



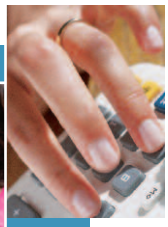

BreakingThrough

Helping Low-Skilled Adults Enter and Succeed in College and Careers

Advancing Adults into Community College Programs:

Data Tools from Breaking Through

By **Peter T. Ewell,** National Center for Higher Education Management Systems

APRIL 2008

JOBS FOR THE FUTURE
CREATING STRATEGIES
for Educational and Economic Opportunity

NCWE
National Council *for*
Workforce Education

Advancing Adults into Community College Programs:

Data Tools from Breaking Through

TABLE OF CONTENTS

Preface

The goal of *Breaking Through* is to enable community colleges to restructure their programs and instructional practices to create clear pathways for low-skill adults into professional/technical certificate and degree programs. The initiative is jointly led by Jobs for the Future and the National Council for Workforce Education, within funding from the Charles Stewart Mott, North Carolina Glaxo-SmithKline, and Ford foundations.

Since the initiative began in 2004, data collection and retrieval have been major challenges to helping low-skilled adults succeed. That year, when JFF and NCWE staff visited 25 community colleges and related programs across the country, we found that the different programs at a single college frequently would use incompatible software and collect different kinds of data. As a result, it was impossible for colleges to know whether low-skilled adults were advancing and completing degree or certificate programs. And the leaders of the colleges knew it: many who later applied to participate in *Breaking Through* told us that incompatible and insufficient data was a significant internal hurdle they needed to overcome.

Recognizing the urgency surrounding the data challenge, we invited Dr. Peter Ewell of the National Center for Higher Education Management Systems to work with the colleges selected in 2005 to participate in *Breaking Through*. He has provided insight into the challenges confronting community colleges that serve low-skilled adults, and he has gone further, suggesting solutions that would be useful not just to *Breaking Through* participants but to any college seeking to serve low-skilled adults.

The challenges—and tactics for confronting them—are central to Dr. Ewell's data tools. These tools provide guidance in areas central to improving data tracking in order to determine whether projects like *Breaking Through* are meeting their goals, and they can assist colleges that are designing new ways to build in accountability systems from the beginning.

Other community colleges—as well as other adult-centered programs—will find value in these tools. For example, the "Common Core of Data Elements" can provide institutions with the minimal set of data elements they should collect to support effective tracking and program evaluation. The "Campus Data Toolkit" is a short, easy-to-follow primer to acquaint programs with the key processes and principles of quantitative analysis.

Dr. Ewell has performed a major service for the nation's community colleges.

Marlene Seltzer
Jobs for the Future

Dr. James Jacobs
National Council for Workforce Education

Overview

Colleges participating in *Breaking Through* have demonstrated a strong interest in the related issues of data types, data collection, and data analysis. Many factors account for this interest. The colleges understand, for example, that program staff can benefit from good data systems to monitor participation, track progress towards program goals, contact students as needed, and provide necessary supports.

Another example concerns program improvement—"continuous quality improvement" in the language of the private sector. Data are scrutinized to figure out which program elements are producing their intended results, which elements are not, and where changes in the program design or implementation can be made to improve outcomes. Of course, consistent and reliable data are essential for this purpose.

Finally, there is a widespread concern among *Breaking Through* participants about institutional reform: how can the innovations of *Breaking Through* be preserved after philanthropic funding ceases? For these purposes, data demonstrating improved outcomes are essential.

It's important to remember, however, that *Breaking Through* is a demonstration project whose goal is to connect adults to programs outside the degree-granting structure (i.e., "adult basic ed," workforce development and customized training, and developmental education). Most if not all colleges in *Breaking Through* are attempting to create pathways that connect multiple programs—each with its own data systems, data elements, even software—whose variety and lack of connection make it virtually impossible to coordinate at the program and department levels.

Understanding that this situation created a formidable barrier to collecting detailed data from funded colleges (Leadership Colleges), Memoranda of Understanding were negotiated with each college that specified the level and type of data that were required for reporting purposes. In a nutshell, Leadership Colleges must collect and report the total number of people enrolling in each identified stage of the pathway, the number that complete, and the number that transition to the next stage.

Nevertheless, a number of Leadership Colleges are seeking to develop systems that meet the needs of program management and improvement, institutional reform and evaluations. In addition, a number of Learning Colleges are seeking to develop programs that connect low-skill adults to degree programs, and are interested in finding out about data-related issues in advance.

As a result of the strong interest in data issues evinced by the participating colleges, *Breaking Through* engaged the National Center for Higher Education Management Systems to survey the colleges' current data collection efforts, identify data elements essential for the purposes described above, to define benchmarks for use in analyzing colleges' strengths and challenges, and to develop a "toolkit" for the colleges' use.

This report begins with a toolkit describing various methods for capturing or collecting data. It then presents recommendations for collecting a common core of data; this core was originally developed for *Breaking Through* colleges, but it is directly relevant for any college seeking to track outcomes for pre-college students. While the common core represents data that colleges *should* collect, the larger "data map" points to a variety of data that colleges *could* collect. Appendix 1 provides a number of benchmark state and national statistics that can help colleges put their own performance in context. Appendix 2 presents the common core, a subset of a much larger set of data elements captured by *Breaking Through* colleges, as a "data map," with descriptions of each element.

Campus Data Toolkit

Breaking Through community colleges need data on the students they serve in order to manage their programs effectively and evaluate the effectiveness of their programs in moving students to higher levels of skills and achievement. This Data Toolkit briefly reviews approaches to collecting, archiving, and analyzing student-level data that colleges may find useful as they develop their programs. The Data Toolkit also provides references or links to more complete documentation that colleges can access if they wish. None of the techniques or approaches reviewed here is mandatory for participants in *Breaking Through*, and campuses should fit what they do to local circumstances and needs.

The Data Toolkit has four main sections. The first distinguishes among several different types of student data that are useful for different purposes. The second addresses the need for, and possible ways to organize, a project management database. The third reviews various ways to collect data, concentrating on the need to collect data about interventions and treatments that are not captured through the college's regular registration records system. The final section examines ways to manipulate and analyze data to monitor student progress and overall program effectiveness.

Different Types of Data

Colleges maintain data on students primarily to: document their progress through academic programs; maintain background information for mandatory state and federal reporting; and contact students and manage transactions related to their enrollment (e.g., tuition payments, financial aid).[1] The vast majority of the data elements they collect for these purposes are contained in the college's registration or student records system. These data are typically termed "unit records" because they consist of a single data record, containing a range of standard data elements, for each student enrolled at the college for a particular block of time (generally a semester or term). In addition to these data resources, *Breaking Through* colleges collect a range of additional unit record data on the students they serve. These additional data elements are intended to help project staff manage their programs or to document impacts on students drawn from particular populations (e.g., first-generation students, low-income students, students with specific challenges or disabilities). For the most part, program administrators use data drawn from these two sources on a day-to-day basis to ensure that requirements are met, document interventions and learning experiences, and maintain contact with individual students and keep track of their progress. But subsets of these data elements are frequently used to support research on the impact of the program—for example, retention and completion studies that look at student success for different types of students or analyses that seek to determine the root causes of success.

These variations in use lead to a basic distinction between two types of data: transactional and analytic.[2]

Transactional Data

A *transactional record* is created in a student records database each time a student interacts with the institution in a way that needs to be documented. Examples include applying for admission, registering for a course, or receiving a grade, financial aid payment, or a degree. Transactional records of this kind constitute the bulk of records generated by colleges about students, and they are used to manage the day-to-day actions that govern enrollment. Registrars use transactional data to produce up-to-date transcripts of student work and to take actions about particular students like placing them on academic probation, placing them in courses, or restricting their ability to register. Faculty and advisors use transactional data to examine student academic histories and relevant demographic and behavioral data to monitor progress and intervene appropriately. Transactional data can be generated at the college or the project level. Most data will be drawn from the former, using the institution's registration records system. But some projects may collect further transactional records only for students served by the project. Examples of the latter might include participation in special courses or study sessions, or specific student demographics or characteristics of interest.

The more up-to-date transactional data are, the better. Advisors and faculty will want to know what a given student's condition is *right now*, and such information is typically updated daily. Transactional data are also most frequently used on a student-by-student basis. That is, the user wants to know something about a particular student and is less interested in questions about patterns of behavior across students. Occasionally, though, a project administrator might want to know the current condition of a particular type of students—for example, how many female students are about to graduate from the medical technician program or how many students are currently placed in developmental mathematics. These are also real-time administrative questions that can be answered by directly querying a transactional database.

Analytical Data

Analytical data, in contrast, are optimized to support systematic research about groups of students whose behavior and success are compared over time. Examples include the completion rates of men as opposed to women, or the persistence rates of students in the *Breaking Through* project and students not participating. While most analytical data are derived from transactional records, some projects may collect additional data elements through special questionnaires. Wherever they come from, analytical data elements must be captured *regularly* at the same points in time each term in order to assemble consistent data for research. For example, all colleges have a "census date" for each term, the point at which official enrollment counts are taken. Establishing an end-of-term census date, in turn, allows the college to collect consistent, official statistics about things like grade distributions and degree awards. Without census dates, analytical data about student progress would be hard to interpret because the populations being compared are allowed different amounts of time to accomplish the same goals. There is no limit to the number or placement of these "data capture" points, so long as they are maintained consistently over time.

