 percent. We can also see that there is a slight difference by gender: the participation rate is 16.7 percent for males, and 19.2 percent for females.

Another way to conduct a descriptive analysis is to look at proportions of student sub-populations among dual enrollment students. A good example of this analysis is found in the study of dual enrollment in Florida and New York City conducted by Karp and her colleagues (2007). Tables 1 and 17 of their report describe and compare the demographic characteristics of dual enrollment students and non-dual enrollment students. We reproduce a part of Table 1 below. The first row, for example, shows that almost 51 percent of students statewide are female. However, the proportion of dual enrollment participants who are female is 62 percent, while the proportion of non-participants who are female is 49 percent.

Student characteristics	All students	Dual Enrollment	
		Non-participants	Participants
Female	50.74%	49.19%	62.03%
White & non-Hispanic	54.93%	52.00%	76.28%
Black & non-Hispanic	23.96%	25.75%	10.95%
Hispanic	17.76%	19.10%	8.05%

American Indian	0.24%	0.23%	0.29%

Sample Study 2: *What are the effects of state- or LEA-sponsored incentive programs on student enrollment in dual enrollment?*

When a government agency implements a new program aimed at encouraging student participation in dual enrollment, stakeholders want to know if participation rates increase. One way to address the question is to look at the enrollment rates before and after the new initiative (sometimes called a pre-/post- design); if participation rates go up immediately after the new incentives, it is likely that the initiative worked. However, this approach does not account for other factors that might have increased participation. It gives us a clue as to the efficacy of the change but cannot demonstrate that the initiative worked.

Sometimes, a "natural experiment" occurs in a state—a new program or set of incentives is only made available to some LEAs. In this case, we can use the fact that some LEAs received the incentive while others did not to test whether or not the incentive actually leads to increased participation. This is called a differences-in-differences approach. Essentially, we compare participation rates before and after the initiative in LEAs that are similar except that some are subject to the new policy and some are not. If participation rates increase in the LEAs receiving the incentive (the treatment group) but not in the LEAs not receiving the incentive (the control group), there is evidence that the incentive has met its goals. This approach is most effective when the control and treatment LEAs are similar on most characteristics (student population, etc.) and/or were randomly chosen to participate in the incentive program. When such a natural experiment occurs, we can be more confident that any outcomes are caused by the policy change, though we still cannot be certain that other influences were not also in play.

The required data elements for this type of analysis include longitudinal dual enrollment participation rates and LEA intervention flags. Suppose that state *s* has two groups of LEAs: (a) those that implement the incentives (treatment) and (b) those that do not (comparison). The first thing to do is to run a simple cross-tab of student DE participation rates, using a statistical software program such as SPSS or Stata. The output of this cross-tab might look like this:

Group	Before implementation (I)	After implementation (II)	Difference = (I)–(II)
Treatment	20.3%	23.3%	3.0%
Comparison	19.5%	20.5%	1.0%

The last column presents the changes in student enrollment before and after the incentives for two groups. A 2 percent difference in the changes, (3%–1%) can be interpreted as the impact of the incentives. We can use SPSS or another statistical program to test whether this difference is statistically significant or not (for example, by conducting a *t* test). If it is, then we can assume that the incentive had an influence on participation rates. Remember, though, that this type of analysis does not completely control for other things that might also be simultaneously influence participation rates, so we need to interpret the results with some caution.

B. Implementation research: How do we ensure program quality?

Another category of research, implementation research, examines the ways that a program is structured or put into practice. Many different types of questions can be addressed by this category of research, but they generally have the goal of ascertaining program quality or seeking program improvement. Thus, rather than seeking to justify a program to stakeholders, this type of research can be used to assure funders that the program is being implemented as intended. Dual enrollment is not always embraced by all stakeholders, so in locales in which dual enrollment has been accused of being inferior to other credit-based transition programs, implementation research can also be useful, as it can provide evidence as to the program's rigor. In addition, this type of research is especially useful to those implementing dual enrollment programs in institutions, as it can help them modify their activities to deliver services more effectively and efficiently.

Implementation research is most appropriate for examining "how," "what," and "where" questions. For example:

- How much does it cost an LEA to offer dual enrollment?
- Where do students take their dual enrollment courses?
- What are the qualifications of instructors teaching dual enrollment courses? Does this vary by course location?
- How are dual enrollment programs structured in various school districts around the state?
- What do the curricula and pedagogical methods in dual enrollment courses look like? If offered at a high school, how do they compare to those in courses offered to regularly matriculated college students?
- What are the unanticipated barriers to implementing dual enrollment programs?

As implied by the questions above, implementation research may focus on specific programs, or it may be used to compare different programs in a state or LEA. For example, gathering information on implementation in each LEA in the state allows for an understanding of patterns (i.e., poor districts use one type of teacher while wealthy districts use another) that might be relevant to policy and funding decisions. Or, in another instance, comparing syllabi, assignments, and exams of dual enrollment courses offered on high school campuses to the same course materials used on college campuses can help determine whether the rigor of dual enrollment is appropriate. Comparing the pass rates of students in dual enrollment and "regular" freshman courses might serve a similar purpose.

This type of research is also useful if a state or LEA would like to move from a dual enrollment program that provides wide institutional autonomy (and therefore encourages a range of experiences) to a program that is more uniform. Understanding the features of current programs is a first step in figuring out which institutions may need to modify their offerings. And tracking institutional features over time can help states ensure that desired program changes are underway.