Using Transactional and Analytical Data

These two different kinds of data are intended for different purposes. Transactional data are optimized for viewing the current condition of individual students or groups of students to inform an immediate decision. Analytical data are optimized for supporting causal and impact research, where the up-to-the-minute condition of every student is less important than the consistency with which these records are maintained over time.

Because these two kinds of data are different and are easily confused, analytical data elements are usually maintained in a separate analytical database that is updated at each census date. When this census date is reached, the value of each data element is extracted from the transactional file and placed in the analytical database to yield a growing longitudinal record. For example, the transactional file will indicate the academic program that the student is enrolled in at the moment, even though (s)he may have been enrolled in a different academic program yesterday. The analytical file, in contrast, can reveal that the student changed academic programs three times in the course of their enrollment.

Institutional research offices frequently maintain fairly extensive analytical databases in order to compile official statistics for the college and to study such topics as completion rates, persistence, and the success of developmental education. *Breaking Through* projects should take advantage of this preassembled analytical data when creating their own databases: these data are clean and organized to support statistical analyses. If an institutional research office already has a particular data element of interest to *Breaking Through* staff, it makes little sense to collect it directly from project students through a special admissions form or survey. Instead, everything relevant that is already available through a central institutional resource should be accessed from these sources.

The result of all this is that *Breaking Through* Projects deal with four different types of data elements *(see Figure 1):*

- "Type A" data elements are regularly collected by the institution and typically reside in its registration and records system. Examples include course grades, demographics, and student goal information collected through college admissions forms.
- "Type B" data elements are regularly collected by projects to monitor student progress and guide individual interventions, but they are not collected at the institutional level. Examples include the availability of day care, how the student was referred to the project, and counseling or tutoring sessions in which the student has participated.

Type A and Type B data elements can be combined in a locally maintained project management database, with Type A data elements downloaded from the institution's registration records system or obtained from the office of Institutional Research.

- "Type C" data elements are extracted according to a defined set of census dates and typically reside in the analytical databases maintained by the institutional research office. Examples include grade point average for the last term, degree awards, and financial aid support.
- "Type D" data elements are collected by projects to inform local analyses of effectiveness or impact. Examples may include such things as additional student descriptors (e.g., single parent) or additional outcomes (e.g., job placement).

Type C and Type D data elements can be combined in a locally maintained project analytical database, with Type C data elements provided each term.

Table 1.
Types of Data Elements

Rows in this table represent the distinction between transactional and analytical data elements—the first live and up-to-the-minute, and the second captured regularly and consistently to support statistical analyses. Columns represent the locus of responsibility for capturing and maintaining these data elements—the institution or the project.

	Institution Maintained	Project Maintained
Transactional	Type A	Type B
Analytical	Type C	Type D

Another important distinction between types of data elements is based on how much they have been manipulated. A simple example is a student's age, which is typically calculated from the date-of-birth data element carried in most student registration records databases. Data elements like age, which are the result of a calculation or manipulation of the raw data entered, are usually called "derived" data elements; they can be quite complicated when created for analytical purposes. For example, many colleges construct a data element that indicates a student's "dropout status," where students are counted as a dropout if they have not been enrolled for a given number of terms. Others create a combined "student success" flag that is based on completing a degree or credential, remaining enrolled in the current term, or having transferred to another institution. The advantage of creating derived data elements of this kind is that they can be calculated according to known rules on a regular schedule. Derived data elements like this are used primarily in analytical databases.

Finally, data elements of all types must be documented so that project staff know where they came from and what they mean. Database administrators typically do this in the form of a "data element dictionary" (DED) that lists each data element, its definition, its coding structure (e.g., M=male and F=female), and its source (e.g., student registration records system, entering student survey, advisor input). Some DEDs go further, noting how the data element should be used or providing any cautions or limits about its use. While the preparation of a formal DED is probably not necessary for *Breaking Through* project management and analytical databases, some documentation of this kind in the form of an annotated list of data elements is highly recommended, especially if more than one person will be accessing and using the data. Documentation is particularly needed for any derived data elements, where the underlying definition and calculation rules used to create the data element may be unclear.

Project Management Databases

Most *Breaking Through* projects are creating project management databases to house all the data they need on participating students in order to run the project. With a single source of data, kept in an accessible location, and containing current data on every participating student, project staff can monitor each student's progress and, where necessary, intervene to address emerging difficulties or known deficiencies. In this sense, the project management database functions for project students much as the college's student registration and records system functions for other students. The difference is that *Breaking Through* students have distinctive educational experiences, and they benefit from special services that regular students do not have access to, and that are not recorded in registration records. Moreover, project staff may collect additional background or demographic information about *Breaking Through* students for monitoring and reporting—for example, if the student needs special services or is an employed single parent. All these data are better used if they are kept in one place that is easy for project staff to access and examine.

Because their contents are used to monitor student progress and inform intervention, project management databases will mostly contain Type A and B data elements, as defined above. Some of these data elements will be derived directly from existing student records maintained by the registrar or accessed by project staff from the college's registration records system. Some will be obtained by project staff directly—either through admissions forms or surveys administered only to *Breaking Through* students or through transactions or activities that are only available to *Breaking Through* students.

When constructing project management databases, it is important to emphasize the principle of one-way data flow. This means that regular institutional records are the preferred

source for any given data element. For example, data on student gender, race/ethnicity, grades earned in particular classes, or program of study would be accessed from existing records instead of being collected independently for *Breaking Through* students by project staff. This avoids duplicate data collection, although it may involve some initial effort on the part of project staff to map existing data resources at the college and to develop a routine way to obtain the required data on a regular basis. More importantly, it avoids the potentially challenging problem of having contradictory data drawn from different sources.

The principle of one-way data flow also means that project staff must determine up front what data elements are available from other sources. Conversations with the registrar and the institutional research office are a good first step, and will probably identify the vast majority of easily accessible, centrally maintained data elements that will be useful for project management. Some colleges may also have conducted a recent "data audit" to determine the entire array of data collected about students by various offices, making it possible to share these data across units for multiple purposes.[3] If so, the project team should carefully examine the results of the audit to identify additional data elements that might be useful for project management or for analyzing effectiveness.

Another consideration that arises in building a project management database is the medium in which data will be kept, as well as how project staff will access and manipulate the data. The simplest approach is a combination of paper and electronic records. The primary drawbacks of this approach are that several sources must be consulted to determine the current status of a given student in all areas of interest and that aggregate analyses are difficult using paper records. As noted below, most of the *Breaking Through* Leadership Colleges took the combination approach initially, but all are moving toward a situation where project management data is all in one place and stored electronically.

An electronic spreadsheet (e.g., Microsoft Excel) is the most basic medium for a project management database. Each student unit record is contained in a row of the spreadsheet, with columns corresponding to individual data elements. With Excel, different kinds of data can be loaded into different worksheets within the same file, and this feature can be employed to organize data about students. For example, the values of a standard set of data elements involving student participation and enrollment (e.g., classes enrolled for, grades earned, tutorial sessions, testing milestones) can be organized in a standard format that is repeated on a different sheet for each successive term.

Spreadsheets allow useful, although limited, manipulations of the data they contain. All entries on a given worksheet can be sorted by data element. This allows project staff to quickly locate individual students by name to consult their latest records. It also allows users to identify groups of students who may be scheduled for some activity or who are in need of intervention. For example, all students who scored below a particular point on a specific placement test or whose mid-term grade is below a particular level can be identified so they can be contacted for special treatment. Spreadsheets also have limited analytical capacities, such as the ability to generate summary statistics (e.g., percentages, totals, averages) for particular populations.

The main drawback to using spreadsheets as project management tools is that Type A data elements are usually entered by hand instead of being downloaded from the original sources in which they reside. Double entry of this kind is inefficient and, more importantly, may introduce errors when data are reentered. The need for hand entry can be avoided when the data records for each new group of students

(or cohort) are first established, because the registrar or institutional research office can provide project staff with a file containing all Type A data elements at the beginning of the enrollment period. This file can be used as the basis for starting the spreadsheet record of this cohort. However, it is usually impractical to download additional performance information like test scores, classes taken, and grades: these are posted piecemeal by the registrar, and project staff will probably not want to wait until the end of the term to receive such information. As a result, the most feasible approach is to periodically retrieve each student's record from the registration system, obtain the needed information, and re-post it in the appropriate place in the spreadsheet.

A more sophisticated project management database can use dedicated database software packages (e.g., Microsoft Access). An environment like this is much more sophisticated with respect to manipulating and analyzing data. It also allows users to construct an easy-to-interpret data record for each student, showing key information in an intuitive array, instead of asking users to search across multiple columns in a given row of a spreadsheet.

Displays can be customized, and many specialized views of an individual student record can be created through the database's programming language. In addition, regular reports and analyses can be pre-programmed so that project staff can quickly produce them for any given group of students. The programming language used by such databases is relatively straightforward, and many project staff will be in a position to learn it. But it is also sufficiently well known that project staff will typically be able to find someone at the institution who can write the needed code. Moreover, with regard to getting data into the system, database systems can be set up to receive new data regularly from external sources in order to update the student records that they contain. And customized input screens can be created to make it easier for project staff to update records manually from the sources for which they are responsible.

Finally, several commercial project management tools are configured especially for initiatives like *Breaking Through*. For example, Efforts to Outcomes (ETO), a software suite for social service providers, enables users to record a wide range of individualized data for case management purposes; it is focused on tracking progress toward a project's or an individual's objectives.[4] Amatrol, an electronic course management system, was developed especially for vocational technical training. It incorporates Student Database and Class Database modules for recording and managing a range of data elements about individual students.[5] Both databases can accommodate user-defined data elements, as well as a range of standard student descriptors, and they can be configured to fit the flexible, non-term-based formats that are typical of *Breaking Through* curricula. More conventional electronic course management systems (e.g., Blackboard) also have limited case management capabilities but are more suited to traditional classroom environments.[6]

Data Collection and Data Capture

The most basic requirement in building a database is to obtain the necessary data. Once a project team has decided which data elements it wants to collect, the first task is to determine whether the data already exist. For example, the registrar and other offices regularly collect transactional Type A data; the principal challenge is finding the data and creating a channel to obtain them on a regular basis.[7] Liaison with the registrar's office is typically the first step in developing a plan for regularly accessing transactional data of this kind. Most registration records systems have a utility that allows users to download selected data elements; project staff should familiarize themselves with these procedures to populate their own project management databases. Alternatively, they should work with the registrar's

office to obtain the needed data regularly in electronic form through a file transfer or portable mass storage device. The key, as noted earlier, is to avoid hand entering data that are already in electronic form.

In contrast, project staff must collect project-maintained data (Type B). One of the most common approaches is to administer a survey. For example, many projects design an application or intake form that *Breaking Through* students complete as they enter the program. Many projects also periodically survey participants to determine their experiences in and reactions to project-related activities. Finally, some projects survey former participants to determine if they are employed in their field of training or are otherwise putting their skills to use.[8]

Although such surveys are usually fairly short and straightforward, they must be designed effectively in order to maximize response and ensure that the information collected is accurate. One useful place to start is to consult with the college's institutional research office: such offices regularly design and conduct student surveys. IR staff may be helpful in reviewing drafts of any contemplated surveys and suggesting improvements based on their considerable experience. In addition, if the survey or form is not completed face-to-face, IR staff may have methodological advice on how best to administer it. Finally, they may be aware of other offices that collect similar information or they may be doing so themselves. If this is the case, it may be useful to partner with IR or another office in administering the survey. Partnering will be particularly helpful in following former students; the IR office may already target graduates with its own questions, and it may be willing to include project-related questions as an addendum to an existing survey.

Regardless of length or complexity, there are a number of considerations to bear in mind when constructing surveys or data collection forms:[9]

- *Keep it simple.* Short, straightforward forms and questions are more likely to be answered accurately (or, in the case of surveys, answered at all). Ask only what you need to know and keep the language simple.
- *Think about how you will use the data.* One way to keep surveys short is to carefully think through in advance exactly what will be done with the answer to each question developed. Who will be interested in the answer and why? What kinds of interventions or analyses will each piece of data support? Not only does this limit length, but it also starts those involved thinking about needed action from the outset.
- *Ask what people can answer.* When posing a question, it is always helpful to think concretely about whether those responding are in a position to answer accurately. For example, students generally give accurate reports about their behaviors (e.g., how many hours a week they study, how long it takes them to commute to school), but they are much less reliable when it comes to reporting how much they have learned. This is not to say that asking about learning should be avoided, but keep in mind the inherent limits on students' answers. On opinion questions, moreover, it is important to word questions so they do not solicit answers that students think you want to hear.
- *Build trust and assure confidentiality.* Many people get nervous when faced with a form asking personal questions or a survey whose purpose is unclear. Given this, it is always important to assure students that their replies will be treated as confidential and accorded the same degree of respect and protection as their academic and personal records. It may also help to briefly describe why the information is being sought, who will have access to it, and how it will be used. A final good practice is to offer to share the overall survey results with respon-

dents when it is completed so they can see how others responded and how the data will be used.

- *Try it out first.* Even very simple forms or questionnaires can benefit from a brief pilot test with a few people. What a given question or conversation means may seem obvious to the person who thought it up, but it may be read differently by someone seeing the material for the first time. Asking three or four people to "talk through" their responses to any proposed form or survey is a good way to identify potential problems.
- *Consider capture in electronic form.* A major source of overhead in collecting data from students is the need to input the data to a database from a paper form. While in many cases this cannot be avoided, consider a couple of alternatives. First, most institutions have electronic scanning capability—most likely in a testing center. If the survey or form can be constructed on a form that can be scanned, this will save a lot of work.[10] Second, the form or survey can be completed on line. This saves on the labor of input and on printing and distribution costs.

Another common challenge for collecting data is documenting student encounters with support services and other interventions intended to improve their chances of success. This is especially relevant for *Breaking Through* projects because advising, tutoring, and similar services are frequently prominent aspects of these initiatives. As a result, determining when and how intensively these interventions are experienced by participating students may be as important as documenting courses taken to determine specific factors responsible for success. Unfortunately, few institutions systematically collect data about these encounters. If records exist, they are usually collected by those offering the service, and frequently they are collected only as paper records.

If such interventions are important elements of a *Breaking Through* project, its staff should think carefully about how they might be documented. One approach, as above, is to periodically survey students to ask if, when, and how many times they have experienced these encounters. More reliably, project staff can establish a regular process for collecting data using a common template in Excel or a similar spreadsheet package. This should be set up to record each time a student engages a particular relevant service or encounter. The record should include the student's name, ID number, the time and date of the encounter, and the type of encounter (e.g., study skills session, math tutoring, counseling). It may also be a good idea to include an open-ended comment field for project or office staff to record additional detail where this is warranted.

Such templates can be loaded on desktop computers at each point of service and periodically sent to project staff as an email attachment or using a portable mass storage device. They can then be aggregated into a single spreadsheet, sorted by name or ID number, and loaded into the project management database. At the highest level of sophistication, some institutions have instituted electronic student identification cards that can be used for bookstore purchases, checking out books at the library, and a number of other services. If this is the case, it may be useful to explore installing a card-swipe capability at key points of service to document when and who is using the service.

Finally, additional data that may be helpful in documenting project impact may be available from state or community college system databases through the institutional research office. For example, many states are beginning to match student enrollment records with state employment records to obtain information on graduates' subsequent employment in field and/or increases in earnings that can be attributed to enhanced training. Match rates are generally quite good using this procedure, and

they are usually superior in response rates to a graduate follow-up questionnaire. Subsequent enrollments and credentials earned at other public institutions can also be obtained from these sources.[11]

Data Analysis and Reporting

Breaking Through project staff will use much of the data that they assemble about student participants for case management. While such data must be accessible on a case-by-case basis, the underlying structure of the database—how data elements are organized and manipulated—does not matter greatly. However, this is not the case for data elements that will be used in impact analyses or statistical reporting. These purposes may require the creation of specially configured data files, as well as new data elements calculated based on existing data. Moreover, to properly document impact, best practice is to compare the performances of *Breaking Through* students with those of similar students who are not participants in the program.

Many reporting statistics can be generated directly from project records stored in spreadsheets (e.g., Excel) or more sophisticated database packages (e.g., Access). These include such data as enrollment counts, average placement test score performance, and course completion rates. Spreadsheets also have built-in functions that can calculate sums, means, frequencies by category, and other common summary statistics. In addition, they give the user a limited ability to manipulate cases through sorting by column, as well as to create graphics. And users can write "macros" (mini-programs) to perform multi-step calculations or procedures; this is a good idea because it ensures that standard reports are generated consistently.[12] Finally, if spreadsheets are the primary medium used for analysis, users can create additional worksheets in the same workbook within which to sort and manipulate the data; this ensures that the underlying data contents remain intact if cases are re-ordered, deleted, or moved.

Database software provides a more sophisticated environment for data analysis. The "relational structure" of these programs allows the data to be manipulated without the need to physically sort cases.[13] For example, users can cross-tabulate multiple variables to examine outcomes statistics for different student populations (e.g., credit-completion ratios by gender or remediation status). Also, the statistical procedures and reporting capabilities of database software are much more sophisticated. While much of this can be accomplished in a point-and-click mode, regular reports and calculations are best created through pre-written programming. Because this is sometimes challenging, project staff may want to set up these procedures once, with the help of someone familiar with the database programming language on campus.