Given the array of possible questions that implementation research can investigate, the data necessary for these activities is highly variable. As with selecting the appropriate research questions for a program's goal, it is important that the data used to answer implementation questions are determined carefully. Unlike descriptive and outcomes research, which rely mainly on quantitative data, implementation research can

use quantitative, qualitative, and survey data. We primarily focus on research using quantitative data here, since states are currently building their capacity to gather and use such data. However, it is important to note that surveys and qualitative methods such as interviews, document reviews, and observations can glean important implementation data as well, and states and LEAs should consider engaging in such activities if quality assurance is a goal of their research activities.

The data needed to conduct implementation analyses can include variables commonly collected by institutions. This includes student demographics, course offerings, and to some extent, the cost to states, institutions, and students and their families (for example, state ADA or FTE funding, expenditures on teachers, and tuition rates). However, common data collection activities will likely need to be augmented with more specific information. Additional data necessary for implementation analyses include information on course location (high school or college campus); teacher credentials (college professor or high school teacher certified as adjunct; the credentials used to deem high school teachers eligible for adjunct status); course content; and pass rates. Additional funding variables (who pays for tuition and textbooks; type of ADA/FTE reimbursement; additional program costs such as transportation) may also need to be collected in order to conduct analyses focusing on costs and return on investment.

Most of these can be generated by adding variables to existing datasets, for example, adding a flag indicating the location of each course section offered through dual enrollment. This does require additional communication between programs and data coordinators, however, as those collecting data need to be given information on program structure at a level of specificity that they might not otherwise think of. This also will require datasets that have the capacity to include information about specific students *and* specific courses. For example, a college may offer dual enrollment English 101 on high school campuses, but also permit students to take English 101 on the college campus as dually-enrolled students. They, or their partner high schools, will need to develop a system that not only flags students taking English 101 as dual enrollment participants, but also identifies whether they were high school- or college-based dual enrollment participants. And, in the case of the high school course, the dataset will also need to include a third field indicating what type of instructor taught the course.

Exhibit A shows a sample data collection table, including coded variables and a sample data dictionary. It illustrates that the data typically collected by institutions (course enrollments, grades) must be augmented with additional variables. This dataset includes only information on high school students in dual enrollment courses. If researchers want to conduct comparative analyses (say, to compare course grades of dually-enrolled and regularly-matriculated students in English 101), the dataset would need to be augmented with student records from a postsecondary institution. Likewise, if researchers want to conduct analyses that calculate the return on investment from dual enrollment, they will need data that address the costs of dual enrollment (tuition paid by the state, FTE/ADA paid by the state, other costs) *as well as* similar data for other programs and/or data on the costs that would be incurred if dual enrollment was not an option (lost wages due to longer postsecondary careers, for example).

Exhibit A

Dataset

Student	Fa07course	grade	duenrol	loc	instr
F1x	Math4	B+	0		
F1x	English 101	A	1	1	0
F1x	History 3	A	0		
F1x	Biology	C	0		
M1y	Chemistry	A-	0		
M1y	History 4	B+	0		
M1y	English 101	C-	1	0	2
F2x	English 101	A	1	1	0
F2x	Math 3	A	0		
F2x	Biology	B-	0		
F2x	History 3	B	0		

Data Dictionary

Student   Individual student identifier (identifying information removed; note that each student has multiple records since she/he takes multiple courses)

Fa07course   Name of courses taken in Fall 2007

Grade   Grade earned in course

Duenrol   Dual enrollment course flag
           0 = no (high school course)
           1 = yes

Loc   Dual enrollment course location
           1 = college campus
           0 = high school campus

Instr   Dual enrollment instructor type
           0 = regular college professor
           1 = high school teacher certified as adjunct (master's in subject)
           2 = high school teacher certified as adjunct (departmental approval)

In order to collect such data, LEAs and state policymakers with an interest in dual enrollment will have to work with the officials overseeing the state's data system to ensure that the necessary variables are included in data collection efforts. Otherwise, they will have to work with individual districts and schools to collect the data as an addendum to other data efforts—a difficult (though not impossible) task. In states without data systems, it is up to individual institutions to ensure that these types of data are collected in a systematic manner and on a regular basis.

Implementation analyses, like descriptive analyses, can be quite simple. They may consist of tabulating the frequency with which a program characteristic occurs within a state or LEA or conducting cross-tabs examining program characteristics by high school type. You might also compare pass rates in order to determine if they vary by location, sponsoring institution, or whether they are dual enrollment or "regular" sections.

With these types of data, however, researchers also have the capability to conduct more complicated analyses. For example, you could examine the return on investment, which requires calculating the obvious and hidden costs of dual enrollment, the benefits of the program, and the rate of return; doing the same for other programs; and comparing

the two. You can also conduct qualitative analyses; for example, researchers may code course syllabi and then compare the rigor of courses across institutional types.

Keep in mind that, as with the descriptive analyses described earlier, these analyses do not reveal causation. For example, they may show that students in dual enrollment courses taught on a college campus earn higher grades than those offered on high school campuses. This does not necessarily mean that college-based courses are of higher quality; there may be differences in student characteristics that cause this difference. But such a finding would be an important clue that further investigation is warranted, either by conducting outcomes analyses (discussed in the next section) that control for course location, or by conducting further qualitative implementation analyses.

Sample Study 3: *How do high school-based dual enrollment courses compare to college freshmen courses with regard to instruction and student success?*
You can think of courses as including three types of components: Inputs (faculty, students), Process (curriculum, instruction), and Outputs (student learning, dropout rates, grades). If course sections are similar in all three areas, they are generally comparable. Thus, in order to determine if dual enrollment courses offered at a high school are similar to freshmen courses on a college campus, you would describe each in terms of their components. For each component, you can use quantitative and/or qualitative methods to compare dual enrollment and regular courses.