Whatever software environment is used, all *Breaking Through* projects will likely find it useful to collect the following summary statistics for each enrollment period (e.g., term, quarter, semester):

- Total headcount enrollment in the project;
- Average credits enrolled for;
- Average credits completed;
- Credit completion ratio;
- Average GPA or current grade distribution;
- Number of students achieving key project milestones (e.g., successful completion of ABE work, successful completion of developmental work, enrollment for first college-level credit, successful completion of 12 college-level credits or equivalent, credential earned);
- Students enrolled in and completing key classes (e.g., developmental, gatekeeper); and
- Numbers of contacts with project activities or staff.

All of these are most useful if broken down by demographic factors, including gender, race/ethnicity, age, income level (where available), level of study, field of study (where appropriate), or full-time/part-time status.

Database structures can become a good deal more complex when the objective is to track students over time.[14] Here, the approach is to create, for each potential enrollment period, blocks of identically defined variables that address such things as credits attempted and earned, as well as various critical events, such as earning a credential—together with a full range of demographic descriptors that enable project staff to compare the progress of different student populations. Additional detail is sometimes added in the form of performance on key courses. If additional data elements capture document student experiences and interventions, these can be included in the tracking record.

Analyses using such longitudinal files are generally based on groups of students tracked from the point in time, and this cohort-based design is also typical of many *Breaking Through* programs. Useful summary performance indicators that can be compiled and reported for each starting cohort using longitudinal files include:

- Percent completing degrees or certificates within X years;
- Percent retained to next term;
- Percent retained to next year (and multiple years, where appropriate); and
- Percent achieving key project milestones within X years.

Once again, these will be most informative if broken down by appropriate student descriptors.

Performance indicators like these are useful for both informing project improvement (formative evaluation) and for documenting project impact (summative evaluation); both kinds of evaluation are important. Analyses intended for formative purposes might include examining the effectiveness of particular courses or course-taking patterns on student progress for different kinds of students, looking at whether student participation in particular program-related activities makes a difference in student success, or investigating the relationships between performance in prior courses and those taken later on.[15] All of these would involve creating the needed performance measures through recodes and calculations using data elements available in the project's analytical files, and then breaking these measures down by "treatment" variables representing presumed causes of these results derived in a similar fashion. As these examples illustrate, the heart of any analysis involves comparison: results mean little in themselves without something to compare them to—either across different student populations or against previously established benchmarks or comparison groups.

The notion of comparison becomes even more important when the purpose is summative—to indicate whether the project makes a substantive difference for the students who participated in it. Important here are two types of comparisons regarding outcomes for students who participated in the *Breaking Through* project:

- Comparisons to similar students who did not participate in *Breaking Through*; and
- Comparisons with external benchmarks at other community colleges or on a national basis.

Comparing outcomes for students who participated in the Breaking Through project and for similar students who did not: Identifying the needed comparison ("control") groups can be a challenge because it is impossible under

real-world conditions to randomly assign students to participate or not to do so in a manner that satisfies the rigorous conditions of educational research. Instead, naturally occurring groups of students must be identified, which are as similar as possible to student participants.[16] For cohort-based programs, for example, a comparison group might be students seeking the same credential as *Breaking Through* participants but who were in the program and who started at the same time. For non-cohort programs involving enhanced student support, the comparison group might simply be students who did not use this support. However, both of these examples are susceptible to misunderstanding because the "treatment" and "control" groups may not be the same.

A first step here is to document any known way that they differ by examining their respective characteristics. If this comparison reveals major differences, a second step might be to match cases drawn from the comparison group with those in the "treatment" group to yield an equal number of comparison students that have a similar demographic profile to that of *Breaking Through* participants. Of course, neither approach can overcome the fact that there may be unobservable differences in attitude or motivation between the two groups. For example, if participation in *Breaking Through* is voluntary, as it generally is in projects involving enhanced student support, those who participate will probably be more engaged simply because they decided to do so. This does not negate the comparison, but it does demand some caution in interpretation.

Another thing to watch out for in making comparisons is the fact that observed differences in performances between two populations may be due to a factor that is not under investigation but is related to a factor that is. For example, an analysis may find that *Breaking Through* students have a higher GPA than non-*Breaking Through* students seeking the same credential, suggesting that participation in the program relates to better performance. But further analysis may reveal that *Breaking Through* students are twice as likely to be female than those not participating in the program, and that women in general earn higher GPAs than men. Sorting this out may require an analysis that makes the GPA comparison between program participants and non-participants separate for men and women.[17]

Comparing Breaking Through outcomes with external benchmarks at other community colleges or on a national basis: One reason for developing the "Common Core Data Elements" for *Breaking Through* (see below) was to enable projects to benchmark performance among themselves. Additionally, a *Breaking Through* evaluation team will compile common outcomes statistics toward the end of the initiative, using data resources supplied by participating colleges. Another source of comparison for the project is national benchmark statistics about community colleges available through such sources as the Community College Research Center at Teachers College, Columbia University; the American Association of Community Colleges; or the Community College Survey of Student Engagement.[18] The Appendix of this report summarizes a range of such benchmarking statistics for community colleges.

Another point about benchmarks is the need for caution in interpretation. For example, observed differences in performance between two groups of students may not be meaningful if small numbers are involved; it is important to ensure that such differences are statistically significant. Confidence in results also builds when the same outcome occurs consistently for the same population over time and when the outcome is confirmed by multiple sources. Finally, benchmark statistics communicate best when presented in graphic form, which makes the underlying story about impact hard to miss.[19]

Common Core Data Elements from *Breaking Through*

This chapter proposes a set of common core data elements that all colleges seeking to track the progress of pre-college students should consider using. These were originally developed as recommendations to colleges in the *Breaking Through* initiative, but they are valid for a wide range of community colleges seeking to benchmark their own performance. These data elements represent the core of a much larger set of data that *Breaking Through* colleges collect, based on a survey conducted in 2006. This latter set of data elements is presented in Appendix 2 as a "Data Map" describing each category and illustrating what each college collects for its participating students. While the data map suggests what colleges could potentially collect, the common core presented here represents a strong recommendation for what colleges should collect.

It will be useful for research and evaluation purposes to define a "common core" of data elements that all *Breaking Through* Leadership Colleges should collect in compatible ways. This would enable the evaluation team to conduct consistent analyses across participating colleges on student persistence and performance, broken down by important subgroups of students—a critical ingredient in documenting the project's impact.

"Compatible" does not mean "identical." For the most part, what is needed is reasonably common definitions for common core data elements and a coding structure that enables categorization at a fairly high level of generality. It does not necessarily mean that identical codes must used. For example, all colleges collect data on student race/ethnicity in order to report these to state offices and the U.S. Department of Education. But some colleges collect race/ethnicity categories that exceed the level of detail required for state/federal reporting and different student registration systems code this data element in different ways. All that is required to achieve "compatibility" is that the categories for race/ethnicity that the college uses can be concatenated to yield the six federal reporting categories and that each category be uniquely coded.

This chapter proposes a set of common core data elements for *Breaking Through*. For each data element, information is provided about a) the basic definition of the data element; b) the source of the definition; and c) the minimum set of categories or codes that should be maintained. Note that all that is required for the latter is that the coding structure and categories that the college currently uses for any given data element can be mapped into the categories shown.

Student Identification Number. A locally assigned number or code that is used to uniquely identify each record for each student enrolled. This is typically the Social Security Number, but it may be any identifier used by the student registration system. This is for local tracking and linking purposes only. All data files used for research will be de-identified or cases will be assigned a different project number.

Gender. The gender of the student. These data are defined in the Integrated Postsecondary Education Data System (IPEDS).[20]

- Male
- Female

Race/Ethnicity. Categories used to describe groups to which individuals belong, identify with, or belong in the eyes of the community. The categories do not demote scientific definitions of anthropological origins. A person may be counted in only one group (IPEDS).

- White, Non-Hispanic
- Black, Non-Hispanic
- Hispanic
- Asian or Pacific Islander
- American Indian or Alaska Native
- Other

Date of Birth/Age. The date of birth of the student or the student's current age in years. The actual year of birth or the current age in years is carried as the code.

Employed. Describes the current employment situation of the student).

- Employed Full-Time (35 hours/week or more)
- Employed Part-Time (less than 35 hours/week)
- Not Employed

Reasons for Enrollment. The primary reason a student reports for attending college (Local or State Definition). The student can select more than one category. The categories need not entirely match the minimum set below.

- Employment
- Earn a Degree or Certificate
- Transfer to Another Institution
- Personal Enrichment/Development

High School Diploma. The type of award granted to the student on completion of a high school curriculum (Local or State Definition).

- High School Graduate
- GED (if not a separate data element)
- Not a High School Graduate

GED. Indicates whether a student has earned a GED award. Used only if this data is not included in High School Diploma (Local or State Definition).

- Earned GED
- Did not earn GED

High School Credential Date. Year of the receipt of the student's high school diploma or equivalent (Local or State Definition). If the student has no high school credential, may be used to record the year of last attendance at high school.

Years of Education Completed. The level of past educational experience obtained prior to enrollment at the college (Local or State Definition).

- Below High School
- Some High School
- High School Graduate
- Some College

Previous College Attendance. Indicates that the student previously attended a postsecondary institution (Local or State Definition).

- No previous postsecondary attendance
- Previously attended another postsecondary institution

Credits/Units Enrolled For. The total number of student credit hours a student is enrolled for in the term of record as of the designated census date for state reporting (State Definition). This can be either semester or quarter hours, so long as this is indicated. Similarly, this can be clock hours, so long as this is indicated. All data will be aggregated for research purposes to a common academic year and converted to semester credit hour equivalents.

Credits/Units Earned. The total number of student credit hours attempted by the student for the term of record for which a passing grade was received (State Definition). This can be either semester or quarter hours, so long as this is indicated. Similarly, this can be clock hours, so long as this is indicated. All data will be aggregated for research purposes to a common academic year and converted to semester credit hour equivalents.

GPA. The cumulative grade-point average of the student as officially recorded at the end of the term of record (Local or State Definition). This can be in any format so long as it can be reasonably equated to a standard 0 to 4.0 scale.

Program. The academic program in which the student is enrolled as of the term of record (IPEDS). The IPEDS Classification of Instructional Programs code is preferred.

Degree/Certificate. The type of award attained, if any, during the term of record (IPEDS).

- None
- Postsecondary Award or Certificate (less than one year)
- Postsecondary Award of Certificate (more than one year)
- Academic Associate's Degree (AA or AS)
- Applied Associate's Degree (AAS or equivalent)

Assessment Information. The student's performance on the college's standard placement examination for basic skills (Local Definition). Separate data elements should be maintained for reading, writing, and mathematics. Separate data elements should be maintained for each time the student takes the examination. The actual score may be carried and the college should indicate what test was used. The objective is to reduce the data for each of the three areas to:

- Placement at College Level
- Placement Below College Level (standard remedial level)
- Placement Below Remedial Level

Subsequent Employment. The employment status and field of employment of program completers or former students at the college (Local Definition).

- Employed Full-Time in Field of Training
- Employed Part-Time in Field of Training
- Employed Not in Field of Training
- Not Employed

Conclusion

Basing decisions on sound data is an axiom of *Breaking Through*. While it is not the intent of this document to turn project staff into full-time data compilers and methodologists, it will provide them with guidance on using data to improve their on-the-ground work, provide evidence of impact—and, ultimately, improve outcomes for low-skilled adults who are seeking entrance to family-supporting careers. The footnotes and links provided point to good treatments of relevant topics, and there are many on community college campuses (most prominently in institutional research departments) who already know a lot about these resources. Identify and cultivate these knowledgeable people.

And above all, project staff should develop a sensitivity to evidence and the need for it in managing their projects. Look constantly at what works and what doesn't—based on the facts—and be willing to act promptly to fix what you find.

Appendix 1: Community College Benchmark Statistics

Community colleges associated with *Breaking Through* frequently collect statistics on various aspects of student progress and performance. Sometimes, they are pleased with what they find, and sometimes they are disappointed. But it is often difficult to determine whether results are good or bad in the absence of comparative information. For such purposes, this appendix presents a number of common community college benchmark statistics. Some of these are national statistics, compiled for all community colleges. Some are state- or system-level statistics from states that have particularly sound data systems. Comparing local results with these benchmark statistics can help project leaders put their own performance in context.

Percent of first-time, full-time students completing an Associate's degree within three years

In 2005, 29.3 percent of first-time, full-time fall entering students graduate within three years with an Associate's degree.

Source: National Center for Education Statistics (NCES), Integrated Postsecondary Education Data System (IPEDS) Graduation Rate Survey

Percent of community college students completing an Associate's degree within periods of time longer than three years

39 percent of students who first enrolled in a community college had attained a credential (certificate, Associate's degree, or Bachelor's degree) within six years. An additional 12 percent had not attained a credential but had transferred to a four-year institution.

Source: NCES, Beginning Postsecondary Students (BPS) Longitudinal Survey [http://nces.ed.gov/surveys/bps/]

50 percent of students who first enrolled in a community college had attained a credential (certificate, Paulson) within six to eight years. An additional 13 percent had not attained a credential but had transferred to a four-year institution.

Source: NCES, National Education Longitudinal Survey (NELS) [http://nces.ed.gov/surveys/nels88/]

Percent of first-time, full-time, students enrolled the following year

In 2004, 58.3 percent of first-time, full-time students enrolled in public two-year colleges were enrolled the next year.

Source: NCES, IPEDS Enrollment Survey (See \"Special Analyses\" for detailed spreadsheets)

In 2002, 54.8 percent of students enrolled in the fall who were first-time, full-time freshmen in two-year colleges returned the following fall semester.

Source: National Center for Public Policy and Higher Education, ACT \"Institutional Data Questionnaire\", 2002

Nationwide, the retention rate of two-year college students, based on how many entering

students in a fall semester enroll in the following fall semester, is 54.8 percent. Of the states, Florida has the highest two-year retention rate at 70.2 percent; Oregon has the lowest at 42.8 percent.

Source: ACT, "Institutional Data Questionnaire," unpublished analysis, 2002

Percent of entering students requiring remediation in reading, writing, and math

61.1 percent of new entrants to community colleges in the national longitudinal studies enrolled for at least one remedial course.

Source: Adelman, Clifford. 2005. Moving into Town—And Moving On. *Washington, DC: U.S. Department of Education*

70 percent of students entering community college in California place in remedial math and 42 percent place in remedial English.

Source: Research and Planning Group for California Community Colleges (2005); Environmental Scan: A Summary of Key Issues Facing California Community Colleges Pertinent to the Strategic Planning Process

Percent of students placed in remedial classes who successfully complete these classes

About two-thirds of the underprepared students enrolled in remedial courses complete their preparatory studies in less than a year.

Source: Phillip R. Day & Robert M. McCabe. 1997. Remedial Education: A Social and Economic Imperative. *Washington, DC: American Association of Community Colleges*

Percent of students completing remediation (in reading, writing, and math) who successfully complete at least one college-level class

91 percent of the students who complete developmental writing succeed in freshman composition.

Source: Boylan, Hunter & Barbara Bonham (1992) as cited in Day and McCabe (1997)

83 percent of those who complete remedial reading succeed in their initial social science courses.

Source: Boylan & Bonham (1992)

77 percent of those who complete developmental math succeed in college math.

Source: Boylan & Bonham (1992)

A Maryland study of post-remediation pass rates in college-level courses found that 80 percent of community college students who passed remedial mathematics also passed their first college-level mathematics course. The study also found that 81 percent of those who passed remedial English also passed college-level English (Maryland Higher Education Commission, 1996).

Texas community colleges found that those who successfully completed remediation were very likely to pass college-level courses in the same subject area. Between 75 percent and 85 percent of those who passed remedial courses in English or mathematics and took college-level English or mathematics within the next year also passed their first college-level courses in these subjects.

Source: As cited by Boylan, Hunter R. & D. Patrick Saxon, National Center for Developmental Education. A Paper prepared for The League for Innovation in the Community College. [www.ncde.appstate.edu/reserve_reading/Outcomes_of_Remediation.htm]

Percent of students completing remediation achieving key performance benchmarks

In the Florida Community College System, of the fall 1997 cohort that needed any amount of remediation, after five years:

17 percent had earned an award

11.8 percent had transferred

27.7 percent were still enrolled

Source: Florida Department of Education. Student Success, Data Trend (#27). March 2005

Percent of community college students enrolling who successfully complete College Algebra

For the Florida Community College System, 38 percent of students pass college algebra after completing remedial math; 47 percent pass college algebra among students who did not need remedial math.

Source: Florida State Board of Community Colleges; Grades in Selected first Year Courses by ELT Status: Data Trend 2005 (based on 1997 data)

46 percent of California community college students successfully completed college algebra in 2001.

Source: www.algebrapathways.org/history.asp

Percent of community college students who successfully complete English Composition

For the Florida Community College System, 56 percent of students pass English composition after completing remedial writing. 64 percent pass English composition among students who did not need remedial writing.

Source: Florida State Board of Community Colleges; Grades in Selected first Year Courses by ELT Status: Data Trend 2005 (Based on 1997 data)

Percent of GED recipients returning to the college to register for college-level work

30 to 35 percent of GED holders obtain any postsecondary education and only 5 to 10 percent complete an entire year of college courses.

Source: Tyler, 2002: The Economic Benefits of the GED: Lessons from Recent Research Brown University and National Bureau of Economic Research, July 2002.

Percent of ABE completers returning to the college to register for college-level work

31 percent of the students who started in ABE or GED courses at community colleges in the state of Washington went on to enroll in at least one college-level course.

Source: Research Report No. 06-2. Washington State Board for Community and Technical Colleges (April 2005)

Percent of ESL completers returning to the college to register for college-level work

12 percent of students at community colleges in the state of Washington who began as ESL student went beyond ESL to enroll for college-credit courses

Source: Research Report No. 06-2. Washington State Board for Community and Technical Colleges (April 2005)

52 percent of basic skills (GED/ABE/ESL) students at community colleges in the state of Washington leave after one quarter of attendance. About 16 percent attend for three or more quarters over a two-year period. The rest (32 percent) attend two or three quarters.