To conduct these analyses, the first step is to match a dual enrollment course (or group of dual enrollment courses) to a regular freshmen course that should be comparable. For example, an LEA could compare all calculus dual enrollment courses it offers with the calculus courses offered at a nearby community college. Data would then be collected from all course sections offered through dual enrollment and the college. The data could be arrayed in a table like this one:

Component	DE courses (I)	Regular courses (II)	Difference (I)–(II)
Inputs			
% of faculty w/ advanced degrees	55.0%	65.0%	–10.0%
Average faculty years of teaching experience	5	6	–1
Class size (average number of students)	45	40	5
Instructional expenditures per pupil	$1,000	$1,100	–$100
Outputs			
Pass rate	60.0%	65.5%	–5.5%
Average grade	2.91	3.11	–0.20

The differences column gives us a sense of whether there are areas in which courses are dissimilar. We can then use this knowledge to investigate whether the differences influence student achievement or course quality through other methods such as outcomes research or further qualitative exploration.

Keep in mind that some components are easier to quantify than others. In this example, the quality of faculty is somewhat challenging to define; we have used the percentage of faculty with advanced degrees and their years of teaching experience as a way to do so. This measure may not perfectly represent faculty's educational attainment, but at least it

---

can serve as a proxy.

Finally, in order to compare the process component, it may be necessary to audit the courses. The College Board is now reviewing the courses offered through its Advanced Placement program; this process can serve as a model for a state- or LEA-wide dual enrollment audit.

Sample Study 4: *How do dual enrollment programs vary across the state? Do different types of students participate in different types of programs?*

Because dual enrollment programs vary in their structure and because these variations may influence equity or effectiveness, states and LEAs may want to understand how programs differ throughout a state.

The data elements required for the analysis might include program location, faculty, course content, number of courses offered, college credit awarded, tuition/fees, eligibility requirements and target populations, and other program characteristics. Once obtaining the data, you can summarize them in a variety of ways. One is to use a chart similar to that in exhibit A, in which program types are coded to allow for quantitative analysis. Another approach is to use the following chart, which provides more detail:

Program	Local agency agreement	Where	Who	Whom (target populations)	…
School District #1	Yes	High school	HS teacher	Advanced students	
School District #2	…				
…	…				
College #1	No	College campus	College faculty	Technical students	
College #2					
…					

You can analyze this chart in a variety of ways, depending upon the questions you are asking. For example, you could compare the student participation or outcomes of programs with different features to see if one type appears more equitable or is related to more positive outcomes (remember, though, that in these analyses we can't be sure of causation; we can just get a sense of potential relationships). You could also run a cross-tab of program characteristics and student demographics, as in descriptive studies, in order to compare student characteristics in different types of programs.

A good example of this type of data analysis is found in Karp and her colleagues' report (2005) summarizing state dual enrollment policies.

C. Outcomes research: Does our program lead to student success?

At the end of the day, stakeholders are most interested in whether dual enrollment contributes to student success. They want to know that the program meets its intended outcomes. If it does, they can use this information to encourage the spread of the program or advocate for additional funding. If it does not, then they must figure out why— implementation research can help in this regard, as can re-thinking the program goals. In

either case, it is better to have some evidence than to make guesses or assume it is working as intended.

Outcomes research is the most methodologically complicated research discussed in this toolkit and requires the most resources. The data required are more complex and harder to collect and the analytic processes demand a higher level of statistical prowess than the other forms of research we have discussed. These types of studies also cost more and take longer to conduct than the other types. However, they are the only ones that can provide evidence of a causal relationship between dual enrollment participation and outcomes; thus, states and LEAs should seriously consider devoting the resources necessary to undertake this work.

The first step in conducting outcomes analyses is to determine which outcomes the research will investigate. This circles back to our earlier discussion of program goals. If a given outcome is not a stated goal of a state's dual enrollment program, it is not reasonable to make it the focus of research activities. The outcomes should stem from the goals.

Generally, outcomes analyses will focus on student achievement resulting from participation in dual enrollment. These can be short-term outcomes, such as completion of the dual enrollment course, high school grades, graduation rates, or college application rates. They can also be long-term outcomes, such as college entry rates, college persistence, college grades, or college completion. In some cases, they may also address the cost of college (i.e. differences in tuition costs or loan debt) or changes in education costs to a state.

Outcomes analyses seek to understand whether students who take dual enrollment courses do better on the relevant measures of success than students who do not, *over and above* any differences that might have existed between these two groups of students in the absence of a dual enrollment program. For example, in locales in which dually-enrolled students must meet eligibility criteria, it is likely that participants and non-participants would have different college-going rates even without dual enrollment. Just by virtue of the fact that they have higher GPAs, students eligible for dual enrollment probably would go to college in greater numbers than their peers who were not eligible for the program. The question is, does dual enrollment increase participating students' likelihood of attending college even more?

This means that the second step in conducting outcomes analyses is to figure out what other variables might influence outcomes aside from dual enrollment participation. These variables will be the controls in the analysis. They might include students' previous academic performance or their socioeconomic background. They might also include school-related factors, since some schools might encourage positive outcomes, regardless of whether they offer dual enrollment. It is essential that this information is collected, along with outcomes information, in order to ensure that the analyses can actually indicate causal relationships.

Another approach to controlling for pre-existing characteristics is to construct comparison groups of similar students. These students should be as close as possible on all characteristics (grades, race/ethnicity, family background, previous academic performance, etc.) as possible, with the exception of their participation in dual enrollment. If students participating in dual enrollment (the treatment group) have more positive outcomes than those not participating (the comparison), we might be able to say

that dual enrollment is related to those outcomes. There is a *very important* caveat here, however. The two groups of students, no matter how carefully constructed, may differ on important characteristics that are not measurable. For example, students who take dual enrollment classes may be more academically motivated than students who do not, even if both types of students have the same high school grades. In such a case, it would not be surprising if the treatment group went to college in greater numbers; this positive outcome would be the result of their motivation however, not the result of the dual enrollment program. As a consequence, we *strongly* recommend conducting analyses that statistically control for as many variables as possible—as described in the previous paragraph and in the two text boxes later in this section—rather than merely comparing two groups of students.

The one exception to this caveat is when similar students can be randomly assigned to treatment and control conditions. This type of research, called a randomized experiment, is actually the most rigorous outcomes research that one can conduct. In this type of research, students who are similar on an array of measured characteristics register for dual enrollment. Usually due to a lack of space, half of these students are randomly assigned to participate (the treatment) and the other half are not (the control). Because all of the students wanted to participate in the program, we can assume that they are similar on unmeasured characteristics such as motivation as well as their measured characteristics and that the only difference between the two groups is whether or not they took a dual enrollment class. Any differences in outcomes under this condition can be seen as caused by participation in dual enrollment.

Outcomes research using random assignment is ideal. It is also, however, costly, time-consuming, and difficult to pull off in practice. There may also be important reasons to allow all students wishing to participate in dual enrollment to do so. Thus, most states and LEAs will not be able to conduct random assignment. We therefore, again, suggest that those engaging in outcomes research rely on well-designed studies that control for as many possible influences on student outcomes as possible.

To conduct outcomes analyses, states and LEAs need to collect a wide range of data, all of which must be linked to individual students and cover multiple years of schooling. Data for these analyses are, necessarily, the most complicated to collect, organize, and use, as they must cross multiple institutional sectors, include multiple records per student, and contain a wealth of information on pre- and post-dual enrollment student status.

Ideally, researchers will have access to data on all students in a given cohort, whether or not they participated in dual enrollment. This will enable you to create the most robust comparison group because it will include all students in the state. A second-best option is to collect data on a subset of students including those who participated in dual enrollment and a comparable group of students who did not. In this case, the number of students in both groups (treatment and comparison) should be large enough to draw statistically significant conclusions.

In order to control for students' characteristics before entering a dual enrollment program, datasets used in outcomes analyses should include variables for students' grades and standardized test scores prior to the year they become eligible for dual enrollment, as well as indicators for race/ethnicity and socioeconomic status. Ideally, this will include

high school transcript data. To create comparison groups, datasets should include indicators for dual enrollment participants and flag each dual enrollment course taken.

Short-term outcomes data include grades received in all dual enrollment courses, as well as cumulative high school grade point averages, scores on standardized tests taken after dual enrollment eligibility, and a flag for high school graduation. In addition, students' enrollment status in postsecondary education should be indicated. Long-term outcomes data include grades in college courses, credit accumulation for at least four semesters (ideally, for eight semesters or more), and degree completion flags. Other possible outcomes variables to examine include enrollment status (full- or part-time), eligibility for state or federal scholarship programs, and financial aid status. States and LEAs interested in examining these long-term outcomes need to set up mechanisms for collecting appropriate data well in advance of students' enrollment in postsecondary education, so that they are not retroactively trying to recreate the data. Ideally, all long-term outcomes data will be generated through students' complete postsecondary transcripts records.

Sample Study 5: *What are the short-term outcomes of dual enrollment students as compared to their non-participating peers?*

As we have noted, including control variables in outcomes analyses can increase confidence that positive impacts are the result of program participation, rather than pre-existing differences between dual enrollment and non-dual enrollment students. One way to do this is to use a regression approach that allows you to adjust for observable differences between dual enrollment students and non-participants. This means that you include variables stemming from student experiences prior to dual enrollment participation in your analyses. This statistically eliminates the influence of these experiences on students' subsequent outcomes. For example, if dual enrollment students typically have higher pre-dual enrollment grade point averages than non-dual enrollment students, including pre-dual enrollment grades in a regression analysis minimizes the chance that any difference in outcomes is a result of the pre-existing differences. As a result, any difference in outcomes between the two types of students is likely attributable to program participation, not differences between students that were there to begin with.

In this case, we will examine whether dual enrollment participation improves students' high school graduation rates (this example is taken from Karp et al., 2007). The data elements required are the outcome variable of interest (whether or not students graduated) and variables to control for student's socioeconomic background (race/ethnicity, gender, eligibility for free/reduced lunch) and prior academic achievement (ninth or tenth grade GPA, standardized test scores). We might also choose to include high school characteristics to make sure that any differences aren't due to the fact students in dual enrollment might also attend high schools that generally produce better outcomes than students who do not participate.

To conduct this type of analysis, you need to use a program like SPSS or Stata to create statistical models using these variables. You will want to run the models a few different times, each time adding additional control variables. First, see if there are differences in outcomes for students who do and do not participate in dual enrollment (see column 1 in

---

the table below; this is Table 5 in the Karp report). If there are significant differences, continue to add control variables. Models 2 and 3 add in student demographics and achievement, followed by high school characteristics. You can see that in all cases, there is a significant difference in outcomes between participants and non-participants (row 1), though that difference gets smaller as more variables are added.

| Variables	Model 1 dy/dx	Model 2 dy/dx	Model 3 dy/dx
DE participation	0.118*	0.054*	0.043*
Student demographics	Yes	Yes	Yes
High school GPA	No	Yes	Yes
High school characteristics	No	No	Yes

Note: * Statistically significant at the level of 1 percent.