Source: *Research Report, No. 01-2. Washington State Board for Community and Technical Colleges. November 2001*

Demographics

Percentage of community college credit enrollment:

Female

57 percent of community college students were women in fall 1997.

Source: American Association of Community Colleges, Fact Sheet, 2004

In each race/ethnic category

White
64.8 percent

Hispanic
11.8 percent

Black
11.1 percent

Asian/Pacific Islander
5.8 percent

Nonresident (alien)
3.8 percent

Race/ethnicity unknown
1.5 percent

Native American
1.3 percent

Source: American Association of Community Colleges, Fact Sheet, 2004; data from fall 1997

25 years old or older

46 percent of community college students are at least 25 years old.

Source: AACC, National Profile of Community Colleges: Trends & Statistics, 3rd Edition. 1999

42 percent of community college students were at least 25 years old in 1990

39 percent of community college students were at least 25 years old in 2000

Note: The number of adult learners is not going down, as these figures might suggest. Instead, there have been substantial increases in the under 25 age group attending community colleges.

Source: U.S. Department of Education (See ACE publication)

Entered directly out of high school (within one year)

In 2004, 55.7 percent of high school graduates went directly to college. (This is for either a four-year or two year institution.)

Source: NCES; Common Core Data, Private High Schools Survey, Fall Residency and Migration Survey; (additional data provided by KY, TN, and UT)

34 percent of the incoming freshman at community colleges in the state of Washington entered directly out of high school.

Source: Research Report No 04-2, Washington State Board for Community Colleges, December 2004

Attending part time

63 percent of community college students attend college on a part-time basis.

Source: AACC, National Profile of Community Colleges: Trends & Statistics, 3rd Edition

First generation

52 percent of community college students are in the first generation of their families to attend college.

Source: AACC, National Profile of Community Colleges: Trends & Statistics, 3rd Edition

Receives financial aid

The percentage of community college students receiving various forms of financial aid in 2006 was:

Any aid	47 percent
Federal grants	23 percent
Federal loans	11 percent
State aid	12 percent

Source: AACC, Fast Facts; data from January 2006 www.aacc.nche.edu/Content/NavigationMenu/AboutCommunityColleges/Fast_Facts1/Fast_Facts.htm

Working full or part time

80 percent of community college students worked either full or part time in 1996.

Source: AACC, National Profile of Community Colleges: Trends & Statistics, 3rd Edition

57 percent of community college students responding to the CCSSE worked more than 20 hours per week in 2006.

Source: CCSSE Website [www.ccsse.org]

Appendix 2: *Data Map* for *Breaking Through* Leadership Colleges

This *Data Map* provides a listing of all the data elements collected by the Leadership Colleges in *Breaking Through*. These data elements are organized according to various topics (e.g., demographics) and include most of the data elements collected by any participating campus. The listing is fairly exhaustive, because it is intended to illustrate the various kinds of data that *Breaking Through* participants currently collect. Colleges elect which of these data elements they will collect. In contrast, the "Common Core" suggests a set of data elements for *all Breaking Through* Leadership Colleges to collect.

Table 1, on pages 24–25, lists the data elements collected by the Leadership Colleges in *Breaking Through*. These data elements are organized by topical heading and include most of the data elements collected by any participating campus. The listing is fairly exhaustive: it is intended to illustrate the various kinds of data that participants currently collect. However, except for elements in the common core, participating colleges can elect to collect or not to collect any of these data elements. Common core data elements are flagged with an asterisk in the chart.

Student Characteristics

These data elements address characteristics of individual students. Their primary purpose is to provide aggregate descriptive information about program participants and to disaggregate outcomes and treatment information for different groups. Evaluations of *Breaking Through* may use these factors as control variables in multivariate studies of program impact.

Data elements under this heading include:

Student Identification Number. This is the tracking number that each college or program uses to identify each student record. For *Breaking Through*, it is used only as an identifier, kept secure, and not shared. Many campuses use the Social Security Number for this purpose, but any unique identifier will work. This data element is typically supplied through the college's student registration system.

Gender. This data element is typically supplied through the college's student registration system.

Race/Ethnicity. This data element is typically supplied through the college's student registration system.

Age or Date of Birth. If age is not calculated, it can be obtained through date of birth. For research purposes, ages are generally categorized into ranges, but it is best to maintain the actual number from which age is derived so that different "cut points" are possible. This data element is typically supplied through the college's student registration system.

Citizenship. The current citizenship or visa status of the student is typically supplied through the college's student registration system.

Address. Most programs maintain full address records so that students can be easily contacted. For research purposes, only ZIP codes are required because they can be mapped to different districts, regions, or Census tracts to obtain information on such topics as commute distances, typical income and housing, and access to social services. This data element is typically supplied through the college's student registration system.

Married. The marital status of the student is not usually supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Children at Home. The number of school-age children that the student is responsible for caring for in her or his household. This information is a valuable indicator of the additional responsibilities that the student must meet as she or he engages in further education. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Native Language. The language spoken by the student at home. This information is useful as an indicator of the challenges involved in providing instruction. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Employed. This indicates whether or not the student is employed. It is usually maintained in terms of full-time or part-time employment. This is useful as an indicator of the additional responsibilities that the student needs to meet as she or he engages in further education. For *Breaking Through*, it is also useful as an outcome variable because a principal purpose of the initiative is to improve current employment. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Employer. The actual employer of the student. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

How Long at This Job. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Income. The current income of the student. Tracking income can be particularly useful for impact studies. Income is usually collected in terms of specified dollar ranges. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Looking for Job. This indicates whether or not the student is actively seeking a new or first job. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Previous Job. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Public Assistance Participation. This indicates whether or not the student is participating in welfare, Food Stamps, or other forms of public assistance. Programs collecting this information typically specify which of many forms of public assistance the student is participating in. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Text continues on page 26.

Table 1.
Data Element Coverage for Breaking Through Leadership Campuses

Data Elements	Community College of Denver	Cuyahoga Community College	Southeast Arkansas Community College	Durham Technical Community College	Central New Mexico Community College	Owensboro Community and Technical College	Portland Community College
Demographics							
Student ID*	X	X	X	X	X	X	X
Gender*	X	X	X	X	X	X	X
Race/Ethnicity*	X	X	X	X	X	X	X
Age/DOB*	X	X	X	X	X	X	X
Citizenship	X	X	X	X	X	X	X
Address (Zip)	X	X	X	X	X	X	X
Married	X				X		X
Children at Home	X				X		X
Native Language	X			X	X		
Employed*	X	X	X	X			X
Employer			X		X		
How Long at this Job?	X		X		X		
Income				X	X		
Looking for Job	X			X			
Previous Job			X				
Public Assistance Participation	X			X			
Reason for Enrollment*		X	X	X	X		X
Career Goals	X			X			X
Educational Plans	X			X			
First Generation	X	X			X		X
Case/Program Management							
Source of Information About Program	X			X			
Daycare Available	X				X		X
Daycare Needed?			X	X			
Reliable Transportation?	X		X		X		X
Drivers License?				X	X		
Handicap/Disability	X			X	X		
Health Problems?	X						
Drug/Alcohol Problems?	X						
Services Needed	X			X	X		
Referrals to Services	X			X			X
Free Comment Fields	X		X				
Academic Background							
High School Diploma*	X	X	X	X	X	X	X
GED*	X	X	X	X	X	X	X
HS Credential Date*	X	X	X	X	X	X	X
High School Attended	X	X	X	X	X	X	X
High School Withdrawal				X			

Data Elements	Community College of Denver	Cuyahoga Community College	Southeast Arkansas Community College	Durham Technical Community College	Central New Mexico Community College	Owensboro Community and Technical College	Portland Community College
Reason for HS Withdrawal				X			
Years of Education Completed*	X	X		X	X		X
Previous College Attendance*	X	X	X		X	X	
Institution of Previous Attendance	X	X	X		X	X	
Enrollment							
Credits/Units Enrolled For*	X	X	X	X	X	X	X
Credits/Units Earned*	X	X	X	X	X	X	X
GPA*	X	X	X	X	X	X	X
Program Application	X	X	X	X	X	X	X
Program*	X	X	X	X	X	X	X
Specific Classes Taken	X	X	X	X	X	X	X
Grades for Specific Classes	X	X	X	X	X	X	X
Test Results for Specific Classes						X	
FAFSA		X	X		X		
Financial Aid Detail	X	X	X	X	X	X	X
Employer Support?	X		X			X	X
TRIO/Other Federal Support		X	X	X	X	X	X
Career Pathways	X		X				
College Readiness Classes				X			X
Learning Community/Group							
Tutoring	X		X		X		X
Career Services			X		X		
Counseling			X		X		
Degree/Certificate*	X	X	X	X	X	X	X
Assessment Information*							
TABE	X	X	X	X	X?	X	
COMPASS		X	X	X		X	X
ASSET				X		X	
ACCUPLACER	X				X		
Work Keys					X		
GED (Practice Test)				X	X	X	
Skills Goals Inventory				X	X		
Self-Assessment of Skills				X			
Career/Interest Inventory		X		X			
Learning Styles Inventory		X		X			
Student Rating of Experience	X	X		X			
Resume/Self Evaluation				X			
Outcomes Information							
Licensure Test Scores	X		X				
Employment*	X	X	X	X	X		
Income		X	X		X		
Societal Achievements		X		X			

Reason for Enrollment. This indicates the primary reason why the student is enrolled, and is usually collected officially by the college via an application form. This data element is typically supplied through the college's student registration system.