We do need to remember that although this analysis is fairly rigorous and tries to control for preexisting differences between groups of students, there may be differences that are not accounted for, such as motivation. Therefore, we can't say for certain that dual enrollment *causes* the higher graduation rates, although there is definitely a relationship between participation and graduation.

Sample Study 6: *What are the postsecondary outcomes of dual enrollment students as compared to non-dual enrollment students?*
A more rigorous approach that still does not require conducting an experiment is to use a method called regression discontinuity. Essentially, this approach uses the fact that, in most places, there is a cutoff score to determine program eligibility. Students who are right above and right below the score are really quite similar on most characteristics, but only some of them can take dual enrollment classes. If, after participating in the program, students right above the cutoff score have more positive outcomes than students right below the cutoff, we can say that the program caused these outcomes.

The following graphs use hypothetical data to describe this approach and show visually the impact of dual enrollment on the postsecondary outcome, college GPA. The left-hand graph highlights a discontinuity at a high school grade point average of 3.0, above which students are allowed to enroll in dual enrollment programs. The right-hand side graph demonstrates another discontinuity on the postsecondary outcome, college grade point average. The difference in the outcome between those right above and below a high school GPA of 3.0 represents the impact of dual enrollment.
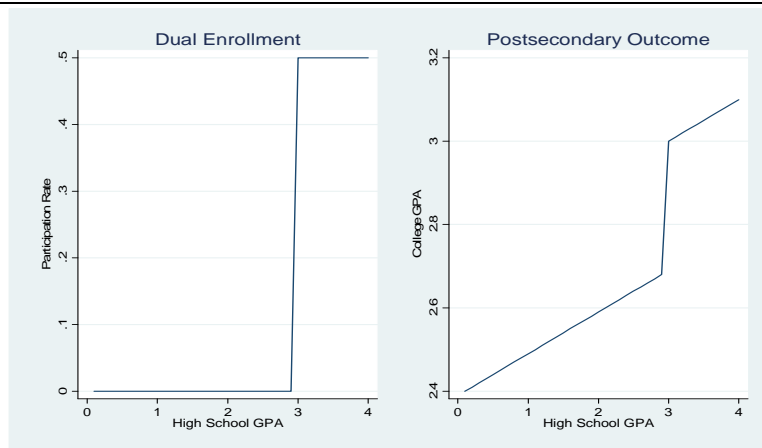
The data elements required for this type of analysis include 1) high school GPA or standardized test scores used to determine program eligibility, 2) program cutoff scores, and 3) student outcomes such as postsecondary enrollment, persistence, college GPA, and degree attainment.

To conduct a regression discontinuity analysis, you first need to make sure that the program environment is appropriate for this technique. One way to do this is to compare the demographic characteristics of students right below and above the cutoff point. For the analysis to be valid, there should be a big jump, or discontinuity, in participation around a cutoff point, and there should be few differences between the groups of students on either side of the jump. You can use a statistical software program such as SPSS or Stata to conduct a *t* test of a mean difference to see whether or not the groups are the same on demographic characteristics. If they are not the same, then narrow the range of scores surrounding the cutoff until you have similar groups.

The next step is to estimate a discontinuity on the postsecondary outcome at the cutoff as presented in the figure above. In practice, you can use SPSS or Stata to create a statistical model where college GPA is regressed on high school GPA or test scores plus a dummy indicating whether the student is above the cutoff. We enter as a control variable high school GPA to eliminate its contribution to the outcome difference between two groups. This step is similar to the regression analysis we conducted in the previous example. The remainder of the outcome difference is interpreted as the impact of dual enrollment. Then we see if there is any significant difference between students right below and above the cutoff.

A good example of this type of analysis is found in Calcagno's dissertation that examines the effectiveness of developmental education (2007).

IV. Setting up a data system

As we noted in the introduction of this report and have emphasized repeatedly throughout, the analyses described here can only be conducted if the appropriate data are

collected by states and LEAs. Moreover, because some analyses, particularly those focused on outcomes, require a relatively high degree of statistical sophistication, states and LEAs need to ensure that staff capacity to conduct dual enrollment research is developed. Most critically, as we have said, data systems need to be developed that can follow individual students from high school into the postsecondary education sector. This section provides suggestions for setting up such a data system and developing the staff know-how to make use of it.

A Statewide Data System: K-16 Education Data Warehouse

The ideal way to collect, store, and organize the data necessary for dual enrollment research is to create a statewide data warehouse. Under this arrangement, a state funds and staffs the system, to which all LEAs submit K-12 data and all postsecondary institutions submit higher education data. The state develops a standard system for gathering, processing, and disseminating the data, as well as a standardized list of variables to collect and a standard data dictionary defining each variable. The state also ensures that each student is assigned an individual identifier and that the identifier follows the student across educational institutions and sectors within the state.

There are many advantages to this system. First, all variables are the same across institutions, eliminating one common barrier to research. More importantly, such a system ensures that all information necessary for analyses on dual enrollment effectiveness are available to researchers. A statewide system allows us to track individual students from high school to college, no matter where they end up after high school, as long as they remain in the state system. Additionally, such a data system includes not only all individuals who participated in dual enrollment programs but also those who did not.

Because a state data system covers dual enrollment programs throughout the state, it allows for analyses comparing the relative effectiveness of different programs and/or program components. Finally, because state data systems can include information on programs other than dual enrollment, such as Advanced Placement or the International Baccalaureate, such a system allows for comparative analyses of the effectiveness of these or other programs and policies. This is particularly useful for return on investment analyses; comparing the returns of dual enrollment versus other credit-based transition programs can help policymakers prioritize which programs to fund in the future.

As we have noted, the Data Quality Campaign has identified 10 essential elements for a data system, and few states have systems that include all elements or even the elements most necessary for dual enrollment research. We recommend that states interested in conducting such research, or states in which LEAs are interested in doing so, develop such a system. We recognize, however, that this is a complicated undertaking.

What is required when setting up such a data system? First, the various entities involved need to understand their roles and commit staff and financial resources to fulfilling them. Generally, this includes state education agencies, state legislatures, LEAs, and postsecondary institutions (ideally, all institutions in the state, but at the very least, all public postsecondary institutions).

State legislatures have a key role in that they are the ones who can require the development of such a system and ensure funding for it. Without adequate funds, it is impossible for other entities to do their jobs. Legislators must understand that developing

a system requires initial funding outlays—to pay for data systems, staff time developing protocols and initiating data collection, and staff training—as well as ongoing funding, to pay for future staff time. Data systems are not a one-time outlay; they require long-term commitment from all involved.

State education agencies or other state officials are the ones who should take the lead in designing and implementing a data system. One way to do so is to create a task force that works exclusively on the project. This team, which should consist of education experts from both K-12 and higher education as well as database managers, is charged with identifying the components of the system. These components include the database infrastructure, the variables to collect and the timeframe in which to do so, and the protocols for institutions to submit their data. The task force team is also charged with addressing issues of data privacy. Members of the team should have their participation be included in their job descriptions and be given ample release time from other duties to devote to the project.

Upon completion of their work, this task force should have made decisions regarding the following:

- Database infrastructure, including programming decisions and mechanisms for linking the state system to the systems used by institutions and LEAs across the state;
- Processes for ensuring complete, timely, and accurate data collection;
- Standardized processes for gathering, storing, processing and disseminating data;
- Mechanisms for creating unique student identifiers that span all educational institutions in the system;
- Definitions of variables to be collected, including a data dictionary and manual to guide institutions in their data collection process;
- Protocols for the use of data by outside researchers, in order to ensure that student privacy is protected; and,
- Mechanisms for providing technical assistance for LEAs and postsecondary institutions.

Once the task force has completed its work, state education agencies continue to play an essential role in a data system. This includes the identification and ongoing support of at least one database technician to oversee data collection activities and the continual update of the database (in reality, a state will probably need to devote more than one staff person to this task). Ideally, state agencies will also hire and support research staff to make use of the database on a regular basis by conducting rigorous research on a range of educational topics important to the state, including, of course, dual enrollment.

Clearly, the entities responsible for generating, collecting, and submitting data to the database—LEAs and postsecondary institutions—are the most important participants in a data system. Within each, at least one individual must be identified who takes responsibility for generating the data and submitting it to the state in a timely manner and in accordance with the established protocols. At the K-12 level, this may be one individual per LEA, or it may be one individual per school. At the postsecondary level, the obvious candidate is a member of the institutional research department. For both, the individual should have some familiarity with database management and research

principles. States should also consider providing technical assistance and training to these individuals, to help ensure that they understand what is expected of them. And LEAs and postsecondary institutions should budget staff time for these activities and account for them in staffing decisions.

As we have noted, the K-12 system is responsible for collecting data on student elementary and secondary achievement, in accordance with state requirements. Most likely, this will include submitting student transcript information. At the postsecondary level, institutions will be expected to submit information on student enrollment, financial aid receipt, and achievement, again most likely including transcript information. At both levels, the data must be collected, organized, and submitted in accordance with the demands of the state system. In some cases, this may require significant modifications in the way that an institution has previously collected data and may require a high level of effort during the first year of the data system.

A comprehensive data system is predicated on qualified personnel to build up and use a system, and all entities involved must ensure that they have adequate human resources for the task. There are two types of staff roles in this system, administrative and research. Administrative staff members include individuals responsible for managing the database and developing data collection procedures. These individuals should have a bachelor's and/or advanced degree in data programming, information management systems, or statistics. While database management is the requirement for these staff members, they should also have a basic statistical background. Research staff members include those who design and use the data. These individuals should have an advanced degree in statistics and quantitative methods and experience conducting large-scale data analyses.

In small institutions and LEAs, it may not be possible to have one individual devoted to these tasks or to find individuals with the required qualifications. In this case, professional development, in the form of college courses or state-sponsored training, should be budgeted for and provided to staff members involved with the data system. The key is to ensure that all staff involved in the data collection process understand basic principles of statistical analyses and database management.

Obviously, developing such a data system takes significant resources. Moreover, it does not happen overnight, so even in states that have committed to such a system, the data will not be ready for analysis right away. This does not mean that no dual enrollment research can be conducted in the interim, however. Many school districts and higher education institutions have invested in collecting and reporting data on student achievement, funding, and faculty qualifications. As we have noted, these data can be used to begin research on dual enrollment. A particularly promising approach is for LEAs and postsecondary institutions to engage in data-sharing agreements in which the LEA provides high school transcript data to the colleges. The higher education institution can then match the transcripts of those students enrolled in the college to their postsecondary transcripts and use the resulting datasets to conduct a variety of analyses that control for pre-existing characteristics.

This type of research comes with some caveats, however. Importantly, the sample will be limited to those students who attended the two institutions in the agreement and so are not necessarily representative. For example, if an LEA partners with a community college, the resulting analyses will only be relevant to students who matriculate into a

two-year institution, not to students graduating from the LEA's high schools more generally.

A second approach for conducting dual enrollment analyses in absence of a state data system is for individual institutions or districts to use their own data to conduct descriptive or implementation analyses of their programs. They might, for example, explore participation rates or course grades or conduct an audit for quality. Of course, as we have noted, this simple approach provides useful information but cannot address causality.