Career Goals. This indicates specific job or employment-related goals that the student is seeking through enrollment. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Educational Plans. This indicates how far the student would like to progress with respect to further education. It is usually expressed in terms of degree levels. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

First Generation. This indicates whether or not the student's parents attended college or were college graduates. It is usually maintained separately for each parent. The aggregate measure "first generation" is usually assigned if neither of the student's parents attended college. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Case/Program Management

Data elements included here are used by program administrators to advise individual students, to maintain contact with students, to trigger interventions, or to track progress. Unlike data elements used for research or evaluation purposes, most data elements under this heading are used to provide information about individual decisions, transactions, or interventions. This means that they do not have to be coded in common, because they will rarely be aggregated. They are usually maintained and updated independently from the student registration system and kept in a separate project management system. None of these elements are part of the proposed common core for *Breaking Through*.

Source of Information About Program. This indicates how the student found out about the *Breaking Through* program and is kept by program managers to monitor which sources of information appear to be the most productive. Formats vary, but multiple sources are usually listed, customized to the program or campus. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Day Care Available. This indicates whether the student currently has access to day care and is used for intervention purposes to determine if services are needed. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Day Care Needed. This indicates whether or not the student needs day care services. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Reliable Transportation. This indicates whether the student currently has a reliable way to attend classes through a car, ride, or the use of public transportation. Some programs collect the actual means of transportation, while others simply treat this as a yes/no variable. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Drivers License. This is collected for similar purposes to the above, although it also indicates that the student has a means to docu-

ment identity. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program application.

Handicap/Disability. This indicates any of a number of potential physical or learning disabilities that could hinder the student from completing her or his studies. Most colleges collect information about a specific range of disabilities that are then combined to yield a summary disability flag. Many disabilities are maintained on the college's student registration system, but projects frequently collect information on additional learning or emotional disabilities through a special questionnaire or program application.

Health Problems. As above, these are usually specific questions about the student's current health status. They are collected by a program through a special questionnaire or interview.

Drug or Alcohol Problems. As above, these are usually specific questions about substance abuse that are collected by the program through a special questionnaire or interview.

Services Needed. This provides a detailed listing of the kinds of support services that the student may need to enroll and succeed in college. The information is used for case management. Individual services are listed by type, and each project has its own array. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or interview.

Referrals to Services. As above, this refers to the actual referrals to services that a program makes for individual students. This data element is not typically supplied through the college's student registration system. Typically, it is drawn from project records.

Free Comment Fields. Projects sometimes find it useful to maintain a number of free comment fields as data elements in project management databases. These are entirely unstructured and can be used to note unanticipated critical events that may affect student retention or performance, such as illness, losing a job, a death in the family, or emotional distress. Project administrators provide comments in these fields to indicate that an event has occurred and the nature of that event. Although such data cannot be aggregated directly, they can be useful in longitudinal research studies on the factors affecting student success.

Academic Background

These data elements provide information about each student's previous educational experiences and attainments. They are most frequently used for project and case management purposes, but they are also important group or control variables for research and evaluation.

High School Diploma. This indicates whether a student has graduated from high school or has otherwise earned a high school credential. Some colleges carry GED as an entry in this field. This data element is typically supplied through the college's student registration system.

GED. This indicates that the student has earned a GED credential, as above. Some colleges maintain this as a separate data element, but this is not necessary as long as the data are recorded somewhere. This data element is typically supplied through the college's student registration system.

High School Credential Date. This indicates the date associated with the high school credential. For students who do not have a high school credential, it is often used to record the last date of high school attendance. This information is useful for research purposes in indicating how long it has been since the student was engaged in educational practice, and therefore how much basic skills reinforcement may be needed. This data

element is typically supplied through the college's student registration system.

High School Attended. This indicates the actual name or number of the high school the student attended. This data element is typically supplied through the college's student registration system.

High School Withdrawal. Some programs want to know whether or not the student formally withdrew from high school without having earned a degree. This information is typically collected through a special questionnaire or interview.

Reason for High School Withdrawal. As above, this indicates the reason why the student withdrew from high school. This information is typically collected through a special questionnaire or interview.

Years of Education Completed. This indicates the student's previous academic attainment, usually reported in terms of degrees or grade levels. Most colleges collect this on their admissions forms, so the data element is typically supplied through the college's student registration system. Some programs include it on a special questionnaire.

Previous College Attendance. This indicates whether the student has previously attended a postsecondary institution. It is typically included on the application form, so the data element is supplied through the student registration system.

Institution of Previous Attendance. This indicates the institution or institutions associated with the above. It is typically included on the application form, so the data element is supplied through the student registration system.

Enrollment

These data elements provide information on the details of attendance at the college, including academic activity and participation in services or support activities. This information is important for project and case management. It enables program managers to track progress, advise individual students, and manage interventions. It is equally important for research and evaluation because these data elements can be used to construct a range of indicators of student success. For the latter, it is important to have separate sets of records for all of these data elements for each term that the student is enrolled. This makes it possible to construct longitudinal data files.[21]

Credits/Units Enrolled For. This indicates the student's academic load in terms of credits or clock hours. For non-credit programs, the same data element can be used to record participation in numbers of classes or activities. This information is usually collected at the college's official "census date," the time at which enrollment information is collected for federal and state reporting. This data element is typically supplied through the college's student registration system.

Credits/Units Earned. This is typically reported as of the end of the term or enrollment period. It is typically supplied through the college's student registration system.

GPA. This is the official current cumulative Grade Point Average for the student at the point at which the data are collected. This data element is typically supplied through the college's student registration system.

Program Application. This indicates whether the student has officially applied for admission to an academic program at the college. This information can be useful as a success indicator for programs such as nursing. This data element is typically supplied through the college's student registration system.

Program. This indicates the program of study in which the student is enrolled at the college. This data element is typically supplied through the college's student registration system.

Specific Classes Taken. All colleges have detailed records of student enrollment in each formal class, but not all of these records are required for *Breaking Through* research and evaluation. Most commonly, project teams want to know class details for each remedial class the student takes (especially if it is contextualized in the course of the project), "gatekeeper" classes (e.g., college algebra, other early math courses, and English composition), and key vocational coursework required by the program. These data are typically supplied through the college's student registration system.

Grades for Specific Classes. As above.

Test Results for Specific Classes. Some classes have particular assessments or tests associated with them, and they are useful for project management or research purposes. Because they may include external examinations not administered by the college, these data may or may not be included in the student registration system.

Free Application for Federal Student Aid. This indicates that the student has completed a FAFSA form and has therefore applied for financial aid. This data element is typically supplied through the college's student registration system.

Financial Aid Detail. These data elements include information about the specific form of aid received by the student. Sometimes a separate data element is included for each individual source of aid, and sometimes flags are provided for categories of aid (e.g., grant, scholarship, loan). Although aid information is carried by student registration systems, access to it is limited, so any such data must be obtained through the office of institutional research or the financial aid office.

Employer Support. This indicates whether the student is receiving tuition assistance or other support from her or his employer. It is typically not included in the college's student registration system. It is supplied by the student.

TRIO/Other Federal Support. This indicates whether the student is officially designated as a participant in the federally supported TRIO program or a similar program that is supported through federal dollars. Each college participating in these programs designates the applicable students, so in most cases this information must be obtained from the program administrator.

Career Pathways. Some states maintain explicit career pathways programs, supported with state dollars and offered by community colleges. Each college participating in such a program designates the applicable students, so the information must be obtained from the program administrator.

Learning Community/Group Learning. This indicates whether the student is a participant in a learning community or other collaborative learning experience organized by the program. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or program records.

Tutoring. This indicates whether or not (and sometimes how often) the student has been served through tutoring support. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or tutoring center records.

Career Services. This indicates whether or not (and sometimes how often) the student has been served through the college's career services. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or career services records.

Counseling. This indicates whether or not (and sometimes how often) the student has been served through counseling support. This data element is not typically supplied through the college's student registration system. Typically, it is collected through a special questionnaire or counseling center records.

Degree/Certificate. This indicates that the student has earned a degree or certificate from the college during the term, together with the level and field of study of this credential. It is used as a measure of success for the program. This data element is typically supplied through the college's student registration system.

Assessment Information

These data elements vary based on the actual examinations and assessments that a college uses to place and otherwise evaluate students. Most of these examinations are intended to test basic skills in order to determine if remediation is needed. But some are outcomes or special purpose tests associated with vocational education, while some are surveys or inventories that the local project managers have deemed useful. These scores are valuable for project and case management, and they can also be used as control variables in research and evaluation.