Principles of data collection

In developing data systems, these principles should guide decisions regarding which variables to include and the structure of the overall system.

1. *Follow the educational careers of individual students.* As we have noted, the key to conducting dual enrollment research is to be able to track individual students from prior to program enrollment until well after program completion. This means that data systems need to be longitudinal, following students over many years. Data should be collected starting in at least students' ninth grade year and continue to be collected for at least six years after high school, to allow for them to complete a bachelor's degree and/or enter post-baccalaureate programs. Data systems must also have a mechanism for keeping tabs on individual students—ideally through the creation of a unique student identifier that follows students from one educational institution to another.

2. *Include cross-sector information.* As we have noted, dual enrollment research requires data from both high school and college transcripts. However, it is important to include private, as well as public, institutions within this system wherever possible. Although including only public institutions is easier politically and will encompass most students in a state, it does create some limitations. First, excluding students in private institutions limits the generalizability of the study, as not all students in a state are represented in the data. Second, students will inevitably be lost as they move from public to private institutions. This is particularly true at the postsecondary level. For example, if only public postsecondary institutions are included in a dataset, dual enrollment students who attend private colleges will look the same as students who do not attend college at all, as both will lack college transcripts. The two groups have very different outcomes, however. In reality, an ideal system will also include data-sharing with other states, in order to capture those students who cross state lines to attend school, but we recognize that this is far beyond the scope of most states' data capabilities at this time. Working with the National Student Clearinghouse is one way to augment state datasets with information about students who leave the state to attend college that does not require a state to collect such data on its own.

3. *Consistent data entry procedures.* Though it sounds self-evident, it is important to emphasize that, in research, missing and not-properly-entered data complicate and significantly limit the data analyses. This is because students for whom data are incomplete must be excluded from analyses, thereby making the sample smaller and because inaccurate data entry leads to measurement error. Thus, accurate entry of data is essential. One way to ensure accuracy is to create instructions for data collection and entry that are to be used across the state. Documenting all data-collection procedures helps researchers detect errors. States should also consider establishing a

data verification process separately from data collection to further ensure accuracy of the database. They should also invest in professional development to ensure that staff throughout the state understand these procedures.

4. *Consistent definitions.* Institutions often define the same variable in different ways. For example, "dual enrollment participant" can actually mean a number of different things: a student who ever enrolled in a dual enrollment course; a student who remained enrolled in a dual enrollment course past a certain date in the semester; a student who completed a dual enrollment course with any grade; or a student who completed a dual enrollment course with a passing grade. If different institutions submit data using different definitions, the dataset will be inconsistent and results of analyses will be skewed. Thus, it is important for the state to define all variables before embarking upon data collection efforts. By creating and disseminating a standardized data dictionary, states can ensure that each LEA and institution report the same piece of information for each variable.

5. *Ensure student confidentiality.* This is a key issue for states to contend with, as it is absolutely essential that any data system protect the identity of students. Although we have advocated being able to track students from one institution to another, it is imperative that this system be constructed in a way that students <u>cannot actually be identified in real life</u>. Thus, student identifiers should not be things such as social security numbers, which have real-world implications. Instead, they should be created through algorithms that combine student names, date of birth, and system entry year. Any location in which the identifier is linked to a student name should be password protected and limited in who has access. In addition, it is important to remember that federal and state laws and regulations limit the release of individually identifiable information about students under 18 years of age without the consent of their families. There should be statewide agreements and protocols that comply with these laws and regulations, particularly if outside researchers will be using the data. States should establish protocols similar to those used by university Institutional Review Boards to monitor the use of data, and researchers should be trained in methods for human subjects' protection.

Three Approaches to Systems-Building

Building a statewide data system is, as we have noted, a time-consuming effort. It also requires states to navigate the political landscape and ensure the buy-in of multiple stakeholders. This can pose challenges outside of the technical ones we have discussed in this section. Three states, Oregon, Florida and California, have used different approaches to build their data systems.

Due to its strong history of local control, Oregon has taken a regionalized approach to creating a data system. The Pre-K through Grade 16 Integrated Data System (KIDS) (www.ode.state.or.us) relies on six regional data warehouses to collect and store student unit record data. The state provides leadership for this system but the regions and individual districts retain control over privacy (including student identifiers), data collection software, data storage, and data definitions. Currently, the system only contains data for K-12 institutions; connections to postsecondary are still in development.

Given the regionalized nature of the system, there are some challenges following students across regions. The state is currently providing leadership to overcome these challenges. Given the strong local control tradition in this state, however, a regionalized system has served to encourage data sharing more quickly than could occur under a more centralized approach.

Cal-PASS (http://www.cal-pass.org) is a voluntary consortium of over 2,000 California institutions. The state does not mandate participation; rather institutions or LEAs decide to become a part of the system. System members include individual elementary and high schools, K-12 school districts, community colleges, and four-year colleges and universities. Participating institutions sign a Memorandum of Understanding and commit to sharing five years of student demographic, course enrollment, and performance data. Student identifiers are encrypted numbers so that no individual can be identified in the database. Because not every institution in the state system is included, the data available through Cal-Pass are somewhat limited. For example, if students matriculate in a postsecondary institution that does not participate in the consortium, their records are not available for inclusion in research. Still, in a state as large as California, the system provides for a high degree of data sharing.

Florida has taken a centralized approach. Its data system, known as the PK-20 Education Data Warehouse (EDW), is a mandated system in which all public education institutions are required to submit data on students, faculty, and facilities. These institutions include all K-12 LEAs, community colleges, and public universities in the state. Data from adult career and technical centers, finance and accounting systems, student assessment databases (such as the tests used for placement into remedial coursework), and student financial aid information are also included in the system. The data system is housed within the state's centralized Department of Education. All participating institutions load their records into the EDW, and the records are integrated and matched into a single database. The system assigns randomly generated unique identification numbers to individuals, allowing for longitudinal tracking of students while maintaining their privacy. Because the system is mandated and such a vast array of institutions participate, it is a comprehensive dataset that allows for complex research on student success. More information can be found at: http://edwapp.doe.state.fl.us.

V. Using research effectively

Dedicating funds toward research, though expensive in the short-term, can create efficiencies in the future and help ascertain that funds are being used wisely and programs are being implemented well. Once policymakers have committed to supporting and using research, the following actions can help ensure that the findings are put to good use.

- *Meet with researchers to understand the activities being conducted.* As we have said, different types of research are useful for answering different questions. Policymakers and research staff need to work together to make sure that the most relevant questions are being answered and to ensure that research findings are not misinterpreted.

Ideally, these meetings should occur prior to the start of a research project, to determine the appropriate questions to ask, and upon completion of the project, to clarify the interpretation of and limitations to the findings.

- *Work with other policymakers to codify research activities into budgets and legislation.* Because research, especially when focused on long-term outcomes, takes time to complete, research activities that are not supported over a number of years are unlikely to be useful for decision-makers. Thus, policymakers should work to ensure that research and research-related activities, like a state database, are supported. This includes sponsoring enabling legislation and dedicating funds for research, thereby ensuring that projects that can be used to guide policy are actually completed.

- *Use research to build on success and learn from failure.* Policymakers should use research findings to identify the programmatic elements that appear to promote desired outcomes and those that do not. They should also seek evidence from research on unintended program consequences, such as inequitable access or varied implementation. These findings can then be used to modify policies and regulations, change funding streams, or increase staff dedicated to program quality, as appropriate.

- *Create new policies that are supported by research findings.* Sometimes, research findings may demand that entirely new policies be developed, and policymakers should be open to, and take advantage of, such a situation. For example, New York State has long lacked a comprehensive dual enrollment policy or dual enrollment funding stream. Research conducted on New York City's College Now dual enrollment program showing positive results influenced state education officials and led them to propose a state dual enrollment program targeted at low-income students. This initiative garnered significant media attention and provided a positive image for the state education community.

- *Use research to create an argument.* Policymakers should not let research findings exist in a vacuum. Rather, they should use them to craft an argument, for example, in favor of including new students in dual enrollment programs. Presenting these arguments, rather than just a list of facts or findings, to stakeholders can engender support for the proposed changes.

VI. Conclusions

This paper has provided a blueprint for conducting research on dual enrollment. We have described the data systems required and the capacities that states and LEAs need in order to make use of them. We have also described three types of research— descriptive, implementation, and outcomes—and the benefits and shortcomings of each.

Throughout, we have emphasized that dual enrollment research should stem from state and/or LEA goals for the program. Research questions should address those goals, and the methods should follow from the questions. If this is not done, it is likely that the findings will not be relevant to the local policy and programmatic landscape.

For states and LEAs that seek to conduct dual enrollment research, what are the next steps? They should consider the following, using the sample research presented here as a guide.

1. *Develop consensus on establishing a data system.* States and LEAs should invest the time and resources in developing a data system that will allow them to conduct the types of analyses that they would like. This will also require them to think strategically about what those analyses might be.
2. *Construct the human and financial infrastructure necessary for a data system.* This includes budgeting for and training personnel as well as the mechanics of the system.
3. *Enhance research capacity.* This includes hiring and training staff to conduct the actual research at the state, LEA, and school levels.

Once these steps have been undertaken, states and LEAs will be ready to conduct the types of research discussed in this report and to disseminate the findings to decision-makers and the public.

Works Cited

American Association of State Colleges and Universities. (2002). The open door: Assessing the promise and problems of dual enrollment. *AASCU State Policy Briefing, 1*(1), 1–10.

Bailey, T. R., & Karp, M. M. (2003). *Promoting college access and success: A review of credit-based transition programs.* Washington, DC: U.S. Department of Education.

Calcagno, J. C. (2007). *Evaluating the impact of remedial education in community colleges: A quasi-experimental regression discontinuity design.* Unpublished doctoral dissertation, Columbia University.

Florida Department of Education. (2004). *Impact of dual enrollment on high performing students.* Data Trend #26. Tallahassee, FL: Author. Retrieved October 5, 2007, from http://www.fldoe.org/cc/OSAS/DataTrendsResearch/Data_Trends.asp

Karp, M. M., Bailey, T. R., Hughes, K. L., & Fermin, B. (2004). *State dual enrollment policies: Addressing access and equity.* Washington, DC: U.S. Department of Education.

Karp, M. M., Calcagno, J. C., Hughes, K. L, Jeong, D. W., & Bailey, T. R. (2007). *The postsecondary achievement of participants in dual enrollment: An analysis of student outcomes in two states.* St. Paul, MN: National Research Center for Career and Technical Education.

 Lerner, J. B., & Brand, B. (2006). *The college ladder: Linking secondary and postsecondary education for success for all students.*

Michalowski, S. (2006). *Positive effects associated with College Now participation for students from New York City high schools: Fall 2003 first-time freshman cohort.* Retrieved October 5, 2007, from http://collaborativetest.cuny.edu/pdf/special/cn_participation_positive.pdf

Skadberg, I. (2005). *Quantitative data analysis of the CUNY College Now program, 2001–2003.* New York, NY: City University of New York, Office of Academic Affairs.