Test of Adult Basic Education. The TABE is a measure of basic skills that is used in Adult Basic Education programs. Because many *Breaking Through* projects have an ABE component, scores on the TABE are frequently collected and maintained. This data element may or may not be included in the college's student registration system. If it is, it should be obtained from there. In some cases, test scores are maintained as a separate set of records kept by the college's testing center and must be obtained from there.

COMPASS/ASSET/ACCUPLACER. These are basic skills test batteries, and they include separate tests (and scores) for reading, writing, and math. These data elements may or may not be included in the college's student registration system. If they are, they should be obtained from there. In some cases, test scores are maintained as a separate set of records kept by the college's testing center and must be obtained from there.

WorkKeys. This work-related assessment, offered by ACT, is designed especially to examine skills associated with the workplace. Assessments are provided in many areas, but the most commonly used are reading for information, applied mathematics, business writing, and locating information—each of which is scored separately. Several *Breaking Through* projects are using WorkKeys as an assessment tool. This data element may or may not be included in the college's student registration system. If it is, it should be obtained from there. In some cases, test scores are maintained as a separate set of records kept by the college's testing center and must be obtained from there.

GED Practice Test. This is a test associated with preparation for GED; if earning a GED is a part of a college's project, these test scores are useful. This data element may or may not be included in the college's student registration system. If it is, it should be obtained from there. In some cases, test scores are maintained as a separate set of records kept by the college's testing center and must be obtained from there.

Inventories. Many projects have found it useful to have students complete one or more self-assessment instruments that tap interests or skills related to the project. Some of these allow students to critically examine what they can do and how well. Others provide them with the opportunity to explore career interest and opportunities. Still others ask them to examine how they learn best. Data elements relating to these inventories are almost never included in the college's student registration system. They are administered directly by the program.

Student Rating of Experience. This is a satisfaction item or set of items asking students to rate the college or program. In some cases, this is part of a general survey administered to all students by the office of institutional research or the student services office. In some *Breaking Through* projects, it is a special question or set of questions referring only to the project, asked by project administrators. In either case, the data must be obtained directly from those responsible for obtaining the rating.

Resume/Self-Evaluation. Some colleges or projects ask students to engage in a formal exercise that involves a detailed self-evaluation of skills and dispositions, together with the preparation of a resume that can be used to support job applications. Some projects keep records on this as part of their project management database.

Outcomes

Outcomes data elements are used in *Breaking Through* to record the ultimate attainments and successes that students experience through enrollment at the college and their participation in *Breaking Through*. They are especially useful for research and evaluation.

Licensure Test Scores. Some occupations that students are seeking either require or encourage students to be certified or licensed through examinations. Prominent examples include the health professions, auto mechanics, and various construction trades. Where these are the target of a college's project, data about performance are frequently maintained. These scores are typically not maintained as part of the college's student registration system; they may not even be collected by the college at all. In most cases, they are obtained directly from students.

Employment. This refers to employment after completing the program, although some colleges try to keep track of employment each term. Because it is used as a measure of program impact, employment in the field for which a participant trained is usually what is collected, including the actual job title and industry. These data can be obtained in a variety of ways; the most typical is a follow-up survey of program graduates and former students.[22]

Income. Because *Breaking Through* is intended to "add value" for the students participating, knowing something about subsequent student earnings, or additional earnings, is valuable information about outcomes. Such data can also be obtained in a variety of ways; the most typical is a follow-up survey.

Societal Achievements. Sometimes, outcomes for program participation include such things as community service, voting, or other forms of civic participation. These data are almost always obtained through a follow-up survey.

Endnotes

[1] Good background resources here are Knight 2003, and McLaughlin & Howard 2004.

[2] See Jones 1982 for an expanded discussion of different types of data in higher education.

[3] Paulson 2003 provides detailed instructions about how to conduct a data audit.

[4] See www.socialsolutions.com.

[5] See www.Amatrol.com.Products.

[6] See www.Blackboard.com.

[7] As noted earlier, a data audit is a useful technique for determining whether a particular set of data elements is already collected or maintained on campus and where it is located (see Paulson 2003).

[8] See Appendix A for examples of *Breaking Through* project intake forms or surveys.

[9] For a fuller treatment of survey design and construction, see Suskie 1996.

[10] Once again, staff in IR or in the testing center itself may be helpful in adapting a current survey or form so it can be scanned.

[11] See Ewell & Boeke 2007 for a description of these databases, what states they are located in, and their contents.

[12] IR staff may be helpful in setting up these reporting templates and routines the first time.

[13] "Relational" database programs store data in the form of multiple tables that are connected by common descriptive categories, allowing users to cross-reference information across tables.

[14] A basic resource here is Ewell, Parker, & Jones 1988.

[15] An excellent example of such an analysis for students beginning in Adult Basic Education and transitioning to vocational study in the state of Washington is Jenkins & Prince 2005.

[16] See Campbell & Stanley 1963 for the classic treatment of "quasi-experimental" designs in sociological research.

[17] Researchers typically use multivariate methods like multiple regression to investigate such effects, but it is unlikely that *Breaking Through* project staff will need to do this. The External Evaluation of the project will do so and, again, those interested are advised to seek help from Institutional Research.

[18] See http://ccrc.tc.columbia.edu, www.aacc.nche.edu, www.ccsse.org.

[19] A useful source on indicators and benchmarks is Ewell & Jones 1996.

[20] Definition provided by the Integrated Postsecondary Education Data System, maintained by the National Center for Education Statistics, U.S. Department of Education.

[21] See the *Data Toolkit* for information about the construction of longitudinal data files.

[22] See the *Data Toolkit* for information about various follow-up options.

References

Campbell, Donald T. & Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Ewell, Peter T. & Marianne Boeke. 2007. *Critical Connections: Liking States' Unit Record Systems to Track Student Progress*. Indianapolis, IN: Lumina Foundation for Education.

Ewell, Peter T. & Dennis P. Jones. 1996. *Indicators of Good Practice in Undergraduate Education*. Boulder, CO: National Center for Higher Education Management Systems.

Ewell, Peter T., Ronald Parker, & Dennis P. Jones. 1988. *Establishing a Longitudinal Student Tracking System: An Implementation Handbook*. Boulder, CO: National Center for Higher Education Management Systems.

Jenkins, Davis & David Prince. 2005. *Building Pathways to Success for Low Skill Adult Students: Lessons for Community College Policy and Practice from a Longitudinal Student Tracking Study*. New York: Community College Research Center, Teachers College, Columbia University.

Jones, Dennis P. 1982. *Data and Information for Executive Decisions in Higher Education*. Boulder, CO: National Center for Higher Education Management Systems.

Knight, William E. 2003. *The Primer for Institutional Research*. Tallahassee, FL: Association for Institutional Research.

McLaughlin, Gerald W. & Richard D. Howard. 2004. *People, Processes, and Managing Data, Second Edition*. Tallahassee, FL: Association for Institutional Research.

Paulson, Karen. 2003. *A Data Audit and Analysis Toolkit to Support the First College Year*. Boulder, CO: National Center for Higher Education Management Systems.

Suskie, Linda. 1996. *Questionnaire Survey Research: What Works, Second Edition*. Tallahassee, FL: Association for Institutional Research.

ABOUT THE AUTHOR

Peter T. Ewell is vice president of the National Center for Higher Education Management Systems, an independent research and development organization on higher education policy. His primary research areas are the assessment of student learning in higher education and improving student success in postsecondary settings. He has undertaken major projects in these areas with support from Lumina Foundation for Education, the Pew Charitable Trusts, and the W. K. Kellogg Foundation. He has consulted with 23 state systems of higher education and with over 400 colleges and universities.

ABOUT BREAKING THROUGH

Around the country, innovative community colleges are helping low-skilled adults gain the valuable skills and credentials that are the gateway to family-supporting careers. *Breaking Through*, a multi-year demonstration project, promotes and enhances the efforts of community colleges to help low-literacy adults prepare for and succeed in occupational and technical degree programs.

Jobs for the Future

Jobs for the Future seeks to accelerate the educational and economic advancement of youth and adults struggling in today's economy. JFF partners with leaders in education, business, government, and communities around the nation to: strengthen opportunities for youth to succeed in postsecondary learning and high-skill careers; increase opportunities for low-income individuals to move into family-supporting careers; and meet the group economic demand for knowledgeable and skilled workers.

National Council for Workforce Education

The National Council for Workforce Education is a private, nonprofit, professional organization committed to promoting excellence and growth in occupational education at the postsecondary level. NCWE, an affiliate council of the American Association of Community Colleges, provides a national forum for administrators and faculty in occupational, vocational, technical, and career education as well as representatives of business, labor, military, and government, to affect and direct the future role of two-year colleges in work-related education.

JOBS FOR THE FUTURE
CREATING STRATEGIES
for Educational and Economic Opportunity

88 Broad Street
Boston, MA 02110
www.jff.org

NCWE
National Council *for*
Workforce Education

1900 Kenny Road
Columbus, OH 43210
www.ncwe.org